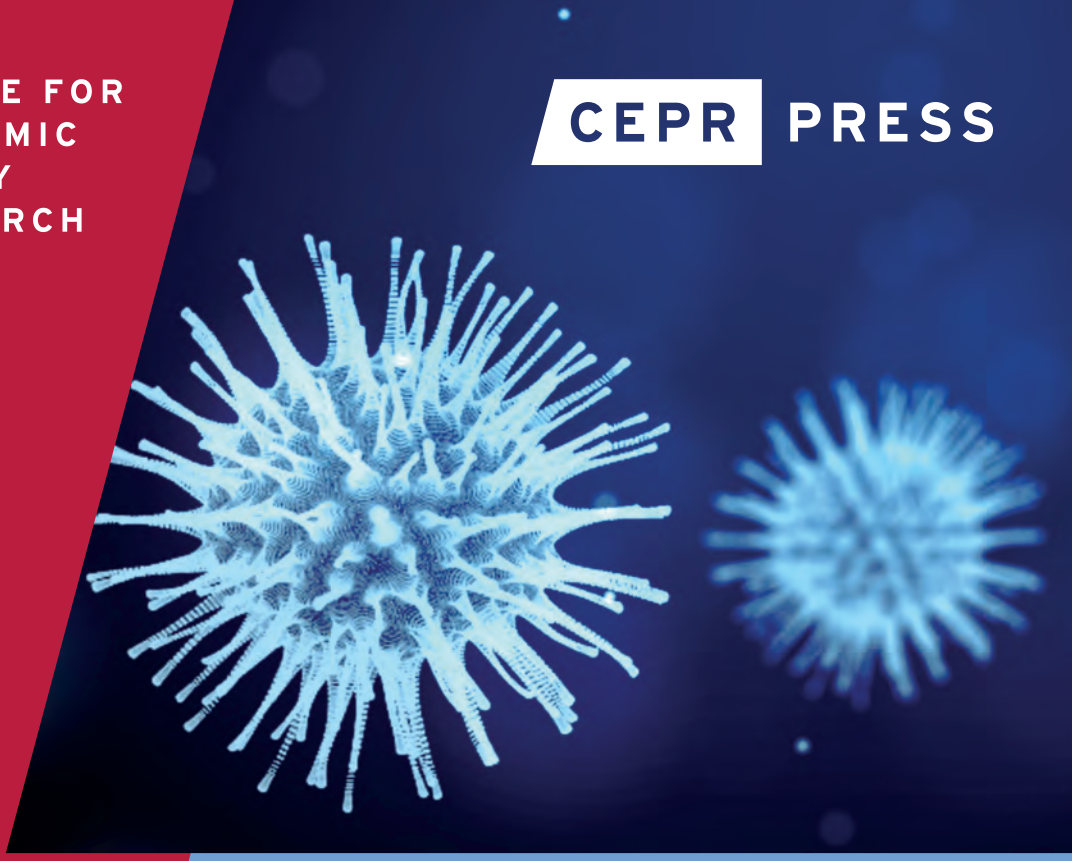
CENTRE FOR ECONOMIC POLICY RESEARCH

CEPR PRESS



COVID ECONOMICS

VETTED AND REAL-TIME PAPERS

ISSUE 30
19 JUNE 2020

Covid Economics
Vetted and Real-Time Papers

Covid Economics, Vetted and Real-Time Papers, from CEPR, brings together formal investigations on the economic issues emanating from the Covid outbreak, based on explicit theory and/or empirical evidence, to improve the knowledge base.

Founder: Beatrice Weder di Mauro, President of CEPR
Editor: Charles Wyplosz, Graduate Institute Geneva and CEPR

Contact: Submissions should be made at https://portal.cepr.org/call-papers-covid-economics. Other queries should be sent to covidecon@cepr.org.

The Centre for Economic Policy Research (CEPR)

The Centre for Economic Policy Research (CEPR) is a network of over 1,500 research economists based mostly in European universities. The Centre's goal is twofold: to promote world-class research, and to get the policy-relevant results into the hands of key decision-makers. CEPR's guiding principle is 'Research excellence with policy relevance'. A registered charity since it was founded in 1983, CEPR is independent of all public and private interest groups. It takes no institutional stand on economic policy matters and its core funding comes from its Institutional Members and sales of publications. Because it draws on such a large network of researchers, its output reflects a broad spectrum of individual viewpoints as well as perspectives drawn from civil society. CEPR research may include views on policy, but the Trustees of the Centre do not give prior review to its publications. The opinions expressed in this report are those of the authors and not those of CEPR.

Editorial Board

Ethics

Covid Economics will feature high quality analyses of economic aspects of the health crisis. However, the pandemic also raises a number of complex ethical issues. Economists tend to think about trade-offs, in this case lives vs. costs, patient selection at a time of scarcity, and more. In the spirit of academic freedom, neither the Editors of *Covid Economics* nor CEPR take a stand on these issues and therefore do not bear any responsibility for views expressed in the articles.

Submission to professional journals

The following journals have indicated that they will accept submissions of papers featured in *Covid Economics* because they are working papers. Most expect revised versions. This list will be updated regularly.

American Economic Review

American Economic Review, Applied Economics

American Economic Review, Insights

American Economic Review, Economic Policy

American Economic Review, Macroeconomics

American Economic Review, Microeconomics

American Journal of Health Economics

Canadian Journal of Economics

Economic Journal

Economics of Disasters and Climate Change

International Economic Review

Journal of Development Economics

*Journal of Econometrics**

Journal of Economic Growth

Journal of Economic Theory

*Journal of the European Economic Association**

Journal of Finance

Journal of Financial Economics

Journal of International Economics

*Journal of Labor Economics**

Journal of Monetary Economics

Journal of Public Economics

Journal of Political Economy

Journal of Population Economics

*Quarterly Journal of Economics**

Review of Economics and Statistics

*Review of Economic Studies**

Review of Financial Studies

(*) Must be a significantly revised and extended version of the paper featured in *Covid Economics*.

Covid Economics
Vetted and Real-Time Papers

Issue 30, 19 June 2020

Contents

Covid Economics Issue 30, 19 June 2020

Power laws in superspreading events: Evidence from coronavirus outbreaks and implications for SIR models[1]

Masao Fukui[2] and Chishio Furukawa[3]

Date submitted: 11 June 2020; Date accepted: 12 June 2020

While they are rare, superspreading events (SSEs), wherein a few primary cases infect an extraordinarily large number of secondary cases, are recognized as a prominent determinant of aggregate infection rates (R_0). Existing stochastic SIR models incorporate SSEs by fitting distributions with thin tails, or finite variance, and therefore predicting almost deterministic epidemiological outcomes in large populations. This paper documents evidence from recent coronavirus outbreaks, including SARS, MERS, and COVID-19, that SSEs follow a power law distribution with fat tails, or infinite variance. We then extend an otherwise standard SIR model with the fat-tailed power law distributions, and show that idiosyncratic uncertainties in SSEs will lead to large aggregate uncertainties in infection dynamics, even with large populations. That is, the timing and magnitude of outbreaks will be unpredictable. While such uncertainties have social costs, we also find that they on average decrease the herd immunity thresholds and the cumulative infections because per-period infection rates have decreasing marginal effects. Our findings have implications for social distancing interventions: targeting SSEs reduces not only the average rate of infection (R_0) but also its uncertainty. To understand this effect, and to improve inference of the average reproduction numbers under fat tails, estimating the tail distribution of SSEs is vital.

1 We thank the Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group for making their dataset available. We are grateful to Hiroshi Nishiura for detailed guidance on epidemiology literature, and Quentin Leclerc for responding to our inquiries; Abhijit Banerjee, Andrew Foster, Alex Toda, Hiroaki Matsuura, Anna Mikusheva, and Charles Wyplosz for their encouragement and helpful advice; Philip MacLellan for his thoughtful editorial support that substantially improved our exposition. All errors are ours.

2 Department of Economics, Massachusetts Institute of Technology.
3 Department of Economics, Massachusetts Institute of Technology.

Covid Economics 30, 19 June 2020: 1-43

1 Introduction

On March 10th, 2020, choir members were gathered for their rehearsal in Washington. While they were all cautious to keep distance from one another and nobody was coughing, three weeks later, 52 members had COVID-19, and two passed away. There are numerous similar anecdotes worldwide.[1] Many studies have shown that the average basic reproduction number ($\mathcal{R}_0$) is around 2.5-3.0 for this coronavirus (e.g. Liu et al., 2020), but 75% of infected cases do not pass on to any others (Nishiura et al., 2020). The superspreading events (SSEs), wherein a few primary cases infect an extraordinarily large number of others, are responsible for the high average number. As SSEs were also prominent in SARS and MERS before COVID-19, epidemiology research has long sought to understand them (e.g. Shen et al., 2004). In particular, various parametric distributions of infection rates have been proposed, and their variances have been estimated in many epidemics under an assumption that they exist (e.g. Lloyd-Smith et al., 2005). On the other hand, stochastic Susceptible-Infectious-Recovered (SIR) models have shown that, as long as the infected population is moderately large, the idiosyncratic uncertainties of SSEs will cancel out each other (Roberts et al., 2015). That is, following the Central Limit Theorem (CLT), stochastic models quickly converge to their deterministic counterparts, and become largely predictable. From this perspective, the dispersion of SSEs is unimportant in itself, but is useful only to the extent it can help target lockdown policies to focus on SSEs to efficiently reduce the average rates $\mathcal{R}_0$ (Endo et al., 2020).

In this paper, we extend this research by closely examining the distribution of infection rates, and rethinking how its dispersion influences the uncertainties of aggregate dynamics. Using data from SARS, MERS, and COVID-19 from around the world, we provide consistent evidence that SSEs follow a power law, or Pareto, distribution with fat tails, or infinite variance. That is, the true variance of infection rates cannot be empirically estimated as any estimate will be an underestimate however large it may be. When the CLT assumption of finite variance does not hold, many theoretical and statistical implications of epidemiology models will require rethinking. Theoretically, even when the infected population is large, the idiosyncratic uncertainties in SSEs will persist and lead to large aggregate uncertainties. Statistically, the standard estimate of the average reproduction number ($\mathcal{R}_0$) may be far from its true mean, and the standard errors will understate the true uncertainty. Because the infected population for COVID-19 is already large, our findings have immediate implications for statistical inference and current policy.

We begin with evidence. Figure 1 plots the largest clusters reported worldwide for COVID-19 from data gathered by Leclerc et al. (2020). If a random variable follows a power law distribution with an exponent α, then the log of its scale (e.g. a US navy vessel had 1,156 cases tested positive) and the log of its severity rank (e.g. that navy case ranked 1st in severity) will have a linear relationship, with its slope indicating $-\alpha$. Figure 1 shows a fine fit of the power law dis-

[1] See Table A.2 in Appendix for a list of several examples.

Covid Economics 30, 19 June 2020: 1-43

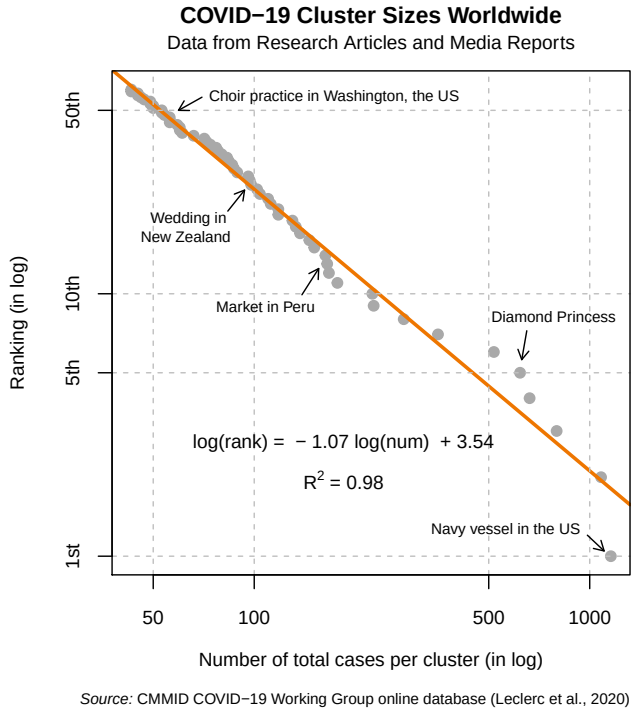


Figure 1: Log cluster size vs log rank for COVID-19 worldwide

Notes: Figure 1 plots the number of total cases per cluster (in log) and their ranks (in log) for COVID-19, last updated on June 3rd. It fits a linear regression for the clusters with size larger than 40. The data are collected by the Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group (Leclerc et al., 2020).

tribution.[2] Moreover, the slope is very close to 1, indicating a significant fatness of the tail to the extent that is analogous to natural disasters such as earthquakes (Gutenberg and Richter, 1954) that are infrequent but can be extreme[3]. While data collection through media reports may be biased towards extreme cases, analogous relationships consistently hold for other SARS, MERS, and COVID-19 data based on surveillance data, with exponents often indicating fat tails. Note that other distributions, including the negative binomial distributions commonly applied in

[2]In Appendix A.2.2, we also estimate the exponent with a small sample bias correction proposed by Gabaix and Ibragimov (2011), which shows the exponent is 1.16, and the R^2 is 0.98. With maximum likelihood estimation, the exponent is 1.01. When using the Kolmogorov-Smirnov test (Clauset et al., 2009), the p-value given $\alpha = 1.01$ is 0.75, failing to reject the null hypothesis that the empirical observation arises from the power law distribution. On the other hand, the p-value given $\alpha = 2$ is 0.000, rejecting the null hypothesis that the distribution is observed from power law distribution with a finite variance.

[3]The power law distribution with $\alpha = 1$ is called the Zipf's law.

Covid Economics 30, 19 June 2020: 1-43

epidemiology research, cannot predict these relationships, and significantly underestimate the risks of extremely severe SSEs.

Using fat-tailed power law distributions, we show that stochastic SIR models predict substantial uncertainties in aggregate epidemiological outcomes. Concretely, we consider a stochastic model with a population of one million, whereby a thousand people are initially infected, and apply epidemiological parameters adopted from the literature. We consider effects of tails of distribution while keeping the average rate ($\mathcal{R}_0$) constant. Under thin-tailed distributions, such as the estimated negative binomial distribution or power law distribution with $\alpha = 2$, the epidemiological outcomes will be essentially predictable. However, under fat-tailed distributions close to those estimated in the COVID-19 data worldwide ($\alpha = 1.1$), there will be immense variations in all outcomes. For example, the peak infection rate is on average 14%, but its 90the percentile is 31% while its 10th percentile is 4%. Under thin-tailed distribution such as negative binomial distribution, the average, 90th percentile and 10th percentile of the peak infection is all concentrated at 26-27%, generating largely deterministic outcomes.

While our primary focus was on the effect on aggregate uncertainty, we also find important effects on average outcomes. In particular, under a fat-tailed distribution, the cumulative and peak infection, as well as the herd immunity threshold, will be lower, and the timing of outbreak will come later than those under a thin-tailed distribution, *on average*. For example, the average herd immunity threshold is 66% with thin-tailed distribution, it is 39% with a fat-tailed distribution.These observations suggest that the increase in aggregate uncertainty over $\mathcal{R}_0$ has effects analogous to a decrease in average $\mathcal{R}_0$. This relationship arises because the average future infection will be a *concave* function of today's infection rate: because of concavity, mean preserving spread will lower the average level. In particular, today's higher infection rate has two countering effects: while it increases the future infection, it also decreases the susceptible population, which decreases it. We provide theoretical interpretations for each outcome by examining the effect of mean-preserving spread of $\mathcal{R}_0$ in analytical results derived in deterministic models.

Our findings have critical implications for the design of lockdown policies to minimize the social costs of infection. Here, we study lockdown policies that target SSEs. We assume that the maximum size of infection rate can be limited to a particular threshold (e.g., 50, 100, or 1000 per day) with some probabilities by banning large gatherings. Because both the uncertainty and mean of the infection rate in the fat-tailed distribution are driven by the tail events, such policies substantially lower the uncertainty and improve the average outcomes. Because the cost of such policy[4] is difficult to estimate reliably, we do not compute the cost-effectiveness of such policy. Nonetheless, we believe this is an important consideration in the current debates on how to re-open the economy while mitigating the risks of subsequent waves.

Finally, we also show the implications of a fat-tailed distributions for the estimation of the

[4]For example, it is prohibitively costly to shut down daycare, but it is less costly to prevent a large concert.

average infection rate. Under such a distribution with small sample sizes, the sample mean yields estimates that are far from the true mean and standard errors that are too small. To address such possibility, it will be helpful to estimate the power law exponent. If the estimate indicates a thin-tailed distribution, then one can be confident with the sample mean estimate. If it indicates a fat-tailed distribution, then one must be aware that there is much uncertainty in the estimate not captured by its confidence interval. While such fat-tailed distributions cause notoriously difficult estimation problems, we explore a "plug-in" method that uses the estimated exponent. Such estimators generate median estimates closer to the true mean with adequate confidence intervals that reflect the substantial risk of SSEs.

Related Literature. First, our paper belongs to a large literature on stochastic epidemiological models. The deterministic SIR model was initiated by Kermack and McKendrick (1927), and later, Bartlett (1949) and Kendall (1956) developed stochastic SIR models (see Britton (2010, 2018) , Britton et al. (2015) for surveys). The traditional view of the stochastic SIR model is that while useful when the number of infected is small, once the infected population is moderately large, it behaves similarly to the deterministic model due to the CLT. Britton (2010) writes "Once a large number of individuals have been infected, the epidemic process may be approximated by the deterministic counter-part." Roberts (2017) also considers an SIR model with small fluctuations of epidemiological parameters, but shows that deterministic models approximate its average reasonably. Here, we consider large aggregate fluctuations arising from idiosyncratic shocks and show that even the average deviates significantly from preditions of deterministic models. There are recent applications of stochastic SIR models that study the very beginning of COVID-19 outbreaks when the number of infection is small (for example, Abbott et al. (2020), Karako et al. (2020), Simha et al. (2020) and Bardina et al. (2020)). However, the major modeling effort has been to use deterministic models based on the common justification above. Our point is that when the distribution is fat-tailed, which we found an empirical support for, the CLT no longer applies, and hence the stochastic model behaves qualitatively differently from its deterministic counterpart even with a large number of infected individuals.

Second, the empirical importance of SSEs is widely recognized in the epidemiological literature before COVID-19 (Lloyd-Smith et al., 2005; Galvani and May, 2005) and for COVID-19 (Frieden and Lee, 2020; Endo et al., 2020). These papers fit the parametric distribution that is by construction thin-tailed, such as negative binomial distribution. It has been common to estimate "the dispersion parameter k" of the negative binomial distribution. We argue that the fat-tailed distribution provides a better fit to the empirical distribution of SSEs, in which a tail parameter, α, parsimoniously captures the fatness of the tail. A recent contribution by Cooper et al. (2019) consider Pareto rule in the context of malaria transmission, but they nonetheless estimate the dispersion with finite variance for the entire infections.

Third, our paper also relates to studies that incorporate heterogeneity into SIR models, in-

Covid Economics 30, 19 June 2020: 1-43

Covid Economics 30, 19 June 2020: 1-43

corporating differences in individual characteristics or community structures. Several recent papers point out that the permanent heterogeneity in individual infection rates lower the herd immunity threshold (Gomes et al., 2020; Hébert-Dufresne et al., 2020; Britton et al., 2020). Although we obtain a similar result, our underlying mechanisms are distinct from theirs. In our model, there is no ex-ante heterogeneity across individuals, and thus their mechanism is not present. Zhang et al. (2013) and Szabó (2020) consider a model in which individuals have heterogeneous infection rates that follow power laws in scale-free networks, but their heterogeneity is permanent (i.e. due to individual characteristics). Instead, what matters for us is the aggregate fluctuations in $\mathcal{R}_0$ (i.e. due to idiosyncratic variations in environments), which their models do not exhibit. Some recent papers emphasize the importance of age-dependent heterogeneity and its implications for lockdown policies (Acemoglu et al., 2020; Davies et al., 2020; Gollier, 2020; Rampini, 2020; Glover et al., 2020; Brotherhood et al., 2020). We emphasize another dimension of targeting: targeting toward large social gatherings, and this policy reduces the uncertainty regarding various epidemiological outcomes. Roberts (2013) analyzes a deterministic model in which basic reproduction number is estimated with noise, and derives probability distributions over epidemiological outcomes due to the uncertainty of the estimates.

Finally, it is well-known that many variables follow a power law distribution. These include the city size (Zipf, 1949), the firm size (Axtell, 2001), income (Atkinson et al., 2011), wealth (Kleiber and Kotz, 2003), consumption (Toda and Walsh, 2015) and even the size of the earthquakes (Gutenberg and Richter, 1954), the moon craters and solar flares (Newman, 2005). Regarding COVID-19, Beare and Toda (2020) document that the cumulative number of infected population across cities and countries is closely approximated by a power law distribution. They then argue that the standard SIR model is able to explain the fact. We document that the infection at the individual level follows a power law. We are also partly inspired by economics literature which argue that the fat-tailed distribution in firm-size has an important consequence for the macroeconomics dynamics, originated by Gabaix (2011). We follow the similar route in documenting that the SSEs are well approximated by a power law distribution and arguing that such empirical regularities have important consequences for the epidemiological dynamics.

Roadmap. The rest of the paper is organized as follows. Section 2 documents evidence that the distribution of SSEs follows power law. Section 3 embed the evidence into an otherwise standard SIR models to demonstrate its implications for the epidemiological dynamics. Section 4 studies estimation of the reproduction numbers under fat-tailed distribution. Section 5 concludes by discussing what our results imply for ongoing COVID-19 pandemic.

Covid Economics 30, 19 June 2020: 1-43

2 Evidence

We present evidence from SARS, MERS, and COVID-19 that the SSEs follow power law distributions. Moreover, our estimates suggest the distributions are often fat-tailed, with critical implications for the probabilities of extreme SSEs. Evidence also suggests a potential role of policies in reducing the tail distributions.

2.1 Statistical model

Let us define the SSEs and their distribution. Following the notations of Lloyd-Smith et al. (2005), let $z_{it} \in \{0, 1, 2, ...\}$ denote the number of secondary cases[5] an infected individual i has at time t. Then, given some threshold $\underline{Z}$, an individual i is said to have caused SSE at time t if $z_{it} \geq \underline{Z}$. To make the estimation flexible, suppose the distribution for non-SSEs, $z_{it} < \underline{Z}$, needs not follow the same distribution as those for SSEs.

In this paper, we consider a power law (or Pareto) distribution on the distribution of SSE. Denoting its exponent by α, the countercumulative distribution is

$$\mathbb{P}\left(z_{it} \geq Z\right) = \pi \left(Z/\underline{Z}\right)^{-\alpha} \quad \text{for } Z \geq \underline{Z}, \tag{1}$$

where π is the probability of SSEs. Notably, its mean and variance may not exist when α is sufficiently low: while its mean is $\frac{\alpha}{\alpha-1}\underline{Z}$ if $\alpha > 1$, it is ∞ if $\alpha \leq 1$. While its variance is $\frac{\alpha}{(\alpha-1)^2(\alpha-2)}\underline{Z}^2$ if $\alpha > 2$, it is ∞ if $\alpha \leq 2$. In this paper, we formally call a distribution to be fat-tailed if $\alpha < 2$ so that they have infinite variance. While non-existence of mean and variance may appear pathological, a number of socioeconomic and natural phenomenon such as city sizes ($\alpha \approx 1$), income ($\alpha \approx 2$), and earthquake energy ($\alpha \approx 1$) have tails well-approximated by this distribution as reviewed in the Introduction. One concrete example[6] that can explain a power law distribution is due to the result in Beare and Toda (2019): suppose each participant can invite some others with some probability. Conditional on inviting, the number of people each paritipant invites follows some distributions such as log-normal distribution. Then, the resulting distribution of all participants follows a power law.

This characteristics stands in contrast with the standard assumption in epidemiology literature that the full distribution of z_{it} follows a negative binomial (or Pascal) distribution[7] with finite mean and variance. The negative binomial distribution has been estimated to fit the data

[5]Note that the number of "secondary" cases include only direct transmissions and exclude indirect transmissions. This is how the COVID-19 data in Figure 1 were also collected (Leclerc et al., 2020).

[6]Another theoretical reason why this distribution could be relevant for airborne diseases is that the number of connections in social networks often follow a power law (Barabasi and Frangos, 2014).

[7]Denoting its mean by R and dispersion parameter by k, the distribution is

$$\mathbb{P}\left(z_{it} \geq Z\right) = 1 - \sum_{z=0}^{Z} \frac{\Gamma(z+k)}{z!\Gamma(k)} \left(\frac{\mathcal{R}}{k}\right)^{z} \left(1 + \frac{\mathcal{R}}{k}\right)^{-(z+k)}$$

better than Poisson or geometric distribution for SARS (Lloyd-Smith et al., 2005), and given its theoretical bases from branching model (e.g. Gay et al., 2004), it has been a standard distributional assumption in the epidemiology literature (e.g. Nishiura et al., 2017).

2.2 Data

This paper uses five datasets of recent coronavirus outbreaks for examining the distribution of SSEs: COVID-19 data from (i) around the world, (ii) Japan, and (iii) India, and (iv) SARS data, (v) MERS data.

(i) COVID-19 data from around the world: this dataset contains clusters of infections found by a systematic review of academic articles and media reports, conducted by the Centre of the Mathematical Modelling of Infectious Diseases COVID-19 Working Group (Leclerc et al., 2020). The data are restricted to first generation of cases, and do not include subsequent cases from the infections. The data are continuously updated, and in this draft, we have used the data downloaded on June 3rd. There were a total of 227 clusters recorded.

(ii) COVID-19 data from Japan: this dataset contains a number of secondary cases of 110 COVID-19 patients across 11 clusters in Japan until February 26th, 2020, reported in Nishiura et al. (2020). This survey was commissioned by the Ministry of Health, Labor, and Welfare of Japan to identify high risk transmission cases.

(iii) COVID-19 data from India: this dataset contains the state-level data collected by the Ministry of Health and Family Welfare, and individual data collected by covid19india.org.[8] We use the data downloaded on May 31st.

(iv) SARS from around the world: this dataset contains 15 incidents of SSEs from SARS in 2003 that occured in Hong Kong, Beijing, Singapore, and Toronto, as gathered by Lloyd-Smith et al. (2005)[9] through a review of 6 papers. The rate of community transmission was not generally high so that, for example, the infections with unknown route were only about 10 percent in the case of Beijing. The data consist of SSEs, defined by epidemiologists (Shen et al., 2004) as the cases with more than 8 secondary cases. For Singapore and Beijing, the contact-tracing data is available from Hsu et al. (2003) and Shen et al. (2004), respectively. When compare the fit to the negative binomial distribution, we compare the fit of power law to that of negative binomial using these contact tracing data.

(v) MERS from around the world: this dataset contains MERS clusters reported up to August 31, 2013. The cases are classified as clusters when thee are linked epidemiologically. The

The variance of this distribution is $\mathcal{R}\left(1+\frac{\mathcal{R}}{k}\right)$. The distribution nests Poisson distribution (as $k \to \infty$) and geometric distribution (when $k = 1$.)

[8] https://www.kaggle.com/sudalairajkumar/covid19-in-india. covid19india.org is a volunteer-based organization that collects information from municipalities.

[9] Even though Lloyd-Smith et al. (2005) had analyzed 6 other infectious diseases, SARS was the only one with sufficient sample sizes to permit reliable statistical analyses.

Covid Economics 30, 19 June 2020: 1-43

data come from three published studies were used in Kucharski and Althaus (2015). Total of 116 clusters are recorded.

We use multiple data sets in order to examine the robustness of findings.[10] Having multiple data sets can address each other's weaknesses in data. While data based on media reports is broad, they may be skewed to capture extreme events; in contrast, data based on contact tracing may be reliable, but are restricted to small population. By using both, we can complement each data's weaknesses.

2.3 Estimation

The datasets report cumulative number of secondary cases, either $\sum_i z_{it}$ (when a particular event may have had multiple primary cases) or $\sum_t z_{it}$ (when an individual infects many others through multiple events over time). Denoting these cumulative numbers by Z, we consider this distribution for some $Z \geq \underline{Z}^*$. As discussed in Appendix A.1, we can interpret the estimates of this tail distribution as approximately the per-period and individual tail distribution and therefore map directly to the parameter of the SIR model in the next section. The thresholds for inclusion, $\underline{Z}^*$, will be chosen to match the threshold for SSEs when possible, but also adjust for the sample size. For COVID-19 in the world, we apply $\underline{Z} = 40$ to focus on the tail of the SSE distribution. For SARS, we apply $\underline{Z} = 8$ as formally defined (Shen et al., 2004). For other samples, we apply $\underline{Z} = 2$ because the sample size is limited.

To assess whether the distribution of Z follows the power law, we adopt the regression-based approach that is transparent and commonly used. If Z follows power law distribution, then by (1), the log of Z and the log of its underlying rank have a linear relationship: $\log rank(Z) = -\alpha \log Z + \log(N\pi\underline{Z}^{\alpha})$. This is because, when there are N individuals, the expected ranking of a realized value Z is $\mathbb{E}rank(Z) \simeq \mathbb{P}(z \geq Z)N$ for moderately large N. Thus, when N is large, we obtain a consistent estimate of α by the following regression:

$$\log rank(Z) = -\alpha \log Z + \log(N\pi\underline{Z}^{\alpha}) + \varepsilon \tag{2}$$

When N is not large, however, the estimate will exhibit a downward bias because log is a concave function and thus $\mathbb{E}\log rank(Z) < \log \mathbb{E}rank(Z)$. While we present the analysis according to (2) in Figures 1 and 2 for expositional clarity, we also report the estimates with small sample bias correction proposed by Gabaix and Ibragimov (2011) in Appendix A.2.2.[11] We also

[10] he infectious diseases considered here share some commonalities as SARS-CoV that causes SARS, MERS-CoV that causes MERS, and SARS-CoV-2 that causes COVID-19 are human coronaviruses transmitted through the air. They have some differences in terms of transmissibility, severity, fatality, and vulnerable groups (Petrosillo et al., 2020). But overall, as they are transmitted through the air, they are similar compared to other infectious diseases.

[11] Their approach is to turn the dependent variable into $\log\left[rank(Z) - \frac{1}{2}\right]$ instead of $\log[rank(Z)]$. We examine the performance of their bias correction method through a estimating regression given random variables generated from power law distributions. While their bias correction almost eliminates bias when N is moderately large, it

Covid Economics 30, 19 June 2020: 1-43

estimate using the maximum likelihood in Appendix A.2.2. Note that when there are ties (e.g. second and third largest had 10 infections), we assigned different values to each observation (e.g. assigning rank of 2 and 3 to each observation).

Next, we also compare the extent to which a power law distribution can approximate the distribution of SSEs adequately relative to the negative binomial distribution. First, we plot what the predicted log-log relationship in (2) would be given the estimated parameters of negative binomial distribution.[12] Second, to quantify the predictive accuracy, we compute the ratio of likelihood of observing the actual data.

2.4 Results

Our analysis shows that the power law finely approximates the distribution of SSEs. Figure 1 visualizes this for COVID-19 from across the world, and Figure 2 for SARS, MERS, and COVID-19 in Japan and India. Their R^2 range between 0.93 and 0.99, suggesting high levels of fit to the data. Because our focus is on upper-tail distribution, Figure 1 truncates below at the cluster size 40, Figure 2 truncates at 8 for SARS and at 2 for MERS and COVID-19 in India and Japan. Figure A.1 in Appendix presents a version of Figure 1 truncated below at 20.

In addition, the estimates of regression (2) suggest that the power law exponent, α, is below 2 and even close to 1. Table 1 summarizes the main findings. The estimated exponents near 1 suggest that extreme SSEs are not uncommon. For COVID-19 in Japan and India, the estimated exponents are larger than 1 but often below 2. Since applying the threshold of $\underline{Z}^* = 2$ is arguably too low, we must interpret out-of-sample extrapolation from these estimates with caution. When higher thresholds are applied, the estimated exponents tend to be higher. For example, when applying the threshold of $\underline{Z}^* = 8$ as in SARS 2003 to COVID-19 in India, the estimated exponent is 1.85 or 2.25. This pattern is already visible in Figure 2. Table A.1 in Appendix A.2.2 presents results using bias correction technique of Gabaix and Ibragimov (2011) as well as maximum likelihood. The results are very similar.

Notably, the estimated exponent of India is higher than those of other data. There are two possible explanations. First, the lockdown policies in India have been implemented strictly relative to moderate approaches in Japan and some other parts of the world during the outbreaks. By discouraging and prohibiting large-scale gatherings, sometimes by police enforcement, they may have been successful at targeting SSEs. Second, contact tracing to ensure data reliability may have been more difficult in India until end of May than in Japan until end of February.[13]

has an upward bias of α whereas the equation (2) has a downward bias. The magnitude of bias is similar when $N = 10$ or $N = 15$. Thus, our preferred approach is to refer to both methods for robustness.

[12] This approach stands in contrast with a common practice to plot the probability mass functions. Unlike such approaches where differences in tail densities are invisible since it is very close to zero, this approach highlights the differences in tail densities.

[13] Concretely, there were only 248 cases of more than one secondary infections reported in the data among 27,890 primary cases in the data from India. That is, only 0.8 percents of primary cases were reported to have

Covid Economics 30, 19 June 2020: 1-43

Covid Economics 30, 19 June 2020: 1-43

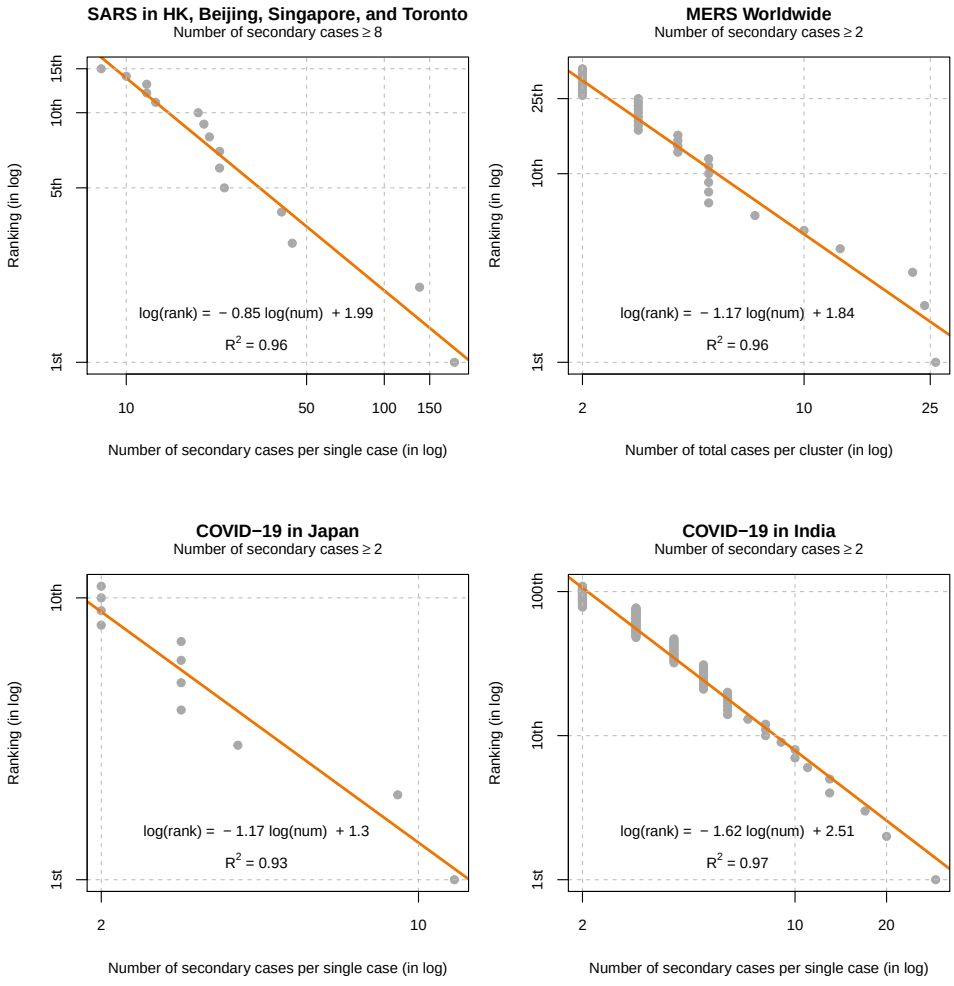


Figure 2: Log size vs log rank for COVID-19

Notes: Figure 2 plots the number of total cases per cluster (in log) and their ranks (in log) for MERS, and the number of total cases per cluster (in log) and their ranks (in log) for SARS and COVID-19 in Japan and India. The data for SARS are from Lloyd-Smith et al. (2005), and focus on SSEs defined to be the primary cases that have infected more than 8 secondary cases. The data for MERS come from Kucharski and Althaus (2015). The data for Japan comes from periods before February 26, 2020, reported in Nishiura et al. (2020). The data for India are until May 31, 2020, reported by the Ministry of Health and Family Welfare, and covid19india.org. The plots are restricted to be the cases larger than 2.

	COVID-19			SARS			MERS
	World	Japan	India	World	Singapore	Beijing	World
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\hat{\alpha}$	1.07	1.17	1.62	0.85	0.75	0.75	1.17
	(0.04)	(0.10)	(0.03)	(0.06)	(0.08)	(0.06)	(0.07)
$\underline{Z}$	40	2	2	8	2	2	2
Obs.	60	11	109	15	19	8	36
R^2	0.98	0.93	0.97	0.96	0.91	0.94	0.96
$\log_{10} LR$	-	11.39	-	-	19.51	8.04	40.89

Table 1: Estimates of power law exponent ($\hat{\alpha}$) and their fit with data

Notes: Table 1 summarizes the estimates of power law exponent ($\hat{\alpha}$) given as the coefficient of regression of log of number of infections (or size of clusters) on the log of their rankings. Heteroskedasticity-robust standard errors are reported in the parenthesis. $\underline{Z}$ denote the threshold number of infection to be included. $\log_{10}$(LR) denotes "likelihood ratios", expressed in the log with base 10, of probability of observing this realized data with power law distributions relative to that with estimated negative binomial distributions. Columns (1)-(3) report estimates for COVID-19; columns (4)-(6) for SARS, and column (7) for MERS.

Covid Economics 30, 19 June 2020: 1-43

While missing values will not generate any biases if the attritions were proportional to the number of infections, large gatherings may have dropped more than in Japan where the SSEs were found through contact tracing. Nonetheless, these estimates suggest that various environments and policies could decrease the risks of the extreme SSEs. This observation motivates our policy simulations to target SSEs.

Next, we compare the assumption of power law distribution relative to that of a negative binomial distribution. Figure 3 shows that the negative binomial distributions would predict that the extreme SSEs will be fewer than the observed distribution: while it predicts the overall probability of SSEs accurately, they suggest that, when they occur, they will not be too extreme in magnitude. Table 1 reports the relative likelihood, in logs, of observing the data given the estimated parameters. It shows that, under the estimated power law distribution relative to the estimated negative binomial distribution, it is $10^8 - 10^{20}$ times more likely to observe the SARS data (10^{40} times more for MERS, and 10^{11} times more COVID-19 data in Japan). Such large differences emerge because the negative binomial distribution, given its implicit assumption of finite variance, suggests that the extreme SSEs are also extremely rare when estimated with

infected more than one persons. In contrast, there were 27 cases with more than one secondary infections among 110 primary cases in Japan. That is, 25 percent of primary cases were infectious. This difference in ration likely reflects the data collection quality than actual infection dynamics.

Covid Economics 30, 19 June 2020: 1-43

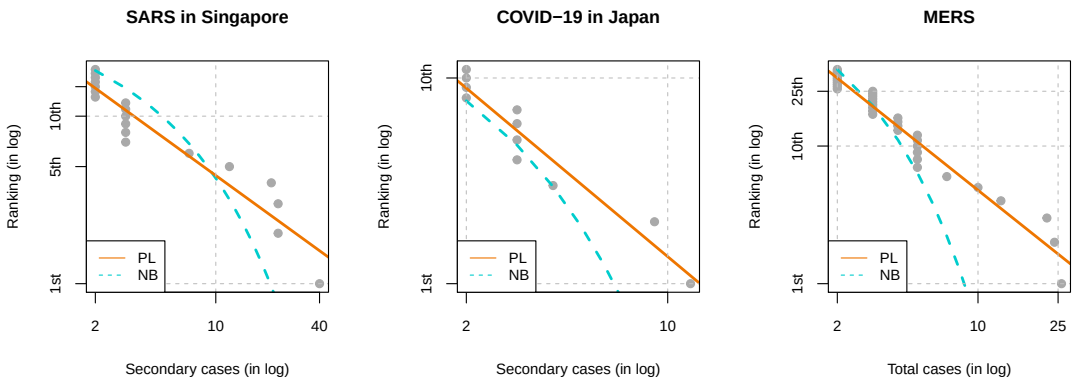


Figure 3: Comparison of power law and negative binomial distributions

Notes: Figure 3 plots the predicted ranking of infection cases given the estimated negative binomial (NB) distribution, in addition to the log-log plots and estimated power law (PL) distributions. The negative binomial distribution is parameterized by (R, k), where R is mean and k is the dispersion parameter with the variance being $R(1 + R/k)$. The estimates for SARS Singapore come from our own estimates using the maximum likelihood ($R = 0.88$, $k = 0.09$); MERS come from the world ($R = 0.47, k = 0.26$) estimated in Kucharski and Althaus (2015); and COVID-19 in Japan were from our own estimates using the maximum likelihood ($R = 0.56, k = 0.21$). The estimates of Singapore is slightly different from Lloyd-Smith et al. (2005) because we pool all the samples.

entire data sets[14]. If our objective is to predict the overall incidents of infections parsimoniously, then negative binomial distribution is well-validated and theoretically founded (Lloyd-Smith et al., 2005).[15] However, if our goal is to estimate the risks of extreme SSEs accurately, then using only two parameters with finite variance to estimate together with the entire distribution may be infeasible.

These distributional assumptions have critical implications for the prediction of the extreme SSEs. Table 2 presents what magnitude top 1%, top 5%, and top 10% among SSEs will be given each estimates of the distribution. Given the estimates of the negative binomial distribution, even the top 1% of SSEs above 8 cases will be around the magnitude of 19-53. However, given a range of estimates from power law distribution, the top 1% could be as large as 569. Thus, it is no longer surprising that the largest reported case for COVID-19 will be over 1,000 people.

[14]For example, the binomial distribution estimate suggests an incidence of 185 cases (residential infection in Hong Kong) only has a chance of 9.5×10^{-10} occurring for any single primary case.

[15]Since the power law distribution is fitted only to SSEs, estimated power law distribution may fit the data better than the estimated negative binomial distribution that was meant to fit the entire data set. Rather than making such comparison, this estimation is intended to illustrate the magnitude of difference between the two distributional assumptions. Because of significant missing values for the low number of infections in the COVID-19 from across the world and India, we will not use the data sets for estimation of negative binomial distributions.

	Power Law					Negative Binomial		
	$\alpha = 1.08$	$\alpha = 1.1$	$\alpha = 1.2$	$\alpha = 1.5$	$\alpha = 2$	SARS	MERS	COVID-19
1%	569	526	371	172	80	44	18	19
5%	128	122	97	59	36	31	15	15
10%	67	65	55	37	25	25	13	14

Table 2: Probabilities of extreme SSEs under each distribution

Notes: Table 2 shows the size of secondary cases at each quantile, top 1 percentile, 5 percentile, and 10 percentile, given each distributions. The negative binomial distribution's estimates for SARS are from Singapore, for COVID-19 are from Japan, and for MARS is from around the world.

In contrast, such incidents have vanishingly low chance under binomial distributions. Since the SSEs are rare, researchers will have to make inference about their distribution based some parametric methods. Scrutinizing such distributional assumptions along with the estimation of parameters themselves will be crucial in accurate prediction of risks of extreme SSEs.

Covid Economics 30, 19 June 2020: 1-43

3 Theory

Motivated by the evidence, we extend an otherwise standard stochastic SIR model with a fat-tailed SSEs. Unlike with thin-tailed distributions, we show that idiosyncratic risks of SSEs induce aggregate uncertainties even when the infected population is large. We further show that the resulting uncertainties in infection rates have important implications for average epidemiological outcomes. Impacts of lockdown policies that target SSEs are discussed.

3.1 Stochastic SIR model with fat-tailed distribution

Suppose there are $i = 1, ..., N$ individuals, living in periods $t = 1, 2, ...$ Infected individuals pass on and recover from infection in heterogeneous and uncertain ways. Let β_{it} denote the number of new infection in others an infected individual i makes at time t. Let $\gamma_{it} \in \{0, 1\}$ denote the recovery/removal, where a person recovers ($\gamma_{it} = 1$) with probability $\gamma \in [0, 1]$. Note that, whereas z_{it} in Section 2 was a stochastic analogue of "effective" reproduction number, β_{it} here is such analogue of "basic reproduction number." Assuming enough mixing in the population, these two models are related by $z_{it} = \beta_{it} \frac{S_t}{N}$, where S_t is a number of susceptible individuals in the population.

This model departs from other stochastic SIR models only mildly: we consider a fat-tailed, instead of thin-tailed, distribution of infection rates. Based onthe evidence, we consider a power

Parameter	Description	Value	Source
A. Common parameters			
γ	recovery & death rate	$7/18$	Wang et al. (2020)
N	total population	10^5	
I_0	initially infected populatoion	10^3	1% of population
$\mathcal{R}_0 \equiv \mathbb{E}[\beta_{it}]/\gamma$	mean basic reproduction number	2.5	Remuzzi and Remuzzi (2020)
B. Power law			
π	probability of infecting	0.25	Nishiura et al. (2020)
α	tail parameter	$\{1.08, 1.1, 1.2, 1.5, 2\}$	
C. Negative binomial			
k	overdispersion parameter	0.16	Lloyd-Smith et al. (2005)

Table 3: Parameter values

Covid Economics 30, 19 June 2020: 1-43

law distribution of β_{it}: its countercumulative distribution is given by

$$\mathbb{P}\left(\beta_{it} \geq \beta\right) = \pi(\beta/\underline{\beta})^{-\alpha}$$

for the exponent α and a normalizing constant $\underline{\beta}$, and $\pi \in [0,1]$ is the probability that $\beta \geq \underline{\beta}$. Note that the estimated exponent α can be mapped to this model, as discussed in Appendix A.1. If we assume β_{it} is distributed according to exponential distribution or negative binomial distribution, we obtain a class of stochastic SIR models commonly studied in the epidemiological literature (see Britton (2010, 2018) for surveys). We will compare the evolution dynamics under this power law distribution against those under negative binomial distribution as commonly assumed, keeping the average basic reproduction number the same. To numerically implement this, we will introduce normalization to the distributions.

The evolution dynamics is described by the following system of stochastic difference equations. Writing the total number of infected and recovered/removed populations by I_t and R_t, we have

$$S_{t+1} - S_t = -\sum_{i=1}^{I_t} \beta_{it} \frac{S_t}{N} \tag{3}$$

$$I_{t+1} - I_t = \sum_{i=1}^{I_t} \beta_{it} \frac{S_t}{N} - \sum_{i=1}^{I_t} \gamma_{it} \tag{4}$$

$$R_{t+1} - R_t = \sum_{i=1}^{I_t} \gamma_{it}. \tag{5}$$

This system is a discrete-time and finite-population analogue of the continuous-time and continuous-population differential equation SIR models.

Parametrization: we parametrize the model as follows. The purpose of simulation is a proof of concept, rather than to provide a realistic numbers. We take the length of time to be one week. We set the sum of the recovery and the death rate per day is $1/18$ following Wang et al. (2020), so that $\gamma = 7/18$. The total population is set to $N = 10^5$, and initially infected population is 1% of the total population. As a benchmark case, we set $\alpha = 1.1$, which is in line with the estimates for the COVID-19 data worldwide, but we explore several other parametrization, $\alpha \in \{1.08, 1.2, 1.5, 2\}$. As documented in Nishiura et al. (2020), 75% of people did not infect others. We therefore set $\pi = 0.25$. This number is also in line with the evidence from SARS reported in Lloyd-Smith et al. (2005), in which 73% of cases were barely infectious. We choose $\underline{\beta}$, which controls the mean of β_{it}, so that the expected $\mathcal{R}_0 \equiv \mathbb{E}\beta_{it}/\gamma$ per day is 2.5, corresponding to the middle of the estimates obtained in Remuzzi and Remuzzi (2020). This leads us to choose $\underline{\beta} = 0.354$ in the case of $\alpha = 1.1$.

We will contrast the above model to a model in which β_{it} is distributed according to negative binomial, $\beta_{it}/\gamma \sim$ negative binomial$(\mathcal{R}_0, k)$. The mean of this distribution is $\mathbb{E}\beta_{it}/\gamma = \mathcal{R}_0$, ensuring that it has the same mean basic reproduction number as in the power law case, and the variance is $\mathcal{R}_0(1 + \mathcal{R}_0/k)$. The smaller values of k indicate greater heterogeneity (larger variance). We use the estimates of SARS by Lloyd-Smith et al. (2005), $k = 0.16$. The mean is set to the same value as power law case, $\mathcal{R}_0 = 2.5$,

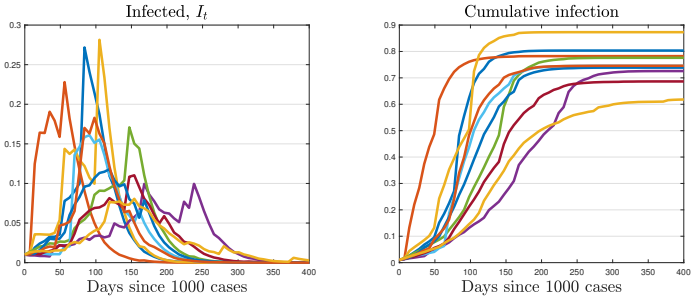
3.2 Effects of fat-tailed distribution on uncertainty

Figure 4a shows 10 sample paths of infected population generated through the simulation of the model with $\alpha = 1.1$. One can immediately see that even though all the simulation start from the same initial conditions under the same parameters, there is enormous uncertainty in the timing of the outbreak of the disease spread, the maximum number of infected, and the final number of susceptible population. The timing of outbreak is mainly determined by when SSEs occur. To illustrate the importance of a fat-tailed distribution, Figure 4b shows the same sample path but with a thin-tailed negative binomial distribution. In this case, as already 1,000 people are infected in the initial period, the CLT implies the aggregate variance is very small and the model is largely deterministic. This is consistent with Britton (2018). Britton (2018) shows that when the total population is as large as 1,000 or 10,000, the model quickly converges to the deterministic counterpart.

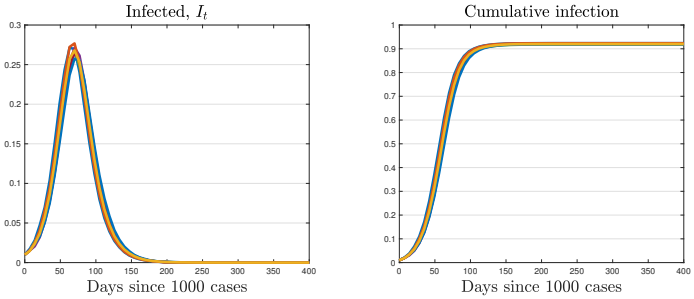
Figure 5 compares the entire distribution of the number of cumulative infection (top-left), the herd immunity threshold (top-right), the peak number of infected (bottom-left), and the days it takes to infect 5% of population (bottom-right). The herd immunity threshold is defined as the cumulative number of infected at which the number of infected people is at its peak. The histogram contrast the case with power law distribution with $\alpha = 1.1$ to the case with negative binomial distribution. It is again visible that uncertainty remains in all outcomes when the

Covid Economics 30, 19 June 2020: 1-43

Covid Economics 30, 19 June 2020: 1-43



(a) Power law ($\alpha = 1.1$)



(b) Negative binomial

Figure 4: Ten sample paths from simulation

Note: Figure 4 plots 10 sample path of the number of infected population from simulation, in which we draw $\{\beta_{it}, \gamma_{it}\}$ randomly every period in an i.i.d. manner. Figure 4a plots the case with power law distribution, and Figure 4b plots the case with negative binomial distribution.

distribution of infection rate is fat-tailed. For example, the cumulative infection varies from 65% to 100% in the power law case, while the almost all simulation is concentrated around 92% in the case of negative binomial distribution.

Table 4 further shows the summary statistics for the epidemiological outcomes for various power law tail parameters, α, as well as for negative binomial distribution. With fat-tails, i.e. α close to one, the range between 90th percentile and 10th percentile for all statistics is wide, but this range is substantially slower as the tail becomes thinner (α close to 2). For example, when $\alpha = 1.08$ the peak infection rate can vary from 6% to 32% as we move from 10the percentile to 90th percentile. In contrast, when $\alpha = 2$, the peak infection rate is concentrated at 26–27%. Moreover, when $\alpha = 2$, the model behaves similarly to the model with negative binomial distribution because the CLT applies to both cases.

Covid Economics 30, 19 June 2020: 1-43

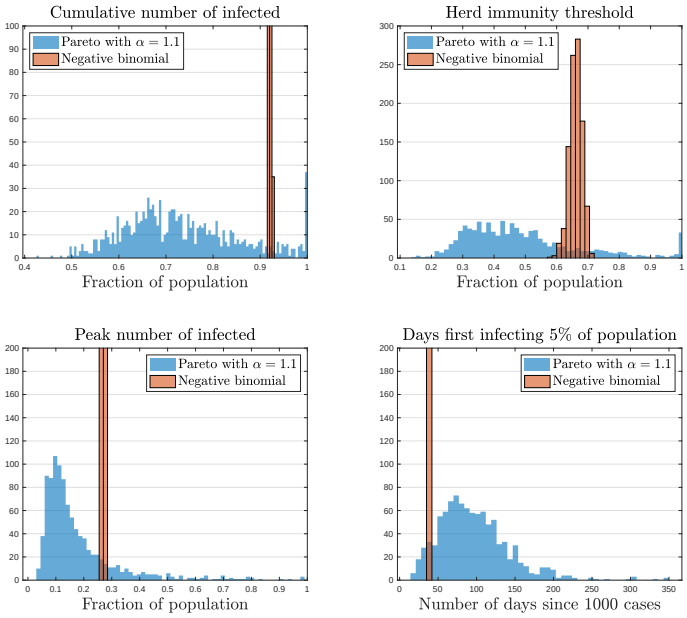


Figure 5: Histogram from 1000 simulation

Note: Figure 5 plots the histogram from 1000 simulations, in which we draw $\{\beta_{it}, \gamma_{it}\}$ randomly every period in an i.i.d. manner. The cumulative number of infected is S_T, where we take $T = 204$ weeks. The herd immunity threshold is given by the cumulative number of infected, at which the infection is at the peak. Formally, S_{t^*} where $t^* = \arg\max_t I_t$. The peak number of infected is $\max_t I_t$.

3.3 Effects of fat-tailed distribution on average

While our primary focus was the effect on the uncertainty of epidemiological outcomes, Figure 5 also shows significant effects on the mean. In particular, fat-tailed distribution also lowers cumulative infection, the herd immunity threshold, the peak infection, and delays the time it takes to infect 5% of population, *on average*. Why could such effects emerge?

To understand these effects, we consider a deterministic SIR model with continuous time and continuum of population. In such a textbook model, we consider the effect of small uncertainties (i.e. mean-preserving spread) in $\mathcal{R}_0$. Such theoretical inquiry can shed light on the effect because the implication of fat-tailed distribution is essentially to introduce time-varying fluctuation in aggregate $\mathcal{R}_0$. We can thus examine how the outcome changes by $\mathcal{R}_0$, and invoke Jensen's inequality to interpret the results.[16]

[16]This assumes that $\mathcal{R}_0$ is drawn at time 0, and stay constant thereafter for each simulation. This exercise is not

CEPR PRESS

	Power law					Negative
	$\alpha = 1.08$	$\alpha = 1.1$	$\alpha = 1.2$	$\alpha = 1.5$	$\alpha = 2$	binomial
1. Cumulative infected						
mean	60%	73%	89%	92%	92%	92%
90th percentile	85%	91%	95%	93%	92%	92%
50th percentile	59%	71%	88%	92%	92%	92%
10th percentile	39%	59%	84%	91%	92%	92%
2. Herd immunity threshold						
mean	39%	49%	62%	65%	66%	66%
90th percentile	65%	75%	78%	71%	69%	69%
50th percentile	35%	45%	59%	65%	66%	66%
10th percentile	17%	29%	51%	60%	62%	64%
3. Peak infection						
mean	14%	18%	25%	27%	27%	27%
90th percentile	31%	34%	36%	29%	28%	27%
50th percentile	9%	13%	22%	26%	27%	27%
10th percentile	4%	7%	18%	25%	26%	26%
4. Days infecting 5%						
mean (days)	137	93	47	37	35	35
90th percentile	252	147	56	42	35	35
50th percentile	119	84	49	35	35	35
10th percentile	49	42	35	35	35	35

Table 4: Summary statistics for epidemiological outcomes

Note: Table 4 shows the summary statistics from 1000 simulations for five different tail parameters for the case of power law distribution, and for the negative binomial distribution.

Covid Economics 30, 19 June 2020: 1-43

1. **Effect on cumulative infection:** note that the cumulatively infected population is given by $1 - S_\infty/N$, where S_∞ is the ultimate susceptible population as $t \to \infty$. Taking the standard derivation, S_∞ satisfies the following equation:[17]

$$\log(S_\infty/N) = -\mathcal{R}_0(1 - S_\infty/N) \tag{6}$$

In Appendix B, we prove that S_∞ is a convex function of $\mathcal{R}_0$ if $\mathcal{R}_0 > 1.125$, , which is likely to be met in SARS or COVID-19.[18] Thus, the cumulative infection is concave in $\mathcal{R}_0$, and the mean-preserving spread in $\mathcal{R}_0$ lowers the cumulative infection.

exactly the same as our original SIR model because there $\mathcal{R}_0$ fluctuates over time within a simulation. Thus this is for providing intuition, rather than a proof.

[17] Here, we set the initially recovered population to zero, $R_0 = 0$.

[18] Numerically, we did not find any counterexample even when $\mathcal{R}_0 \in [1, 1.125]$.

2. **Effect on herd immunity threshold:** denoting the number of recovered/removed and infected population by R, the infection will stabilize when $\mathcal{R}_0\left(\frac{N-R}{N}\right) = 1$. Rearranging this condition, the herd immunity threshold, R^* is given by

$$\frac{R^*}{N} = 1 - \frac{1}{\mathcal{R}_0}, \tag{7}$$

where $\mathcal{R}_0 \equiv \beta/\gamma$. Since R^* is concave in $\mathcal{R}_0$, the mean-preserving spread in $\mathcal{R}_0$ lowers the herd immunity threshold.

3. **Effect on timing of outbreak:** let us consider the time t^* when some threshold of outbreak $\left(\frac{I}{N}\right)^*$ is reached. Supposing $S/N \approx 1$ at the beginning of outbreak, t^* satisfies

$$\left(\frac{I}{N}\right)^* \approx \frac{I_0}{N}\exp(\frac{1}{\gamma}(\mathcal{R}_0 - 1)t^*) \tag{8}$$

Thus, t^* is convex in $\mathcal{R}_0$, and the mean-preserving spread in $\mathcal{R}_0$ delays the timing of the outbreak.

4. **Effect on peak infection rate:** the peak infection rate, denoted by $\frac{I^{\max}}{N}$, satisfies

$$\frac{I^{\max}}{N} = 1 - \frac{1}{\mathcal{R}_0} - \frac{1}{\mathcal{R}_0}\log(\mathcal{R}_0 S_0), \tag{9}$$

where S_0 is initial susceptible population. We show in the Appendix that (9) implies that the peak infection, $I^{\max}/N$, is a concave function of $\mathcal{R}_0$ if and only if $\mathcal{R}_0 \geq \frac{1}{S_0}\exp(0.5)$. If we let $S_0 \approx 1$, this implies $\mathcal{R}_0 \geq \exp(0.5) \approx 1.65$. This explains why we found a reduction in peak infection rate, as we have assumed $\mathcal{R}_0 = 2.5$. Loosely speaking, since the peak infection rate is bounded above by one, it has to be concave for sufficiently high $\mathcal{R}_0$.

Overall, we have found that the increase in the uncertainty over $\mathcal{R}_0$ has effects similar to a decrease in the level of $\mathcal{R}_0$. This is because the aggregate fluctuations in $\mathcal{R}_0$ introduce negative correlation between the future infection and the future susceptible population. High value of today's $\mathcal{R}_0 \equiv \mathbb{E}\frac{\beta_{it}}{\gamma}$ increases tomorrow's infected population, I_{t+1}, and decreases tomorrow's susceptible population, S_{t+1}. That is, $Cov(S_{t+1}, I_{t+1}) < 0$. Because the new infection tomorrow is a realization of β_{t+1} multiplied by the two (that is, $\beta_{t+1}I_{t+1}\frac{S_{t+1}}{N}$) this negative correlation reduces the spread of the virus in the future on average, endogenously reducing the magnitude of the outbreak.

This interpretation also highlights the importance of intertemporal correlation of infection rates, $Cov(\beta_t, \beta_{t+1})$. When some individuals participate in events at infection-prone environments more frequently than others, the correlation will be positive. Such effects can lead to a sequence of clusters and an extremely rapid rise in infections (Cooper et al., 2019) that over-

Covid Economics 30, 19 June 2020: 1-43

whelm the negative correlation between S_{t+1} and I_{t+1} highlighted above. On the other hand, when infections take place at residential environments (e.g. residential compound in Hong Kong for SARS, and dormitory in Singapore for COVID-19), then the infected person will be less likely to live in another residential location to spread the virus. In this case, the correlation will be negative. In this way, considering the correlation of infection rates across periods will be crucial.

Note that the mechanism we identified on herd immunity thresholds is distinct from the ones described in Gomes et al. (2020); Hébert-Dufresne et al. (2020); Britton et al. (2020). They note that when population has permanently heterogenous activity rate, which captures both the probability of infecting and being infected, the herd immunity can be achieved with lower threshold level of susceptible. They explain this because majority of "active" population becomes infected faster than the remaining population. Our mechanism does not hinge on the permanent heterogeneity in population, which could have been captured by $Cov(\beta_{it}, \beta_{it+1}) = 1$. The fat-tailed distribution in infection rate alone creates reduction in the required herd immunity rate in expectation.

3.4 Lockdown policy targeted at SSEs

How could the policymaker design the mitigation policies effectively if the distribution of infection rates is fat-tailed? Here, we concentrate our analysis on lockdown policy. Unlike the traditionally analyzed lockdown policy, we consider a policy that particularly targets SSEs. Specifically we assume that the policy can impose an upper bound on $\beta_{it} \leq \bar{\beta}$ with probability ϕ. The probability ϕ is meant to capture some imperfection in enforcements or impossibility in closing some facilities such as hospitals and daycare[19]. Here, we set $\phi = 0.5$. For tractability, we assume that the government implements targeted lockdown policies for entire periods. We experiment with $\bar{\beta}$ for various values: 1000 cases per day, 100 cases per day, and 50 cases per day.

While Table B.3 in Appendix presents results in detail, we briefly summarize the main results here. First, the policy reduces the mean of the peak infection rate if and only if the distribution features fatter tails. Second, the targeted lockdown policy is effective in reducing the volatility of the peak infection rate in the case that such risks exist in the first place. For example, consider the case with $\alpha = 1.1$. Moving from no policy to the upper-limit of 100 cases reduces the 90th percentile of peak infection from 31% to 17%.[20] In contrast, when $\alpha = 2$ or

[19]Note that, even though the theoretical variance is infinite, the realized variance in numerical simulations will always be finite. Therefore, such stochastic reductions can still reduce the simulated variance even though the theoretical variance remains infinite.

[20]We may be concerned that the unbounded support of power law distribution is unrealistic; at the extreme case, one cannot infect more than 8 billion people since that will exceed the world population. Imposing some upperbound on the distribution of infection rate will be equivalent to imposing a lockdown policy with perfect implementation ($\phi = 1$). As shown in the results of lockdown policy, imposing such upperbounds can significantly

Covid Economics 30, 19 June 2020: 1-43

with negative binomial distribution, the policy has virtually no effect. Therefore the policy is particularly effective in mitigating the upward risk of overwhelming the medical capacity. This highlights that while the fat-tailed distribution induces the aggregate risk in the epidemiological dynamics, the government can partly remedy this by appropriately targeting the lockdown policy.

We conclude this section by discussing several modeling assumptions. First, we have assumed that $\{\beta_{it}\}$ is independently and identically distributed across individuals and over time. This may not be empirically true. For example, a person who was infected in a big party is more likely to go to a party in the next period. This introduces ex ante heterogeneities as discussed in (Gomes et al., 2020; Hébert-Dufresne et al., 2020; Britton et al., 2020), generating positive correlation in $\{\beta_{it}\}$ along the social network. Or, a person who tends to be a superspreader may be more likely to be a superspreader in the next period. This induces a positive correlation in $\{\beta_{it}\}$ over time. If the resulting cascading effect were large, then the average effects on the epidemiological outcomes we have found may be overturned. Second, we have exogenously imposed power law distributions without fully exploring underlying data generation mechanisms behind them. The natural next step is to provide a model in which individual infection rate follows a power law. We believe SIR models with social networks along the line of Pastor-Satorras and Vespignani (2001), Moreno et al. (2002), Castellano and Pastor-Satorras (2010), May and Lloyd (2001), Zhang et al. (2013), Gutin et al. (2020), and Akbarpour et al. (2020) are promising avenue to generate endogenous power law in individual infection rates.

4 Estimation methods

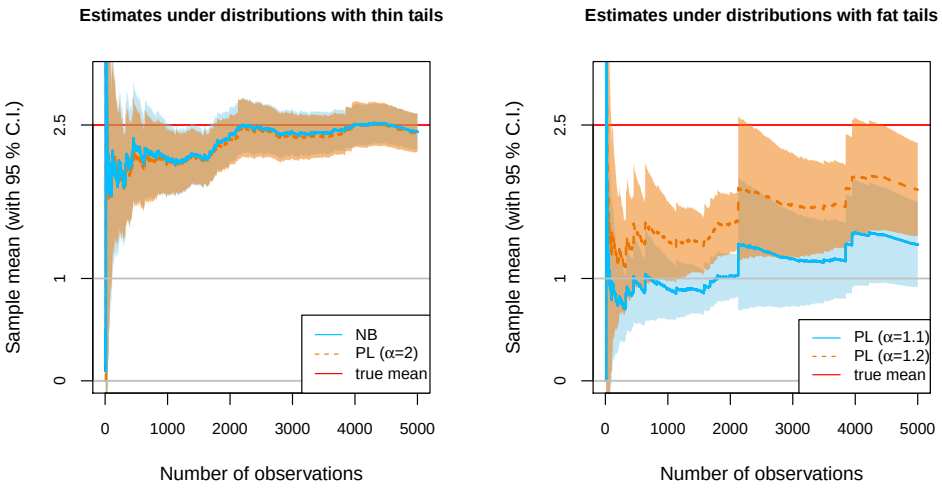
We began with the evidence that SSEs follow a power law distribution with fat tails in many settings, and showed that such distributions substantively change the predictions of SIR models. In this Section, we discuss the implications of power law distributions for estimating the effective reproduction number.

4.1 Limitations of sample means

Estimation of average reproduction numbers ($\mathcal{R}_t$) has been the chief focus of empirical epidemiology research (e.g. Becker and Britton, 1999). Our estimates across five different data sets suggest that the exponent satisfies $\alpha \in (1,2)$ in many occasions: that is, the infection rates have a finite mean but an infinite variance. Since the mean exists, by the Law of Large Numbers, the sample mean estimates (see e.g. Nishiura, 2007) that have been used in the epidemiology

reduce the volatility relative to the unbounded case, and nonetheless, some uncertainties will persist and remain much larger than the predictions of negative binomial distributions.

Covid Economics 30, 19 June 2020: 1-43



(a) Thin tails

(b) Fat tails

Figure 6: An example of sample mean estimates

Notes: Figure 6 depicts an example of sample mean estimates for thin-tailed and fat-tailed distributions. The draws of observations are simulated through the inverse-CDF method, where the identical uniform random variable is applied so that the sample means are comparable across four different distributions. All distributions are normalized to have the mean of 2.5. The negative binomial (NB) distribution has the dispersion parameter $k = 0.16$ taken from (Lloyd-Smith et al., 2005). The range of power law (PL) parameters is also taken from the empirical estimates.

research will be consistent (i.e. converge to the true mean asymptotically) and also unbiased (i.e. its expectation equals the true mean with finite samples.)

Due to the infinite variance property, however, the sample mean will converge very slowly to the true mean because the classical CLT requires finite variance. Formally, while the convergence occurs at a rate $\sqrt{N}$ for distributions with finite variance, or thin tails, it occurs only at a rate $N^{1-\frac{1}{\alpha}}$ for the power law distributions with fat tails, $\alpha \in (1,2)$ (Gabaix, 2011).[21] Under distributions with infinite variance, or fat tails, the sample mean estimates could be far from the true mean with reasonable sample sizes, and their estimated 95 confidence intervals will be too tight. Figure 6 plots a Monte Carlo simulation of sample mean's convergence property. For thin-tailed distributions such as the negative binomial distribution or the power law distribution with $\alpha = 2$, even though the convergence is slow due to their very large variance, they still converge to the true mean reasonably under a few 1,000 observations. In contrast, with fat-

[21] For $\alpha = 1$ exactly, the convergence will occur at rate $\ln N$.

Covid Economics 30, 19 June 2020: 1-43

tailed distributions such as power law distribution with $\alpha = 1.1$ or $\alpha = 1.2$, the sample mean will remain far from the true mean. Their sample mean estimates behave very differently as the sample size increases. Every so often, some extraordinarily high values occur that significantly raises the sample mean and its standard errors. When such extreme values are not occurring, the sample means gradually decrease. With thin tails, such extreme values are rare enough not to cause such sudden increase in sample means; however, with fat tails, the extreme values are not so rare.

4.2 Using power law exponents to improve inference

What methods could address the concerns that the sample mean may be empirically unstable? One approach may be to exclude some realizations as an outlier, and focus on subsamples without extreme values[22]. However, such analysis will neglect major source of risks even though extreme "outlier" SSEs may fit the power law distributions as shown in Figure 1. While estimating the mean of distributions with rare but extreme values has been notoriously difficult[23], there are some approaches to address this formally.

With power law distributions, the estimates of exponent have information that can improve the estimation of the mean. Figure 7 shows that the exponents α can be estimated adequately with reasonable sample sizes.[24] If $\alpha > 2$, as may be the case for the India under strict lockdown, then one can have more confidence in the reliability of sample mean estimates. However, if $\alpha < 2$, the sample mean may substantially differ from the true mean. At the least, one can be aware of the possibility.

One transparent approach is a "plug-in" method: to estimate the exponent $\hat{\alpha}$, and plug into the formula of the mean $\frac{\hat{\alpha}}{\hat{\alpha}-1}\underline{Z}$. This method yields a valid 95 confidence intervals (C.I.) of the median[25] since the estimated $\hat{\alpha}$ has valid confidence intervals.[26] Figure 7 shows the estimation results for the same data with $\alpha = 1.1, 1.2$ as shown in Figure 6. First, while the sample mean in Figure 6 had substantially underestimated the mean, this estimated median is close to

[22]In Japan, the case of over 620 infections in the cruise ship Diamond Princess was excluded from all other analyses.

[23]Consider, for example, a binary distribution of infection rates such that one infects N others with $1/N$ probability, and 0 others with $1 - 1/N$ probability. In this case, the true mean $R_t = 1$. Suppose a statistician observes 10 infected cases for each estimation. If N were 1,000, then with $99(\approx 0.999^{10})$ percent chance, nobody becomes infected so that $\hat{R}_t = 0$, and the estimates' confidence interval will be $[0, 0]$. But with less than 1 percent chance when any infection occurs, $\hat{R}_t$ will be larger than 100. Thus, the 95 percent confidence interval contains the true mean in less than 1 percent of the time. To the best of our knowledge, there is no techniques that can help us completely avoid this problem given the fundamental constraint of small sample size.

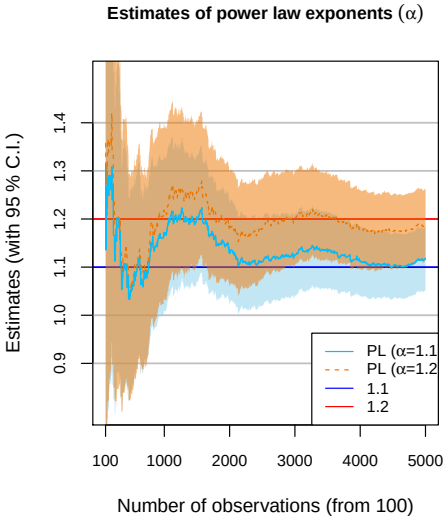
[24]The standard errors are computed by the maximum likelihood approach, as the linear regressions are known to underestimate the standard errors (see Gabaix and Ibragimov, 2011).

[25]Note that the estimate corresponds to the median estimate because $\frac{\hat{\alpha}}{\hat{\alpha}-1}$ is a non-linear transformation of $\hat{\alpha}$.

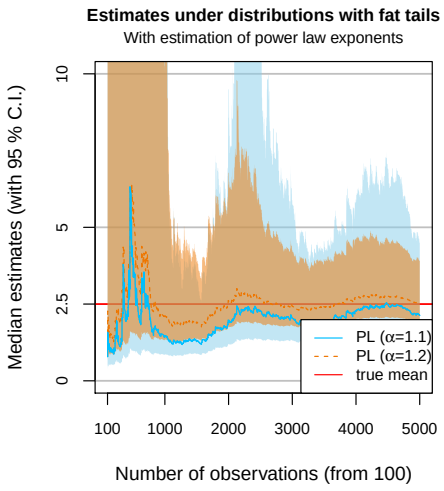
[26]To be more formal, the correct C.I. will be to consider the uncertainties with the mean of observations below $\underline{Z}$. To focus on the uncertainty from upper tail, we construct the 95 percent C.I. from that of the estimate of α here.

Covid Economics 30, 19 June 2020: 1-43

Covid Economics 30, 19 June 2020: 1-43



(a) Power law exponents estimates



(b) Sample median using the estimated exponents

Figure 7: An example of "plug-in" estimates

Notes: Figure 7 plots the estimates of power law exponents and the resulting estimates of sample median, using the same data as in Figure 6. Note that while the number of observations contains all observations, the data points contributing to the estimates are only above some thresholds: only less than 25 percents of the data contribute to the estimation of the exponents.

the true mean. Second, while the sample mean estimation imposed symmetry between lower and upper bounds of 95 percent confidence intervals, this estimate reflects the skewness of uncertainties: upward risks are much higher than downward risks because of the possibility of extreme events. Third, the standard errors are much larger, reflecting the inherent uncertainties given the limited sample sizes.[27] Fourth, the estimates are more stable and robust to the extreme values[28] than the sample mean estimates that have sudden jumps in the estimates after the extreme values.

Table 5 demonstrates the validity of the "plug-in" method through a simulation experiment. The table shows the comparison of the probability that the constructed 95% C.I. covers the true mean using the 1,000 Monte-Carlo simulation. When the estimate is unbiased and has correct

[27]When the number of observations is less than 1000, the estimated confidence interval of α contains values less than 1.0, turning the upper bound of the mean to be ∞. This does not mean that a correct expectation is ∞ infections in the near future, but that there is serious upward risks in infection rates.

[28]This is because the estimation through log-likelihood will take the log of the realized value, instead of its level.

	$\alpha = 1.08$	$\alpha = 1.1$	$\alpha = 1.2$	$\alpha = 1.5$	$\alpha = 2$
1. $N = 100$					
Sample means	21%	26%	42%	74%	89%
Plug-in	98%	98%	98%	94%	87%
2. $N = 500$					
Sample means	24%	29%	45%	78%	90%
Plug-in	98%	98%	95%	94%	84%
3. $N = 1000$					
Sample means	24%	26%	48%	78%	92%
Plug-in	97%	97%	93%	93%	86%

Table 5: Coverage probability of 95% confidence interval

Note: Thable 5 reports the probability that the 95% confidence interval, constructed in two different ways, covers the true value in 1000 simulation. "Sample means" is simply uses the sample mean. "Using power laws uses" first estimates the Pareto exponent using the maximum likelihood, and then convert it to the mean estimates.

Covid Economics 30, 19 June 2020: 1-43

standard errors, this coverage probability is 95%. When the power law exponent is close to one, the traditional "sample means" approach has the C.I. that covers the true mean only with 20-40% for all sample sizes. By contrast the "plug-in" method covers the true estimates close to 95%. As the tail becomes thinner toward $\alpha = 2$, the difference between the two tends to disappear, with "sample mean" approach performing better some times. When the underlying distribution has fat-tails, however, estimation using the plug-in method is preferred.

While the C.I. in the plug-in method has adequate coverage probabilities, it is often very large and possibly infinite. Figure 7 visualizes this. This large C.I. occurs especially when $\alpha \simeq 1$ because the mean of a power law distribution is proportional to $\frac{\alpha}{1-\alpha}$. How could the policymakers plan their efforts do given such large uncertainty in $\mathcal{R}_0$? Given the theoretical results in Section 3 that the epidemiological dynamics will be largely uncertain even when $\alpha \simeq 1$ is perfectly known, we argue that applying the estimated $\mathcal{R}_0$ into a deterministic SIR model will not lead to a reliable prediction. Instead of focusing on the mean, it will be more adequate and feasible to focus on the distribution of near-future infection outcomes. For example, using the estimated power law distribution, policymakers can compute the distribution of the future infection rate. The following analogy might be useful: in planning for natural disasters such as hurricanes and earthquakes, policymakers will not rely on the estimates of average rainfall or average seismic activity in the future; instead, they consider the probabilities of some extreme events, and propose plans contingent on realizations. Similar kinds of planning may be also constructive regarding preparation for future infection outbreaks.

To overcome data limitations, epidemiologists have developed a number of sophisticated methods such as backcalculation assuming Poisson distribution (Becker et al., 1991), and ways

to account for imported cases. There are also a number of methods developed to account for fat-tailed distributions (see e.g. Stoyanov et al., 2010, for a survey), such as tail tempering (Kim et al., 2008) and separating the data into sub-groups (Toda and Walsh, 2015). In the future, it will be important to examine what power law distributions will imply about existing epidemiological methods, and how statistical techniques such as plug-in methods can be combined with epidemiological techniques to allow more reliable estimation of risks.

5 Conclusion: implications for COVID-19 pandemic

Most research on infection dynamics has focused on deterministic SIR models, and have estimated its key statistics, the *average* reproduction number ($\mathcal{R}_0$). In contrast, some researchers have concentrated on SSEs, and estimated the *dispersion* of infection rates using negative binomial distributions. Nonetheless, stochastic SIR models based on estimated distributions have predicted that idiosyncratic uncertainties in SSEs would vanish when the infected population is large, and thus, the epidemiological dynamics will be largely predictable. In this paper, we have documented evidence from SARS, MERS, and COVID-19 that SSEs actually follow a power law distribution with the exponent $\alpha \in (1, 2)$: that is, their distributions have infinite variance, or fat tails. Our stochastic SIR model with these fat-tailed distributions have shown that idiosyncratic uncertainties in SSEs will persist even when the infected population is large, inducing major unpredictability in aggregate infection dynamics.

Since the currently infected population is estimated to be around 3 million in the COVID-19 pandemic,[29] our analysis has immediate implications for policies of today. For statistical inference, the aggregate unpredictability suggests caution is warranted on drawing inferences about underlying epidemiological conditions from observed infection outcomes. First, large geographic variations in infections may be driven mostly by idiosyncratic factors, and not by fundamental socioeconomic factors. While many looked for underlying differences in public health practices to explain the variations, our model shows that these variations may be more adequately explained by the presence of a few, idiosyncratic SSEs. Second, existing stochastic models would suggest that, keeping the distribution of infection rates and pathological environments constant, recent infection trends can predict the future well. In contrast, our analysis shows that even when the average number of new infections may seem to have stabilized at a low level in recent weeks, subsequent waves can suddenly arrive in the future.

Such uncertainties in outbreak timing and magnitude introduce substantial socioeconomic difficulties, and measures to assess and mitigate such risks will be invaluable. The death rate is shown to increase when the medical capacity binds. Thus, reducing uncertainties can reduce average fatality. Furthermore, uncertainties can severely deter necessary investments and im-

[29] According to worldometers.info, the cumulative infection worldwide is 7 million, among which 4 million have already recovered or died, as of June 9, 2020.

Covid Economics 30, 19 June 2020: 1-43

pede planning for reallocation and recovery from the pandemic shocks. To assess such risks, we can estimate the tail distributions to improve our inference on the average number. To address such risks, social distancing policies and individual efforts can focus on large physical gatherings in infection-prone environments. Our estimates suggest, like earthquakes, infection dynamics will be largely unpredictable. But unlike earthquakes, they are a consequence of social decisions, and efforts to reduce SSEs can significantly mitigate the uncertainty the society faces as a whole.

Covid Economics 30, 19 June 2020: 1-43

References

ABBOTT, S., J. HELLEWELL, J. MUNDAY, CMMID NCOV WORKING GROUP, AND S. FUNK (2020): "The Transmissibility of Novel Coronavirus in the Early Stages of the 2019-20 Outbreak in Wuhan: Exploring Initial Point-Source Exposure Sizes and Durations Using Scenario Analysis," *Wellcome Open Research*, 5, 17.

ACEMOGLU, D., V. CHERNOZHUKOV, I. WERNING, AND M. D. WHINSTON (2020): "Optimal Targeted Lockdowns in a Multi-Group SIR Model," Working Paper 27102, National Bureau of Economic Research.

AKBARPOUR, M., C. COOK, A. MARZUOLI, S. MONGEY, A. NAGARAJ, M. SACCAROLA, P. TEBALDI, S. VASSERMAN, AND H. YANG (2020): "Socioeconomic Network Heterogeneity and Pandemic Policy Response," 77.

ATKINSON, A. B., T. PIKETTY, AND E. SAEZ (2011): "Top Incomes in the Long Run of History," *Journal of Economic Literature*, 49, 3–71.

AXTELL, R. L. (2001): "Zipf Distribution of U.S. Firm Sizes," *Science*, 293, 1818–1820.

BARABASI, A.-L. AND J. FRANGOS (2014): *Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*, Basic Books.

BARDINA, X., M. FERRANTE, AND C. ROVIRA (2020): "A Stochastic Epidemic Model of COVID-19 Disease," *arXiv:2005.02859 [q-bio]*.

BARTLETT, M. S. (1949): "Some Evolutionary Stochastic Processes," *Journal of the Royal Statistical Society. Series B (Methodological)*, 11, 211–229.

BEARE, B. K. AND A. A. TODA (2019): "Geometrically Stopped Markovian Random Growth Processes and Pareto Tails," *arXiv:1712.01431 [econ, math, stat]*.

——— (2020): "On the Emergence of a Power Law in the Distribution of COVID-19 Cases," *arXiv:2004.12772 [physics, q-bio]*.

BECKER, N. G. AND T. BRITTON (1999): "Statistical Studies of Infectious Disease Incidence," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61, 287–307.

BECKER, N. G., L. F. WATSON, AND J. B. CARLIN (1991): "A Method of Non-Parametric Back-Projection and Its Application to Aids Data," *Statistics in Medicine*, 10, 1527–1542.

BRITTON, T. (2010): "Stochastic Epidemic Models: A Survey," *Mathematical Biosciences*, 225, 24–35.

Covid Economics 30, 19 June 2020: 1-43

——— (2018): "Basic Stochastic Transmission Models and Their Inference," *arXiv:1801.09594 [stat]*.

BRITTON, T., T. HOUSE, A. L. LLOYD, D. MOLLISON, S. RILEY, AND P. TRAPMAN (2015): "Five Challenges for Stochastic Epidemic Models Involving Global Transmission," *Epidemics*, 10, 54–57.

BRITTON, T., P. TRAPMAN, AND F. G. BALL (2020): "The Disease-Induced Herd Immunity Level for Covid-19 Is Substantially Lower than the Classical Herd Immunity Level," Preprint, Infectious Diseases (except HIV/AIDS).

BROTHERHOOD, L., P. KIRCHER, C. SANTOS, AND M. TERTILT (2020): "An Economic Model of the Covid-19 Epidemic: The Importance of Testing and Age-Specific Policies," SSRN Scholarly Paper ID 3594329, Social Science Research Network, Rochester, NY.

CASTELLANO, C. AND R. PASTOR-SATORRAS (2010): "Thresholds for Epidemic Spreading in Networks," *Physical Review Letters*, 105, 218701.

CLAUSET, A., C. R. SHALIZI, AND M. E. J. NEWMAN (2009): "Power-Law Distributions in Empirical Data," *SIAM Review*, 51, 661–703.

COOPER, L., S. Y. KANG, D. BISANZIO, K. MAXWELL, I. RODRIGUEZ-BARRAQUER, B. GREENHOUSE, C. DRAKELEY, E. ARINAITWE, S. G. STAEDKE, P. W. GETHING, P. ECKHOFF, R. C. REINER, S. I. HAY, G. DORSEY, M. R. KAMYA, S. W. LINDSAY, B. T. GRENFELL, AND D. L. SMITH (2019): "Pareto Rules for Malaria Super-Spreaders and Super-Spreading," *Nature Communications*, 10, 3939.

DAVIES, N. G., P. KLEPAC, Y. LIU, K. PREM, M. JIT, C. C.-. WORKING GROUP, AND R. M. EGGO (2020): "Age-Dependent Effects in the Transmission and Control of COVID-19 Epidemics," *medRxiv*, 2020.03.24.20043018.

ENDO, A., CENTRE FOR THE MATHEMATICAL MODELLING OF INFECTIOUS DISEASES COVID-19 WORKING GROUP, S. ABBOTT, A. J. KUCHARSKI, AND S. FUNK (2020): "Estimating the Overdispersion in COVID-19 Transmission Using Outbreak Sizes Outside China," *Wellcome Open Research*, 5, 67.

FRIEDEN, T. R. AND C. T. LEE (2020): "Identifying and Interrupting Superspreading Events-Implications for Control of Severe Acute Respiratory Syndrome Coronavirus 2," *Emerging Infectious Diseases*, 26, 1059–1066.

GABAIX, X. (2009): "Power Laws in Economics and Finance," *Annual Review of Economics*, 1, 255–294.

Covid Economics 30, 19 June 2020: 1-43

——— (2011): "The Granular Origins of Aggregate Fluctuations," *Econometrica*, 79, 733–772.

GABAIX, X. AND R. IBRAGIMOV (2011): "Rank - 1 / 2: A Simple Way to Improve the OLS Estimation of Tail Exponents," *Journal of Business & Economic Statistics*, 29, 24–39.

GALVANI, A. P. AND R. M. MAY (2005): "Dimensions of Superspreading," *Nature*, 438, 293–295.

GAY, N. J., G. DE SERRES, C. P. FARRINGTON, S. B. REDD, AND M. J. PAPANIA (2004): "Assessment of the Status of Measles Elimination from Reported Outbreaks: United States, 1997-1999," *The Journal of Infectious Diseases*, 189 Suppl 1, S36–42.

GLOVER, A., J. HEATHCOTE, D. KRUEGER, AND J.-V. RÍOS-RULL (2020): "Health versus Wealth: On the Distributional Effects of Controlling a Pandemic," *CEPR Covid Economics: Vetted and Real-Time Papers*, 6, 22–64.

GOLLIER, C. (2020): "Cost-Benefit Analysis of Age-Specific Deconfinement Strategies," *CEPR Covid Economics: Vetted and Real-Time Papers*, 24, 1–31.

GOMES, M. G. M., R. M. CORDER, J. G. KING, K. E. LANGWIG, C. SOUTO-MAIOR, J. CARNEIRO, G. GONCALVES, C. PENHA-GONCALVES, M. U. FERREIRA, AND R. AGUAS (2020): "Individual Variation in Susceptibility or Exposure to SARS-CoV-2 Lowers the Herd Immunity Threshold," *medRxiv*, 2020.04.27.20081893.

GUTENBERG, B. AND C. RICHTER (1954): *Seismicity Of The Earth And Associated Phenomena*.

GUTIN, G., T. HIRANO, S.-H. HWANG, P. R. NEARY, AND A. A. TODA (2020): "The Effect of Social Distancing on the Reach of an Epidemic in Social Networks," *arXiv:2005.03067 [physics, q-bio]*.

HÉBERT-DUFRESNE, L., B. M. ALTHOUSE, S. V. SCARPINO, AND A. ALLARD (2020): "Beyond R_0: Heterogeneity in Secondary Infections and Probabilistic Epidemic Forecasting," *arXiv:2002.04004 [physics, q-bio]*.

HSU, L.-Y., C.-C. LEE, J. A. GREEN, B. ANG, N. I. PATON, L. LEE, J. S. VILLACIAN, P.-L. LIM, A. EARNEST, AND Y.-S. LEO (2003): "Severe Acute Respiratory Syndrome (SARS) in Singapore: Clinical Features of Index Patient and Initial Contacts," *Emerging infectious diseases*, 9, 713.

JESSEN, A. H. AND T. MIKOSCH (2006): "Regularly Varying Functions," *Publications de l'Institut Mathématique*, 80(94), 171–192.

KARAKO, K., P. SONG, Y. CHEN, AND W. TANG (2020): "Analysis of COVID-19 Infection Spread in Japan Based on Stochastic Transition Model," *BioScience Trends*, advpub.

Covid Economics 30, 19 June 2020: 1-43

KENDALL, D. G. (1956): "Deterministic and Stochastic Epidemics in Closed Populations," in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 4: Contributions to Biology and Problems of Health*, The Regents of the University of California.

KERMACK, W. O. AND A. G. MCKENDRICK (1927): "A Contribution to the Mathematical Theory of Epidemics," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 115, 700–721.

KIM, Y. S., S. T. RACHEV, M. L. BIANCHI, AND F. J. FABOZZI (2008): "Financial Market Models with Lévy Processes and Time-Varying Volatility," *Journal of Banking & Finance*, 32, 1363–1378.

KLEIBER, C. AND S. KOTZ (2003): *Statistical Size Distributions in Economics and Actuarial Sciences*, John Wiley & Sons.

KUCHARSKI, A. J. AND C. L. ALTHAUS (2015): "The Role of Superspreading in Middle East Respiratory Syndrome Coronavirus (MERS-CoV) Transmission," *Eurosurveillance*, 20, 21167.

LECLERC, Q. J., N. M. FULLER, L. E. KNIGHT, CMMID COVID-19 WORKING GROUP, S. FUNK, AND G. M. KNIGHT (2020): "What Settings Have Been Linked to SARS-CoV-2 Transmission Clusters?" *Wellcome Open Research*, 5, 83.

LIU, Y., A. A. GAYLE, A. WILDER-SMITH, AND J. ROCKLÖV (2020): "The Reproductive Number of COVID-19 Is Higher Compared to SARS Coronavirus," *Journal of Travel Medicine*, 27.

LLOYD-SMITH, J. O., S. J. SCHREIBER, P. E. KOPP, AND W. M. GETZ (2005): "Superspreading and the Effect of Individual Variation on Disease Emergence," *Nature*, 438, 355–359.

MAY, R. M. AND A. L. LLOYD (2001): "Infection Dynamics on Scale-Free Networks," *Physical Review E*, 64, 066112.

MORENO, Y., R. PASTOR-SATORRAS, AND A. VESPIGNANI (2002): "Epidemic Outbreaks in Complex Heterogeneous Networks," *The European Physical Journal B - Condensed Matter and Complex Systems*, 26, 521–529.

NEWMAN, M. E. J. (2005): "Power Laws, Pareto Distributions and Zipf's Law," *Contemporary Physics*, 46, 323–351.

NISHIURA, H. (2007): "Time Variations in the Transmissibility of Pandemic Influenza in Prussia, Germany, from 1918–19," *Theoretical Biology & Medical Modelling*, 4.

NISHIURA, H., K. MIZUMOTO, AND Y. ASAI (2017): "Assessing the Transmission Dynamics of Measles in Japan, 2016," *Epidemics*, 20, 67–72.

Covid Economics 30, 19 June 2020: 1-43

NISHIURA, H., H. OSHITANI, T. KOBAYASHI, T. SAITO, T. SUNAGAWA, T. MATSUI, T. WAKITA, M. C.-. R. TEAM, AND M. SUZUKI (2020): "Closed Environments Facilitate Secondary Transmission of Coronavirus Disease 2019 (COVID-19)," *medRxiv*, 2020.02.28.20029272.

PASTOR-SATORRAS, R. AND A. VESPIGNANI (2001): "Epidemic Spreading in Scale-Free Networks," *Physical Review Letters*, 86, 3200–3203.

PETROSILLO, N., G. VICECONTE, O. ERGONUL, G. IPPOLITO, AND E. PETERSEN (2020): "COVID-19, SARS and MERS: Are They Closely Related?" *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, 26, 729–734.

RAMPINI, A. A. (2020): "Sequential Lifting of COVID-19 Interventions with Population Heterogeneity," Working Paper 27063, National Bureau of Economic Research.

REMUZZI, A. AND G. REMUZZI (2020): "COVID-19 and Italy: What Next?" *The Lancet*, 395, 1225–1228.

ROBERTS, M., V. ANDREASEN, A. LLOYD, AND L. PELLIS (2015): "Nine Challenges for Deterministic Epidemic Models," *Epidemics*, 10, 49–53.

ROBERTS, M. G. (2013): "Epidemic Models with Uncertainty in the Reproduction Number," *Journal of Mathematical Biology*, 66, 1463–1474.

——— (2017): "An Epidemic Model with Noisy Parameters," *Mathematical Biosciences*, 287, 36–41.

SHEN, Z., F. NING, W. ZHOU, X. HE, C. LIN, D. P. CHIN, Z. ZHU, AND A. SCHUCHAT (2004): "Superspreading SARS Events, Beijing, 2003," *Emerging Infectious Diseases*, 10, 256–260.

SIMHA, A., R. V. PRASAD, AND S. NARAYANA (2020): "A Simple Stochastic SIR Model for COVID 19 Infection Dynamics for Karnataka: Learning from Europe," *arXiv:2003.11920 [math, q-bio]*.

STOYANOV, S. V., S. RACHEV, B. RACHEVA-IOTOVA, AND F. J. FABOZZI (2010): "Fat-Tailed Models for Risk Estimation," SSRN Scholarly Paper ID 1729040, Social Science Research Network, Rochester, NY.

SZABÓ, G. M. (2020): "Propagation and Mitigation of Epidemics in a Scale-Free Network," *arXiv:2004.00067 [physics, q-bio]*.

TODA, A. A. AND K. WALSH (2015): "The Double Power Law in Consumption and Implications for Testing Euler Equations," *Journal of Political Economy*, 123, 1177–1200.

Covid Economics 30, 19 June 2020: 1-43

WANG, H., Z. WANG, Y. DONG, R. CHANG, C. XU, X. YU, S. ZHANG, L. TSAMLAG, M. SHANG, J. HUANG, Y. WANG, G. XU, T. SHEN, X. ZHANG, AND Y. CAI (2020): "Phase-Adjusted Estimation of the Number of Coronavirus Disease 2019 Cases in Wuhan, China," *Cell Discovery*, 6, 1–8.

ZHANG, H., Z.-H. GUAN, T. LI, X.-H. ZHANG, AND D.-X. ZHANG (2013): "A Stochastic SIR Epidemic on Scale-Free Network with Community Structure," *Physica A: Statistical Mechanics and its Applications*, 392, 974–981.

ZIPF, G. K. (1949): *Human Behavior and the Principle of Least Effort*, Human Behavior and the Principle of Least Effort, Oxford, England: Addison-Wesley Press.

Covid Economics 30, 19 June 2020: 1-43

Appendix

A Empirical Appendix

A.1 Relating empirical distribution of Z to theoretical distribution of β_{it}

In this paper, we have used the estimates from the data to simulate the evolution dynamics of the epidemiological model. The key step in our argument is that the tail distribution of $\sum_i z_{it}$ or $\sum_t z_{it}$, the *cumulative* "effective" number of infections, is equivalent to the tail distribution of β_{it}, the *individual and per-period* "basic" number of infection. However, in general, this needs not hold: for example, even if β_{it} were normally distributed (i.e. thin tailed), Z may follow a t-distribution (i.e. fat-tailed). Under what conditions is our interpretation about the relationship between distribution of Z and distribution of β_i valid? Are they plausible in the settings of the coronaviruses?

To clarify this question, let us lay out a model. Formally, Z is a *mixture distribution* of the *weighted sum* of β_{it}. Here, we provide notations for $\sum_t z_{it}$ but the identical argument will also apply to $\sum_i z_{it}$. Specifically, suppose i stays infected for $\bar{t}$ periods, and let the probability mass be $\delta\left(\bar{t}\right)$. In the case of exponential decay as in the SIR model, $\delta\left(\bar{t}\right) = \gamma^{\bar{t}}$. Denoting the counter-cumulative distribution of Z_i by Φ, and that of β_{it} by F, we have

$$\Phi\left(Z_i\right) = \sum_{\bar{t}=1}^{\infty} \delta\left(\bar{t}\right) G_{\bar{t}}\left(\sum_{t=1}^{\bar{t}} \frac{S_t}{N}\beta_{it}\right), \ \beta_{it} \sim F,$$

where $G_{\bar{t}}$ denotes the distribution of $\sum_{t=1}^{\bar{t}} \frac{S_t}{N}\beta_{it}$.

A.1.1 Empirical evidence on causes of SSEs

First, we may be concerned that, even if Φ is a power law distribution, F may not be a power law distribution. A counterexample is that a geometric Brownian motion with stochastic stopping time that follows exponential distribution can also generate power law distributions of the tail (Beare and Toda, 2020). That is, the tail property of Φ needs not be due to tails of F: for $\sum_t z_{it}$, it could also due to some individuals staying infectious for an extremely long periods. For $\sum_i z_{it}$, it could also be due to some events having extremely high number of infected primary cases.

While we acknowledge such possibilities, we argue that for superspreaders or SSEs of the coronaviruses, the main mechanism of extremely high number of cumulative infection is primarily due to some extreme events at particular time t. Let us be concrete. If the counterexample's reasoning were true for $\sum_t z_{it}$, then a superspreader is someone who goes, for example, to

Covid Economics 30, 19 June 2020: 1-43

a restaurant and infect two other people at time t, and then goes to a shopping mall and infects three other people at time $t+1$, and then goes to meet her two friends and infect them, and so on. However, this interpretation is inconsistent with numerous anecdotes. Instead, a super-spreader infects many people because he attends a SSE that has infection-prone environment at a particular time t. Conferences, parties, religious gatherings, and sports gyms are a particular place that can infect many at the same time. Moreover, Nishiura et al. (2020) paper whose data we use has identified particular environment that has caused SSEs. This interpretation is important because, if the extremely high cumulative number of infection were due to some staying infectious for a long time or some events having extremely high number of primary cases, then our model's prediction of sudden outbreak due to SSE is no longer a valid prediction.

A.1.2 Theoretical analysis on interpretation of exponents

Second, we may be concerned that the exponent of $\Phi(Z_i)$ may be different than the exponent of $F(\beta_{i\tau})$, even if both have tails that follow power laws. We use two steps to show that this is not a concern:

(i) if a random variable has a power law distribution with exponent α, then its weighted sum also has a tail distribution that follows a power law with exponent α (see e.g. Jessen and Mikosch (2006) or Gabaix (2009)). Thus, neither summation over multiple periods nor the weights of $\frac{S_\tau}{N}$ will change this.

(ii) the tail property of distribution can be examined by considering $\alpha_F(Z) = \frac{f(Z)}{f(cZ)}$ for some $c \neq 1$ and taking its limit. In particular, if F has a power law distribution, then $\alpha_F(Z) = c^\alpha$.[30] Denoting the probability mass of $G_{\bar{t}}(\cdot)$ by $g_{\bar{t}}(\cdot)$, and the normalizing constant of each $\bar{t}$ by $A_{\bar{t}}$,

$$\lim_{Z\to\infty} \alpha_\Phi(Z) = \frac{\sum_{\bar{t}=1}^{\infty} \delta(\bar{t}) \lim_{Z\to\infty} g_{\bar{t}}(Z)}{\sum_{\bar{t}=1}^{\infty} \delta(\bar{t}) \lim_{Z\to\infty} g_{\bar{t}}(cZ)} = \frac{\sum_{\bar{t}=1}^{\infty} \delta(\bar{t}) A_{\bar{t}} Z^{-\alpha}}{\sum_{\bar{t}=1}^{\infty} \delta(\bar{t}) A_{\bar{t}} (cZ)^{-\alpha}} = c^\alpha.$$

Thus, the exponent of $\Phi(Z_i)$ will be identical to the exponent of $F(\beta_{i\tau})$ asymptotically.

This discussion suggests that whenever possible, it is desirable to take the estimates from the tail end of the distribution instead of using moderate values of Z. For the COVID-19 from the world, the distributions are estimated from the very extreme tail. But when the sample size of SSEs is limited, choice of how many observations to include thus faces a bias-variance trade-off. Nonetheless, as many statistical theories are based on asymptotic results, these arguments show that it is theoretically founded to interpret the exponent of $\Phi(Z_i)$ as the exponent of $F(\beta_{i\tau})$, at least given the data available.

[30] This capture the essence of power laws – that whatever the value of Z, its frequency and frequency of cZ has the same ratio.

A.2 Robustness

We present several robustness checks on our empirical results.

A.2.1 Figure 1 with a different cut-off

In Figure 1, we truncated the size of cluster from below at 40. Figure A.1 instead show results with a cut-off of 20. The fit is worse at the lower tail of the distribution, which suggests that the lower tail may not be approximated by power law distribution. This is a common feature among many examples. However, what matters for the existence of variance is the upper tail distribution, we do not think this is a concern. Moreover, given that the data partly come from media reports, the clusters of small sizes likely suffer from omission due to lack of media coverage.

A.2.2 Robustness of power law exponents estimates

Gabaix and Ibragimov (2011) show that an estimate of 2 is biased in a small sample and propose a simple bias correction method that replace the dependent variable with $\ln(rank - 1/2)$. Panel A of Table A.1 show the results with this bias correction method. The results are broadly very similar to our baseline results in Table 1.

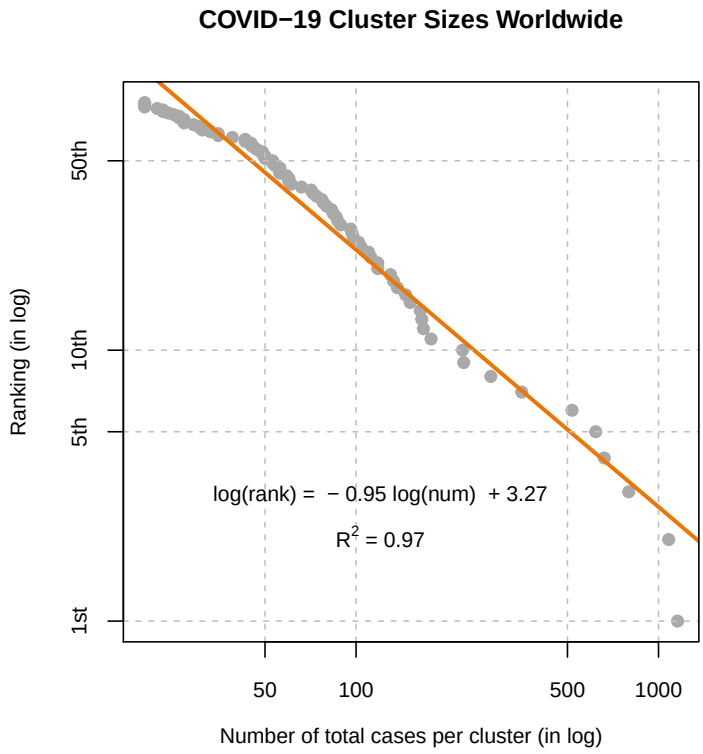
Panel B of Table A.1 conduct another robustness check, where we estimate using the maximum likelihood. Again, the point estimates are overall similar to the baseline results, although standard errors are larger.

A.3 Additional Tables and Figures

Table A.2 shows fseveral examples of superspreading events during COVID-19 pandemic.

Covid Economics 30, 19 June 2020: 1-43

Covid Economics 30, 19 June 2020: 1-43



Source: CMMID COVID−19 Working Group online database (Leclerc et al., 2020)

Figure A.1: Log size vs log rank for Superspreading Events in SARS 2003

Notes: Figure A.1 plots the number of total cases per cluster (in log) and their ranks (in log) for COVID-19, last updated on June 3rd. It fits a linear regression for the clusters with size larger than 20. The data are collected by the Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group (Leclerc et al., 2020).

	Panel A. Bisas corrected regression estimates						
	COVID-19			SARS			MERS
	World	Japan	India	World	Singapore	Beijing	World
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\hat{\alpha}$	1.16	1.45	1.70	1.02	0.86	0.96	1.29
	(0.07)	(0.16)	(0.06)	(0.10)	(0.12)	(0.10)	(0.11)
$\underline{Z}$	40	2	2	8	2	2	2
Obs.	60	11	109	15	19	8	36
R^2	0.97	0.93	0.96	0.95	0.89	0.93	0.95
$\log_{10} LR$	-	11.73	-	-	19.92	8.05	41.19
	Panel B. Maximum likelihood estimates						
	COVID-19			SARS			MERS
	World	Japan	India	World	Singapore	Beijing	World
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\hat{\alpha}$	1.01	1.96	1.71	0.89	1.21	0.87	1.49
	(0.13)	(0.59)	(0.16)	(0.23)	(0.28)	(0.31)	(0.25)
$\underline{Z}$	40	2	2	8	2	2	2
Obs.	60	11	109	15	19	8	36
$\log_{10} LR$	-	11.93	-	-	20.34	8.07	46.93

Table A.1: Estimates of power law exponent: robustness

Notes: Table A.1 summarizes two robustness check exercises of power law exponent ($\hat{\alpha}$). Panel A. bias corrected estimates take $\log(\text{rank}-\frac{1}{2})$ as the dependent variable. This is a small sample bias correction proposed by Gabaix and Ibragimov (2011). Heteroskedasticity-robust standard errors are reported in the parenthesis. Panel B. presents the maximum likelihood estimates. Standard errors are reported in the parenthesis. In both panels, $\log_{10}$(LR) denotes "likelihood ratios", expressed in the log with base 10, of probability of observing this realized data with power law distributions relative to that with estimated negative binomial distributions. Columns (1)-(3) report estimates for COVID-19; columns (4)-(6) for SARS, and column (7) for MERS.

Covid Economics 30, 19 June 2020: 1-43

Major super-spraeding evernts	Confirmed cases	Date
Choir practice in Washington, the US	52	03/10
Conference in Boston, the US	89	02/26
Religious gathering in Daegu, South Korea	49	02/19
Religious gathering in Frankfurt, Germany	49	02/19
Wedding ceremony in New Zealand	76	03/21
Prison in IL, the US	351	04/23
Food processing plant in Ghana	533	05/11
Dormitory in Singapore	797	04/09

Table A.2: Examples of superspreading events

Noes: Table A.2 summarizes some examples of superspreading events, their dates and the number of confirmed cases for COVID-19. *Source:* COVID-19 settings of transmission - database (accessed, June 4, 2020)

B Theory Appendix

B.1 Proof that S_∞ is convex in $\mathcal{R}_0$ if $\mathcal{R}_0 > \frac{9}{8(1-R_0)}$

We show that S_∞ is a concave function in $\mathcal{R}_0$. Recall that S_∞ is a solution to

$$\log S_\infty = -\mathcal{R}_0(1 - S_\infty).$$

By the implicit function theorem,

$$\begin{aligned}\frac{dS_\infty}{d\mathcal{R}_0} &= -\frac{1}{\left(\frac{1}{S_\infty} - \mathcal{R}_0\right)}(1 - S_\infty) \\ &< 0.\end{aligned}$$

because $S_\infty < 1/\mathcal{R}_0$. Applying the implicit function theorem again,

$$\underbrace{\left(\frac{1}{S_\infty} - \mathcal{R}_0\right)}_{>0}\frac{d^2S_\infty}{d\mathcal{R}_0^2} = \underbrace{\frac{dS_\infty}{d\mathcal{R}_0}}_{<0}\left(2 - \frac{1/S_\infty - 1}{1 - \mathcal{R}_0 S_\infty}\right).$$

It remains to show that $\left(2 - \frac{1/S_\infty - 1 - R_0}{1 - \mathcal{R}_0 S_\infty}\right) < 0$. We can rewrite this as

$$f(S_0) \equiv 2\mathcal{R}_0 S_\infty^2 - 3S_\infty + 1 > 0.$$

Note that $f(\cdot)$ is minimized at $S_\infty^* = \frac{3}{4\mathcal{R}_0}$. The minimum value is

$$\min_{S_0} f(S_\infty) = -\frac{9}{8\mathcal{R}_0} + 1.$$

Therefore $f(S_\infty) > 0$ for all S_∞ if and only if $\mathcal{R}_0 > \frac{9}{8}$. This implies that when $\mathcal{R}_0 > \frac{9}{8}$, S_∞ is a concave function of $\mathcal{R}_0$.

B.2 Proof that $I^{\max}$ is concave in $\mathcal{R}_0$ if and only if $\mathcal{R}_0 > \frac{1}{S_0}\exp(0.5)$

Recall that the peak infection rate is given by

$$I^{\max}/N = 1 - \frac{1}{\mathcal{R}_0} - \frac{1}{\mathcal{R}_0}\log(\mathcal{R}_0 S_0).$$

Covid Economics 30, 19 June 2020: 1-43

The derivative is

$$\frac{dI^{\max}/N}{d\mathcal{R}_0} = \frac{1}{(\mathcal{R}_0)^2}\log(\mathcal{R}_0 S_0).$$

The second derivative is

$$\frac{d^2(I^{\max}/N)}{d\mathcal{R}_0^2} = \frac{1}{(\mathcal{R}_0)^3}\left(1 - 2\log(\mathcal{R}_0 S_0)\right),$$

which is negative if and only if $\mathcal{R}_0 > \frac{1}{S_0}\exp(0.5)$.

B.3 Results for targeted lockdown policy experiment

Table B.3 shows the simulation results with lockdown policies targeted at SSEs. $\bar{\beta}$ is the daily upperbound of infection rates due to policies, and we consider cases of $\bar{\beta} = 1000, 100, 50$. As already discussed in the main text, when the distribution is fat-tailed, the targeted policy is not only effective in reducing the mean of the peak infection rate, but also its volatility (the interval between 90 percentile and 10 percentile).

Covid Economics 30, 19 June 2020: 1-43

CEPR PRESS

Covid Economics 30, 19 June 2020: 1-43

	Power law					Negative
	$\alpha = 1.08$	$\alpha = 1.1$	$\alpha = 1.2$	$\alpha = 1.5$	$\alpha = 2$	binomial
1. $\bar{\beta}$: 1000 cases per day						
mean	11%	15%	23%	27%	27%	27%
90th percentile	19%	23%	29%	29%	28%	27%
50th percentile	8%	12%	21%	26%	27%	27%
10th percentile	4%	7%	17%	25%	26%	26%
3. $\bar{\beta}$: 100 cases per day						
mean	9%	12%	20%	26%	27%	27%
90th percentile	17%	20%	26%	27%	28%	27%
50th percentile	5%	8%	18%	26%	27%	27%
10th percentile	3%	5%	16%	24%	26%	26%
3. $\bar{\beta}$: 50 cases per day						
mean	8%	11%	19%	26%	27%	27%
90th percentile	14%	19%	26%	27%	28%	27%
50th percentile	4%	8%	17%	25%	27%	27%
10th percentile	2%	5%	14%	24%	26%	26%

Table B.3: Peak infection under targeted lockdown policy

Note: Table B.3 shows the summary statistics for peak infection rates from 1000 simulations with various policy parameters $\bar{\beta}$, where $\bar{\beta}$ is the upperbound on the infection imposed by the policy.

Mass gathering contributed to early COVID-19 spread: Evidence from US sports[1]

Alexander Ahammer,[2] Martin Halla[3] and Mario Lackner[4]

Date submitted: 13 June 2020; Date accepted: 13 June 2020

Social distancing is important to slow the community spread of infectious disease, but it creates enormous economic and social cost. It is thus important to quantify the benefits of different measures. We study the ban of mass gatherings, an intervention with comparably low cost. We exploit exogenous spatial and temporal variation in NBA and NHL games, which arise due to the leagues' predetermined schedules, and the suspension of the 2019-20 seasons. This allows us to estimate the impact of these mass gatherings on the spread of COVID-19 in affected US counties. One additional mass gathering increased the cumulative number of COVID-19 deaths in affected counties by 13 percent.

Covid Economics 30, 19 June 2020: 44-62

1 Financial support from the Christian Doppler Laboratory "Aging, Health and the Labor Market" is gratefully acknowledged.

2 Assistant Professor of Economics, Johannes Kepler University; Christian Doppler Laboratory on Ageing, Health, and the Labor Market.

3 Professor of Economics, Johannes Kepler University; Christian Doppler Laboratory on Ageing, Health, and the Labor Market; IZA; GöG.

4 Assistant Professor of Economics, Johannes Kepler University; Christian Doppler Laboratory on Ageing, Health, and the Labor Market.

Covid Economics 30, 19 June 2020: 44-62

1. Introduction

Due to the lack of vaccines and effective antiviral drugs, countries have to rely on a set of non-pharmaceutical interventions (NPI) in response to the *Coronavirus Disease 2019* (COVID-19) pandemic. The goal of these measures is to prevent a sharp peak of infections, to take pressure off healthcare systems, and ultimately to save lives. In addition to good personal hygiene and mandatory face masks, social distancing is perhaps the most important NPI (Cowling & Aiello 2020). Since maintaining physical distance inevitably creates enormous economic and social cost, it is crucial to quantify the benefits of different measures in controlling epidemics.

One important public policy to promote physical distancing is to ban mass gatherings (Memish et al. 2019).[1] Such events may foster the transmission of contagious disease as a result of large crowds being in close contact, often for extended periods of time. A temporary mass gathering ban is relatively cheap and easy to implement compared to, for example, school and workplace closures. In response to the spread of SARS-CoV-2, the pathogen leading to COVID-19, several prominent events have been canceled or postponed, even before widespread quarantine measures were enacted (McCloskey et al. 2020). These include religious, cultural, and sporting events.

We quantify how much *National Basketball Association* (NBA) and *National Hockey League* (NHL) games have contributed to the spread of COVID-19 in the United States.[2] Before the leagues suspended play on March 12, up to 12 games per league with an average audience of about 18,000 people were held per day. We analyze how much the number of games held between March 1 and March 11 has contributed to the community spread of COVID-19 in counties surrounding NBA and NHL venues. Since the game schedules were determined long before the first COVID-19 case became public, their spatial and temporal distribution should be unrelated to the initial spread of COVID-19 in the US.[3] In fact, we can show that game schedules are not correlated with observable county characteristics and that game attendance did not systematically change until the NBA and NHL suspended play.

Our results suggest that one additional mass gathering between March 1 and 11 in the form of a NBA or NHL game increased the cumulative number of COVID-19 cases (measured on April 30, 2020) in affected counties by at least 379 per one million population ($p < 0.05$) or 13 percent, and the number of COVID-19 deaths per million by 16 ($p < 0.05$) or 11 percent. These effects are larger in colder regions and in states where shelter-in-place orders (SIPOs) were implemented late. We conclude that banning mass gatherings has an enormous potential to save lives, which is especially important given that such measures are relatively easy and cheap to implement.

[1] The World Health Organization (WHO) describes a mass gathering as "*a planned or spontaneous event where the number of people attending could strain the planning and response resources of the community or country hosting the event. The Olympic Games, The Hajj, and other major sporting, religious, and cultural events are all examples of a mass gathering.*"

[2] We focus on NBA and NHL because their seasons were ongoing when COVID-19 broke out. The *National Football League* (NFL) was in offseason and *Major League Baseball* (MLB) in spring training, which involves scrimmage games in smaller ballparks held in Arizona and Florida.

[3] The first known case in the US was a man in Washington State who returned January 15, 2020 from Wuhan. The NBA 2019/20 schedule had been released on August 12, 2019, the NHL schedule on June 25, 2019.

Our results contribute to the literature evaluating the role of mass gatherings in the spread of infectious disease, and the benefits of social distancing more generally. In a recent survey of the literature, Nunan & Brassey (2020) conclude that the impact of mass gatherings on COVID-19 is still poorly understood. So far, evidence comes almost solely from case reports. For other infectious diseases there is more evidence, but mostly in the form of retrospective observational studies (Rainey et al. 2016, Hoang & Gautret 2018, Karami et al. 2019). We are not aware of any design-based estimation of the impact of mass gatherings on the *community* spread of infectious disease.[4] The best available evidence suggests multiple-day events with crowded communal accommodations are most associated with increased risk of infection (Nunan & Brassey 2020).

Other NPIs have received more attention in the context of COVID-19. These studies differ with respect to outcomes, interventions, and geographic coverage. Gupta et al. (2020) demonstrate how different state- and county-level measures that aim at fostering social distancing have affected people's mobility. The authors proxy mobility with cell signal data, and find SIPOs to have the largest mobility-reducing impact. Two studies examine the impact of SIPOs on COVID-19 cases and deaths. Dave et al. (2020) exploit variation in SIPOs across time and all US states. Their results suggest that approximately three weeks following the adoption of a SIPO, cumulative COVID-19 cases fell by 44 percent. Friedson et al. (2020) focus on California, which was the first state to enact a SIPO. Using a synthetic control design, they find that California's SIPO reduced cases by 125.5 per 100,000 population and deaths by 1,661. Methodologically, all papers use a difference-in-differences approach.[5] Goodman-Bacon & Marcus (2020) provide a critical account of this estimation approach in the context of NPIs.

There is extensive evidence on previous pandemics available. However, the majority of these studies are descriptive in nature. Studying the 1918 influenza pandemic, Markel et al. (2007), Bootsma & Ferguson (2007), and Hatchett et al. (2007), for example, find a strong correlation between excess mortality and how early public health measures were enacted in US cities. It is difficult to infer causality from these results, however, because NPIs are not exogenous and may be enacted in response to preexisting trends in death rates. Barro (2020) attempts to account for this endogeneity by using the distance to army ports in Boston as instrumental variables for NPI introduction. He argues that, because the influenza spread from Boston to other US cities, the farther away cities are from Boston, the more time they had to react and implement NPIs. Barro finds no effect on overall deaths, but that the ratio of peak to average deaths decreased (i.e., a flatter curve). Chapelle (2020) finds a similar pattern using a difference-in-differences model exploiting differences in the timing of NPI introduction. He claims that the lack of herd immunity in subsequent years offset the initial reduction in deaths during the peak of the pandemic, which led to an overall zero effect on deaths.

For recent influenza waves, there is some suggestive evidence that school closures (e.g., Earn 2012, Wheeler et al. 2010) and workplace social distancing (e.g., Ahmed et al. 2018, Miyaki et al. 2011) may be associated with lower disease transmission. However, this literature consists mostly of small case

[4] A notable contribution is Mangrum & Niekamp (2020), who present evidence that college student travel contributed to the spread of COVID-19. Their estimates show that counties with more early spring break students had higher confirmed case growth rates than counties with fewer early spring break students.

[5] There are also a number of non-US studies. Fang et al. (2020) study the case of Wuhan (China), and Hsiang et al. (2020) study localities within China, France, Iran, Italy, South Korea, and the US.

Covid Economics 30, 19 June 2020: 44-62

studies on scheduled school closures (for example, during holidays) or single firms. Viner et al. (2020) conclude that school closures were largely ineffective in controlling *past* Coronavirus outbreaks (i.e., SARS and MERS).

The cost of school and workplace closures are massive. For example, Sadique et al. (2008) estimate that school closures in the US could cost up to £1.2 billion per week. In the early stages of COVID-19, Alexander & Karger (2020) find that people already traveled 9% less and made 13% fewer visits to non-essential businesses. Their preliminary evidence suggests that consumer spending for over 1 million small US business may be reduced by 40%. In a recent survey, respondents reported average wealth losses due to COVID-19 of about $33,000 (Coibion et al. 2020). However, Greenstone & Nigam (2020) find that even a moderate form of social distancing (i.e., isolation of suspect cases and their family members and social distancing of the elderly) can reduce COVID-19 fatalities by almost 1.8 million over the next 6 months, amounting to economic benefits of almost $8 trillion. Similarly, Thunström et al. (forthcoming) estimate the potential benefits of social distancing at around $5.2 trillion.

The remainder of the paper is structured as follows. Section 2 describes our data sources. In Section 3, we present our estimation strategy. Sections 4 and 5 report the main results and a heterogeneity analysis. Section 6 provides concluding comments. Additional figures and tables we delegate to a web appendix.

2. Data

We use information on NBA and NHL games played between March 1 and March 11.[6] During this time span, 78 NBA games (on average about 7 per day) and 57 NHL games (on average 5 per day) were played in US venues. Both leagues suspended all remaining games for the 2019/20 season indefinitely on March 12. The NBA cancelled two games right before tip-off on March 11: Utah Jazz at Oklahoma City Thunder, where Utah player Rudy Gobert tested positive for Sars-Cov-2 prior to the game, and New Orleans Pelicans at Sacramento Kings, due to a suspected infection involving a referee who was part of the officiating crew in a game involving the Utah Jazz earlier the same week.

In our estimation sample, we focus on 38 counties which host either a NBA or a NHL venue, or both, and all their 204 neighboring counties, which we call the 'perimeter' (see Figure A.1 in the Web Appendix for a map). For all affected venue and perimeter counties, we collect information on COVID-19 cases and related deaths.[7] In Figure 1, we show the number of cases (panel a) and deaths (panel b) per million population measured on March 13 (indicated by the left scatter) and on

[6]The information on NBA games is scraped from *Basketball Reference* (see basketball-reference.com). Data on NHL games are collected from *Hockey Reference* (see hockey-reference.com). Since we focus on US territory, we disregard 16 NHL games played in Canada. No NBA game was played in Toronto during the relevant time span.

[7]The information on COVID-19 cases and deaths up to April 30, 2020 is obtained from a database maintained by *The New York Times*, which collects county-level information from reports of state and local health agencies (see nytimes.com/interactive/2020/us/coronavirus-us-cases.html). In our main analysis, we exclude New York City. Data on COVID-19 cases and deaths are not available for city boroughs separately. Adjacent counties in New Jersey and New York are coded as affected by games in New York City.

Covid Economics 30, 19 June 2020: 44-62

Table 1 — Descriptive statistics for main variables

	By county type	
	Venue	Perimeter
Total number of games (NBA + NHL) between March 1–11	3.74	3.48
	(3.06)	(2.71)
Cumulative number of COVID-19 infections†		
On March 13	9.08	4.92
	(14.18)	(13.28)
On April 30	3,846.71	2,834.96
	(4,425.77)	(4,792.83)
Cumulative number of COVID-19 deaths†		
On March 13	0.05	0.02
	(0.20)	(0.22)
On April 30	211.30	139.18
	(329.72)	(249.98)
Total county population (in mio.)	1.79	0.38
	(1.79)	(0.45)
Population density	3,506.84	835.78
	(3,915.39)	(1,588.35)
Population characteristics		
% female	51.20	50.57
	(0.85)	(1.10)
% non-white pop.	33.50	17.03
	(13.43)	(12.89)
% pop. 60+	18.46	21.62
	(2.52)	(5.28)
Number of counties	38	204

Notes: Sample means with standard deviations in parentheses.

† Per one million population.

April 30 (the right scatter) for each venue county, grouped by state, in our data. Additionally, we compute the average number of cases and deaths across each set of neighboring counties. The highest increases are in Essex County, NJ; Orleans Parish, LA; and Suffolk County, MA.

To generate covariates, we collect county-level data on population by age, sex, and ethnicity from the 2016 US census provided by the *National Bureau of Economic Research*.[8] To stratify our analysis, we use, among others, information on population density, climate, and the timing of SIPOs. The information on county land area is collected from the US Census Bureau.[9] Data on historical county climate, including data on April temperature, are collected from the *National Centers for Environmental Information*.[10] Finally, information on the introduction of SIPOs on the state level is taken from Dave et al. (2020). Descriptive statistics for all variables used in our empirical analysis by county type are presented in Table 1.

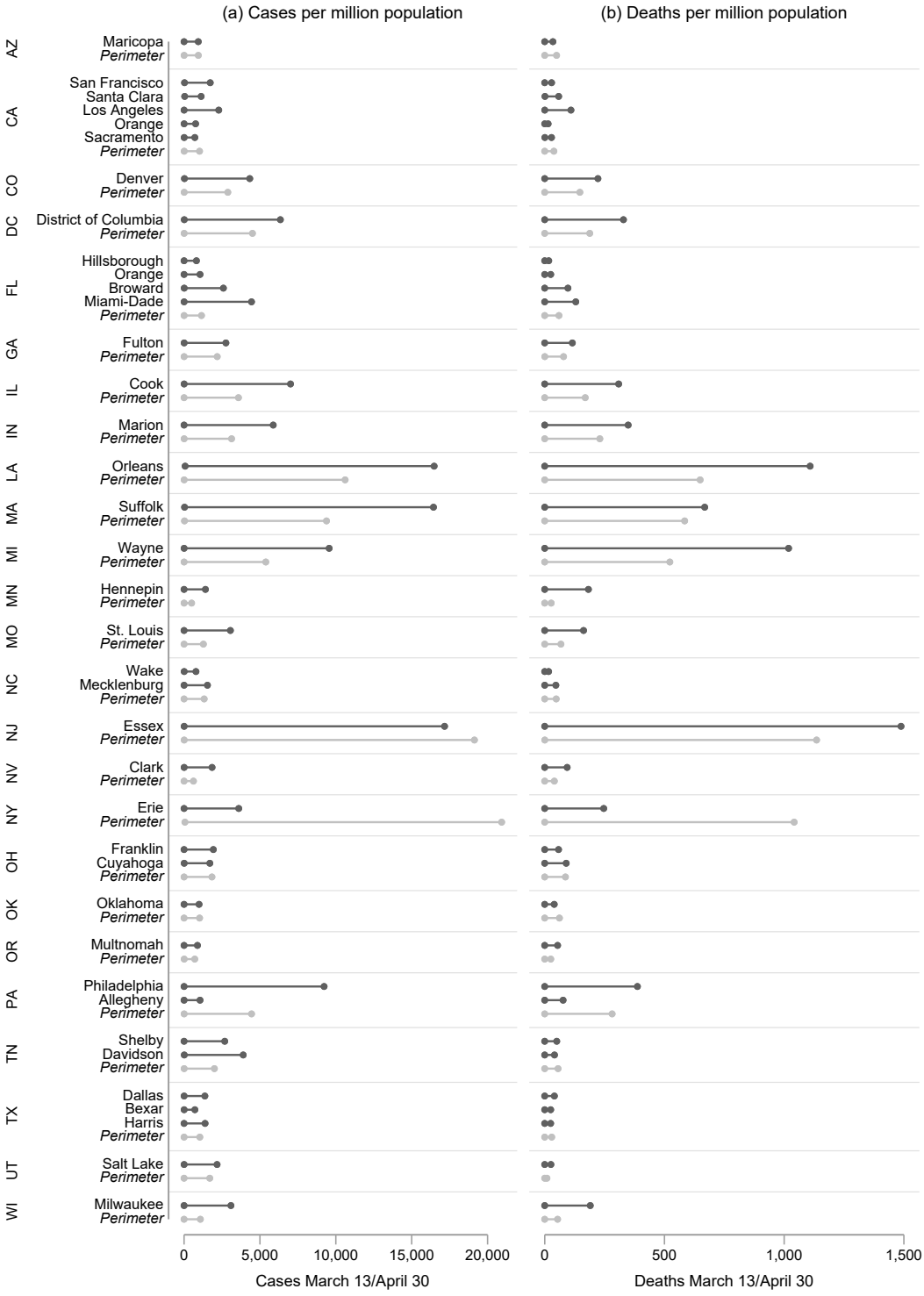
[8]Data are available at data.nber.org/seer-pop/desc/.

[9]Data are available at https://data.census.gov/cedsci.

[10]The NOAA's Climate Divisional Database is available at data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.ncdc:C00005.

Covid Economics 30, 19 June 2020: 44-62

FIGURE 1 — Change in reported COVID-19 cases and deaths per venue and perimeter county over time



Notes: This figures displays the evolution of the number of reported COVID-19 cases (a) and deaths (b) for each venue county in our data. The left dot is always the number of cases per million population on March 13, while the right dot is the number of cases per million population on April 30. Additionally, we calculate averages of cases and deaths from March 13 and April 30 over the counties adjacent to NBA or NHL venues in each state, which we call the 'perimeter.'

Covid Economics 30, 19 June 2020: 44-62

Covid Economics 30, 19 June 2020: 44-62

3. Estimation Strategy

In our estimation analysis, we aim to explain the cumulative number of COVID-19 infections and deaths in a given county c in state s adjacent to, or hosting, venue v. Our sample comprises two types of counties, those which host an NBA or NHL venue (hereafter venue county) and those adjacent to a venue county. This sample definition provides us with a clear match between each county c and venue v. The dependent variable, COVID-19 deaths$_{c,v(s)}$, is defined as the cumulative number of COVID-19 deaths in county c (measured on April 30, 2020) per one million population. The mean of the death rate in venue counties is 211.3 with a standard deviation of 329.7 (see Table 1).

The explanatory variable of primary interest, games$_{c,v(s)}$, varies across venues and measures the cumulative number of games (NBA *and* NHL) at venue v between March 1 and 11. Starting from March 12, both leagues suspended their seasons and all games were canceled.[11] There were on average 12.3 games with considerable variation to exploit. The number of games varies between 0 and 16, with a standard deviation of 3.52. This set-up translates to the following estimation model:

$$\text{COVID-19 deaths}_{c,v(s)} = \beta \cdot \text{games}_{c,v(s)} + \mathbf{X}_c\delta + \sum \gamma_s + \varepsilon_{c,v(s)}, \tag{1}$$

where $\mathbf{X}_c$ are county-level controls, and γ_s are venue-state fixed-effects. Our county-level controls comprise population density and the sex-race-age distribution.

Our main parameter of interest is β, which captures the impact of an additional mass gathering due to a NBA or NHL game on the cumulative number of COVID-19 deaths. Given that the game schedules were determined long before the first COVID-19 case became public, there should be no correlation between games$_{c,v(s)}$ and the error term $\varepsilon_{c,v(s)}$. This identifying assumption is supported by the fact that the number of games does not correlate with observed county characteristics (see Appendix Table A.1). A potential issue for the interpretation of our estimate would be anticipation effects, in the sense that people may had increasingly refrained from visiting games prior to the lockdown. This would lead to an attenuation bias and our results being a lower bound of the actual effect. However, we can show that game attendance did not systematically change before suspension of play (see Appendix Figure A.2).

4. Main Estimation Results

Our estimation results are summarized in Table 2. We estimate the model in equation (1) both on the cumulative number of COVID-19 cases (panel A) and deaths (panel B) per one million population. We find a significant positive effect of the number of mass gatherings on both of these outcomes. Our most conservative estimates indicate that each additional mass gathering between March 1 and 11 increased cases by 379 per one million (column 4) and deaths by approximately 16 per million population (column 3). These are substantial effects. Compared to the average case and death rates

[11] Prior to that only two games on March 11 were cancelled. In both cases, players were tested/suspected for COVID-19.

TABLE 2 — Impact of pre-scheduled mass gatherings on COVID-19 infections and death rate

	(1)	(2)	(3)	(4)	(5)
Panel A. Cumulative number of confirmed COVID-19 infections per million population[a]					
Cum. number of games	485.679**	452.777**	383.526**	379.147**	380.524**
	(208.756)	(202.890)	(183.365)	(174.000)	(174.385)
Population density			0.277*	0.314*	0.320*
			(0.141)	(0.178)	(0.178)
Perc. non-white population			69.612***		
			(17.147)		
Perc. population aged 60+			−0.356		
			(30.150)		
Perc. population female			−288.475		
			(536.027)		
Days since SIPO in place[c]					47.235
					(35.065)
Sex-race-age distribution[e]	No	No	No	Yes	Yes
Venue-state fixed effects	No	Yes	Yes	Yes	Yes
Venue county (1 = yes, 0 = no)	No	No	Yes	Yes	Yes
Panel B. Cumulative number of COVID-19 deaths per million population[b]					
Cum. number of games	29.470***	23.597**	15.825**	16.350**	16.410**
	(11.067)	(9.767)	(7.064)	(6.598)	(6.610)
Population density			0.004	0.012	0.012
			(0.007)	(0.009)	(0.009)
Confirmed COVID-19 cases[d]			4.635***	4.633***	4.656***
			(0.517)	(0.589)	(0.599)
Perc. non-white population			2.997***		
			(0.717)		
Perc. population aged 60+			2.419*		
			(1.316)		
Perc. population female			1.541		
			(9.605)		
Days since SIPO in place[c]					2.858
					(2.126)
Sex-race-age distribution[e]	No	No	No	Yes	Yes
Venue-state fixed effects	No	Yes	Yes	Yes	Yes
Venue county (1 = yes, 0 = no)	No	No	Yes	Yes	Yes

Notes: The number of observations is 242. Robust standard errors are presented in parentheses, stars indicate significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. [a] The dependent variable in Panel A is the number of confirmed COVID-19 infections per million inhabitants on April 30, 2020 with a mean of 2,993.83 (std. dev. 4,742.62). [b] The dependent variable in Panel B is the death rate, defined as the number of COVID-19 deaths per million population on April 30, 2020 with a mean of 150.50 (std. dev. 264.61). The number of games measures all NBA and NHL games which took place between March 1 and March 12. [c] The number of days Shelter-in-Place Orders were active on April 30 2020 in the observed county; source: Dave et al. (2020). [d] The number of confirmed COVID-19 infections per county and 1,000,000 county residents. [e] The sex-race-age distribution is defined as a set of 16 variables capturing the share of the total population of sex g, of race h, and in age-group i, where h is white and non-white, and i is 0 – 19, 20 – 39, 40 – 59, 60+.

across the counties in the data, our estimates correspond to increases of 13 percent and 11 percent per game, respectively. Both are statistically significant at the 5 percent level.

These findings are robust across different specifications. In column (1), we show the unconditional relationship between cases/deaths and games. In column (2), we introduce venue-state fixed effects. In column (3), we additionally include a binary indicator capturing whether the county hosts one or multiple venues, the population density, and the shares of females, people above 60 years of age, and non-whites in the population. In column (4), we alternatively use the full sex-race-age distribution, defined as a set of 16 variables capturing the share of sex g, of race h, and in age-group i in the population; where h is white or non-white, and i is 0–19, 20–39, 40–59, or 60+. In column (5), we

Covid Economics 30, 19 June 2020: 44-62

control also for the number of days a statewide SIPO had been in place in state the observed county c is located in. When analyzing deaths in panel B, columns (3) to (5), we additionally control for the number of confirmed cases by March 13. Our covariates do not have causal interpretations, hence we refrain from interpreting them. However, we note that the negative sign on "Days since SIPO in place" may simply point towards the fact that states with high early case and death counts had to introduce SIPOs sooner.

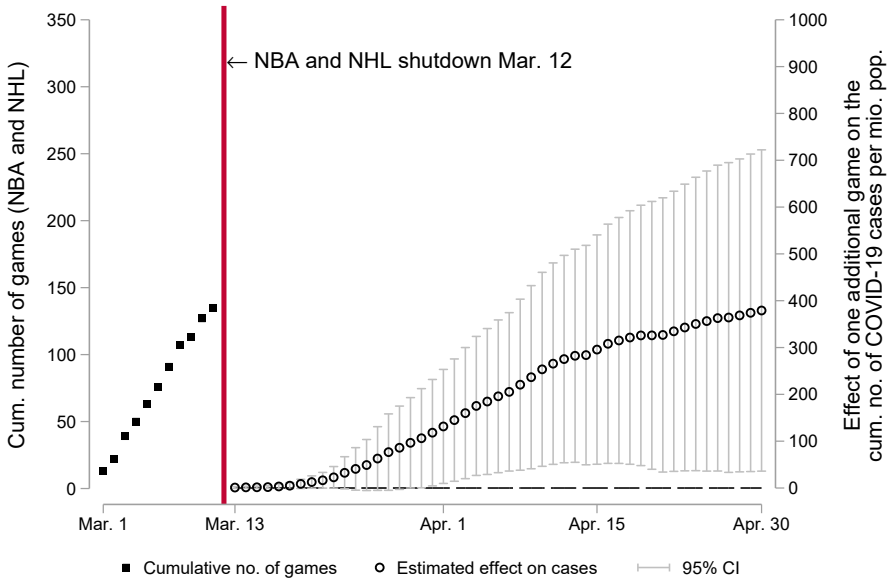
In Figure 2, we provide an overview on the dynamics underlying these effects. The horizontal axis measures time from March 1 to April 30. The squares capture the cumulative number of games (NBA plus NHL) before the leagues suspended play, indicated by the red vertical line. The hollow circles measure the estimated effect of an additional game on the cumulative number of COVID-19 cases (Panel A) and deaths (Panel B) on each day between March 13 and April 30. Each estimate comes from a separate regression, with the dependent variable being measured on different days. The right-most estimate is our baseline. A priori, we expect effects to be strongest around day 14 after the shutdown. This is precisely what we find. The effect of games starts to pick up around March 28 and increases at a decreasing rate since then. This is true for both cases and deaths. Furthermore, we see that cases respond earlier than deaths, which makes sense given the natural lag between diagnosis and death. In terms of magnitudes, estimates for COVID-19 deaths (cases) range between −0.003 (0.996) on March 13 and 16.350 (379.147) on April 30.

Covid Economics 30, 19 June 2020: 44-62

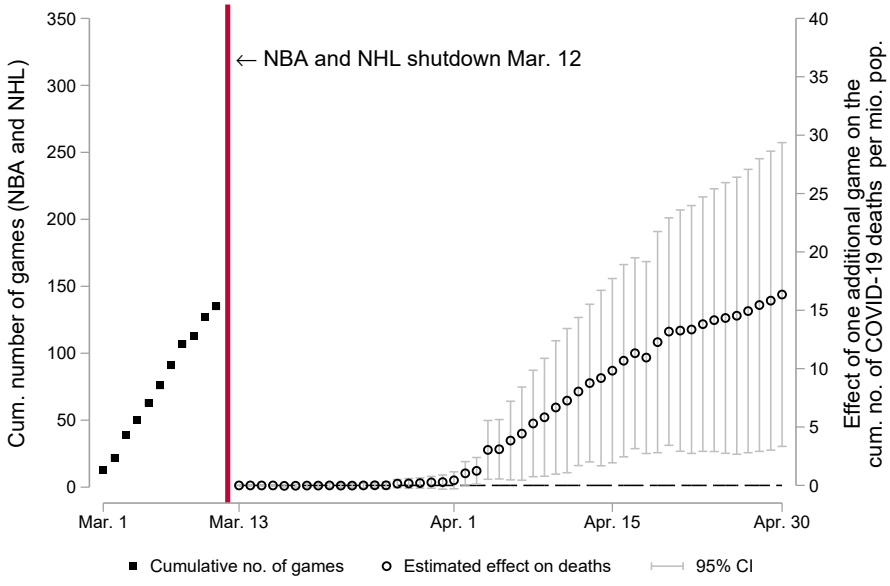
Covid Economics 30, 19 June 2020: 44-62

FIGURE 2 — Estimated effect of non-canceled games on confirmed deaths and cases.

(a) Reported cases per million population

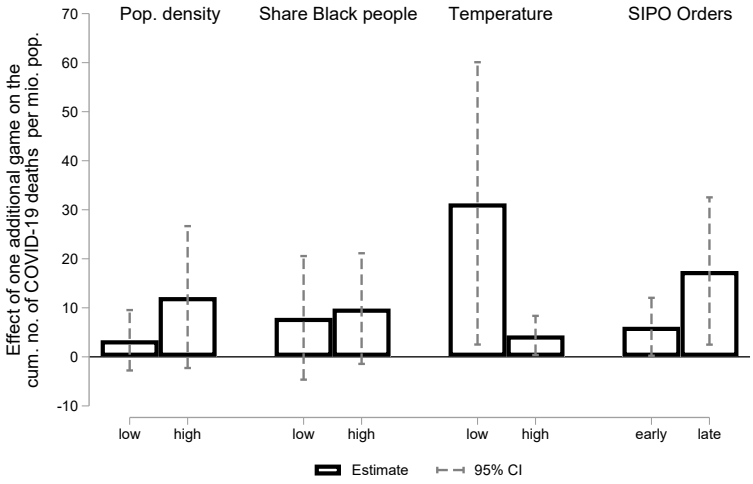


(b) Reported deaths per million population



Notes: The squares capture the cumulative number of games (NBA plus NHL) before suspension of the leagues (red vertical line). The hollow circles measure the estimated effect of one additional game on the cumulative number of COVID-19 cases (Panel A) and deaths (Panel B) on each day between March 13 and April 30. Each estimate comes from a separate regression, with dependent variables measured on different days, and the control variables as in column (4) of Table 2 are used.

FIGURE 3 — Treatment effect heterogeneity



Notes: We replicate our results in sub-samples defined by the median of the respective stratification variable. Estimates base on county level population density are based on the split along the median of the population density distribution of all 242 counties in our data. The county level share of African American population was calculated using 2016 US census data provided by the National Bureau of Economic Research. The third sample split is based on the maximum temperature in April for the 20 most recent years, 1998 – 2019. *low* indicates counties below the median of this long-term temperature median, *high* indicates above the median. Finally, we split along the median of days statewide SIPO regulations were in place by April 30, source: Dave et al. (2020). States without statewide SIPO regulations (MA,MN,OK,TN and UT) are coded 0.

5. Heterogeneity

So far, we have established that mass gatherings in early March increased COVID-19 deaths in counties surrounding NBA and NHL venues by 11 percent. Additionally, we are interested whether there is heterogeneity in these effects by county characteristics. In Figure 3, we therefore stratify our sample by population density, ethnic composition (measured by the share of Black people in the population), average temperature, and policy responsiveness (i.e., when SIPOs were first introduced). We split each variable by its sample median and repeat our regressions from above.

Effects tend to be stronger in counties with a denser population, but the difference to less densely-populated counties is not significant. Splitting the sample by the share of Black people in the population gives a similar picture. This is surprising, given that early reports in the medical literature suggest that Black people tend to be affected more strongly by COVID-19 than other ethnic groups (e.g., Yancy 2020). Splitting by temperature, we find that colder areas clearly drive our effects. In counties with below-median temperatures, the effect on deaths is almost twice as high than in the baseline. This is in line with the idea that the virus replicates more easily in lower-temperature conditions. However, the literature has not yet reached consensus whether this is indeed the case. While some early reports from China document a negative correlation between temperature and COVID-19 spread (e.g., Wang et al. 2020), others find no such (or even a positive) connection (e.g., Yao et al. 2020,

Covid Economics 30, 19 June 2020: 44-62

Ma et al. 2020). Finally, we use the time since statewide SIPO orders were enacted as a measure of policy responsiveness. Here we find that mass gatherings have the largest effects in counties situated in late-adopter states. This is perhaps because the virus can spread more easily without SIPOs in place.

6. Policy Conclusions

In this paper, we present estimates for the impact of mass gatherings in the form of NBA or NHL games on the community spread of COVID-19. We find that one additional game increased the cumulative number of COVID-19 deaths in affected US counties by 11 percent. We conclude that banning mass gatherings is an effective NPI to slow the spread of COVID-19.

Common estimates in NBA circles, for example, suggest that each game yields an average $1.2 million in gate revenue.[12] This figure comprises all game-day revenue, including tickets and concessions, but excludes revenues from TV and sponsoring deals, and the resulting consumer surplus. The latter two components might not be lost if games are played without audience. Thus, even if the full 82 game season in 2020 would have been canceled, this would not have exceeded $100 million in losses, which is considerably less than what estimates suggest school closures in the wake of COVID-19 cost. Since we now know that most of the season is merely postponed instead of canceled completely, we expect the league to recover at least part of these losses.

We suggest that public health officials recommend canceling or postponing mass gatherings during COVID-19 and future pandemics.

Covid Economics 30, 19 June 2020: 44-62

[12] See, for example, nbcsports.com/chicago/bulls/report-nba-could-lose-nearly-500-million-ticket-revenue-wit accessed June 9, 2020.

References

Ahmed, F., Zviedrite, N. & Uzicanin, A. (2018), 'Effectiveness of workplace social distancing measures in reducing influenza transmission: A systematic review', *BMC Public Health* **18**(1), 518.

Alexander, D. & Karger, E. (2020), Do stay-at-home orders cause people to stay at home? Effects of stay-at-home orders on consumer behavior, Report, Federal Reserve Bank of Chicago.

Barro, R. J. (2020), Non-Pharmaceutical Interventions and Mortality in U.S. Cities during the Great Influenza Pandemic, 1918-1919, Working Paper 27049, National Bureau of Economic Research.

Bootsma, M. C. J. & Ferguson, N. M. (2007), 'The effect of public health measures on the 1918 influenza pandemic in U.S. cities', *Proceedings of the National Academy of Sciences* **104**(18), 7588–7593.

Chapelle, G. (2020), The medium-run impact of non-pharmaceutical Interventions: Evidence from the 1918 influenza in US cities, *in* B. W. di Mauro & C. Wyplosz, eds, 'Covid Economics, Vetted and Real-Time Papers', Vol. 18, CEPR Press.

Coibion, O., Gorodnichenko, Y. & Weber, M. (2020), 'The Cost of the COVID-19 Crisis: Lockdowns, Macroeconomic Expectations, and Consumer Spending', *Covid Economics: Vetted and Real-Time Papers* (29), 1–49.

Cowling, B. J. & Aiello, A. E. (2020), 'Public Health Measures to Slow Community Spread of Coronavirus Disease 2019', *Journal of Infectious Diseases* **221**(11), 1749–1751.

Dave, D. M., Friedson, A. I., Matsuzawa, K. & Sabia, J. J. (2020), When Do Shelter-in-Place Orders Fight COVID-19 Best? Policy Heterogeneity Across States and Adoption Time, Working Paper 27091, National Bureau of Economic Research.

Earn, D. J. (2012), 'Effects of School Closure on Incidence of Pandemic Influenza in Alberta, Canada', *Annals of Internal Medicine* **156**(3), 173.

Fang, H., Wang, L. & Yang, Y. (2020), Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-NCOV) in China, NBER Working Paper 26906, National Bureau of Economic Research, Cambridge, MA.

Friedson, A. I., McNichols, D., Sabia, J. J. & Dave, D. (2020), Did California's Shelter-in-Place Order Work? Early Coronavirus-Related Public Health Effects, Working Paper 26992, National Bureau of Economic Research.

Goodman-Bacon, A. & Marcus, J. (2020), Using Difference-in-Differences to Identify Causal Effects of COVID-19 Policies, Working paper, Vanderbilt University.

Greenstone, M. & Nigam, V. (2020), 'Does Social Distancing Matter?', *Covid Economics: Vetted and Real-Time Papers* (7), 1–22.

Gupta, S., Nguyen, T. D., Rojas, F. L., Raman, S., Lee, B., Bento, A., Simon, K. I. & Wing, C. (2020), Tracking Public and Private Responses to the COVID-19 Epidemic: Evidence from State and Local Government Actions, NBER Working Paper 27027, National Bureau of Economic Research, Cambridge, MA.

Hatchett, R. J., Mecher, C. E. & Lipsitch, M. (2007), 'Public health interventions and epidemic intensity during the 1918 influenza pandemic', *Proceedings of the National Academy of Sciences* **104**(18), 7582–7587.

Covid Economics 30, 19 June 2020: 44-62

Hoang, V. & Gautret, P. (2018), 'Infectious Diseases and Mass Gatherings', *Current Infectious Disease Reports* **20**(11), 44.

Hsiang, S., Allen, D., Annan-Phan, S., Bell, K., Bolliger, I., Chong, T., Druckenmiller, H., Huang, L. Y., Hultgren, A., Krasovich, E., Lau, P., Lee, J., Rolf, E., Tseng, J. & Wu, T. (2020), The Effect of Large-Scale Anti-Contagion Policies on the Coronavirus (COVID-19) Pandemic, Working paper, UC Berkeley, Berkeley, CA.

Karami, M., Doosti-Irani, A., Ardalan, A., Gohari-Ensaf, F., Berangi, Z., Massad, E., Yeganeh, M. R., Asadi-Lari, M. & Gouya, M. M. (2019), 'Public Health Threats in Mass Gatherings: A Systematic Review', *Disaster Medicine and Public Health Preparedness* **13**(5–6), 1035–1046.

Ma, Y., Zhao, Y., Liu, J., He, X., Wang, B., Fu, S., Yan, J., Niu, J., Zhou, J. & Luo, B. (2020), 'Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China', *Science of The Total Environment* **724**, 138226.

Mangrum, D. & Niekamp, P. (2020), College Student Contribution to Local COVID-19 Spread: Evidence from University Spring Break Timing, Unpublished manuscript, Vanderbilt University and Ball State University.

Markel, H., Lipman, H. B., Navarro, J. A., Sloan, A., Michalsen, J. R., Stern, A. M. & Cetron, M. S. (2007), 'Nonpharmaceutical Interventions Implemented by US Cities During the 1918-1919 Influenza Pandemic', *Journal of the American Medical Association* **298**(6), 644–654.

McCloskey, B., Zumla, A., Ippolito, G., Blumberg, L., Arbon, P., Cicero, A., Endericks, T., Lim, P. L. & Borodina, M. (2020), 'Mass gathering events and reducing further global spread of COVID-19: A political and public health dilemma', *Lancet* **395**(10230), 1096–1099.

Memish, Z. A., Steffen, R., White, P., Dar, O., Azhar, E. I., Sharma, A. & Zumla, A. (2019), 'Mass gatherings medicine: Public health issues arising from mass gathering religious and sporting events', *Lancet* **393**(10185), 2073–2084.

Miyaki, K., Sakurazawa, H., Mikurube, H., Nishizaka, M., Ando, H., Song, Y. & Shimbo, T. (2011), 'An effective quarantine measure reduced the total incidence of influenza A H1N1 in the workplace: Another way to control the H1N1 flu pandemic', *Journal of Occupational Health* **53**(4), 287–292.

Nunan, D. & Brassey, J. (2020), What is the evidence for mass gatherings during global pandemics? A rapid summary of best-available evidence, Evidence Service, Centre for Evidence-Based Medicine, Nuffield Department of Primary Care Health Sciences, University of Oxford.

Rainey, J. J., Phelps, T. & Shi, J. (2016), 'Mass Gatherings and Respiratory Disease Outbreaks in the United States – Should We Be Worried? Results from a Systematic Literature Review and Analysis of the National Outbreak Reporting System', *PLoS ONE* **11**(8).

Sadique, M. Z., Adams, E. J. & Edmunds, W. J. (2008), 'Estimating the costs of school closure for mitigating an influenza pandemic', *BMC Public Health* **8**(1), 135.

Thunström, L., Newbold, S. C., Finnoff, D., Ashworth, M. & Shogren, J. F. (forthcoming), 'The benefits and costs of using social distancing to flatten the curve for COVID-19', *Journal of Benefit-Cost Analysis* pp. 1–27.

Viner, R. M., Russell, S. J., Croker, H., Packer, J., Ward, J., Stansfield, C., Mytton, O., Bonell, C. & Booy, R. (2020), 'School closure and management practices during coronavirus outbreaks including COVID-19: A rapid systematic review', *Lancet Child & Adolescent Health* **4**(5), 397–404.

Covid Economics 30, 19 June 2020: 44-62

Wang, J., Tang, K., Feng, K., Lin, X., Lv, W., Chen, K. & Wang, F. (2020), High Temperature and High Humidity Reduce the Transmission of COVID-19, SSRN Scholarly Paper ID 3551767, Social Science Research Network, Rochester, NY.

Wheeler, C. C., Erhart, L. M. & Jehn, M. L. (2010), 'Effect of School Closure on the Incidence of Influenza Among School-Age Children in Arizona', *Public Health Reports* **125**(6), 851–859.

Yancy, C. W. (2020), 'COVID-19 and African Americans', *Journal of the American Medical Association* **323**(19), 1891–1892.

Yao, Y., Pan, J., Liu, Z., Meng, X., Wang, W., Kan, H. & Wang, W. (2020), 'No association of COVID-19 transmission with temperature or UV radiation in Chinese cities', *European Respiratory Journal* **55**(5).

Covid Economics 30, 19 June 2020: 44-62

CEPR PRESS

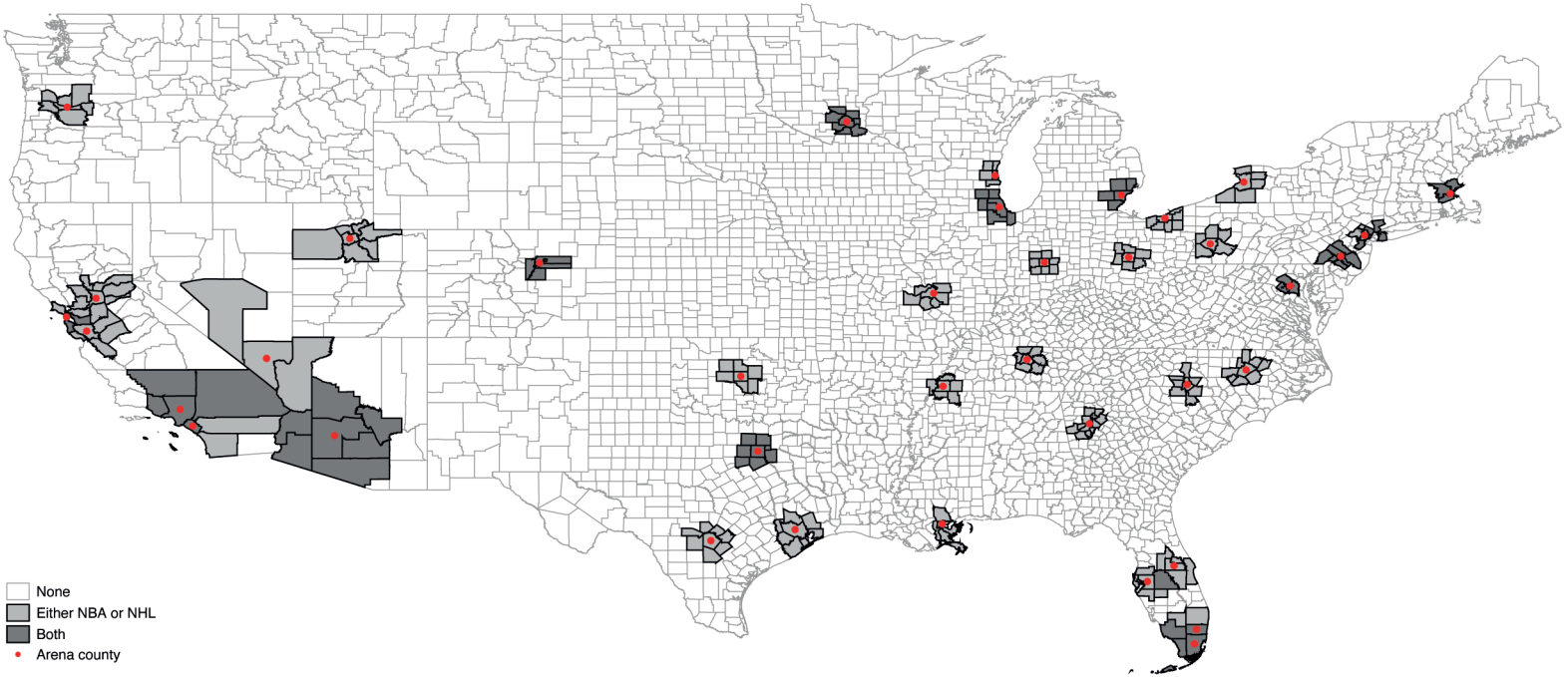
Appendix

This appendix provides additional tables and figures.

Covid Economics 30, 19 June 2020: 44-62

Covid Economics 30, 19 June 2020: 44-62

FIGURE A.1 — NBA and NHL venues and adjacent venues in the United States

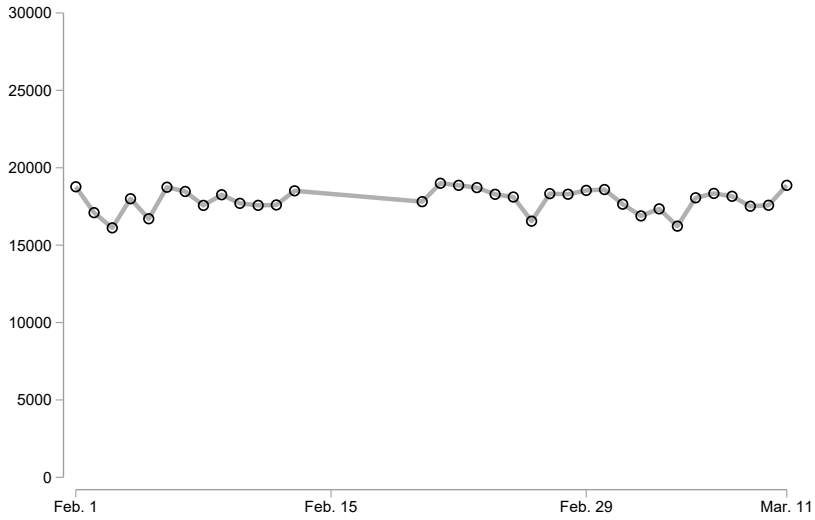


Notes: This map provides an overview on the counties we use in our analysis. The light-gray shaded areas are counties that are adjacent to either a NBA or a NHL venue, the dark-gray shaded counties are in the perimeter of both a NBA and a NHL venue. Counties where venues are located are marked with red dots.

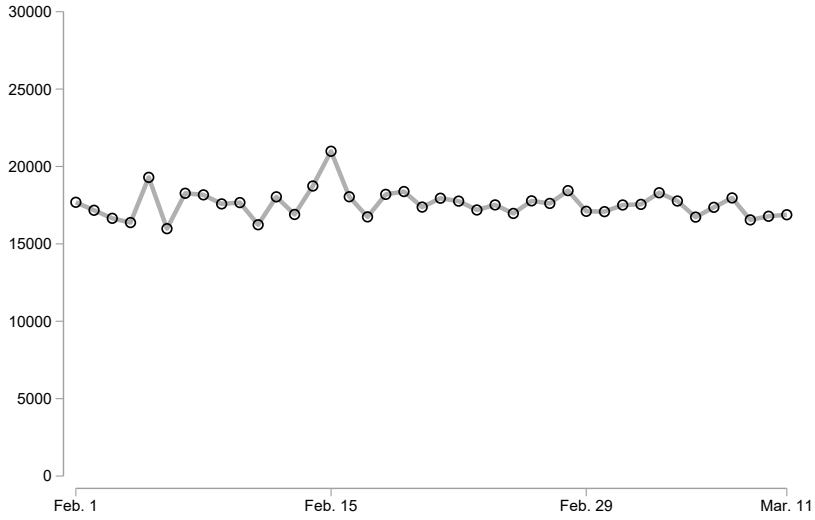
Covid Economics 30, 19 June 2020: 44-62

FIGURE A.2 — Average absolute attendance by NBA and NHL game between February 1 and March 11

(a) NBA



(b) NHL



Notes: This figure displays the average attendance of all NBA and NHL games held between February 1 and March 11 in venues located in the US. On March 12, the NBA cancelled two games before tip-off. After March 11 both leagues suspended the seasons indefinitely.

Covid Economics 30, 19 June 2020: 44-62

TABLE A.1 — Correlation of the number of NBA and NHL games between March 1 and March 11 with observed county characteristics

	(1)	(2)
Population density	0.000**	0.000
	(0.000)	(0.000)
Share female	–0.146	0.249
	(0.160)	(0.160)
Share non-white	0.008	–0.003
	(0.013)	(0.013)
Share elderly	–0.029	–0.050
	(0.030)	(0.035)
Max. Temp. April[a]	–0.021	0.092
	(0.018)	(0.081)
Venue-state fixed effects	No	Yes
Venue county (1 = yes, 0 = no)	No	Yes
Number of Observations	242	242

Notes: The number of observations is 242. Robust standard errors are presented in parentheses, stars indicate significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the number of all NBA and NHL games which took place between March 1 and March 12 in the observed county. Independent variables correspond to those use in specification (4) of Table 2. [a] measures the monthly average maximum temperature in April, based on the 20 most recent years, 1998 – 2019.

Covid Economics Issue 30, 19 June 2020

COVID-19, social distancing, remote work and transport choice

Frank Crowley,[1] Hannah Daly,[2] Justin Doran[3] and Geraldine Ryan[4]

Date submitted: 16 June 2020; Date accepted: 17 June 2020

*Since late 2019, Covid-19 has devastated the global economy, with indirect implications for the environment. As governments' prioritized health and implemented measures such as isolation, the closure of non-essential businesses and social distancing, many workers lost their jobs, were furloughed, or started working from home. Consequently, the world of work has drastically transformed and this period is likely to have major implications for mobility, transportation and the environment. We have estimated the variability of people to engage in remote work and social distancing using O*NET data and Irish Census data. We show that while those who commute by car have a relatively high potential for remote work, they are less likely to be able to engage in social distancing in their workplace. While this may be negative for employment prospects in the short run, this dynamic has the potential for positive environmental implications in the short and long run.*

Covid Economics 30, 19 June 2020: 63-82

1 Lecturer, Spatial and Regional Economics Research Centre, Department of Economics, Cork University Business School, University College Cork.
2 Lecturer, MaREI, the SFI Centre for Climate, Energy and the Marine, Environmental Research Institute, University College Cork.
3 Lecturer, Spatial and Regional Economics Research Centre, Department of Economics, Cork University Business School, University College Cork
4 Senior Lecturer, Department of Accounting and Finance, Cork University Business School, University College Cork.

Covid Economics 30, 19 June 2020: 63-82

1. Introduction

Since late 2019, at the time of writing, Covid-19 has resulted in over 7 million confirmed cases and over 400,000 deaths globally (Worldometer, 2020). Non-essential businesses, schools and public areas closed across many countries, with half of the world's population having spent time under some form of lockdown (Muhammad, Long, & Salman, 2020). Social distancing, self-isolation and travel restrictions have impacted all sectors of the economy (Nicola et al., 2020). While many have lost their jobs, others have switched to working-from-home (Adams-Prassl, Boneva, Golin, & Rauh, 2020; Brynjolfsson et al., 2020; Dingel & Neiman, 2020; Foucault & Galasso, 2020). The related reduction in commuters has had unexpected positive consequences for the environment (Gupta, Tomar, & Kumar, 2020; Helm, 2020; Isaifan, 2020; Muhammad et al., 2020). Now, as governments ease lockdown measures, we examine whether a workers ability to practice socially distancing and to work remotely relates to their commuting choice.

Covid-19 has provided data for an important environmental experiment (Helm, 2020). A consensus has emerged in recent decades about the need to reduce greenhouse gas (GHG) emissions. With the adoption of the Paris Agreement in 2015, 189 countries agreed to limit global warming to between 1.5-2°C above pre-industrial levels. While global integrated assessment modelling (IAM), which underpins the scenarios developed by the Intergovernmental Panel on Climate Change (IPCC) to develop pathways consistent with the Paris agreement, has mainly focussed on CO_2 mitigation from low-carbon fuel and technology measures (Sims et al., 2014), recent focus has turned to the importance of demand-side measures (Faber et al., 2012; Hickel & Kallis, 2020; Mundaca, Ürge-Vorsatz, & Wilson, 2019). Studies have found that down-sizing the global energy system by reducing energy service demands, such as personal transport, will dramatically improve the feasibility of meeting the Paris Agreement temperature targets (Grubler et al., 2018).

Transportation accounts for 23% of global energy-related GHG emissions, is a major source of ambient (outdoor) air pollution, and has many implications for human health.[4] Carbon dioxide (CO_2) emissions from road transport in Ireland amounted to 12 million tonnes in 2018, representing 31% of total emissions (EPA, 2020). Emissions from road transport have grown by nearly 150% since 1990, despite efficiency gains in the vehicle fleet (O'Mahony, Zhou, & Sweeney, 2012). Despite policies to mitigate this growth including a carbon tax, and regulations to mandate and incentivize electric vehicles, biofuels and engine efficiency, emissions from road transport are projected to continue growing (EPA, 2019). This contradicts Ireland's obligation of reduce emissions from transport, heat and agriculture by 30% on 2005 levels by 2030 (European Commission, 2018).

Any changes in commuting behaviour, related to an increase in remote working will have implications for transport emissions and environmental outcomes. The introduction of stay-at-home restrictions, to slow the spread of Covid-19, has greatly reduced transport-related emissions. NASA (2020) and ESA (2020), for example, have shown that lockdown measures have resulted in significant reductions in nitrogen dioxide (NO_2),[5] while the European Environment Agency (2020) found that NO_2 levels have almost halved in many major European cities. As lockdown measures ease, those returning to their workplaces will be required to practice social distancing. They will need to keep at least one metre from anyone outside their household, including their colleagues, customers, and those they interact

[4]The World Health Organisation (2018) estimates that 4.2 million deaths are caused by ambient (outdoor) air pollution each year. While many of these deaths occur in low- and middle-income countries, almost 500,000 occur in Europe, of which 1,180 occur in Ireland (European Environmental Agency, 2019). An additional 1.35 million premature deaths worldwide are caused by road accidents.

[5] Nitrogen dioxide (NO_2) emissions from traffic is a major cause of ambient air pollution (Bishop, Peddle, Stedman, & Zhan, 2010; Muhammad et al., 2020).

with on their daily commute (Block et al., 2020; Galbadage, Peterson, & Gunasekera, 2020; World Health Organization, 2020). Workers who cannot practice social distancing are encouraged to work from home or to remain off work.

The literature teaches us little about the relationship between occupation, commuting choice and the ability to socially distance in the workplace (or to work from home). We know that the majority of workers commute to work by car (90.5% in US (Kopf, 2016); approx. 70% in Europe (Blanckaert, 2019)). We also know that while public transport, walking and cycling are less utilised means of commuting, they are crucial if Europe is to meet its long-term sustainability goals (EEA, 2019). However, we do not know whether workers who drive to work will find it easier to practice social distancing than those who walk, cycle or take public transport. Nor do we know if their occupations can facilitate working from home, or, if they are more likely to lose their jobs. This paper examines the variability of people to engage in remote work and social distancing by commuter type in Ireland.

The rest of the paper is organised as follows. Section 2 presents a brief overview of the related literature. Section 3 presents the data and explains how the social distancing and remote working indices are calculated for Ireland. Section 4 presents the results, while Section 5 concludes the study and presents the implications of our findings.

Covid Economics 30, 19 June 2020: 63-82

2. Literature Review

Covid-19 is severely disrupting labour markets. Since the virus spreads mainly through droplet transmission workers relying on face-to-face communication or close physical proximity in the workplace are particularly vulnerable. While interventions such as social distancing and working from home are deemed necessary responses to the pandemic by governments worldwide, how these labour market disruptions will affect commuting and, in turn, the environment in the longer term are largely unknown.

In commuting choice models, commuters are assumed to make deliberate evaluations around the attributes of the different transport options available to them and to choose the transport mode that will provide them with maximum utility (Ortúzar & Willumsen, 2011). However, it has also been posited that commuting behaviours are habitual and that changes to commuting choice are more likely at the time of major life events such as moving house, job, relationship breakdowns, or the birth of a child (Clark, Chatterjee, & Melia, 2016). The Covid-19 crisis will most likely represent a major life event for many commuters. Stay-at-home and lockdown measures have restricted commuting for all workers, and particularly for those operating in non-essential businesses. Many transport modes such as buses, trains, taxis and underground transport systems represent high-risk environments for the spread of the coronavirus resulting in an increased likelihood of a structural break to commuting habits for many, as they attempt to avoid contracting the virus. Previously, in the US, Kopf (2016) found a link between commuting choice and occupation noting that those working in the *Physical and Social Science* are more likely to cycle to work while those working in *Law*, *Computer Science*, *Finance* and *Media* are more likely to take public transport. However, the pandemic has completely altered normal everyday commuting activities and little is known about the relationship between commuting choice and occupational social distancing, or remote working potential.

Since workplace interactions constitute the majority of social contacts among people of working age (Lewandowski, 2020), those who cannot practice social distancing are advised, where possible, to work from home. The concept of 'social distancing' and the important role it can play in mitigating pandemics is not new to the literature (e.g. Glass, Glass, Beyeler, & Min, 2006; Kelso, Milne, & Kelly, 2009; Rashid et al., 2015). Frequently used interventions include the bans on public events, the closure of schools, universities and non-essential businesses, restrictions on travel, movement, and physical interactions (Block et

al., 2020). The goal of these measures is to delay the spread of influenza while scientists search for a vaccine or treatment. While restrictions on travel and the cancelation of public gatherings have been shown to be effective, the effectiveness of business and school closures is less well understood (Aledort, Lurie, Wasserman, & Bozzette, 2007; Jackson, Vynnycky, Hawker, Olowokure, & Mangtani, 2013; Rashid et al., 2015). Ahmed, Zviedrite, and Uzicanin (2018), focusing on workplace social distancing, note that when the reproduction number (R_0)[6] is low then social distancing in the workplace, combined with other interventions, is effective. More recent studies examining Covid-19 find that workplace social distancing is easier in some occupations (e.g. Agriculture, Forestry and Fishing, Information and Communication) than others (e.g. Retail Trade, Accommodation and Food Services, Human Health) (see Barbieri, Basso, & Scicchitano, 2020; Crowley & Doran, 2020; Koren & Pető, 2020; Mongey, Pilossoph, & Weinberg, 2020). Nevertheless, even if the reproduction number is low and workplace social distancing is feasible, it is likely that people will avoid commuting to work, especially if they have the capacity to conduct their work remotely during Covid-19.

Prior to the pandemic, only 14% of the Irish workforce worked remotely (Redmond & McGuinness, 2020), with higher and lower professional workers being more likely to fall into this category (Fu, Andrew Kelly, Peter Clinch, & King, 2012). Remote working was marginalised by businesses and lacked appropriate regulation and guidelines (Hynes, 2014). The Covid-19 pandemic has shown that a greater proportion of jobs can be done from home, including those in the *Educational Services*, *Professional, Scientific, and Technical Services*, *Finance and Insurance*, and *Information Technology* sectors (Crowley & Doran, 2020; Dingel & Neiman, 2020; Gallacher & Hossain, 2020; Gottlieb, Grobovšek, & Poschke, 2020; Mongey et al., 2020; Mongey & Weinberg, 2020; Montenovo et al., 2020). Restricted movement in Ireland resulting from the pandemic has forced and accelerated digital and remote working adoption by employers and employees, with 34% now stating that they are working from home (CSO, 2020). Until a vaccine to protect against Covid-19 is developed and freely available, many workers will be required to continue working from home. In the long term, many of these workers may choose to continue working from home for family, health, and productivity reasons (Kramer & Kramer, 2020).

Remote work is often suggested by policymakers as a remedy for reducing the environmental and socio-economic impacts of transport and mobility patterns on society (Cerqueira, Motte-Baumvol, Chevallier, & Bonin, 2020; Hynes, 2014; Van Lier, de Witte, Macharis, & Research, 2014). Remote working has been identified to reduce travel distances and number of trips (Choo, Mokhtarian, & Salomon, 2005; Helminen & Ristimäki, 2007). Previously, using the Irish case, Fu et al. (2012) identified that at least an average net saving of 9.33 kWh per day can be achieved from an individual converting to working from home. If the negative 'rebound effects' that have been identified with remote working, such as increases in personal non-work trips and residential relocation, can be minimised (Cerqueti, Correani, & Garofalo, 2013; He & Hu, 2015), then the Covid-19 crisis may represent a strategic opportunity for policymakers to transform transportation networks. This could in turn reduce congestion and the negative externalities associated with transportation, such as greenhouse gas emissions. If those who normally drive can work remotely, then traffic pollution caused by commuters should reduce in the short and the long term, thereby resulting in a positive impact for the environment without harming the economy.

Next, we outline the data we use to examine the variability of people to engage in remote work and social distancing by commuter type in Ireland.

[6] R_0 is the average number of secondary cases produced by a typical infectious case in a fully susceptible population.

Covid Economics 30, 19 June 2020: 63-82

Covid Economics 30, 19 June 2020: 63-82

3. Data

This paper makes use of two datasets. Firstly, similar to many papers in this field of research information is acquired which provides information on worker tasks, context and activities from O*Net. This data facilitates the generation of social distancing and remote working potential indices. Secondly, we utilise 2011 Irish Census data from IPUMS international to examine what types of people are less (or more) exposed from social distancing and remote work. Specifically, we focus on the primary method of transport individuals use to go to work in order to gain a better understanding of the potential impact of social distancing and remote working measure on their jobs and their travel patterns.

3.1 Social Distancing and Remote Working Index

The social distancing and remote working indexes that we use are constructed by Crowley and Doran (2020) and Crowley, Doran, and Ryan (2020) and are detailed in these papers. We present a concise summary of these indexes and their construction here and refer interested readers to the Appendix[7] where we provide details on the precise O*Net questions used in the construction of the indexes by Crowley and Doran (2020) and Crowley et al. (2020).

The indexes are based on O*NET data, which is the primary source of occupational information in the United States and is used to understand the changing world of work and how it impacts on the workforce and the economy. It has been used extensively in occupational studies of automation (Frank, Sun, Cebrian, Youn, & Rahwan, 2018; Frey, Berger, & Chen, 2017; Frey & Osborne, 2017) but recently also in the identification of the occupations which are most likely to be disrupted due to (i) enforced social distancing protocols in workplaces and (ii) the ability/lack of ability to engage in remote working (Delaporte & Peña, 2020; Dingel & Neiman, 2020; Gottlieb et al., 2020; Mongey et al., 2020; Mongey & Weinberg, 2020). This data is based on US occupational codes and therefore it is necessary to translate this to Irish occupational classifications, which is completed by Crowley and Doran (2019) and results in 318 detailed occupations for the Irish case.

The social distancing index is constructed based on work by Koren and Pető (2020), who constructed a social distancing index for the U.S. The index is comprised of information from 15 different questions using O*Net data. A detailed list and the precise questions and coding are displayed in Appendix 2. There are three broad categories to which the underlying questions used in the index relate, which include teamwork requirements, customer orientation and physical presence. A further underlying commonality of the questions is how they relate to the degree to which face-to-face interaction is required for each occupational role and in turn, the ability of workers with the associated occupation to engage in social distancing measures in a workplace. Each variable takes a value ranging from 0 to 100 and an unweighted average of the social distancing indicator is used as a measure of social distancing potential for each occupation. The higher the value of this index then the less team work intensive, customer's contact-intensive or physical presence intensive the job is.[8]

The remote working index is based on work by Dingel and Neiman (2020). The construction of the index again exploits O*Net data using information from 17 questions. The precise questions obtained from O*Net which comprise this index are presented in Appendix 3. In summary, the questions contain data that relates to workers being able to use remote communications such as e-mail, whether the job requires the operation of specialised equipment, the degree to which workers need to use protective equipment and whether the

[7] Appendix 1 details the matching of US occupational codes to Irish occupational codes. Appendix 2 describes the O*Net questions used to generate the social distancing potential index. Appendix 3 describes the O*Net questions used to generate the remote working potential index.

[8] This is the inverted form of the social distancing measure presented by Koren and Peto (2020), where the potential to social distance in a job ranges from low to high.

worker performs for people or directly serves customers. Again, the values range for each indicator from 0 to 100 and the unweighted average of the 17 indicators is used as the indicator for remote working potential value for each occupation. The higher the value of this index then the greater the potential to be able to work from home.

3.2 Sample of Ammonised Individual Level Irish Census Data 2011

For our individual level data, we obtain data from IPUMS International for the Irish Census of 2011. We are specifically interested in those in the sample who are in employment, which results in 156,287 observations of individuals in Ireland in the year 2011. In the 2011 Census, 1,807,369 individuals were identified as being in employment indicating that our sample captures 8.64% of the entire Irish workforce. The sample is a random sample of the total population and is designed to provide a representative sample. In order to match in the social distancing and remote working index for each individual we match these at the three-digit occupational code level with each individual in the Census sample.

The primary explanatory variable of interest in our analysis is the means of transport to work variable. In the 2011 Irish Census respondents were asked "How do you usually travel to work, school or college?" A total of 10 travel options were provided and the IPUMS data condenses these to the categories presented in Table 1. An 'other' category is also included but this is excluded from our analysis as this captures a variety of alternative (non-comparable) transport options.[9] We note that the primary method of transport for individuals is auto (driver), accounting for a total of 71.53% of all individuals in our sample. This is followed by walking, bus or trolley bus, auto (passenger), railroad or train, bicycle, and finally, motorcycle or moped.

Table 1: Descriptive Statistics for Transport Choice

Means of transportation to work	Freq.	Percent
Walking	14,836	11.03
Auto (driver)	96,211	71.53
Auto (passenger)	6,003	4.46
Motorcycle or moped	743	0.55
Bicycle	3,612	2.69
Bus or trolley bus	8,038	5.98
Railroad or train	5,065	3.77

In addition to this, the Irish census provides detailed information on a variety of socio-economic characteristics which may also impact individual's abilities to socially distance in their work place or engage in remote work. Therefore, we also include information on the gender of individuals, marital status, their highest level of educational attainment, whether the individual has a disability or not, the age of the individual, the NUTS3 region in which the individual lives, whether they are a national or non-national and finally the NACE sector in which the individual is employed.

4. Empirical Model

When considering the relationship between individuals usual transport to work choice and their ability to engage in social distancing in their workplace and their ability to work remotely, we estimate equation (1) below. We estimate this model twice, once for each index.

[9] The number of individuals in this 'other' category is 11,645.

Covid Economics 30, 19 June 2020: 63-82

$$Index_i = \beta_0 + \beta_1 TransportChoice_i + \beta_2 Z_i + \varepsilon_i \quad (1)$$

Where, $Index_i$ is the dependent variable, which is the relevant index in equation (either the social distance or remote working index). $TransportChoice_i$ is a series of dummy variables indicating the usual method of transport individual i uses to commute to work. Z_i is a matrix of control variables which were outlined in Section 3 previously. The β values are the coefficients of the model. ε_i is the error term.

The model is estimated using ordinary least squares (OLS) with heteroskedastic robust standard errors. Variance of Inflation (VIF) tests for potential multicollinearity are applied and in all cases report a mean VIF of below 5, suggesting that multicollinearity is not a problem within the model.

5. Results

Table 2 presents the results of our estimations of equation (1).

Table 2: Estimation Results

VARIABLES	(1) Social Distance Index	(2) Remote Work Index
Means of transportation to work - base Auto (driver)		
Walking	0.396***	-0.151**
	(0.0497)	(0.0616)
Auto (passenger)	0.245***	-1.052***
	(0.0727)	(0.0901)
Motorcycle or moped	0.626***	-0.463*
	(0.199)	(0.246)
Bicycle	0.675***	-0.331***
	(0.0929)	(0.115)
Bus or trolley bus	0.850***	0.361***
	(0.0656)	(0.0813)
Railroad or train	1.250***	1.110***
	(0.0810)	(0.100)
Female	1.530***	3.352***
	(0.0322)	(0.0399)
Marital status		
Married	-0.0381	0.329***
	(0.0378)	(0.0469)
Separated (including divorced)	-0.0549	-0.0804
	(0.0713)	(0.0884)
Widowed	0.173	-0.119
	(0.146)	(0.181)
Highest level of education completed		
Lower secondary	0.246***	1.583***
	(0.0837)	(0.104)
Upper secondary	0.921***	4.078***
	(0.0773)	(0.0958)
Third level, non-degree	0.852***	5.885***
	(0.0947)	(0.117)
Third level, degree or higher	1.357***	8.292***
	(0.0798)	(0.0989)
Class of worker		
Self-employed with paid employees	-1.109***	2.048***
	(0.0701)	(0.0869)
Self-employed without paid employees	-0.582***	-0.686***
	(0.0670)	(0.0831)
No disability	-0.0952	-0.0440

Covid Economics 30, 19 June 2020: 63-82

	(0.0641)	(0.0794)
Age		
Age	0.0428***	0.0734***
	(0.00923)	(0.0114)
Age Squared	-0.000271**	-0.000359***
	(0.000106)	(0.000131)
Region		
Dublin	0.245***	1.025***
	(0.0559)	(0.0693)
Mid-East	-0.0260	0.330***
	(0.0639)	(0.0792)
Midlands	-0.335***	-0.142
	(0.0784)	(0.0972)
Mid-West	-0.164**	0.0107
	(0.0714)	(0.0885)
South-East	-0.156**	-0.232***
	(0.0672)	(0.0834)
South-West	-0.116*	-0.00360
	(0.0614)	(0.0761)
West	-0.135**	-0.289***
	(0.0680)	(0.0843)
Irish	0.402***	1.953***
	(0.0415)	(0.0515)
Constant	43.45***	35.34***
	(0.260)	(0.322)
Observations	134,508	134,508
R-squared	0.283	0.441

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Tables 3 and 4 present a summary of a series of t-tests for each transport coefficient against the others to determine a hierarchy in terms of which modes of transport are associated with individuals who have the greatest potential to engage in social distancing in work or work remotely. A plus sign (+) indicates that an individual travelling by mode of transport (listed in column 2) is more likely to be able to engage in social distancing/remote work than an individual in the corresponding mode of transport (listed in columns 3 through 8). A negative sign (–) indicates that the individual travelling by mode of transport (listed in column 2) is less likely to be able to engage in social distancing/remote work than an individual in the corresponding mode of transport (listed in columns 3 through 8). An equals sign (=) indicates that the individual travelling by mode of transport (listed in column 2) is just as likely to be able to engage in social distancing/remote work than an individual in the corresponding mode of transport (listed in columns 3 through 8).

If we take row (5) in Table 3, as an example, we observe that individuals who commute to work by bike are more likely to be able to engage in social distancing in their workplace than those who commute by walking, auto (driver), or auto (passenger), but they are just as likely to be able to socially distance as those who commute by motorcycle or moped. The results clearly show that versus all other modal choices auto (drivers) are the least able to engage in social distancing at work. While those who use public transport in the form of bus or trolley bus, and railroad or train are the most likely to be able to engage in social distancing in their workplace.

Regarding Table 4, which shows the corresponding t-tests for remote working, we observe significant differences in the patterns which are emerging relative to our social distancing index. We note again that individuals who commute by railroad or train commuters are more likely to be able to remote work than any other type of commuter. Further, in this remote working index, auto (drivers) fair reasonable well. However, we

Covid Economics 30, 19 June 2020: 63-82

observe that relative to those who walk, are auto (passengers), use motorcycles or mopeds, or use bicycles; individuals who commute by auto (driver) have higher remote work potential.

Table 3: T-test of differences for transport variables for social distancing index

Row		Walking	Auto (driver)	Auto (passenger)	Motorcycle or moped	Bicycle	Bus or trolley bus
(1)	Walking						
(2)	Auto (driver)	-					
(3)	Auto (passenger)	-	+				
(4)	Motorcycle or moped	+	+	+			
(5)	Bicycle	+	+	+	=		
(6)	Bus or trolley bus	+	+	+	=	=	
(7)	Railroad or train	+	+	+	+	+	+

Table 4: T-test of differences for transport variables for remote working index

Row		Walking	Auto (driver)	Auto (passenger)	Motorcycle or moped	Bicycle	Bus or trolley bus
(1)	Walking						
(2)	Auto (driver)	+					
(3)	Auto (passenger)	-	-				
(4)	Motorcycle or moped	=	-	+			
(5)	Bicycle	=	-	+	=		
(6)	Bus or trolley bus	+	+	+	+	+	
(7)	Railroad or train	+	+	+	+	+	+

6. Conclusions and Discussion

Policies to reduce road transport emissions have to date not addressed structural mobility patterns, which encompass the number and distance of trips people make and their choice of mode. The shutdown of economic and social activity as part of measures to contain the spread of Covid-19 in 2020 have caused a dramatic change in mobility patterns and may represent a game changing situation for policymakers: within two days of the first stage of lockdown on March 12th, the number of cars counted by traffic cameras halved. Ultimately, car traffic fell by nearly 80% in April. Nitrous dioxide pollution detected from air quality monitoring stations halved, while the reduction in road traffic is estimated to reduce Ireland's energy-related CO_2 emissions by 1.5 million tonnes in 2020, representing 5% of all energy-related emissions (Glynn et al., 2020).

Given this dramatic structural break in mobility and work practice, the question is whether emissions savings can be achieved from increased home working in the short and long term. In this paper, we examined, by commuter type, the degree to which there is variability in individuals' ability to engage in social distancing in their workplace and in their ability to remote work. Our analysis used O*Net data to construct social distancing and remote work indexes and Irish Census data to identify commuting choice and other individual factors that may affect social distancing and remote working potential.

Our results suggest short and long run direct implications for commuting outcomes. In the short term, individuals who commute by train or bus are more likely to be able to remote work and to practice occupational social distancing than any other type of commuter. Since

Covid Economics 30, 19 June 2020: 63-82

the perceived risk of contracting the virus is higher by going by train or bus, these commuters may switch to car commuting, which could have a negative impact on the environment. This is something policymakers will need to consider in the short run, as the economy starts opening up. Further, in the short term, where social distancing measures are set to continue in workplaces, individuals who commute by car are less likely to be able to engage in social distancing in their workplace. This means that their physical presence in work may be restricted. As a result, it is likely that the number of commuters by car for work purposes will fall in the short term. While this will be negative in terms of the employment prospects of the individuals, it may have positive environmental impacts due to a reduction in car usage and greenhouse gas emissions. When this finding is coupled with the insights gained from the analysis on remote working, further potential long-term impacts emerge. Our analysis shows that those who commute by car have a relatively high potential for remote work. This suggests that in the short term these individuals, who may not be able to attend work physically due to social distancing regulations, have a greater capacity work from home.

In the longer term, if this trend of working from home was to be incentivized to last beyond the Covid-19 time horizon this would suggest that there would be potential to continue to reduce car commuting for work purposes, without a negative impact on employment outcomes. This would result in significant benefits for the environment as car transport is a major contributor to greenhouse gas emissions and is the only key sector where total emissions have increased in the last three decades (European Environmental Agency, 2018; Hulshof & Mulder, 2020). Previously, in the Irish case, home working was identified as a valuable energy saving option. Fu et al. (2012) estimated that a 5 per cent shift of the regular commuting workforce to full time remote working would result in net energy savings of 0.36 per cent relative to total transport energy use. Although, increased remote working can have immediate and direct environmental benefits, as evidenced by the Covid-19 crisis (European Environment Agency, 2020), in the longer run, indirect costs or 'rebound effects,' associated with non-urban residential relocation, an increase in non-work related trips, car dependency and a modification of consumption patterns for remote workers, may nullify any benefits (Cerqueira et al., 2020; He & Hu, 2015; Ory & Mokhtarian, 2006; Pérez, Sánchez, de Luis Carnicer, & Jiménez, 2004). Consequently, policymakers need to carefully consider policy interventions and incentives that promote remote working, induced by the Covid-19 crisis, to ensure a positive net environmental impact.

There are two limitations with this study that we would like to note. Firstly, the occupational data used for constructing the social distancing and remote working index uses U.S. O*NET data on the description of tasks associated with each occupation. As a result, this should be viewed as an approximation of the workplace environment for the same occupations in Ireland. Secondly, the individual data is not from the most recent Irish Census as ammonised individual level data and individual level transport choice decisions are currently unavailable for the 2016 Irish Census. Lastly, we would like to suggest a potential area for future research. We have identified that Covid-19 is likely to have significant outcomes on transport mobility in the short and long run, which will have implications for GHG emissions. It would be worthwhile for future work to quantify the 'potential' emission savings that could be achieved by remote working in the long run. This would put any future policy interventions in the area of remote working on a more solid foundation. Conversely, any likely 'rebound effects' should also be quantified and mitigation measures identified.

Covid Economics 30, 19 June 2020: 63-82

Covid Economics 30, 19 June 2020: 63-82

References

Adams-Prassl, A., Boneva, T., Golin, M., & Rauh, C. (2020). *Inequality in the Impact of the Coronavirus Shock: Evidence from Real Time Surveys* CEPR Discussion Paper No. DP14665. Retrieved from https://ssrn.com/abstract=3594297

Ahmed, F., Zviedrite, N., & Uzicanin, A. (2018). Effectiveness of workplace social distancing measures in reducing influenza transmission: a systematic review. *BMC public health, 18*(1), 518.

Aledort, J. E., Lurie, N., Wasserman, J., & Bozzette, S. A. (2007). Non-pharmaceutical public health interventions for pandemic influenza: an evaluation of the evidence base. *BMC public health, 7*(1), 208.

Barbieri, T., Basso, G., & Scicchitano, S. (2020). *Italian Workers at Risk during the COVID-19 Epidemic*. Available at SSRN 3572065. Retrieved from https://ssrn.com/abstract=3572065

Bishop, G. A., Peddle, A. M., Stedman, D. H., & Zhan, T. (2010). On-road emission measurements of reactive nitrogen compounds from three California cities. *Environmental science & technology, 44*(9), 3616-3620.

Blanckaert, S. (2019). Car remains primary means of commuting in Western Europe. *Fleet Europe.* Retrieved from https://www.fleeteurope.com/en/smart-mobility/europe/features/car-remains-primary-means-commuting-western-europe?a=SBL09&t%5B0%5D=Traffic&t%5B1%5D=study&t%5B2%5D=France&curl=1

Block, P., Hoffman, M., Raabe, I. J., Dowd, J. B., Rahal, C., Kashyap, R., & Mills, M. C. (2020). Social network-based distancing strategies to flatten the COVID 19 curve in a post-lockdown world. *arXiv preprint arXiv:2004.07052*.

Brynjolfsson, E., Horton, J., Ozimek, A., Rock, D., Sharma, G., & Ye, H. Y. T. (2020). COVID-19 and Remote Work: An Early Look at US Data. *Unpublished work*.

Bureau of Labor Statistics. (2012). *Crosswalk between the International Standard Classification of Occupations (ISCO-08) and the 2010 Standard Occupational Classification (SOC)*. US: Bureau of Labor Statistics.

Cerqueira, E. D. V., Motte-Baumvol, B., Chevallier, L. B., & Bonin, O. (2020). Does working from home reduce CO2 emissions? An analysis of travel patterns as dictated by workplaces. *Transportation Research Part D: Transport and Environment, 83*, 102338. doi:https://doi.org/10.1016/j.trd.2020.102338

Cerqueti, R., Correani, L., & Garofalo, G. (2013). Economic interactions and social tolerance: A dynamic perspective. *Economics Letters, 120*(3), 458-463. doi:https://doi.org/10.1016/j.econlet.2013.05.032

Choo, S., Mokhtarian, P. L., & Salomon, I. (2005). Does telecommuting reduce vehicle-miles traveled? An aggregate time series analysis for the U.S. *Transportation, 32*(1), 37-64. doi:10.1007/s11116-004-3046-7

Clark, B., Chatterjee, K., & Melia, S. (2016). Changes to commute mode: The role of life events, spatial context and environmental attitude. *Transportation Research Part A: Policy and Practice, 89*, 89-105. doi:https://doi.org/10.1016/j.tra.2016.05.005

Crowley, F., & Doran, J. (2019). Automation and Irish towns: who's most at risk? *Spatial and Regional Economics Research Centre, SRERCWP2019-1*.

Crowley, F., & Doran, J. (2020). Covid-19, Occupational Social Distancing and Remote Working Potential in Ireland.

Crowley, F., Doran, J., & Ryan, G. (2020). *The impact of Covid-19 restrictions on workers: Who is most exposed?* Spatial and Regional Economics Research Centre. Department of Economics, Cork University Business School, Ireland. Retrieved from https://www.ucc.ie/en/media/projectsandcentres/srerc/Covid-19_FC_JD_GR.pdf

CSO. (2020). *Employment and Life Effects of COVID-19: CSO statistical release*.
Delaporte, I., & Peña, W. (2020). *Working From Home Under COVID-19: Who Is Affected? Evidence From Latin American and Caribbean Countries*. Retrieved from
Dingel, J. I., & Neiman, B. (2020). *How many jobs can be done at home?* NBER Working Paper No. 26948. National Bureau of Economic Research. Retrieved from https://www.nber.org/papers/w26948
EEA. (2019). *Transport and environment report* Retrieved from EEA Report No 18/2019: https://www.eea.europa.eu/publications/the-first-and-last-mile/
EPA. (2019). Ireland's Greenhouse Gas Emissions Projections 2018-2040. Retrieved from https://www.epa.ie/pubs/reports/air/airemissions/ghgprojections2018-2040/
EPA. (2020). Ireland's National Greenhouse Gas Emissions Inventory for 2018. Retrieved from http://www.epa.ie/pubs/reports/air/airemissions/ghg/nir2018/Ireland%20NIR%202018.pdf
ESA. (2020). Seen from space: COVID-19 and the environment. Retrieved from https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-5P
Regulation (EU) 2018/842 of the European Parliament and of the Council of 30 May 2018 on binding annual greenhouse gas emission reductions by Member States from 2021 to 2030 contributing to climate action to meet commitments under the Paris Agreement, (2018).
European Environment Agency. (2020). Air pollution goes down as Europe takes hard measures to combat coronavirus. Retrieved from https://www.eea.europa.eu/highlights/air-pollution-goes-down-as
European Environmental Agency. (2018). EEA greenhouse gas-data viewer. Retrieved from https://www.eea.europa.eu/data-and-maps/data/data-viewers/greenhouse-gases-viewer
European Environmental Agency. (2019). Air quality in Europe — 2019 report *EEA Report No 10/2019*. Retrieved from https://www.eea.europa.eu/publications/air-quality-in-europe-2019
Faber, J., Schroten, A., Bles, M., Sevenster, M., Markowska, A., Smit, M., . . . Gigli, M. J. D. C. D. (2012). Behavioural climate change mitigation options and their appropriate inclusion in quantitative longer term policy scenarios.
Foucault, M., & Galasso, V. (2020). *Working After Covid-19: Cross-Country Evidence From Real-Time Survey Data. Note on Attitudes towards COVID-19 - A comparative study*. Sciences Po CEVIPOF, Note 9. Retrieved from https://spire.sciencespo.fr/hdl:/2441/5cmk499mce8lvosvi0jdis0dla/resources/note9-foucault-galasso-.pdf
Frank, M. R., Sun, L., Cebrian, M., Youn, H., & Rahwan, I. (2018). Small cities face greater impact from automation. *Journal of the Royal Society Interface, 15*(139), 20170946.
Frey, C. B., Berger, T., & Chen, C. (2017). Political Machinery: Automation Anxiety and the 2016 US Presidential Election: Oxford Martin School.
Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change, 114*, 254-280. doi:https://doi.org/10.1016/j.techfore.2016.08.019
Fu, M., Andrew Kelly, J., Peter Clinch, J., & King, F. (2012). Environmental policy implications of working from home: Modelling the impacts of land-use, infrastructure and socio-demographics. *Energy Policy, 47*, 416-423. doi:https://doi.org/10.1016/j.enpol.2012.05.014
Galbadage, T., Peterson, B. M., & Gunasekera, R. S. (2020). Does COVID-19 Spread Through Droplets Alone? *Frontiers in Public Health, 8*, 163.

Covid Economics 30, 19 June 2020: 63-82

Gallacher, G., & Hossain, I. (2020). Remote Work and Employment Dynamics under Covid-19: Evidence from Canada. *Canadian Public Policy/ Analyse de politiques, forthcoming*. doi:https://doi.org/10.3138/cpp.2020-026

Glass, R. J., Glass, L. M., Beyeler, W. E., & Min, H. J. (2006). Targeted social distancing designs for pandemic influenza. *Emerging infectious diseases, 12*(11), 1671.

Glynn, J., Joshi, S., Brinkerink, M., Daly, H., Deane, P., McDonagh, S., . . . Gallachoir, B. O. (2020). The pandemic & Ireland's Energy System - What are the impacts on Irish greenhouse gasemissions? Retrieved from https://www.siliconrepublic.com/wp-content/uploads/2015/05/Glynn.pdf

Gottlieb, C., Grobovšek, J., & Poschke, M. (2020). Working from home across countries. *Covid Economics, 71*.

Grubler, A., Wilson, C., Bento, N., Boza-Kiss, B., Krey, V., McCollum, D. L., . . . Valin, H. (2018). A low energy demand scenario for meeting the 1.5 °C target and sustainable development goals without negative emission technologies. *Nature Energy, 3*(6), 515-527. doi:10.1038/s41560-018-0172-6

Gupta, N., Tomar, A., & Kumar, V. (2020). The effect of COVID-19 lockdown on the air environment in India. *Global Journal of Environmental Science and Management, 6*(Special Issue (Covid-19)), 31-40.

He, S. Y., & Hu, L. (2015). Telecommuting, income, and out-of-home activities. *Travel Behaviour and Society, 2*(3), 131-147. doi:https://doi.org/10.1016/j.tbs.2014.12.003

Helm, D. (2020). The Environmental Impacts of the Coronavirus. *Environmental and resource economics, 76*(1), 21-38. doi:10.1007/s10640-020-00426-z

Helminen, V., & Ristimäki, M. (2007). Relationships between commuting distance, frequency and telework in Finland. *Journal of transport geography, 15*(5), 331-342. doi:https://doi.org/10.1016/j.jtrangeo.2006.12.004

Hickel, J., & Kallis, G. (2020). Is Green Growth Possible? *New Political Economy, 25*(4), 469-486. doi:10.1080/13563467.2019.1598964

Hulshof, D., & Mulder, M. (2020). Willingness to Pay for CO2 Emission Reductions in Passenger Car Transport. *Environmental and resource economics, 75*(4), 899-929. doi:10.1007/s10640-020-00411-6

Hynes, M. (2014). Telework isn't working: a policy review. *The Economic and Social Review, 45*(4, Winter), 579–602.

Isaifan, R. (2020). The dramatic impact of Coronavirus outbreak on air quality: Has it saved as much as it has killed so far? *Global Journal of Environmental Science and Management, 6*(3), 275-288.

Jackson, C., Vynnycky, E., Hawker, J., Olowokure, B., & Mangtani, P. (2013). School closures and influenza: systematic review of epidemiological studies. *BMJ open, 3*(2), e002149.

Kelso, J. K., Milne, G. J., & Kelly, H. (2009). Simulation suggests that rapid activation of social distancing can arrest epidemic development due to a novel strain of influenza. *BMC public health, 9*(1), 117.

Kopf, D. (2016). Which Professions Have the Longest Commutes? *Priceonomics*. Retrieved from https://priceonomics.com/which-professions-have-the-longest-commutes/

Koren, M., & Pető, R. (2020). Business disruptions from social distancing. *Covid Economics, 2*(8 APRIL 2020), 13-31.

Kramer, A., & Kramer, K. Z. (2020). The potential impact of the Covid-19 pandemic on occupational status, work from home, and occupational mobility: Elsevier.

Mongey, S., Pilossoph, L., & Weinberg, A. (2020). Which Workers Bear the Burden of Social Distancing Policies? *Covid Economics, 12*(1 May 2020), 69-86.

Covid Economics 30, 19 June 2020: 63-82

Mongey, S., & Weinberg, A. (2020). Characteristics of workers in low work-from-home and high personal-proximity occupations. *Becker Friedman Institute for Economic White Paper*.

Montenovo, L., Jiang, X., Rojas, F. L., Schmutte, I. M., Simon, K. I., Weinberg, B. A., & Wing, C. (2020). Determinants of Disparities in Covid-19 Job Losses. *NBER Working Paper No. 27132*.

Muhammad, S., Long, X., & Salman, M. (2020). COVID-19 pandemic and environmental pollution: A blessing in disguise? *Science of The Total Environment*, 138820.

Mundaca, L., Ürge-Vorsatz, D., & Wilson, C. (2019). Demand-side approaches for limiting global warming to 1.5 °C. *Energy Efficiency, 12*(2), 343-362. doi:10.1007/s12053-018-9722-9

NASA. (2020). Nitrogen Dioxide Levels Rebound in China. Retrieved from https://earthobservatory.nasa.gov/images

Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-Jabir, A., Iosifidis, C., . . . Agha, R. (2020). The socio-economic implications of the coronavirus and COVID-19 pandemic: a review. *International Journal of Surgery*.

O'Mahony, T., Zhou, P., & Sweeney, J. (2012). The driving forces of change in energy-related CO2 emissions in Ireland: A multi-sectoral decomposition from 1990 to 2007. *Energy Policy*, 256-267.

Office for National Statistics. (2010). Standard occupational classification (SOC). Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=2ahUKEwi5tLPSv5zfAhXsSBUIHQx_COYQFjABegQICBAC&url=https%3A%2F%2Fwww.ons.gov.uk%2Ffile%3Furi%3D%2Fmethodology%2Fclassificationsandstandards%2Fstandardoccupationalclassificationsoc%2Fsoc2010%2Fug201002soc2010toisco08v2_tcm77-283163.xls&usg=AOvVaw0uKzrFxZBMpOOuWbeFH3QP

Ortúzar, J. d. D., & Willumsen, L. (2011). Modelling Transport (Fourth Edi): John Wiley & Sons, Ltd.

Ory, D. T., & Mokhtarian, P. L. (2006). Which Came First, the Telecommuting or the Residential Relocation? An Empirical Analysis of Causality. *Urban Geography, 27*(7), 590-609. doi:10.2747/0272-3638.27.7.590

Pérez, M. P., Sánchez, A. M., de Luis Carnicer, M. P., & Jiménez, M. J. V. (2004). The environmental impacts of teleworking. *Management of Environmental Quality: An International Journal*.

Rashid, H., Ridda, I., King, C., Begun, M., Tekin, H., Wood, J. G., & Booy, R. (2015). Evidence compendium and advice on social distancing and other related measures for response to an influenza pandemic. *Paediatric respiratory reviews, 16*(2), 119-126.

Redmond, P., & McGuinness, S. (2020). *Who can work from home in Ireland?* Retrieved from Dublin: https://www.esri.ie/publications/who-can-work-from-home-in-ireland

Sims, R., Schaeffer, R., Creutzig, F., Cruz-Núñez, X., D'agosto, M., Dimitriu, D., . . . pdf, N. Y. C. U. P. A. a. h. w. i. c. p. a.-r. a. w. i. w. a. c. (2014). Transport Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change ed O Edenhofer et al.

Van Lier, T., de Witte, A., Macharis, C. J. E. J. o. T., & Research, I. (2014). How worthwhile is teleworking from a sustainable mobility perspective? The case of Brussels Capital region. *14*(3).

World Health Organisation. (2018). Global Health Observatory (GHO) data - Ambient air pollution. Retrieved from http://www.euro.who.int/en/health-topics/environment-

Covid Economics 30, 19 June 2020: 63-82

and-health/air-quality/news/news/2018/5/over-half-a-million-premature-deaths-annually-in-the-european-region-attributable-to-household-and-ambient-air-pollution
World Health Organization. (2020). Rational use of personal protective equipment (PPE) for coronavirus disease (COVID-19): interim guidance, 19 March 2020. Retrieved from https://apps.who.int/iris/handle/10665/331498
Worldometer. (2020, 07 June 2020). COVID-19 Cornoavirus Pandemic. Retrieved from https://www.worldometers.info/coronavirus/?utm_campaign=homeAdUOA?Si

Covid Economics 30, 19 June 2020: 63-82

Covid Economics 30, 19 June 2020: 63-82

Appendix 1: Matching US occupational codes to Irish occupational codes

O*Net provides 968 occupational codes which match to 2010 US Standard Occupational Classifications (SOCs). These occupational codes do not directly match to Irish occupational codes as the Irish Central Statistics Office (CSO) bases their occupational classifications on the UK SOC system. We apply a crosswalk in the same was as Crowley and Doran (2019). The US and UK SOC are not directly comparable and there is no direct conversion available. Therefore, in order to convert the US codes to their UK counterparts (which are approximately identical to the Irish codes used by the CSO) we transform these data using a series of established international classifications. This is accomplished through the use of the International Standard Occupational Classifications (ISOC). The US SOCs can be converted using the Bureau of Labour Statistics official conversion (Bureau of Labor Statistics, 2012). The codes available from O*Net are 6-digit US SOCs. When converting these to the ISOC there is not a one to one match. This is due to the ISOC codes being at a higher aggregation level. Therefore, in some instances, two or more of the US SOC codes are combined into one ISOC code. Where this occurs, any data on occupations is averaged to provide a single value. Once the codes are in ISOC format it is possible to convert these ISOC codes to the UK SOC codes using a conversion framework developed by the Office for National Statistics (2010). In doing so, again there are a small number of occupations which have more than a one for one match and therefore there is a need to average any occupational details associated with these occupations. It is possible, once this process has been completed, to translate these UK SOC codes to Irish SOC codes in a perfect one for one match. When the merge process is complete, out of a possible 327 SOC codes available in Ireland we have occupational data for 318 of these codes at 4-digit level. As the census uses 3-digit occupational codes it is necessary to aggregate these 318 codes to the 88 3-digit codes. When doing so we calculate the weighted average of the 318 codes.

Appendix 2: Definition of elements of Social Distancing Index

Variable	Original Coding	Recoding	Context
How important is it to work with others in a group or team in this job?	0 - Not important at all 25 - Fairly important 50 - Important 75 - Very important 100 - Extremely important	0 - Extremely important 25 - Very important 50 - Important 75 - Fairly important 100 - Not important at all	Face to face discussions several time a week and often more than e-mails, letters, and memos.
Providing guidance and expert advice to management or other groups on technical, systems-, or process-related topics.	0 – Not important 100 – Important	0 – Important 100 – Not important	
Getting members of a group to work together to accomplish tasks.	0 – Not important 100 – Important	0 – Important 100 – Not important	
Providing guidance and direction to subordinates, including setting performance standards and monitoring performance.	0 – Not important 100 – Important	0 – Important 100 – Not important	
Encouraging and building mutual trust, respect, and cooperation among team members.	0 – Not important 100 – Important	0 – Important 100 – Not important	
How important is it to work with external customers or the public in this job?	0 - Not important at all 25 - Fairly important 50 - Important 75 - Very important 100 - Extremely important	0 - Extremely important 25 - Very important 50 - Important 75 - Fairly important 100 - Not important at all	Face to face discussions several times a week
Performing for people or dealing directly with the public. This includes serving customers in restaurants and stores, and receiving clients or guests.	0 – Not important 100 – Important	0 – Important 100 – Not important	
Providing personal assistance, medical attention, emotional support, or other personal care to others such as coworkers, customers, or patients.	0 – Not important 100 – Important	0 – Important 100 – Not important	
Developing constructive and cooperative working relationships with others, and maintaining them over time.	0 – Not important 100 – Important	0 – Important 100 – Not important	
Using hands and arms in handling, installing, positioning, and moving materials, and manipulating things.	0 – Not important 100 – Important	0 – Important 100 – Not important	Density of co-workers like shared office or more
Running, maneuvering, navigating, or driving vehicles or mechanized equipment, such as forklifts, passenger vehicles, aircraft, or water craft.	0 – Not important 100 – Important	0 – Important 100 – Not important	

Covid Economics 30, 19 June 2020: 63-82

CEPR PRESS

Servicing, repairing, adjusting, and testing machines, devices, moving parts, and equipment that operate primarily on the basis of mechanical (not electronic) principles.	0 – Not important 100 – Important	0 – Important 100 – Not important	
Servicing, repairing, calibrating, regulating, fine-tuning, or testing machines, devices, and equipment that operate primarily on the basis of electrical or electronic (not mechanical) principles.	0 – Not important 100 – Important	0 – Important 100 – Not important	
Inspecting equipment, structures, or materials to identify the cause of errors or other problems or defects.	0 – Not important 100 – Important	0 – Important 100 – Not important	
To what extent does this job require the worker to perform job tasks in close physical proximity to other people?	0 - I don't work near other people (beyond 100 ft.) 25 - I work with others but not closely (e.g., private office) 50 - Slightly close (e.g., shared office) 75 - Moderately close (at arm's length 100 - Very close (near touching)	0 - Very close (near touching) 25 - Moderately close (at arm's length 50 - Slightly close (e.g., shared office) 75 - I work with others but not closely (e.g., private office) 0 - I don't work near other people (beyond 100 ft.)	Physical Proximity

Covid Economics 30, 19 June 2020: 63-82

Appendix 3: Definition of elements of Remote Working Index

Variable definition	Original coding	New coding
How often do you use electronic mail in this job?	0-Never 25 - Once a year or more but not every month 50 - Once a month or more but not every week 75 - Once a week or more but not every day 100 - Every day	same as original
How often does this job require working outdoors, exposed to all weather conditions?	0 - Never 25 - Once a year or more but not every month 50 - Once a month or more but not every week 75 - Once a week or more but not every day 100 - Every day	0 – Every day 25 – Once a week or more but not every day 50 – Once a month or more but not every week 75 – Once a year or more but not every month 100 - Never
How often does this job require working outdoors, under cover (e.g., structure with roof but no walls)?	0 - Never 25 - Once a year or more but not every month 50 - Once a month or more but not every week 75 - Once a week or more but not every day 100 - Every day	0 – Every day 25 – Once a week or more but not every day 50 – Once a month or more but not every week 75 – Once a year or more but not every month 100 - Never
How frequently does this job require the worker to deal with physical aggression of violent individuals?	0 - Never 25 - Once a year or more but not every month 50 - Once a month or more but not every week 75 - Once a week or more but not every day 100 - Every day	0 – Every day 25 – Once a week or more but not every day 50 – Once a month or more but not every week 75 – Once a year or more but not every month 100 - Never
How much does this job require wearing common protective or safety equipment such as safety shoes, glasses, gloves, hard hats or life jackets?	0 - Never 25 - Once a year or more but not every month 50 - Once a month or more but not every week 75 - Once a week or more but not every day 100 - Every day	0 – Every day 25 – Once a week or more but not every day 50 – Once a month or more but not every week 75 – Once a year or more but not every month 100 - Never
How much does this job require wearing specialized protective or safety equipment such as breathing apparatus, safety harness, full protection suits, or radiation protection?	0 - Never 25 - Once a year or more but not every month 50 - Once a month or more but not every week 75 - Once a week or more but not every day 100 - Every day	0 – Every day 25 – Once a week or more but not every day 50 – Once a month or more but not every week 75 – Once a year or more but not every month 100 - Never
How much does this job require walking and running?	0 – Never 25 - Less than half the time 50 - About half the time 75 - More than half the time 100 - Continually or almost continually	0 – Every day 25 – Once a week or more but not every day 50 – Once a month or more but not every week 75 – Once a year or more but not every month 100 - Never
How often does this job require exposure to minor burns, cuts, bites, or stings?	0 - Never 25 - Once a year or more but not every month	0 – Every day 25 – Once a week or more but not every day

Covid Economics 30, 19 June 2020: 63-82

Covid Economics 30, 19 June 2020: 63-82

	50 - Once a month or more but not every week 75 - Once a week or more but not every day 100 - Every day	50 – Once a month or more but not every week 75 – Once a year or more but not every month 100 - Never
How often does this job require exposure to disease/infections?	0 - Never 25 - Once a year or more but not every month 50 - Once a month or more but not every week 75 - Once a week or more but not every day 100 - Every day	0 – Every day 25 – Once a week or more but not every day 50 – Once a month or more but not every week 75 – Once a year or more but not every month 100 - Never
Performing physical activities that require considerable use of your arms and legs and moving your whole body, such as climbing, lifting, balancing, walking, stooping, and handling of materials.	0 – Not important 100 – Important	0 – Important 100 – Not important
Using hands and arms in handling, installing, positioning, and moving materials, and manipulating things.	0 – Not important 100 – Important	0 – Important 100 – Not important
Using either control mechanisms or direct physical activity to operate machines or processes (not including computers or vehicles).	0 – Not important 100 – Important	0 – Important 100 – Not important
Running, maneuvering, navigating, or driving vehicles or mechanized equipment, such as forklifts, passenger vehicles, aircraft, or water craft.	0 – Not important 100 – Important	0 – Important 100 – Not important
Performing for people or dealing directly with the public. This includes serving customers in restaurants and stores, and receiving clients or guests.	0 – Not important 100 – Important	0 – Important 100 – Not important
Servicing, repairing, adjusting, and testing machines, devices, moving parts, and equipment that operate primarily on the basis of mechanical (not electronic) principles.	0 – Not important 100 – Important	0 – Important 100 – Not important
Servicing, repairing, calibrating, regulating, fine-tuning, or testing machines, devices, and equipment that operate primarily on the basis of electrical or electronic (not mechanical) principles.	0 – Not important 100 – Important	0 – Important 100 – Not important
Inspecting equipment, structures, or materials to identify the cause of errors or other problems or defects.	0 – Not important 100 – Important	0 – Important 100 – Not important

Macroeconomics of epidemics: Interstate heterogeneity in Brazil

Luan Borelli[1] and Geraldo Sandoval Góes[2]

Date submitted: 17 June 2020; Date accepted: 18 June 2020

We apply the SIR-macro model proposed by Eichenbaum et al. (2020) in its complete version to comparatively study the interaction between economic decisions and COVID-19 epidemics in five different Brazilian states: São Paulo (SP), Amazonas (AM), Ceará (CE), Rio de Janeiro (RJ) and Pernambuco (PE). Our objective is to analyze qualitatively how the main intrinsic differences of each of these states can affect the epidemic dynamics and its consequences. For this purpose, we compute and compare the model for each of the states, both in competitive equilibrium and under optimal containment policy adoption, and analyze the implications of optimal policy adoption. We conclude that the intrinsic characteristics of the five different states imply relevant differences in the general dynamics of the epidemic, in the optimal containment policies, in the effect of the adoption of these policies and the severity of the economic recessions. Our study can serve as an alert for policymakers of countries of huge dimensions and interstate heterogeneity as Brazil for the necessity of discriminating policies by states or regions instead of adopting a single unified policy for the whole country.

Covid Economics 30, 19 June 2020: 83-119

1 Research Assistant, Institute for Applied Economic Research (IPEA-RJ).
2 Researcher, Institute for Applied Economic Research (IPEA-RJ).

Covid Economics 30, 19 June 2020: 83-119

1 Introduction

As COVID-19 spreads throughout Brazil, governors of the country are struggling with the challenge to understand and manage the epidemic that ravages the states they command. When compared to other smaller countries, this challenge seems to be even more complicated in Brazil, due to its continental geographic and demographic dimensions and to the huge heterogeneity of the characteristics of its states. A recent study by Hallal et al. (2020) illustrates how this heterogeneity can affect the epidemic dynamics around the country. The study reports the first wave of seroprevalence surveys relying upon on household probabilistic samples of 133 large sentinel cities in Brazil, including 25.025 participants from all 26 states and the Federal District and find that the seroprevalence of antibodies to SARS-CoV-2, assessed using a lateral flow rapid test, varied markedly across the cities and regions, from below 1% in most cities in the South and Center-West regions to up to 25% in the city of Breves in the Amazon (North) region.

In this work, our objective is to investigate how and how much this interstate heterogeneity affects the dynamics of epidemics in Brazilian states. An answer to this question can provide insights on the consequences of conducting aggregated containment policies for countries as a whole or disaggregated policies by states or regions.

To address this analysis, we use the simple framework of the SIR-macro model proposed by Eichenbaum et al. (2020). Differently of the epidemiology models that have been widely used to predict the course of epidemics (e.g., Ferguson et al. (2020)), the SIR-macro model extends the classical SIR model (an acronym for Susceptible, Infected and Recovered) originally proposed by Kermack and McKendrick (1927) incorporating an important factor that can potentially affect significantly the dynamics of the epidemics: the interaction between economic decisions and infection rates. In this new extended model, the people's decision to cut back on consumption and work reduces the severity of the epidemic as measured by total deaths. The same decisions exacerbate the size of the recession caused by the epidemic.

In this work, we compute the SIR-macro model for five Brazilian states. We consider, for each of these states, two scenarios: one under competitive equilibrium and the other under optimal policy adoption. The criteria for the selection of these five states was the severity of the epidemics at the moment of the research. Thus, we selected the five Brazilian states that presented the worst epidemic situation between all the 26 Brazilian states. The selected states were São Paulo (SP), Amazonas (AM), Ceará (CE), Rio de Janeiro (RJ) and Pernambuco (PE).

The different intrinsic characteristics of these states are incorporated in the results through our calibration strategy for the parameter and variable values of the model. Our calibration allows the results to incorporate interstate differences relative to the following aspects: i) size of the population, ii) per capita income; iii) average time devoted to work; iv) average time spent in public transport; v) average time devoted to household chores; vi) the average number of people

per residence; vii) the number of workers; viii) number of students and ix) infection fatality rates estimated with the evolution of the epidemic until the moment of the research.

We compare the differences in the epidemic dynamics and their consequences between the five states both for the competitive equilibrium scenario (without adopting containment policies) and for the scenario in which optimal containment policies are adopted. We also compare the differences in the optimal containment policies required for each state and evaluate the effects of adopting these policies on each of these states.

As a result of these analyzes, we conclude that the intrinsic characteristics of the five different states imply relevant differences in their epidemic dynamics, in the optimal containment policies to be adopted, in the effect of adopting these policies and in the severity of the economic recessions resulting from the epidemic. This conclusion emphasizes the importance of the disaggregated analysis of countries with huge geographic and demographic dimensions like Brazil in the formulation of policies to combat COVID-19. We warn that the adoption of unique aggregate policies for huge and heterogeneous countries like Brazil can trigger a series of containment policy errors in the states and regions of the country, deepening both the economic recession and the number of deaths resulting from the epidemic. The propagation channel of these errors is the following: if there is large heterogeneity in the optimal level of containment policy required for each state, a single containment policy for the whole country would be unable to adequately deal with the needs of all states simultaneously. Inevitably, the containment rate adopted across the country would be lower than needed for some states, resulting in more deaths, and higher than needed for others, resulting in unnecessarily deeper economic recessions.

Our paper is organized as follows. In Section 2, we describe the SIR-macro model of Eichenbaum et al. (2020) in its complete version. In Section 3, we discuss our calibration strategy for the parameter values used in the model for each state. In Section 4, we present and describe the results of the model for each of the five selected states in the competitive equilibrium scenario. In Section 5, we solve the Ramsey policy problems and analyze the optimal containment rate results for each state. In Section 6, we analyze for each state the implications of adopting optimal containment rates when compared to the baseline scenario of competitive equilibrium. Finally, Section 7 concludes.

Covid Economics 30, 19 June 2020: 83-119

2 The SIR-macro model

In this section, we describe with some adaptations the full version of the SIR-macro model as developed by Eichenbaum et al. (2020). The adaptations were necessary because in Eichenbaum et al. (2020) the complete version of the model is constructed with the gradual inclusion of extensions to the basic SIR-macro model along the sections of the article. As in this work we are only interested in the full version of the model, we have adapted the exposition to present it in a single section. The full version of the model includes three main characteristics: the endogenous mortality rate as a function of the infected population, the possibility of discovering effective treatments that cure the infected population, and the possibility of discovering vaccines.

Covid Economics 30, 19 June 2020: 83-119

2.1 The economic framework and the pre-infection economy

The population of the economy is represented by a continuum of ex-ante identical agents with measure one. These agents maximize the objective function:

$$U = \sum_{t=0}^{\infty} \beta_t u(c_t, n_t). \tag{1}$$

Here $\beta \in (0,1)$ denotes the discount fator and c_t and n_t denote consumption and hours worked, respectively. For simplicity, we assume that momentary utility takes the form

$$u(c_t, n_t) = ln\, c_t - \frac{\theta}{2} n_t^2. \tag{2}$$

The budget constraint of the representative agent is:

$$(1 + \mu_t) c_t = w_t n_t + \Gamma_t. \tag{3}$$

Here, w_t denotes the real wage rate, μ_t is the tax rate on consumption and Γ_t denotes lump-sum transfers from the government. As dicussed in Eichenbaum et al. (2020), μ_t is thought as a proxy for containment measures aimed at reducing social interactions. For this reason, this parameter is referred as the containment rate.

The first-order condition for the representative-agent's problem is:

$$(1 + \mu_t)\theta n_t = c_t^{-1} w_t. \tag{4}$$

There is a continuum of competitive representative firms of unit measure that produce consumption goods (C_t) using hours worked (N_t) according to the technology:

$$C_t = AN_t. \tag{5}$$

The firm chooses hours worked to maximize its time-t profits Π_t:

$$\Pi_t = AN_t - w_t N_t. \tag{6}$$

The government's budget constraint is given by:

$$\mu_t c_t = \Gamma_t. \tag{7}$$

In equilibrium, $n_t = N_t$ and $c_t = C_t$.

2.2 The epidemiological framework and the outbreak of an epidemic

The population is divided into four groups: susceptible, infected, recovered, and deceased. People who have not yet been exposed to the disease are classified as susceptible. People who contracted the disease are classified as infected. People who survived the disease and acquired immunity are classified as recovered. Finally, people who died from the disease are classified as deceased. The fractions of people in these four group, at time-t, are denoted by S_t, I_t, R_te D_t, respectively. The number of newly infected at time-t is denoted by T_t.

Susceptible people can become infected in three ways: i) meeting infected people while purchasing consumption goods; ii) meeting at work or iii) meeting in ways not directly related to consuming or working, for example in the public transport, meeting a neighbor, at home or touching a contaminated surface.

The number of newly infected people that results from the consumption activities is given by $\pi_1(S_t C_t^s)(I_t C_t^i)$. The terms $S_t C_t^s$ e $I_t C_t^i$ represent total consumption expenditures by susceptible and infected people, respectively. The parameter π_1 reflects both the amount of time spent shopping and the probability of becoming infected as a result of that activity.

The number of newly infected people that results from interactions at work is given by $\pi_2(S_t N_t^s)(I_t N_t^i)$. The terms $S_t N_t^s$ and $I_t N_t^i$ represent total hours worked by susceptible and infected people, respectively. The parameter π_2 reflects the probability of becoming infected as a result of work interactions.

The number of newly infected people that results from the activities not directly related to consuming or working is given by $\pi_3 S_t I_t$. The term $S_t I_t$ represent the number of random meetings

Covid Economics 30, 19 June 2020: 83-119

Covid Economics 30, 19 June 2020: 83-119

between infected and susceptible people. The parameter π_3 reflects the probability of becoming infected as a result of these activities.

The total number of newly infected people is given by:

$$T_t = \pi_1(S_t C_t^s)\left(I_t C_t^i\right) + \pi_2(S_t N_t^s)\left(I_t N_t^i\right) + \pi_3 S_t I_t. \quad (8)$$

Note that the Kermack & McKendrick's (1927) canonical SIR model is a special case of the SIR-macro model in which the propagation of the disease is unrelated to economic activity. This case is characterized by parametric values $\pi_1 = \pi_2 = 0$.

The number of infected people at time $t+1$ is equal to the number of infected people at time t (I_t) plus the number of newly infected (T_t) minus the number of infected people that recovered ($\pi_r I_t$) and the number of infected people that died ($\pi_d I_t$):

$$I_{t+1} = I_t + T_t - (\pi_r + \pi_d)I_t, \quad (9)$$

Here, π_r is the rate at which infected people recover from the infection and π_d is the mortality rate, that is the probability that an infected person dies. To incorporate the effect of the efficacy of the healthcare system in the model, this probability is modeled as a convex function of the fraction κ of the population that becomes infected:

$$\pi_{dt} = \pi_d + \kappa I_t^2. \quad (10)$$

The timing convention in equation (9) is as follows. Social interactions happen in the beginning of the period, when infected and susceptible people meet. Then, changes in health status unrelated to social interactions (recovery and death) occur. Finally, at the end of the period the consequences of social interactions materialize: T_t susceptible people become infected.

The number of recovered people at time $t+1$ is the number of recovered people at time t (R_t) plus the number of infected people who just recovered ($\pi_r I_t$):

$$R_{t+1} = R_t + \pi_r I_t. \quad (11)$$

Finally, the number of deceased people at time $t+1$ is the number of deceased people at time t (D_t) plus the number of new deaths ($\pi_{dt} I_t$):

$$D_{t+1} = D_t + \pi_{dt} I_t. \quad (12)$$

The total population evolves according to:

$$POP_{t+1} = POP_t - \pi_{dt} I_t, \tag{13}$$

with $POP_0 = 1$.

The model assumes that at time zero, a fraction ε of susceptible people is infected by a virus through zoonotic exposure, that is the virus is directly transmitted from animals to humans,

$$I_0 = \varepsilon, \tag{14}$$

$$S_0 = 1 - \varepsilon. \tag{15}$$

Everybody is aware of the initial infection and understand the laws of motion governing population heath dynamics. Critically, they take as given aggregate variables like $I_t C_t^i$ and $I_t N_t^i$.

We now describe the optimization problem of different types of people in the economy, as formulated by Eichenbaum et al. (2020). The variable U_t^j denotes the time-t lifetime utility of a type-j agent ($j = s, i, r$). The budget constraint of a type-j person is

$$(1 + \mu_t) c_t^j = w_t \phi^j n_t^j + \Gamma_t, \tag{16}$$

where c_j and n_t^j denote the consumption and hours worked of an agent of type j, respectively. The parameter governing labor productivity, ϕ^j, is equal to one for susceptible and recovered people ($\phi^s = \phi^r = 1$) and less than one for infected people ($\phi^i < 1$).

Susceptible people. The lifetime utility of a susceptible person, U_t^s, is

$$U_t^s = u(c_t^s, n_t^s) + (1 - \delta_v)\left[(1 - \tau_t)\beta U_{t+1}^s + \tau_t \beta U_{t+1}^i\right] + \delta_v \beta U_{t+1}^r. \tag{17}$$

Critically, susceptible people understand that consuming and working less reduces the probability of becoming infected. The variable τ_t represents the probability that a susceptible person becomes infected:

$$\tau_t = \pi_1 c_t^s (I_t C_t^I) + \pi_2 n_t^s (I_t N_t^I) + \pi_3 I_t. \tag{18}$$

The parameter δ_v represents the probability that a vaccine is discovered. With probability $1 - \delta_v$ a person susceptible at time t remains susceptible at time $t + 1$. With probability δ_v this person is vaccinated and becomes immune to the disease. So, at time $t + 1$, this person's health situation is identical to that of a recovered person. The vaccine has no impact on people who were infected or have recovered.

Covid Economics 30, 19 June 2020: 83-119

The first-order conditions for consumption and hours worked are:

$$u_1(c_t^s, n_t^s) - (1+\mu_t)\lambda_{bt}^s + \lambda_{\tau t}\pi_1(I_t C_t^l) = 0, \tag{19}$$

$$u_2(c_t^s, n_t^s) + w_t\lambda_{bt}^s + \lambda_{\tau t}\pi_2(I_t N_t^l) = 0. \tag{20}$$

Here, λ_{bt}^s and $\lambda_{\tau t}$ are the Lagrange multipliers associated with constraints (*16*) and (*18*), respectively.

The first-order condition for τ_t is:

$$\beta(U_{t+1}^i - U_{t+1}^s) - \lambda_{\tau t} = 0. \tag{21}$$

Infected people. The lifetime utility of an infected person, U_t^i, is

$$U_t^i = u(c_t^i, n_t^i) + (1-\delta_c)[(1-\pi_r-\pi_d)\beta U_{t+1}^i + \pi_r\beta U_{t+1}^r] + \beta\delta_c U_{t+1}^r. \tag{22}$$

The expression for U_t^i embodies a common assumption in macro and health economics that the cost of death is the foregone utility of life.

The parameter δ_c represents the probability that an effective treatment that cures infected people is discovered each period. Once discovered, treatment is provided to all infected people in the period of discovery and all subsequent periods transforming them into recovered people. As a result, the number of new deaths from the disease goes to zero.

The first-order conditions for consumption and hours worked are given by:

$$u_1(c_t^i, n_t^i) = \lambda_{bt}^i(1+\mu_t), \tag{23}$$

$$u_2(c_t^i, n_t^i) = -\phi^i w_t \lambda_{bt}^i, \tag{24}$$

where λ_{bt}^i is the Lagrange multiplier associated with constraint (*16*).

Recovered people. The lifetime utility of a recovered person, U_t^r, is

$$U_t^r = u(c_t^r, n_t^r) + \beta U_{t+1}^r. \tag{25}$$

The first-order conditions for consumption and hours worked are:

Covid Economics 30, 19 June 2020: 83-119

$$u_1(c_t^r, n_t^r) = \lambda_{bt}^r(1 + \mu_t), \tag{26}$$

$$u_2(c_t^r, n_t^r) = -w_t\lambda_{bt}^r, \tag{27}$$

where λ_{bt}^r is the Lagrange multiplier associated with constraint (*16*).

Government budget constraint. The government budget constraint is

$$\mu_t\left(S_t c_t^s + I_t c_t^i + R_t c_t^r\right) = \Gamma_t(S_t + I_t + R_t). \tag{28}$$

Equilibrium. In equilibrium, each person solves their maximization problem and the government budget constraint is satisfied. In addition, the goods and labor markets clear:

$$S_t C_t^s + I_t C_t^i + R_t C_t^r = AN_t, \tag{29}$$

$$S_t N_t^s + I_t N_t^i \phi^i + R_t N_t^r = N_t. \tag{30}$$

3 Calibration

In this section we present our calibration strategy used to determine the model parameters for each of the five chosen states, namely, São Paulo (SP), Amazonas (AM), Ceará (CE), Rio de Janeiro (RJ) and Pernambuco (PE).

First of all, we need to state an important consideration concerning the time convention adopted in this paper for the model results. Since the main objective of the model is not to compute quantitative punctual predictions about questions such as the date of the peak of the infected population, we decided to give up the definition of a specific time periodicity for the time of the model. Instead, we chose to normalize the total duration of the time interval for which the exercises were computed to one. Thus, the time interpretation of the model becomes in relation to the progress (%) of the total duration of the epidemics. Since the total duration of the epidemics will be known only ex-post, we exempt this work from any predictive responsibilities that could be attributed to it and we emphasize the qualitative and comparative nature of its results.

To calibrate the value of κ for each state we used infection fatality rate estimates from Mellan's et al. (2020) Report 21 from Imperial College for a series of Brazilian states. We calibrate the values of κ for each of the five states so that the mortality rates in the competitive equilibrium peaks at the values of these estimated infection fatality rates. Although the mortality rates are endogenous

Covid Economics 30, 19 June 2020: 83-119

Covid Economics 30, 19 June 2020: 83-119

to the model, they need an initial value. Following Rabelo and Soares (2020) we use 0.3% as the initial mortality rate for all states. This value is obtained as a weighted average of the mortality rates verified in South Korea (the country that had the world's highest per capita test rates for COVID-19) using weights equal to the percentage of the Brazilian population for different age groups.

Proceeding with our calibration strategy for π_1, π_2, and π_3 for each state, we start adopting the same premise of Eichenbaum et al. (2020) that, as is common in epidemiology, the relative importance of different modes of transmission is similar across viruses that cause respiratory diseases. Ferguson et al. (2006) argue that, in the case of influenza, 30% of transmissions occur in the household, 33% in the general community, and 37% occur in schools and workplaces. Based on this assumption adopted, we use these percentages as the basis for our calibration.

As in Eichenbaum et al. (2020), we calibrate the values of π_1, π_2, and π_3 to satisfy the system

$$\begin{aligned} \frac{\pi_1 C^2}{\pi_1 C^2 + \pi_2 N^2 + \pi_3} &= \alpha_1 \\ \frac{\pi_2 N^2}{\pi_1 C^2 + \pi_2 N^2 + \pi_3} &= \alpha_2 \\ \frac{\pi_3}{\pi_1 C^2 + \pi_2 N^2 + \pi_3} &= \alpha_3 \end{aligned} \tag{31}$$

where C and N are the consumption and hours worked in the pre-epidemic steady state. Note that α_1, α_2, and α_3 are the shares of transmissions that occur in consumption, at work and in other activities, respectively. Having these shares calibrated, it is possible to solve the system and therefore obtain the values of π_1, π_2, and π_3. Thus, it is easy to see that the calibration problem of π_1, π_2, and π_3 is equivalent to the calibration problem of α_1,α_2, and α_3. Also note that as $\alpha_1 + \alpha_2 + \alpha_3 = 1$, it is enough that two of the portions are calibrated for the latter to be obtained residually.

Due to the lack of data available for Brazil, it became impractical to calibrate α_1 directly. We then chose to calibrate α_2 and α_3 and obtain α_1 residually. We calibrated the share of transmissions that occur at work, α_2, using the same approach as Eichenbaum et al. (2020). We weight the number of students and workers in each state by 10 and 4, respectively. These weights refer to the average amount of physical contacts per day at school and at work, obtained from Lee et al. (2010). For the total number of workers, we used the number of persons aged 14 and over employed in the labor force obtained from IBGE's SIDRA[1]. For the total number of students, we used the number of students aged 4 and over, also obtained from IBGE's SIDRA[2].

We calibrated the share of infections that occur in other activities, α_3, using some calibration hypothesis and data on i) the proportion of daily hours dedicated to household chores of persons aged 14 and over; ii) the average number of people per household and iii) the usual time of commuting of individuals to work. Having the values of α_2 and α_3 calibrated, we obtain the value

[1] *Sistema IBGE de Recuperação Automática* (IBGE Automatic Recovery System), Table 4093.
[2] Ibid., Table 983.

of α_1 residually and, with these three calibrated fractions, we finally obtain π_1, π_2, and π_3 solving system (31). More technical details regarding the calibration of α_1, α_2, and α_3 can be found in the Appendix at Section 9.

The initial population is normalized to one. For each state, we calibrate the number of people that are initially infected to represent a number of 100 infected people at $t = 0$. Thus, for each state ε is calculated as:

$$\varepsilon = \frac{100}{POP} \tag{32}$$

where POP is the estimated resident population for 2019 for the state, obtained from IBGE's SIDRA[3]. Calibrating in this way for each state, it was possible to capture and compare the effect of the different population sizes of each state on the local dynamics of the epidemic.

As in Eichenbaum et al. (2020), we calibrate the values of A and θ so that in the pre-epidemic steady state the representative person of each state works the average hours worked by the population of his state and earns the per capita income of the population of his state. For this, we use the average hours usually worked per week in all jobs of people aged 14 and over and the average real monthly household income per capita, at average prices of the year, for each of the states. The first data set was obtained from IBGE's SIDRA[4] and the latter from IBGE's 2019 *PNAD Contínua* survey.

To calibrate the discount factor, β, for all states we use the same value calculated by Rabelo and Soares (2020) for Brazil. This value is equivalent to a statistical value of life of R$ 2,9 million. As indicated by these authors, this value is based on recent estimates for Brazil by Ferrari et al. (2019) and Rocha et al. (2019).

Again, as in Eichenbaum et al. (2020), we calibrate the parameter that controls the relative productivity of infected people, ϕ^i, as 0,8. This value is consistent with the notion that symptomatic agents don't work and the assumption that 80 percent of infected people are asymptomatic according to the China Center for Disease Control and Prevention.

Regarding the probabilities of discovering effective treatments and vaccines, δ_c and δ_v, respectively, following Eichenbaum et al. (2020) we calibrate them to reflect an average Discovery time of one year.

[3] Ibid., Table 6579.
[4] Ibid., Table 6373.

Covid Economics 30, 19 June 2020: 83-119

Covid Economics 30, 19 June 2020: 83-119

Table 1: Main values used for the calibration of the parameters and variables of each state.

Value	SP	AM	CE	RJ	PE
Imperial College's Estimated IFR (%)	0,70	0,80	1,10	0,80	1,10
Number of Workers	22.782.714	1.657.700	3.764.280	7.651.617	3.602.820
Number of Students	10.306.000	1.284.000	2.376.000	3.853.000	2.475.000
Average time devoted to household chores (daily hours)	2,06	1,44	1,87	2,04	2,02
Average number of people per residence	2,80	3,60	3,10	2,70	2,90
Average time spent in public transport (minutes)	37,15	33,95	26,60	43,07	30,52
Pre-epidemic population	45.919.049	4.144.597	9.132.078	17.264.943	9.557.071
Average time devoted to work (daily hours)	8,24	7,30	7,58	8,10	7,74
Per capita income	1.889	838	939	1.809	954

Table 2: Calibrated values for the main variables and parameters of each state.

Parameter/Variable	SP	AM	CE	RJ	PE
κ	0,63	1,10	2,35	1,33	1,90
α_1	0,16	0,28	0,30	0,12	0,25
α_2	0,17	0,13	0,14	0,16	0,14
α_3	0,66	0,60	0,56	0,71	0,61
π_1	$4{,}28 \cdot 10^{-7}$	$3{,}68 \cdot 10^{-6}$	$3{,}17 \cdot 10^{-6}$	$8{,}53 \cdot 10^{-7}$	$2{,}59 \cdot 10^{-6}$
π_2	$5{,}99 \cdot 10^{-5}$	$5{,}54 \cdot 10^{-5}$	$5{,}85 \cdot 10^{-5}$	$5{,}13 \cdot 10^{-5}$	$5{,}33 \cdot 10^{-5}$
π_3	0,39	0,35	0,33	0,33	0,36
ε	$\frac{100}{45.919.049}$	$\frac{100}{4.144.597}$	$\frac{100}{9.132.078}$	$\frac{100}{17.264.943}$	$\frac{100}{9.557.071}$
A	11,46	5,74	6,19	11,17	6,16
θ	$5{,}89 \cdot 10^{-4}$	$7{,}51 \cdot 10^{-4}$	$6{,}96 \cdot 10^{-4}$	$6{,}10 \cdot 10^{-4}$	$6{,}68 \cdot 10^{-4}$
β	0,966	0,966	0,966	0,966	0,966
ϕ^i	0,80	0,80	0,80	0,80	0,80

4 Competitive Equilibrium

In this first exercise, we present results for the competitive equilibrium of each state. In the competitive equilibrium, there are no attempts by the authorities to contain the evolution of the epidemic. The dynamics of the epidemic are affected only by the decisions made by economic agents, that are free to reduce their chances of being infected by reducing consumption and hours worked.

In this and the next sections, we illustrate the differences between states by describing the main results for each variable. To keep text clarity, we describe the results only for the states that presented the highest and lowest values for each variable. These values will be sufficient to assess the size of the heterogeneity of results between states. Nevertheless, the other results can be analyzed in greater detail in the figures and tables that will be presented by the end of the article. All numerical results will be summarized in Table 3, in Section 6. Regarding this section, Figure 1 shows, in level, the dynamics of the evolution of the epidemic for each of the five states, as well as its effects on the aggregates of consumption and hours worked in each of them, for the competitive equilibrium. Figure 2 shows the same results but in proportion (%) to the initial population of each state.

Proceeding to the results, first, we can observe that the state that reaches the highest peak of the infected share of the pre-epidemic population is the state of São Paulo (SP), at 4.95%. At the other extreme, the state that reaches the lowest peak is the state of Ceará (CE), with 3.83% of the pre-epidemic population infected. In relation to the temporal dimension, the peak of infected people occurs first in Amazonas (AM), with 39.33% of the progress of the total time of the epidemic occurred and, only lastly, in São Paulo (SP), with 47.33% of the time progress of the epidemic occurred.

Concerning the macroeconomic consequences of the epidemic, among the five states analyzed, the most severe macroeconomic shock occurs in Ceará (CE), where the aggregates of consumption and hours worked offered suffer a fall of approximately 18.11% in comparison to the pre-epidemic steady state. On the other hand, the least severe occurs in São Paulo (SP), where the falls in aggregate consumption and the aggregate hours of work reached a valley of –13.55%. Regarding time, the recession valley occurs first in Amazonas (AM), when the progress of the total duration of the epidemic reaches 38.67% and, lastly, in São Paulo (SP), when the progress of the total duration of the epidemic reaches 46.00%, 8.67 percentage points of the progress of the epidemic after Amazonas (AM).

The epidemic ends with 52.93% of the population of São Paulo (SP) infected, the largest share of the population infected among the five states analyzed in the competitive equilibrium. On the other hand, Ceará (CE) has the lowest share among the five states, with 47.12% of the population infected at the end of the epidemic. As a consequence of the infections, the state with the largest share of the initial population affected by death is the state of Pernambuco (PE), with 0.36%, while those with the smallest shares are the states of Rio de Janeiro (RJ) and São Paulo (SP), both with 0.27%.

Covid Economics 30, 19 June 2020: 83-119

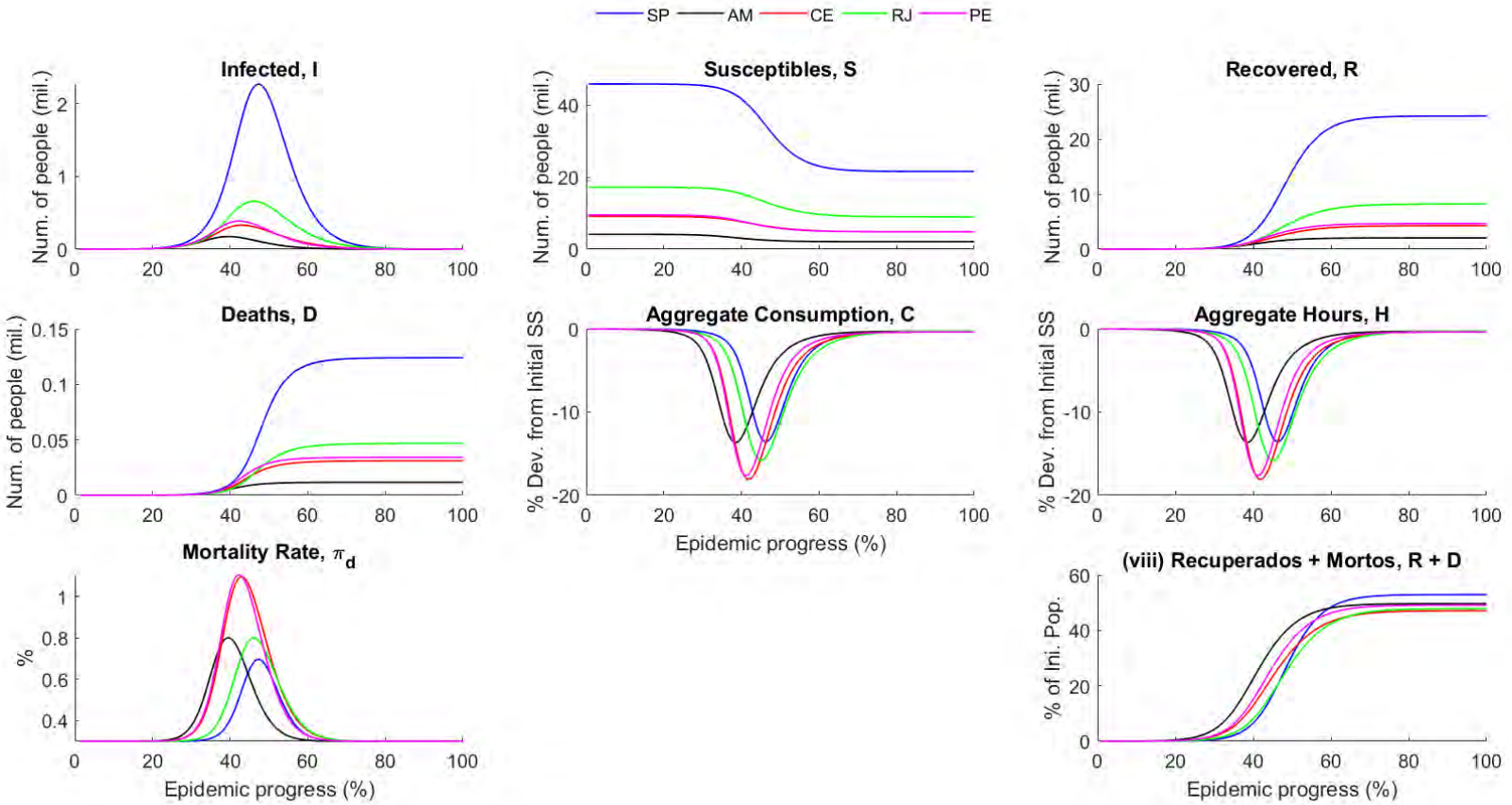
These sample results illustrate evidence of relevant differences in the epidemic dynamics of each state in the competitive equilibrium due to its intrinsic differences. It is possible to note that the size of the infected peak and the time required to reach it are significantly different in each state. The depth and duration of recessions, the share of the total population infected and the share of the population affected by death at the end of the epidemic in each state also differ between them.

In the next section, we will analyze optimal containment policies that can be adopted by states to improve the results obtained in competitive equilibrium. Our main interest will be to analyze how these containment policies differ between these states. In Section 6 we analyze the results under the scenario of adopting these policies and compare them with the results presented in this section to assess the effects of adopting these measures.

Covid Economics 30, 19 June 2020: 83-119

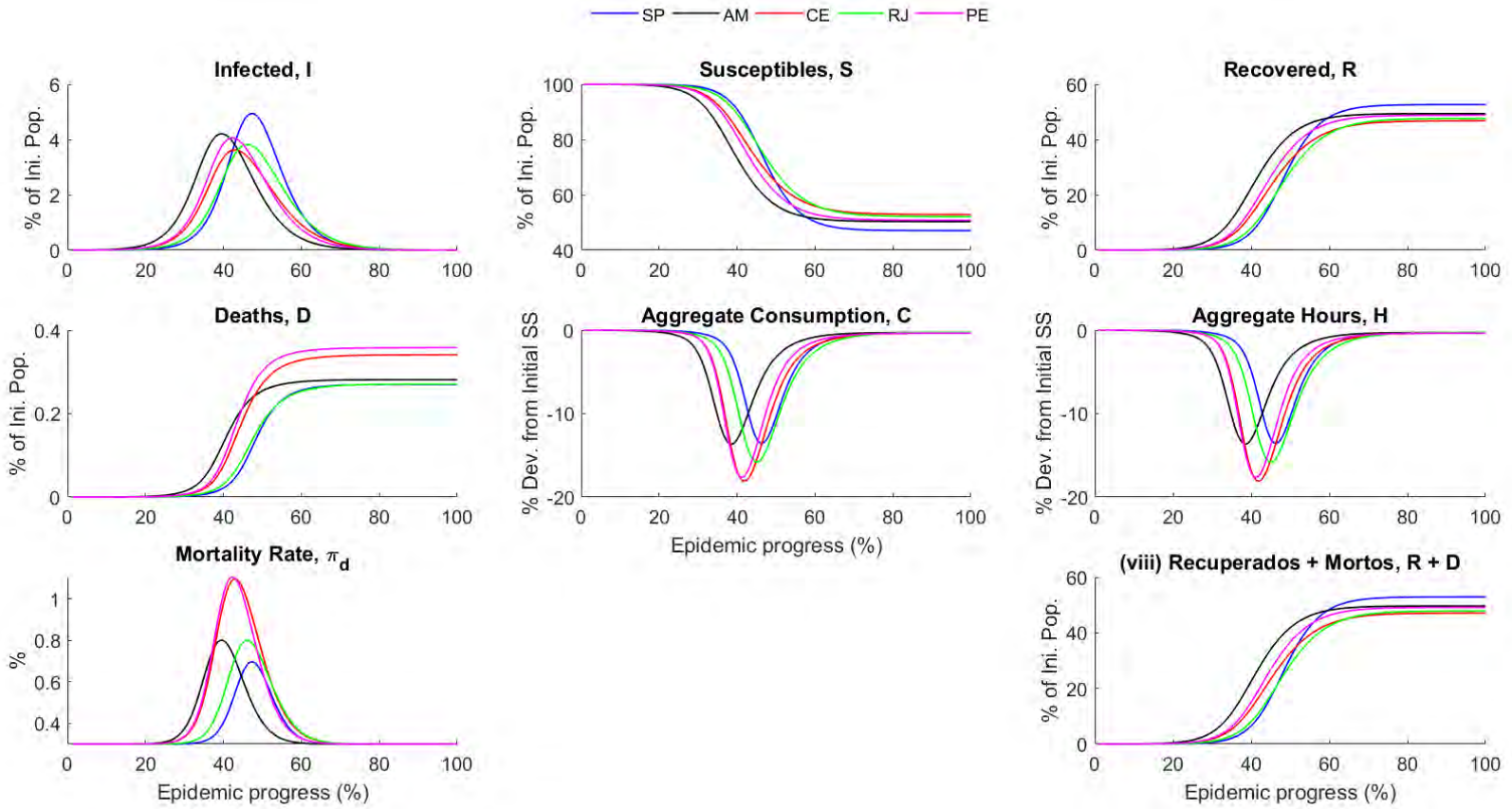
Covid Economics 30, 19 June 2020: 83-119

Figure 1: SIR-macro model - Results for competitive equilibrium, in level.



Covid Economics 30, 19 June 2020: 83-119

Figure 2: SIR-macro model – Results for competitive equilibrium, in percentage of the initial population.



Covid Economics 30, 19 June 2020: 83-119

5 Optimal containment policy

In this section, inspired by Eichenbaum et al. (2020), we consider a simple Ramsey problem to deal with a classic externality associated with the behavior of infected agents in the competitive equilibrium. Because agents are atomistic, they don't take into account the impact of their actions on the infection and death rates of other agents. For this reason, the competitive equilibrium is not socially optimal.

Eichenbaum et al. (2020), analogous to Farhi and Werning's (2012) treatment of capital controls, model the containment measures as a tax on consumption, the proceeds of which are rebated lump sum to all agents. This tax on consumption is referred as the containment rate. We proceed, as well as Eichenbaum et al. (2020), modelling in this way.

In this section we compute, for each state, the optimal sequence of 250 containment rates $\{\mu_t\}_{t=0}^{249}$ that maximize social welfare, U_0, defined as a weighted average of the lifetime utility of the different agents. Since at time zero $t = 0$, $R_0 = D_0 = 0$, the value of U_0 is

$$U_0 = S_0 U_0^s + I_0 U_0^i. \tag{33}$$

Given the sequence of containment rates, we solve for the competitive equilibrium and evaluate the social welfare function. We iterate on this sequence until we find the optimum.

Figure 3 shows the results, in level, for which all states adopt optimal containment policies. Figure 4 shows the same results, but as percentage of the initial population.

First, we can see that in the period $t = 0$, where the infected population in each state reaches the level of 100 individuals, the containment rates are already high for all five states. The highest initial containment rate occurs in Ceará (CE), at 16.12%, while the lowest in São Paulo (SP), at 14.05%.

As the containment rates evolve, the state that reaches the highest peak of the containment rate among the five states is the state of Pernambuco (PE), with a containment rate of 53.40%. On the other hand, São Paulo (SP) is the state with the lowest peak containment rate, at 38.76%. These numbers illustrate the alarming size of heterogeneity between states in terms of the level to which containment should be taken. The state of Pernambuco (PE) requires that the containment rate be raised 14.64 percentage points more than in São Paulo (SP).

Regarding time, the peak of the containment rate occurs first in the state of Amazonas (AM), in 40.67% of the progress of the total time of the epidemic, this being, therefore, the first state to begin to relax the containment measures. On the other hand, the peak occurs lastly in Rio de Janeiro (RJ), with 48.67% of the progress of the epidemic occurred. Rio de Janeiro (RJ) is, therefore, the last state to start the process of gradual relaxation of containment measures. These results illustrate how the optimal duration and speed of containment policies can vary across states.

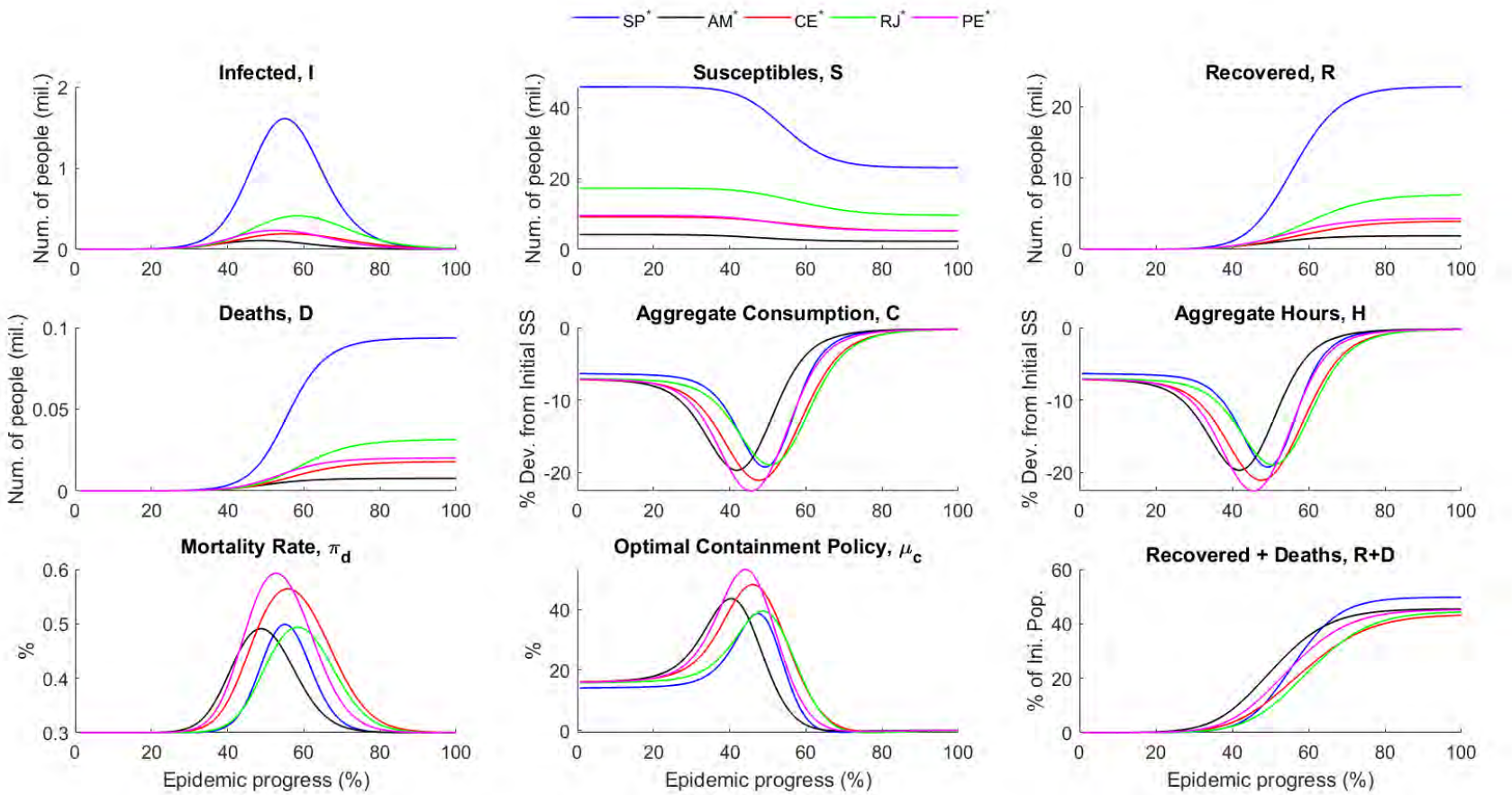
In summary, these results illustrate that the intrinsic differences of each state not only affect the initial severity of the containment measures adopted by each of them but also the dynamics of the evolution of these containment rates over time, implying differences in i) the ideal moment when the measures containment will need to be raised, ii) how far they will need to be raised and iii) when they can finally be relaxed. Notwithstanding these interstate differences in the optimal conduct of containment policies, all results indicate that the adoption of relatively severe containment measures was an optimal decision for each of the five states.

In this section, we studied the optimal containment policies to be adopted by each of the five states analyzed. However, the implications of adopting these optimal policies have not been analyzed. In the next section, these implications (in comparison to the competitive equilibrium set out in the previous section) will be analyzed under various dimensions.

Covid Economics 30, 19 June 2020: 83-119

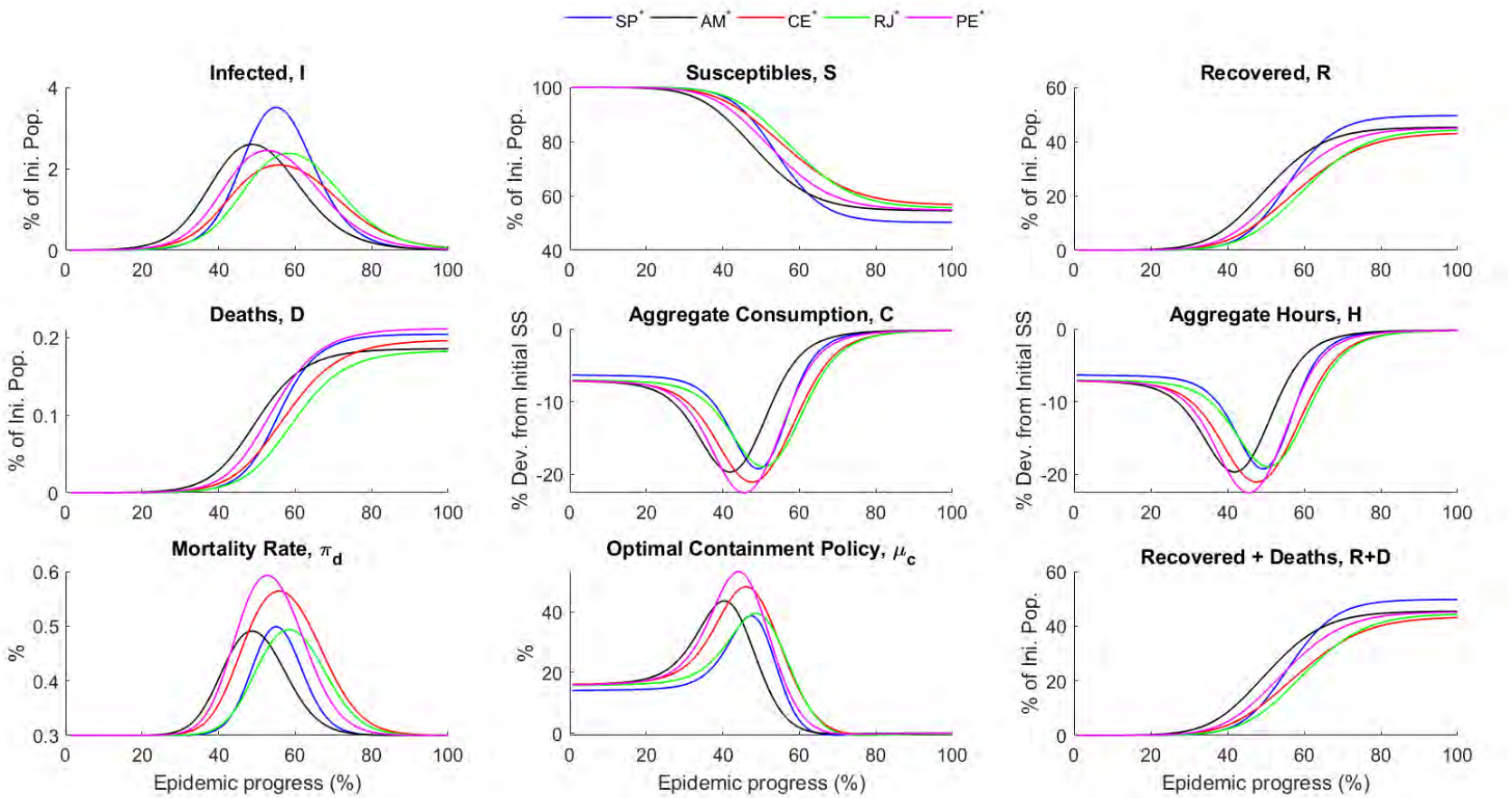
Covid Economics 30, 19 June 2020: 83-119

Figure 3: SIR-macro model - Results under the adoption of optimal containment policies, in level.



Covid Economics 30, 19 June 2020: 83-119

Figure 4: SIR-macro model – Results under the adoption of optimal containment policies, in percentage of the initial population.



6 The implications of optimal containment policies

All the results that will be described in this section are summarized in Table 3. Table 4 summarizes the effects of the adoption of optimal containment policies by the states in comparison to the competitive equilibrium scenario. Table 5 shows, for each scenario, the moments of the progress of the epidemic when the main peaks and valleys in the results of each state occur.

In the scenario under the adoption of optimal containment policies, the state with the largest share of the population infected at the peak is São Paulo (SP), at 3.50%. On the other hand, the lowest is the state of Ceará (CE), at 2.09%. In comparison to the competitive equilibrium, with the adoption of optimal policies, the greatest reduction in the infected peak occurs for the states of Amazonas (AM) and Pernambuco (PE), both with the peak of infected individuals reduced by 1.61 percentage points of the pre-epidemic population. At the other extreme, the smallest reduction occurs in São Paulo (SP), of only 1.44 percentage points.

With regard to time, the peak of infected people occurs first in Amazonas (AM), with 48.67% of the progress of the total time of the epidemic occurred, and lastly in Rio de Janeiro (RJ), with 58.67%. In comparison to the competitive equilibrium, the state that suffers the longest extension of time until the peak of infected due to the adoption of optimal containment policies, is the state of Ceará (CE), which has its peak of infected delayed by 13.33 percentage points of the epidemic's progress. At the other extreme, the one that suffers the least prolongation is the state of São Paulo (SP), with a delay of only 8 percentage points in the progress of the epidemic.

Proceeding for macroeconomic shocks, with the adoption of optimal containment policies there is a worsening recession for all states. The biggest recession happens in the state of Pernambuco (PE), with a drop in aggregate consumption and the aggregate of hours worked of 22.58%. The smallest recession occurs in São Paulo (SP), with a fall of 19.28% of the aggregate consumption and aggregate hours worked. In comparison to the results of competitive equilibrium, the state that suffers the least impact on the recession is Ceará (CE), as a deepening of only 3 percentage points of the aggregates valley. On the other hand, the one that suffers the greatest deepening is the state of Amazonas (AM), of 6.06 percentage points of the pre-epidemic steady state, more than double the deepening suffered by Ceará (CE).

Observing the temporal dynamics, we can notice that the first state to reach the valley is Amazonas (AM), in 42.00% of the progress of the total time of the epidemic, while the one that finally reaches it is Rio de Janeiro (RJ), in 50.67% of the progress of the total time of the epidemic. In comparison to the competitive equilibrium, the states that suffer the greatest increases in the duration of the recession are the states of Ceará (CE) and Rio de Janeiro (RJ), both with an extension of 5.33 percentage points of the time progress of the epidemic in that the valley of recession occurs. On the other hand, the states that experienced the smallest increases in duration were the states of Amazonas (AM) and São Paulo (SP), both with an extension of only 3.33 percentage points.

At the end of the epidemic, the state with the largest share of the population infected becomes the state of São Paulo (SP), with 49.74%. The state with the smallest becomes the state of Ceará

Covid Economics 30, 19 June 2020: 83-119

(CE), with 43.11%. In comparison to the competitive equilibrium, the state that suffers the greatest reduction in this share due to the adoption of optimal containment policies is the state of Amazonas (AM), with a reduction of 4.24 percentage points of the share of the total population that is infected. The state that suffers the smallest reduction is the state of São Paulo (SP), with a reduction of 3.19 percentage points of the share of the total population that is infected.

As a result of infections, the largest share of the initial population that dies occurs in Pernambuco (PE), where it reaches 0.21%. The lowest occurs in the state of Rio de Janeiro (RJ), where it reaches 0.18% of the population. In comparison to the competitive equilibrium, the largest reduction in the portion of the population that dies occurs in the state of Pernambuco (PE), of approximately 0.16 percentage points of the pre-epidemic population, while the smallest reduction occurs in São Paulo (SP), of only 0.07 percentage points of the pre-epidemic population, a reduction of less than half that observed in Pernambuco (PE).

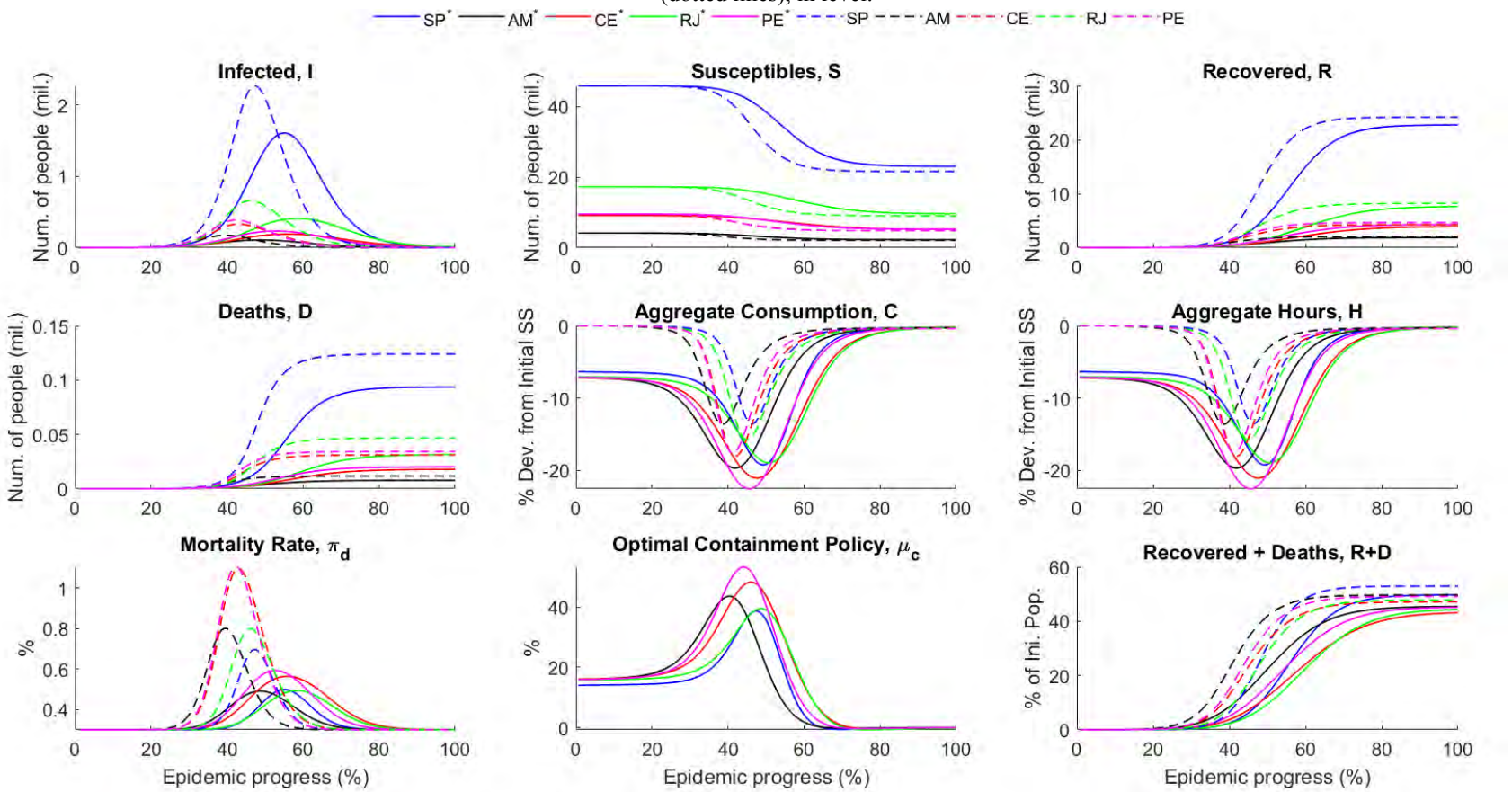
These results show evidence that the intrinsic differences between the states significantly influence the effect of adopting optimal containment policies, heterogeneously affecting the peak of the infected curve and its time dynamics, the severity of economic recessions, the share of the total population that is infected at the end of the epidemic and the percentage of the pre-epidemic population that dies.

Figure 5 shows the comparison of results between the competitive equilibrium and the optimal policy scenario for all states together, in level. Figure 6 shows the same results but as a percentage of the initial population. Figure 7, Figure 8, Figure 9, Figure 10 and Figure 11 show the comparisons of results between the competitive equilibrium and the optimal policy scenario for each of the states separately, in level.

Covid Economics 30, 19 June 2020: 83-119

Covid Economics 30, 19 June 2020: 83-119

Figure 5: SIR-macro model – Comparison of results between the scenarios of optimal containment policy (solid line) and competitive equilibrium (dotted lines), in level.



Covid Economics 30, 19 June 2020: 83-119

Figure 6: SIR-macro model – Comparison of results between the scenarios of optimal containment policy (solid line) and competitive equilibrium (dotted lines), in percentage of the initial population.

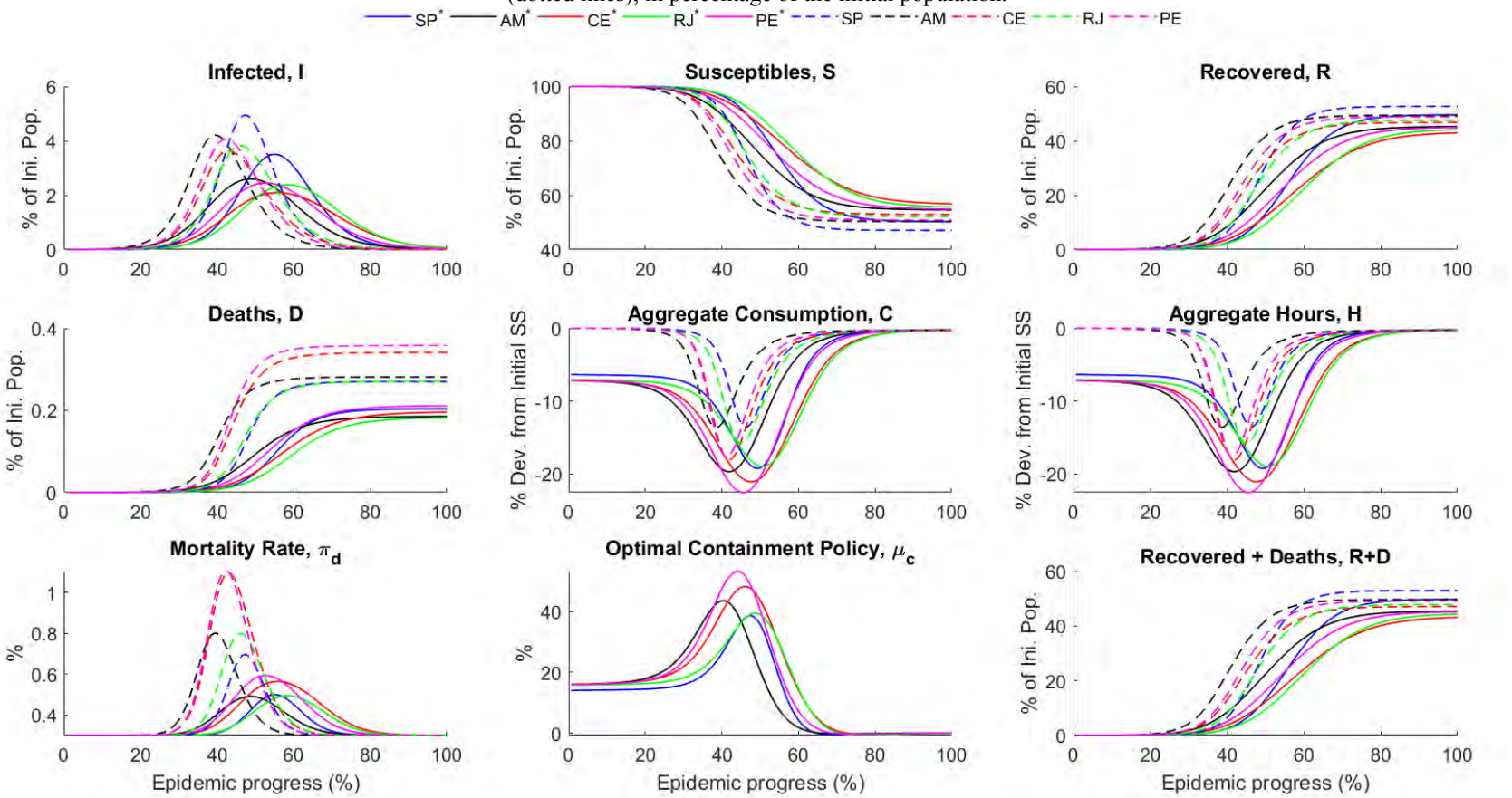


Table 3: Results for the main variables of the model for the competitive equilibrium and optimal policy scenarios.

Variable	Scenario	São Paulo (SP)		Amazonas (AM)		Ceará (CE)		Rio de Janeiro (RJ)		Pernambuco (PE)	
Total uninfected by the end of the epidemic, S	Competitive Eq.	21,614,655.03	(47.07%)	2,086,398.55	(50.34%)	4,828,940.12	(52.88%)	8,999,825.11	(52.13%)	4,858,460.45	(50.84%)
(Number of people. In parentheses: percentage of initial population)	Optimal Policy	23,073,448.23	(50.25%)	2,261,651.74	(54.57%)	5,189,707.29	(56.83%)	9,607,865.98	(55.65%)	5,242,447.63	(54.85%)
Peak infected population, I	Competitive Eq.	2,271,910.63	(4.95%)	174,441.47	(4.21%)	331,309.55	(3.63%)	660,399.78	(3.83%)	387,726.67	(4.06%)
(Number of people. In parentheses: percentage of initial population)	Optimal Policy	1,608,921.78	(3.50%)	107,708.85	(2.60%)	190,939.73	(2.09%)	410,857.91	(2.38%)	234,004.12	(2.45%)
Total recovered by the end of the epidemic, R	Competitive Eq,	24,180,003.96	(52.66%)	2,046,520.78	(49.38%)	4,271,783.72	(46.78%)	8,217,878.32	(47.60%)	4,664,293.18	(48.80%)
(Number of people. In parentheses: percentage of initial population)	Optimal Policy	22,747,425.46	(49.54%)	1,874,748.12	(45.23%)	3,918,618.10	(42.91%)	7,615,853.03	(44.11%)	4,291,798.23	(44.91%)
Total deaths by the end of the epidemic, D	Competitive Eq.	124,101.88	(0.27%)	11,660.84	(0.28%)	31,144.56	(0.34%)	46,800.59	(0.27%)	34,221.33	(0.36%)
(Number of people. In parentheses: percentage of initial population)	Optimal Policy	93,693.42	(0.20%)	7,676.48	(0.19%)	17,884.19	(0.20%)	31,433.62	(0.18%)	20,163.66	(0.21%)
Total infected by the end of the epidemic, R+D	Competitive Eq.	24,304,105.84	(52.93%)	2,058,181.61	(49.66%)	4,302,928.28	(47.12%)	8,264,678.91	(47.87%)	4,698,514.51	(49.16%)
(Number of people. In parentheses: percentage of initial population)	Optimal Policy	22,841,118.88	(49.74%)	1,882,424.60	(45.42%)	3,936,502.30	(43.11%)	7,647,286.65	(44.29%)	4,311,961.89	(45.12%)
Peak mortality rate	Competitive Eq.	0.70%		0.80%		1.10%		0.80%		1.10%	
(%)	Optimal Policy	0.50%		0.49%		0.56%		0.49%		0.59%	
Aggregate consumption in the valley	Competitive Eq.	-13.55%		-13.67%		-18.11%		-15.80%		-17.66%	
(% Dev. from Initial Steady State)	Optimal Policy	-19.28%		-19.72%		-21.10%		-18.98%		-22.58%	
Aggregate of hours worked in the valley	Competitive Eq.	-13.55%		-13.67%		-18.11%		-15.80%		-17.66%	
(% Dev. from Initial Steady State)	Optimal Policy	-19.28%		-19.72%		-21.10%		-18.98%		-22.58%	

Covid Economics 30, 19 June 2020: 83-119

Covid Economics 30, 19 June 2020: 83-119

Table 4: Effects of the adoption of optimal containment policies by the states in relation to the competitive equilibrium scenario.

Value	SP	AM	CE	RJ	PE
Peak containment rate (%)	38.76%	43.68%	48.35%	39.54%	53.40%
Infected people avoided at peak (Number of people. In parentheses: percentage of initial population)	662,989 (1.44%)	66,733 (1.61%)	140,370 (1.54%)	249,542 (1.45%)	153,723 (1.61%)
Total infections avoided (Number of people. In parentheses: percentage of initial population)	1,462,987 (3.19%)	175,757 (4.24%)	366,426 (4.01%)	617,392 (3.58%)	386,553 (4.04%)
Peak mortality rate reduction (%)	-0.20%	-0.31%	-0.53%	-0.31%	-0.51%
Saved lives (Number of people. In parentheses: percentage of initial population)	30,408 (0.07%)	3,984 (0.10%)	13,260 (0.15%)	15,367 (0.09%)	14,058 (0.15%)
Effect on aggregate consumption in the valley (% Dev. From Initial Steady State)	-5.73%	-6.06%	-2.99%	-3.18%	-4.92%
Effect on aggregate hours worked in the valley (% Dev. From Initial Steady State)	-5.73%	-6.06%	-2.99%	-3.18%	-4.92%

CEPR PRESS

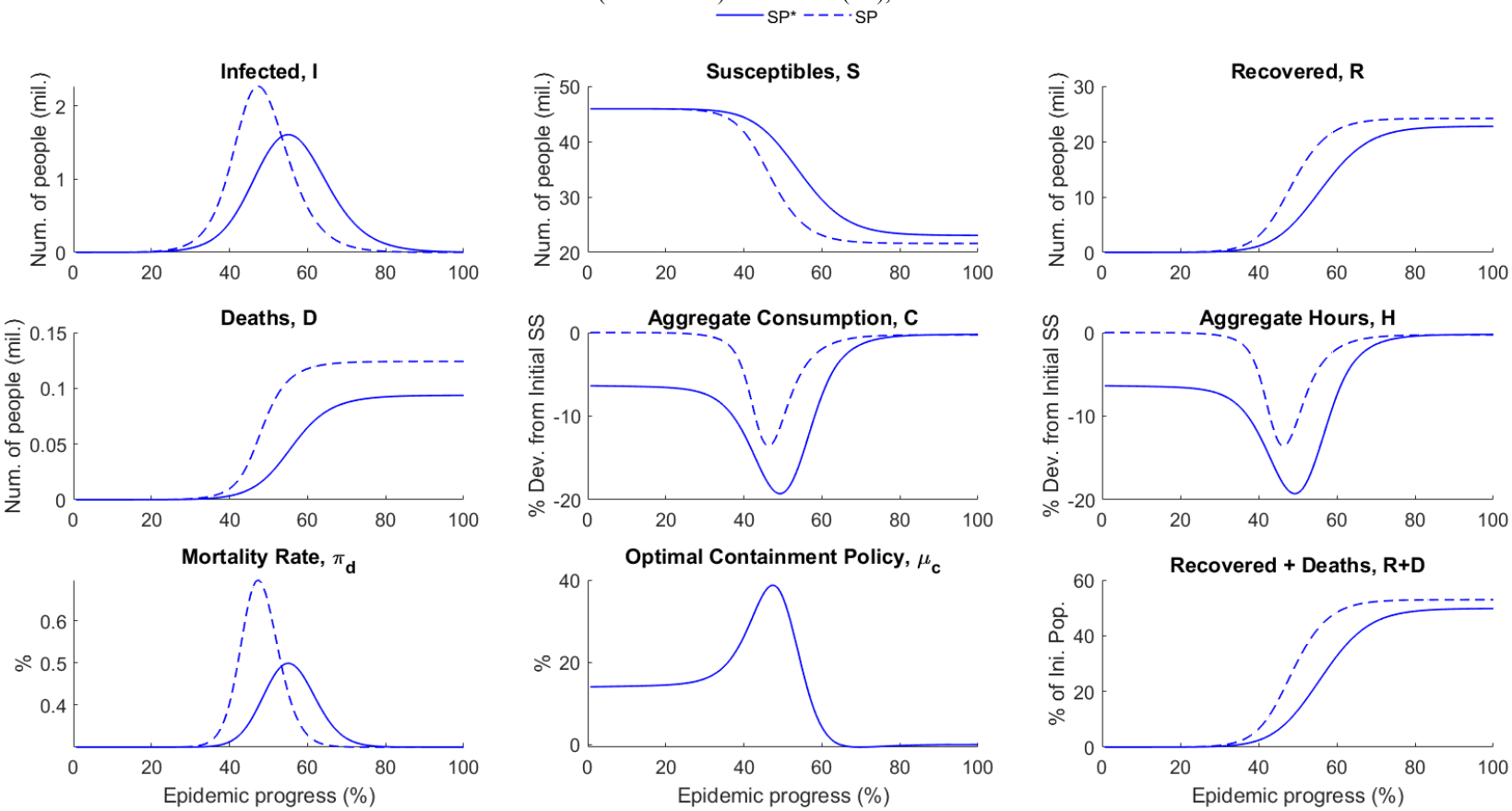
Table 5: Moments of progress of the epidemic in which the peaks and valley of the main variables of the model occur in each scenario, for each state.

Value	Scenario	SP	AM	CE	RJ	PE
Moment of progress of the epidemic in which the peak of infected people occurs, I (% progress of the total time of the epidemic)	**Competitive Eq.**	47.33%	39.33%	42.67%	46.00%	42.00%
	Optimal Policy	55.33%	48.67%	56.00%	58.67%	52.67%
Moment of progress of the epidemic when peak containment rate occurs, μ_t (% progress of the total time of the epidemic)	**Optimal Policy**	47.33%	40.67%	46.00%	48.67%	44.00%
Moment of progress of the epidemic when the peak mortality rate occurs, π_d (% progress of the total time of the epidemic)	**Competitive Eq.**	47.33%	39.33%	42.67%	46.00%	42.00%
	Optimal Policy	55.33%	48.67%	56.00%	58.67%	52.67%
Moment of progress of the epidemic when the valley of recession occurs (% progress of the total time of the epidemic)	**Competitive Eq.**	46.00%	38.67%	42.00%	45.33%	41.33%
	Optimal Policy	49.33%	42.00%	47.33%	50.67%	45.33%

Covid Economics 30, 19 June 2020: 83-119

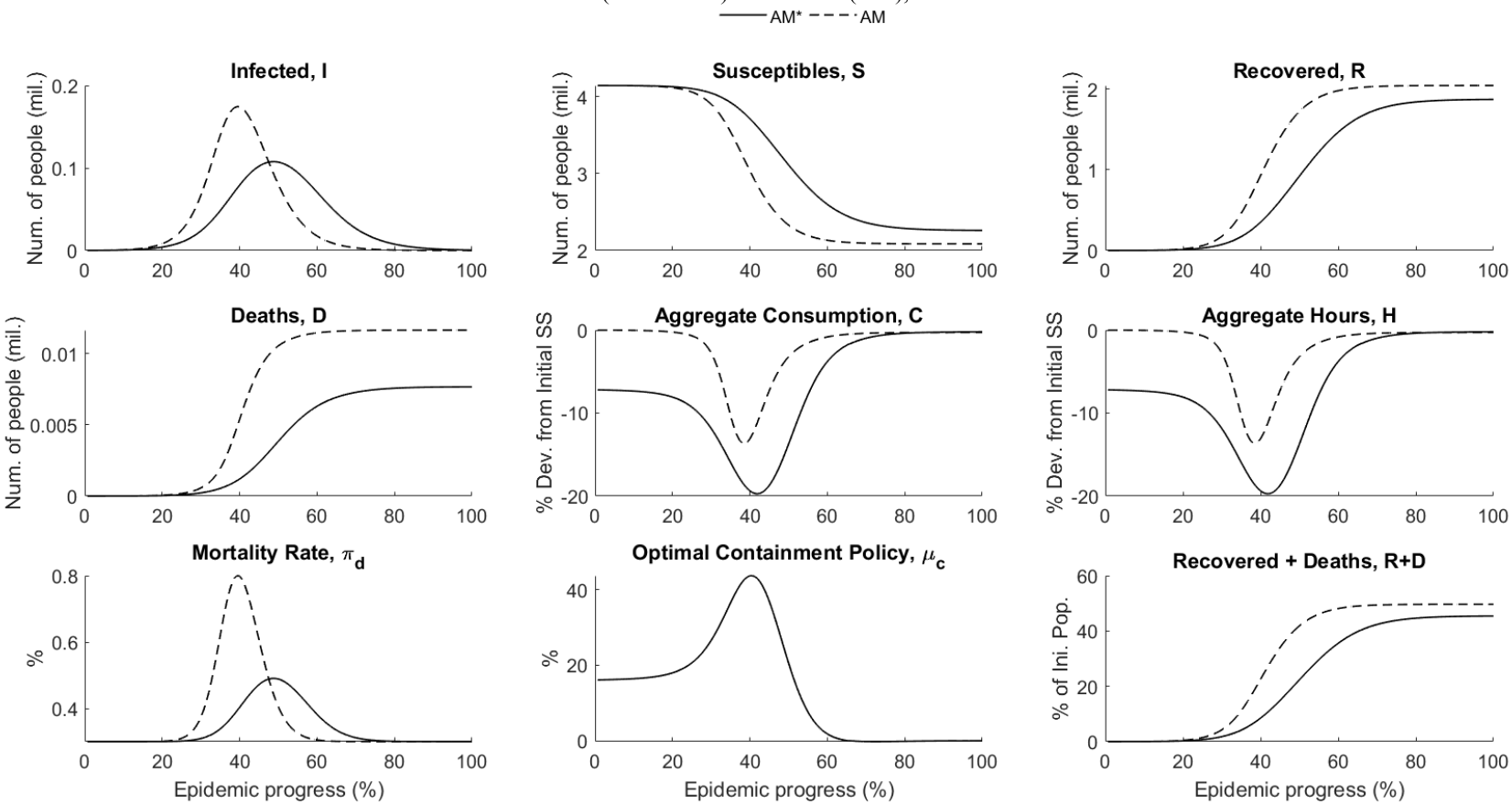
Covid Economics 30, 19 June 2020: 83-119

Figure 7: SIR-macro model – Comparison of results between the scenarios of optimal containment policy (solid line) and competitive equilibrium (dotted lines): São Paulo (SP), in level.



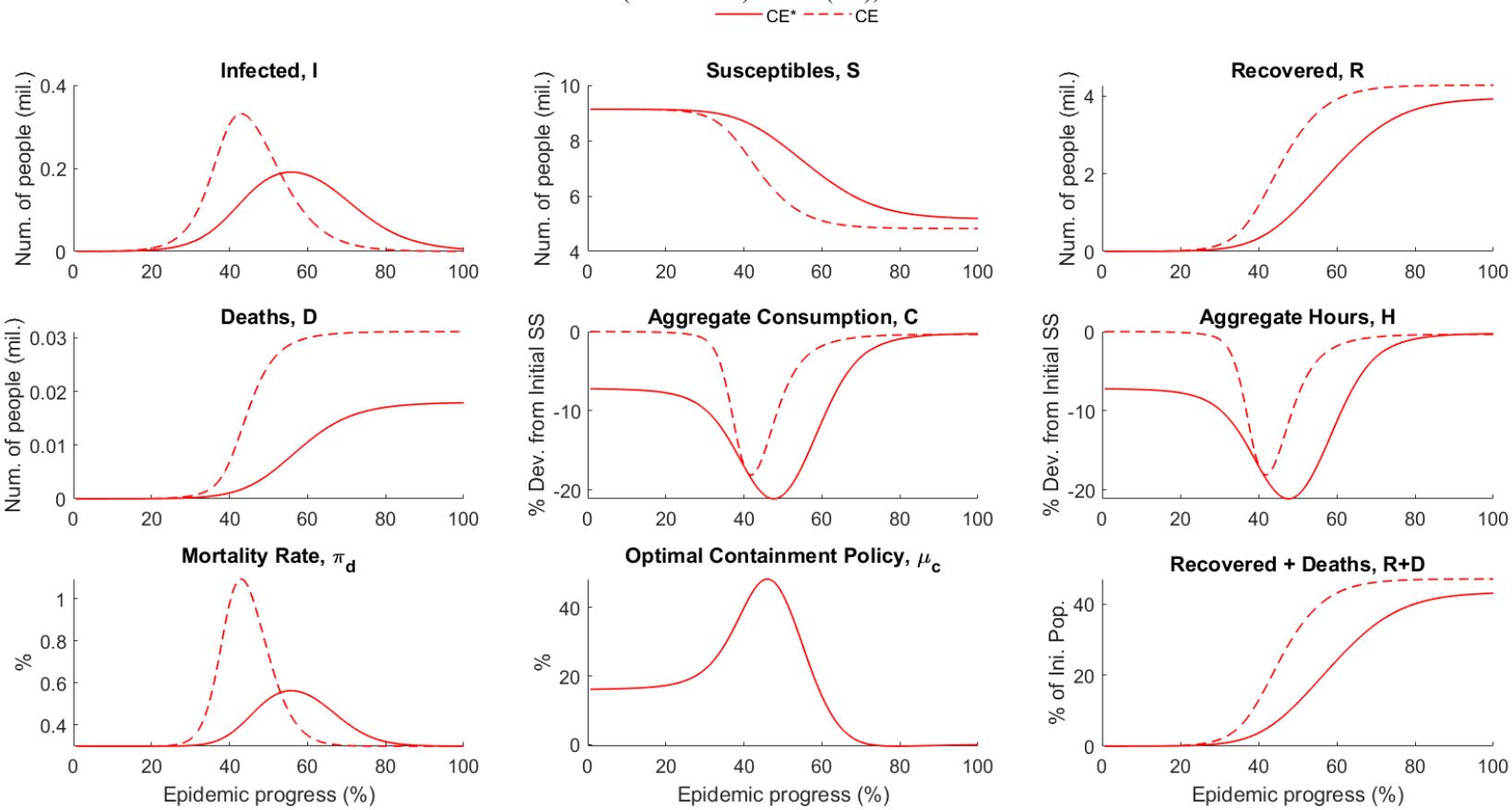
Covid Economics 30, 19 June 2020: 83-119

Figure 8: SIR-macro model – Comparison of results between the scenarios of optimal containment policy (solid line) and competitive equilibrium (dotted lines): Amazonas (AM), in level.



CEPR PRESS

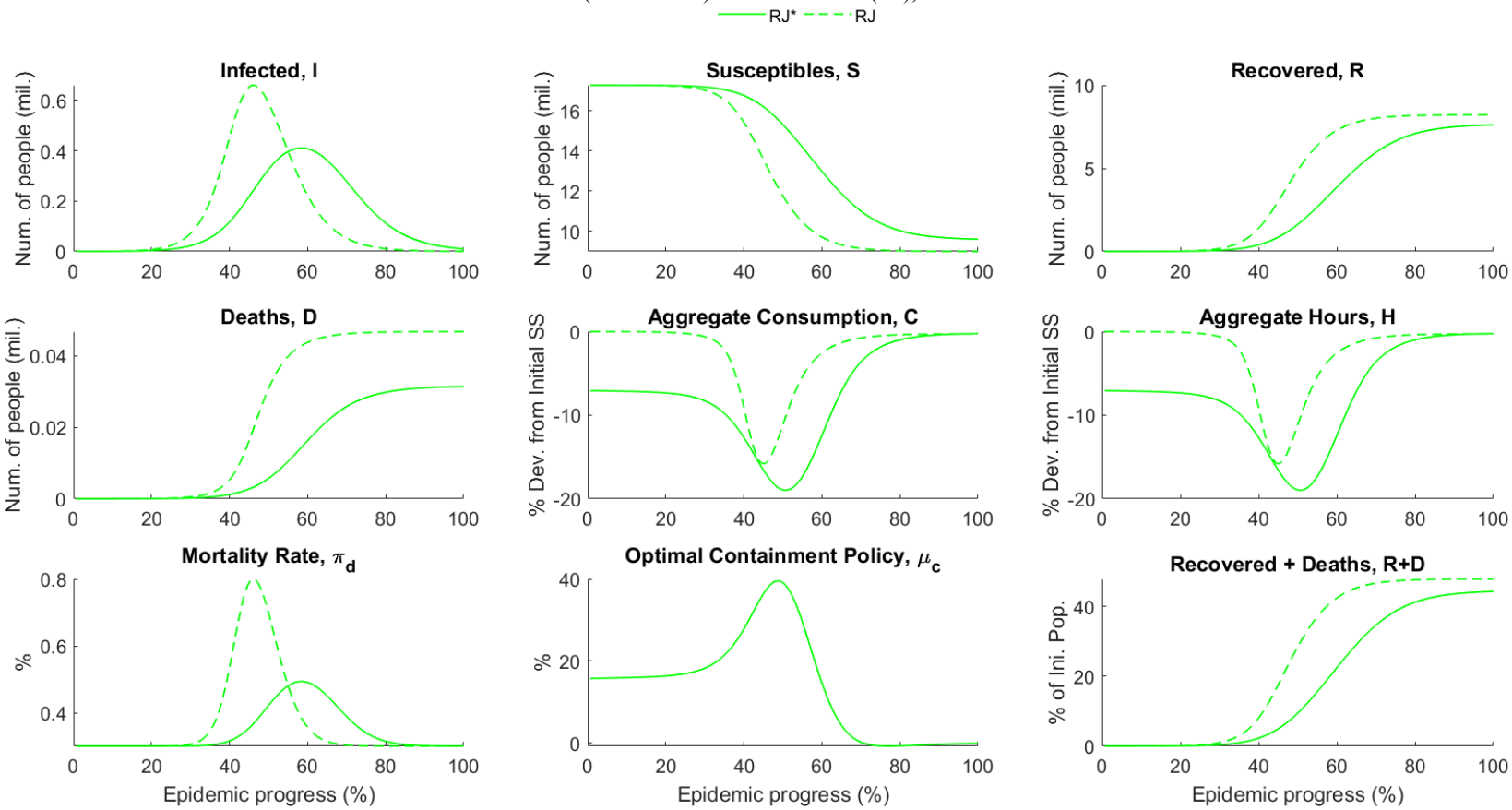
Figure 9: SIR-macro model – Comparison of results between the scenarios of optimal containment policy (solid line) and competitive equilibrium (dotted lines): Ceará (CE), in level.



Covid Economics 30, 19 June 2020: 83-119

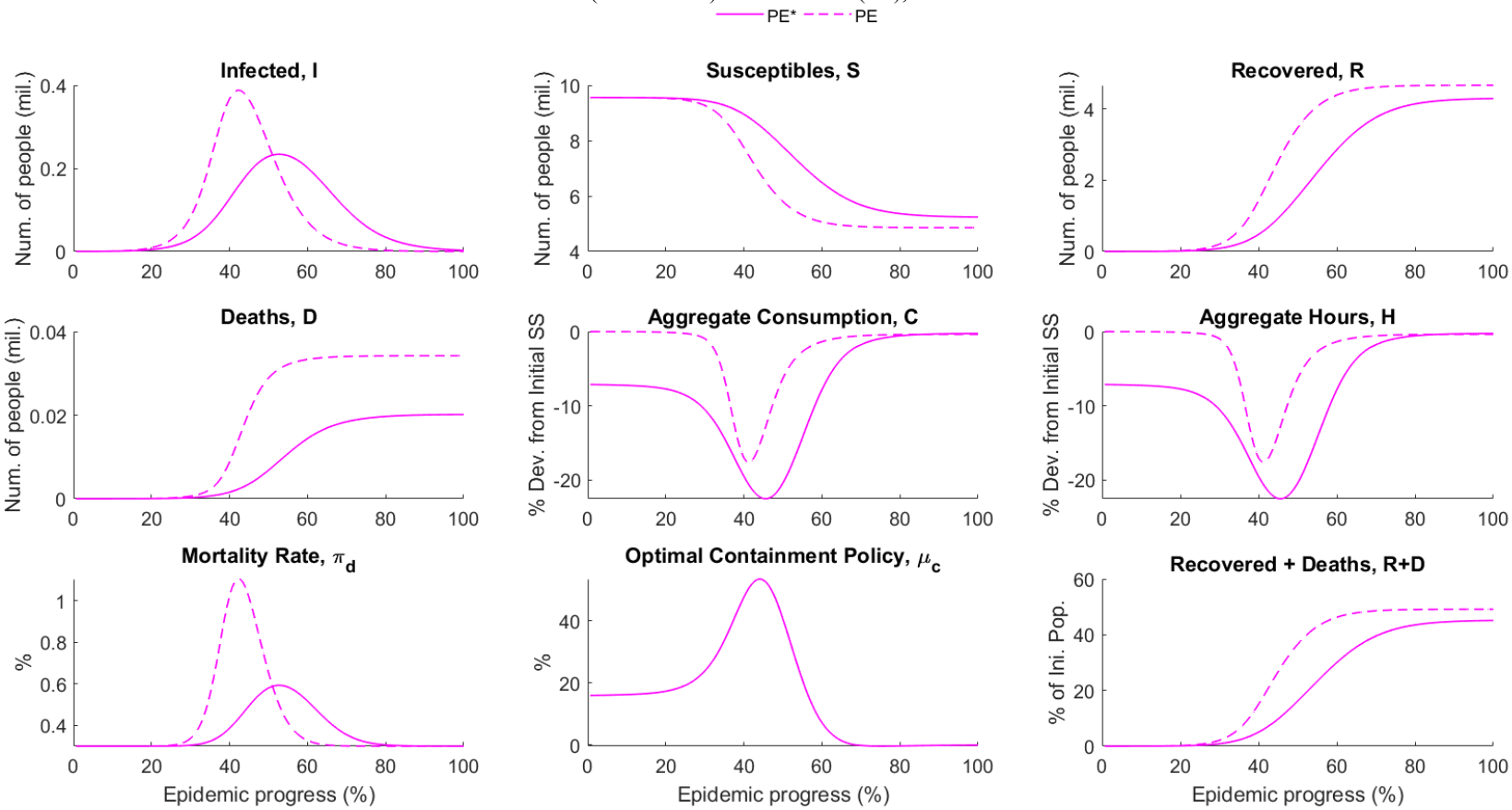
Covid Economics 30, 19 June 2020: 83-119

Figure 10: SIR-macro model – Comparison of results between the scenarios of optimal containment policy (solid line) and competitive equilibrium (dotted lines): Rio de Janeiro (RJ), in level.



Covid Economics 30, 19 June 2020: 83-119

Figure 11: SIR-macro model – Comparison of results between the scenarios of optimal containment policy (solid line) and competitive equilibrium (dotted lines): Pernambuco (PE), in level.



7 Conclusion

Our interest was to analyze qualitatively how the main intrinsic differences of each state can affect its epidemic dynamics and results. In particular, our exercises, carried out for five states selected by the severity of the epidemic they presented at the time of the research, pointed to substantial differences in the evolution and outcome of the epidemic in each state, both in the competitive equilibrium scenario and in the scenario in which optimal containment policies are adopted. In both scenarios, we found evidence of differences in i) the size of the infected population peak, ii) the time required to reach the infected population peak, iii) the depth of the recessions, iv) the duration of the recessions, v) the share of the initial population that has been infected by the end of the epidemic and vi) the share of the initial population that died by the end of the epidemic in each state.

Our results present evidence that the intrinsic differences of each state also imply different optimal trajectories of containment policies for each one of them. These intrinsic differences affect i) the initial severity of the containment measures that ideally should be adopted by each state and the dynamics of the evolution of these containment rates over time, implying ii) differences in the ideal moment that the containment measures will need to be elevated, iii) how far they will need to be elevated and iv) when they can finally be relaxed.

Using the simple framework of the SIR-macro model proposed by Eichenbaum et al. (2020) and analyzing the results obtained for each state we conclude that the main implicit characteristics of the five different states imply relevant differences in

i) The epidemic dynamics and its general epidemiological consequences;
ii) The optimal containment policies to be adopted by each state;
iii) The effect of adopting optimal containment policies;
iv) The severity of the economic recessions from the epidemic.

This conclusion emphasizes the importance of the disaggregated analysis of countries with huge geographic and demographic dimensions like Brazil in the formulation of policies to combat COVID-19. We warn that the adoption of unique aggregate policies for huge and heterogeneous countries like Brazil can trigger a series of containment policy errors in the states and regions of the country, deepening both the economic recession and the number of deaths resulting from the epidemic. The propagation channel of these errors is the following: if there is large heterogeneity in the optimal level of containment policy required for each state, a single containment policy for the whole country would be unable to adequately deal with the needs of all states simultaneously. Inevitably, the containment rate adopted across the country would be lower than needed for some states, resulting in more deaths, and higher than needed for others, resulting in unnecessarily in deeper economic recessions.

Covid Economics 30, 19 June 2020: 83-119

8 References

[1] Eichenbaum, Martin S., Rebelo, S., and Trabandt, M. (2020, March) "The Macroeconomics of Epidemics", *NBER Working Paper* No. 26882.

[2] Farhi, E., and Werning, I. (2020, June) "Dealing with the Trilemma: Optimal Capital Controls with Fixed Exchange Rates", *NBER Working Paper* No. 18199.

[3] Ferguson, N., Cummings, D., Fraser, C., Cajka, J., Cooley, P., and Burke, D. (2006, April) "Strategies for mitigating an influenza epidemic", *Nature*, 442, 448-452.

[4] Ferrari, T., Dusi, L., Lopes, D., and Pompermayer, M. (2019, December) "Estimativa do valor da vida estatística e do valor da economia de tempo em viagens nas rodovias brasileiras com a utilização de pesquisa de preferência declarada", *Texto para Discussão IPEA* No. 2533.

[5] Hallal, P. et al. (2020, May) "Remarkable variability in SARS-CoV-2 antibodies across Brazilian regions: nationwide serological household survey in 27 states", *medRxiv* doi: 10.1101/2020.05.30.20117531.

[6] Kermack, W., and McKendrick, A. (1927) "A Contribution to the Mathematical Theory of Epidemics", *Proceedings of the Royal Society of London*, series A 115, no. 772: 700-721, 1927.

[7] Lee, B., Brown, S., Cooley, P., Zimmerman, R., Wheaton, W., Zimmer, S, Grefenstette, J., Assi, T.-M., Furphy, T., and Wagener, D. (2010, March) "A computer simulation of employee vaccination to mitigate an influenza epidemic", *American Journal of preventive medicine*, 38, 247-257.

[8] Mellan, T. et al. (2020, May) "Report 21: Estimating COVID-19 cases and reproduction number in Brazil", *Imperial College COVID-19 Response Team.*

[9] Rabelo, M., and Soares, J. (2020, April) "The Macroeconomics of Epidemics: results for Brazil", Draft.

[10] Rocha, G., de Morais, R., and Klug, L. (2019, October) "O custo econômico da poluição do ar: Estimativa de valor da vida estatística para o Brasil", *Texto para Discussão IPEA* No. 2517.

Covid Economics 30, 19 June 2020: 83-119

9 Appendix: Calibration of α_1, α_2 e α_3

We calibrate the shares of the transmissions that occur in consumption, at work and in other activities as proportional to the time that individuals spend performing these activities. To calibrate α_1, Eichenbaum et al. (2020) uses data from time use surveys produced by the United States Bureau of Labor Statistics. Given the lack of equivalent surveys for Brazil, directly calibrating the share of transmissions that occur in consumption has become a very complicated task for Brazilian states. We then decided to calibrate the values of α_2 and α_3 and obtain the values of α_1 residually. We calibrated the share of transmissions that occur at work, α_2, using the same approach as Eichenbaum et al. (2020). We weight the number of students and workers in each state at 10 and 4, respectively. These weights refer to the average amount of physical contacts per day at school and at work, obtained at work by Lee et al. (2010). We disaggregate the fraction of transmissions that occur in schools and work environments to obtain the fraction that occurs only in work environments considering, for this, a proportionality involving the average number of contacts, the population of workers and the student population of each state. We calculated the proportion of infections that occurs only at work and, finally, we obtain the value of α_2 by multiplying it by the share of infections that occurs in schools and work environments (37%). Thus, for each state the value of α_2 will be calculated as:

$$\alpha_2 = \frac{4 \cdot (Workers)}{4 \cdot (Workers) + 10 \cdot (Students)} \cdot 37\%, \tag{34}$$

where $Workers$ is the total number of workers in the state and $Students$ is the total number of students in the state. For the total number of workers we used the number of people aged 14 and over employed in the labor force obtained from IBGE's SIDRA[5]. For the total number of students, we used the number of students aged 4 years and over, also obtained from IBGE's SIDRA[6].

Now we proceed to calibrate the share of infections that occurs in other activities, α_3. This share is by definition the portion of infections that occurs in all activities except those related to consumption and work. Obviously, there are many of these activities, so that the need for simplifying hypotheses for its calculation becomes inevitable.

We approach this problem as follows: we assume as a calibration hypothesis that the vast majority of infections that occur outside of consumption and work occur mainly at home, at school or in public transport. This hypothesis is defined according to our intuition regarding the nature of the virus transmission inserted in the Brazilian reality and the availability of data for the states. We know that this hypothesis leaves out a number of other circumstances and environments in which the transmission of the disease could occur, perhaps the most significant of them being the hospitals themselves. However, the absence of specific state data related to these conditions and environments does not allow us to accurately measure these shares.

[5] Ibid., Table 4093.
[6] Ibid., Table 983.

Covid Economics 30, 19 June 2020: 83-119

We deal with this problem as follows: if we calibrate α_3 based only on the time people spend at home, in public transport and at school, the resulting value is very low and, as a consequence, calculating the α_1 portion related to infections that occur in consumption, we observe that this portion is, on the other hand, unrealistically high. This is due to two main factors. Firstly, due to the previously mentioned issue (our calibration hypothesis for α_3 leaves out many environments with a high degree of infection, such as hospitals, causing the resulting value for this portion to be underestimated) and, secondly, because the share of infections should not be explained only by the time spent in each environment, but by that time driven by a multiplying factor related to the characteristics of the environment in question. All of these findings indicate that the value of α_3, if calculated in this way, is underestimated, which gives us room to add it by alternative means to incorporate the effect of these omitted factor.

We have chosen to incorporate the contribution of the omitted environments to infections in the transport multiplier factor. But how big should this multiplier factor be? We calibrated it in order to obtain values for the α_3 shares of the states that revolve around 1/2 to 2/3, magnitudes consistent with the value obtained by Eichenbaum et al. (2020) for the United States, where the availability of data enabled the residual calculation of α_3, thus contributing to the greater reliability of the estimate of the magnitude of this share. As this multiplier factor will be common for all states, we argue that there is no qualitative loss for the model, since the relative differences in calibration between states will not be changed.

To calibrate α_3 we then proceed as follows: we calculate the share of infections that occurs at home, ${\alpha_3}^{HOME}$, in schools, ${\alpha_3}^{SCHOOLS}$, and in transport, ${\alpha_3}^{TRANSP}$, and add them up, so that $\alpha_3 = {\alpha_3}^{HOME} + {\alpha_3}^{SCHOOLS} + {\alpha_3}^{TRANSP}$. Now, note that with the calculation of the share of infections that occurs in the workplace, α_2, it is consequently also obtained the share of infections that occurs in schools:

$$ {\alpha_3}^{SCHOOLS} = 37\% \cdot \frac{10 \cdot (Students)}{4 \cdot (Workers) + 10 \cdot (Students)}. \tag{35} $$

The α_3 calibration problem is therefore reduced to calibrating the fractions of infections that occur at home, ${\alpha_3}^{CASA}$, and in transport, ${\alpha_3}^{TRANSP}$.

To calibrate the share of infections that occurs at home, we use data on the proportion of daily hours dedicated to the care of people or household chores of people aged 14 and over (%), obtained from IBGE's SIDRA[7], to calculate the average proportion of daily time that individuals in each state spend at home in situations susceptible to infection. As sleep time cannot be considered time susceptible to infection, we consider an average sleep time of 8 hours a day and calculate the proportion as the daily hours dedicated to the care of people or household chores in each state divided by 16 (24-8). Remember, the work of Ferguson et al. (2006), that 30% of infections occur at home. We then multiplied this proportion by 30% and further, under the hypothesis that the more people per household, the greater the share of infections within the home, by a multiplying factor characterized by the average number of

[7] Ibid., Table 6730.

Covid Economics 30, 19 June 2020: 83-119

people per household in 2019, obtained from IBGE's SIDRA[8] for each state. The resulting value of ${\alpha_3}^{CASA}$ for each state is therefore described by the equation

$$ {\alpha_3}^{HOME} = 30\% \cdot \frac{(Household\ chores)}{16} \cdot (People\ per\ household) \tag{36} $$

where $Household\ chores$ is the average daily time that individuals spend on household chores in the state and $People\ per\ household$ is the average number of people living per household in the state.

To calibrate the fraction that occurs in transport, ${\alpha_3}^{TRANSP}$, we proceeded using the usual commuting time to work of individuals, obtained from IBGE's SIDRA[9] for each state, to calculate the fraction of the time spent on general activities devoted to the use of transport. We consider the time for general activities to be 24 hours a day minus the time spent on household chores, sleep, work and school. Having obtained the fraction of time devoted to general activities that is intended for the use of transport, we multiply it by the share of infections that occur in general activities (33%) to obtain the share of infections that occur in transport. Finally, we calibrate the transport multiplying factor so that the values of α_3 revolves around 1/2 to 2/3 for all states, a range of values consistent with the values of α_3 calibrated in Eichenbaum et al. (2020) for the United States, where the availability of data for calibration is broader. The multiplier fator is therefore defined as 10. The value of ${\alpha_3}^{TRANSP}$ is then obtained for each state by calculating

$$ {\alpha_3}^{TRANSP} = 33\% \cdot \frac{(Average\ time\ in\ transport)}{16 - (Household\ chores) - (Work\ or\ school)} \cdot 10 \tag{37} $$

With that, finally the value of α_3 is obtained as being

$$ \alpha_3 = {\alpha_3}^{HOME} + {\alpha_3}^{SCHOOL} + {\alpha_3}^{TRANSP}. \tag{38} $$

Having obtained the shares α_1, α_2 e α_3, it will be enough to solve the system (*31*) to obtain the values of π_1, π_2 and π_3.

[8] Ibid., Table 6578.
[9] Ibid., Table 3422.

Covid Economics 30, 19 June 2020: 83-119

Covid Economics Issue 30, 19 June 2020

The COVID-19 pandemic in the EU: Macroeconomic transmission and economic policy response[1]

Philipp Pfeiffer,[2] Werner Roeger[3] and Jan in 't Veld[4]

Date submitted: 17 June 2020; Date accepted: 18 June 2020

This paper uses a macroeconomic model to analyse the transmission of the COVID19-pandemic and its associated lockdown and quantify the stabilising effects of the economic policy response. Our simulations identify firm liquidity problems as crucial for shock propagation and amplification. We then quantify the effects of short-term work allowances and liquidity guarantees - central policy strategies in the European Union. The measures reduce the output loss of COVID19 and its associated lockdown by about one fourth. However, they cannot prevent a sharp but temporary decline in production.

Covid Economics 30, 19 June 2020: 120-145

1 The views expressed in this paper are those of the authors and should not be attributed to the European Commission. We would like to thank Declan Costello, Björn Döhring, Robert Kuenzel, João Leal, José Leandro, Hauke Vierke, Lukas Vogel, Milan Vyskrabka and many colleagues involved in Directorate-General for Economic and Financial Affairs' forecast for useful comments and suggestions.

2 Economic Analyst at the European Commission, Directorate-General for Economic and Financial Affairs.
3 Head of the Unit "Models and databases" at the European Commission, Directorate-General for Economic and Financial Affairs.
4 Head of the Sector "Model-based economic analysis" at the European Commission, Directorate-General for Economic and Financial Affairs.

1. INTRODUCTION

COVID19 is a rare disaster requiring urgent and targeted policy action. The economic fallout of the pandemic has triggered an intense debate on effective policies. Governments have announced and implemented exceptional stabilisation efforts. Our paper contributes to this debate using a dynamic general equilibrium model and an up-to-date collection of fiscal policy measures in the EU. We assess the economic transmission of COVID-19 and quantify gains from stabilisation policies in the European Union (EU). Our model is two-region TANK model (two-agent New Keynesian) consisting of the EU-27 and the rest-of-the world (RoW) based on the European Commission's QUEST III model (Ratto et al., 2009). We extend the baseline model with a parsimonious model of heterogeneous firms and liquidity constraints.

Our analysis proceeds in two steps. First, a set of scenarios analyses the essential economic characteristics of the pandemic.[1] Stylised shocks capture the dynamic adjustment of the economy to (i) supply shocks through precautionary measures and (ii) shortfalls in consumer demand. Precautionary supply-side measures constrain labour input in production. The demand lockdown prevents consumers from undertaking certain consumption activities by either legal restrictions or voluntary protection. We quantify the impact of these shocks on the depth and duration of the recession. In addition, a simple financial accelerator mechanism via firm liquidity constraints amplifies the shocks. Without policy support, the simulations show an output contraction by more than 20% in the second quarter of 2020 compared to a no-pandemic baseline. Firm liquidity constraints almost double the depth of the recession and substantially prolong the slump. This amplification leads to a more U-shaped recovery. Output in 2021 remains markedly below a no-pandemic baseline (almost 40% of the 2020 impact remains in 2021). By contrast, a standard model without borrowing constraints does not generate a (large) fall investment based on the temporary lockdown shocks to demand and supply alone. As a result, the simulations without firm liquidity constraints (and in the absence of other shocks), show a V-shaped recovery.

In a second step, we turn to an early assessment of economic policy. The predicted collapse is unprecedented in post-war history. It warrants systematic comparison and quantification of the main policy strategies based on the identified transmission channels. Concretely, we focus on stabilisation gains from short-term work (STW) allowances and government guarantees. Besides automatic stabilisers, STW allowances and government guarantees are the dominant forms of fiscal response in the EU. Exploiting an up-to-date dataset of planned fiscal measures in the EU-27, we then show that together economic policies likely eliminate about one-fourth of the macroeconomic fallout of the pandemic.

At the heart of our policy analysis are two interacting adjustment inefficiencies, namely employment adjustment costs and liquidity constraints for firms. Measures targeted at reducing these distortions, such as STW allowances and loan guarantees, prevent a considerable revision of investment plans. Alleviating liquidity constraints supports spending in wages, intermediates, investment, and servicing of loans. STW allowances are particularly effective. Since workers stay in the firm, firms avoid matching frictions and hiring or firing costs. This cost-saving channel, in turn, improves the corporate liquidity position by stabilising the gross operating surplus. As a result, fewer firms become constrained, limiting the amplification arising from the occasionally binding constraint. Apart from short-run expansionary effects, these measures may avoid that temporary liquidity problems morph into insolvency issues. Together STW and guarantees reduce the output loss of COVID19 and its

[1] As discussed in Section 1.1, the scenarios are based on stylised versions of the simulations done in the context of the European Commission's Spring Forecast published on May 7, 2020: https://ec.europa.eu/info/business-economy-euro/economic-performance-and-forecasts/economic-forecasts/spring-2020-economic-forecast-deep-and-uneven-recession-uncertain-recovery_en

Covid Economics 30, 19 June 2020: 120-145

associated lockdown by around four percentage points: one fourth of the negative economic impact of the pandemic.

The policies, however, cannot prevent a temporary decline in production because of sickness, restrictions on the mobility of workers or supply chain interruptions. Moreover, additional consumption from transfers can only (fully) materialise after the lockdown period. It cannot prevent a sharp drop in consumer spending during the pandemic.

We analyse our results for a given short-lived pandemic. This optimistic simplification is designed to transparently distil essential model features, in particular the internal propagation mechanism in response to the exogenous pandemic shocks. An additional longer pandemic scenario relaxes this assumption to show the sensitivity in a stylised way.

1.1. RELATED LITERATURE

There have been widespread calls for policies to mitigate the impact of the COVID-19 shock (e.g., Brunnermeier et al., 2020; Gopinath, 2020; and papers in Baldwin and Weder di Mauro, 2020). While monetary policy plays a crucial role in safeguarding liquidity conditions in the banking system and protecting the continued flow of credit to the real economy, the space of central banks is largely constrained. It is generally recognised that the ECB should not be expected to do all the heavy lifting. Economists have therefore emphasised that governments must step in with generous loans and other support programs to prevent mass bankruptcies, and direct fiscal measures to support demand when the lockdowns are lifted (e.g. Bénassy-Quéré et al. (2020), Lane (2020), Claes and Wolff (2020), and others).

In an assessment of the efficacy of fiscal policies in the financial crisis, Coenen et al. (2012) compare transfer multipliers in seven large-scale DSGE models. In these models, spending multipliers are typically largest, often exceeding unity when the zero lower bound (ZLB) constrains monetary policy. By contrast, tax multipliers remain lower. Typically, transfer multipliers are also lower than government consumption and investment shocks. However, targeted transfers to constrained households entail multipliers closer to those of spending shocks. Guerrieri et al. (2020) emphasise that when some sectors are shut down, a traditional fiscal stimulus is less effective, as any money spent cannot go to "closed" sectors, whose workers have the greater marginal propensity to consume. As long as there are sectors shut down, there is a unit government spending multiplier and a transfer multiplier equal to the average marginal propensity to consume. Faria e Castro (2020) analyses different types of fiscal policies and finds that unemployment insurance benefits are the most effective tool to stabilise income for borrowers, who are the hardest hit, while savers may favour unconditional transfers. Liquidity assistance programs are effective if the policy objective is to stabilise employment in the affected sector. Bayer et al. (2020) also emphasise the importance of conditional transfers. Fornaro and Wolf (2020) emphasise the role expectations and the risks of demand-induced growth slowdowns.

Our paper considers firm and household heterogeneity. Yet, the model remains parsimonious. Important related work complements our analysis by considering a more granular economic structure. Hagedorn and Mitman (2020) apply a HANK model to study the interaction of fiscal and monetary policy. Guerrieri et al. (2020) as well as Baqaee and Farhi (2020) use multi-sector models to show amplification effects through complementarities and incomplete markets. Bigio et al. (2020) compare transfers to credit policy and highlight the role of debt.

The economic literature on COVID19 is growing rapidly. Important contributions also link economic models with epidemiological frameworks (e.g., Acemoglu et al., 2020; Eichenbaum et al., 2020; Glover et al., 2020; and Jones et al., 2020; and references therein).

Covid Economics 30, 19 June 2020: 120-145

We contribute to this literature by implementing a lockdown shock in a standard macro model and highlighting the role of liquidity constraints. Exploiting information about size and composition of EU fiscal measures, our paper then shows the depth of the recession and the shape of the recovery with and without fiscal measures.

The paper is closely related to the scenario analysis published in the Spring Forecast of the European Commission (2020). Compared to the present paper, the simulations reported in the forecast document consider a richer modelling environmental with deeper regional and sectoral disaggregation and more transmissions channels (e.g. uncertainty shocks) and time patterns of the pandemic. By contrast, here we stress the role of inefficiencies arising from employment adjustment and liquidity constraints in a more stylised way.

1.2. ROAD MAP

The next two sections present the model, its calibration, and the underlying assumptions of the pandemic shock. Section 4 shows simulations results absent policy intervention, while Section 5 summarises announced policy measures and the mapping into the model. Section 6 conducts robustness analysis and Section 7 concludes.

Covid Economics 30, 19 June 2020: 120-145

2. MODEL

We conduct our analysis in a two-region TANK model (two-agent New Keynesian) consisting of the EU-27 and the RoW. The framework is based on the European Commission's QUEST III model suite (see, e.g., Ratto et al., 2009). Our discussion, therefore, focusses on the main model elements and refers for standard features to Ratto et al. (2009). We extend the baseline model with a parsimonious model of firm liquidity constraints.

The model structure of all regions is symmetric. It includes nominal price and wage rigidities as well as adjustment costs associated with employment and investment. Households provide labour services to domestic firms. A share of households is liquidity constrained. Monopolistic trade unions set sticky wage rates. Governments purchase the local final good; make transfers to households; levy labour, profit, and consumption taxes; and issue debt. We integrate automatic fiscal stabilisation via tax revenues, constant spending in real terms and unemployment insurance. Trade and financial markets link the EU (based on EU-27 shares) and the rest-of-world. A limited interest rate response captures restricted monetary policy. We next present the core of the EU model block. The RoW block has the same structure except for the zero lower bound (ZLB) constraint on monetary policy (but features a block-specific calibration). $*$-superscript denotes RoW variables.[2] To ease notation, the presentation abstracts from linear taxation of consumption, labour, and profits.

2.1. HOUSEHOLDS

The household sector consists of two representative households $h \in \{R, C\}$, of total mass one. The Ricardian household, indexed R, enjoys full access to financial markets. The other household is liquidity-constrained and indexed by C. This household does not trade on asset markets. Instead, she

[2] Parameters such as the degree of openness differ across the EU and RoW. See the discussion below.

consumes her entire disposable wage and transfer income in each period. Both households have this utility function over consumption $C_t^{h,U}$ and leisure $(1 - N_t^i)$[3]:

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\left(\frac{1}{1+\psi_t} \right) \log\left(C_t^{h,U}\right) - \chi \frac{(N_t^i)^{\kappa+1}}{\kappa+1} \right] \quad (1)$$

where β is the subjective discount factor and $\kappa > 0$. $\left(\frac{1}{1+\psi_t}\right) < 1$ captures a self-imposed demand constraint in t as in Eichenbaum et al. (2020). Moreover, $C_t^{h,U} \leq \bar{C}^{h,U}$ represents a regulatory constraint on consumption (binding in t, see below). We denote the multiplier attached to the regulatory constraint by ϕ_t.

The aggregate value of any household-specific variable X_t, in per-capita terms, is given by $X_t = (1 - s^{lc})X_t^r + s^{lc}X_t^l$, where s^{lc} denotes the relative size of the liquidity constrained household.

2.1.1. The Ricardian household

The Ricardian household maximises utility subject to a sequence of budget constraints

$$\Delta L_t + \Delta B_t + \Delta V_t = r_{t-1}B_{t-1} + r_{t-1}^L L_{t-1} + D_t - C_t^{R,U} + W_t^r N_t^R + TR_t^R + ben_t^R - T_t, \quad (2)$$

where W_t^r, N_t, T_t, and TR_t^R denote the real wage rate (same for both households), labour supply, direct taxes paid by households, household-specific transfers, respectively. Assets of the household are made up of loans to firms L_t (the return r_t^L includes a loan default risk premium), bonds B_t (an internationally traded bonds[4] and government bonds) with net return r_{t-1}, and firm shares V_t, yielding dividends D_t. ben_t^R summarises STW allowances, stw_t, and unemployment benefits, $uben_t$:[5]

$$ben_t^R = (uben_t)(1 - npart - N_t^R) + stw_t(N_t^R - N_0^R), \quad (3)$$

where $(1 - npart)$ and N_0 denote the labour force participation rate and baseline labour demand, respectively.

In equilibrium, intertemporal consumption-saving choice satisfies:

$$\frac{C_{t+1}^R}{C_t^R} = \beta\left(1 + r_t^f\right)(1 + \psi_t^R)(1 + \phi_t^R), \quad (4)$$

where r_t^f denotes the risk-free rate. Both constraints imply a negative shock to consumption in t. The consumption constraint $C_t^{R,U} \leq \bar{C}^{R,U}$ becomes binding in period t and $\phi_t^R > 0$ if $C_t^R = \overline{C^R}$. Note that both consumption constraints, self-imposed ($\psi^R{}_t$) and regulatory (ϕ_t^R), have the same effect on the

[3] Households supply differentiated types of labour services i, which we assume to be distributed equally over both household types. Unions bundle the differentiated labour services provided by the two types of households and maximise a joint utility function for each type of labour i. See below.

[4] The international bond features a country risk premium which depends on the net foreign asset position to ensure long-run stability of the model (Schmitt-Grohé and Uribe, 2003).

[5] The unemployed $(1 - npart - N_t)$ receive benefits $uben_t = benrr\ W_t^r$, where $benrr$ is the exogenous benefit replacement rate.

Covid Economics 30, 19 June 2020: 120-145

Euler equation.[6] In the following, we assume $\phi_t > 0$ and $\psi_t = 0$, corresponding to a shock to current consumption in the Euler equation.

2.1.2. The liquidity-constrained household

Voluntary social distancing and regulatory constraints on consumption also apply to the constrained household; namely a self-imposed consumption constraint

$$W_t^r N_t^L - T_t + TR_t^C + ben_t^C > C_t^C \tag{5}$$

and a regulatory constraint

$$W_t^r N_t^L + TR_t^C + ben_t^C > \bar{C}^C. \tag{6}$$

As a result, the constrained household features forced savings:

$$\Delta B_t^C = W_t^r N_t^C + TR_t^C + ben_t^C - T_t - C_t^C \tag{7}$$

and dissaving in periods following the lockdown $C_{t+\tau}^C = W_{t+\tau}^r N_{t+\tau}^C + TR_{t+\tau}^C + ben_{t+\tau}^C - T_{t+\tau} + sB_{t+\tau-1}^C$, where τ denotes the post-lockdown period.

2.2. INTERMEDIATE GOODS FIRMS

There is a continuum of intermediate goods indexed by $j \in [0,1]$. A single firm produces each good. Firms face symmetric decision problems and make identical choices. Firm j has technology $Y_t^j = A\left(N_t^j\right)^{\alpha}(u_t^j K_t^j)^{1-\alpha}$ where Y_t^j, N_t^J, u_t^j and K_t^j are the firm's output, labour input, capacity utilisation and capital stock, respectively. A is a constant common productivity level. The law of motion of firm j's capital stock is $K_t^j = (1-\delta)K_{t-1}^j + I_t^j$, with depreciation rate δ and gross investment I_t^j. The period t dividend of intermediate good firm j is:

$$D_t^j = p_t^j Y_t^j - W_t N_t^j - p_t^K I_t^j - r_{t-1} L_{t-1}^j + L_t^j + \Gamma_t^j, \tag{8}$$

where p_t^j and p_t^K denote the price charged by the firm and the price of production capital, respectively. L_t^j are one-period loans. Γ_t^j summarises quadratic price and factor adjustment costs.[7] Firm j maximises the present value of dividends $V_t^j = D_t^j + \Lambda_{t,t+1} V_{t+1}^j$, where $\Lambda_{t,t+1}$ denotes the discount factor of Ricardian households. For later purpose, it useful to define the (real) gross operating surplus: $GOS_t^j = Y_t^j - W_t^r N_t^j$. We model labour and capital/investment adjustment costs as[8]:

[6] However, consumption responds different to policies for the two shocks. With a self-imposed reduction of consumption, household consumption will respond to fiscal measures (e.g. a reduction of VAT), while the government imposed constraint on consumption is a quantity constraint which cannot be affected by fiscal measures.

[7] Quadratic price adjustment costs imply that the inflation rate of local intermediates obeys an expectational Phillips curve. See Annex B.

[8] Given substitutability between capital and labour allowed by the production technology, firms could increase the utilisation of capital to partly offset the labour input constraint. We find this an unrealistic option in the short run given the scale of the supply constraint. Therefore, we impose a partial short-run complementarity between labour and utilised capital

$$\frac{\Delta N_t}{N_t} \approx \frac{\Delta UC_t}{UC_t}.$$

Covid Economics 30, 19 June 2020: 120-145

$$\Gamma_t^{N,j}(N_t^j) \equiv \frac{\gamma^N}{2}\left(\frac{N_t^j}{N_{t-1}^j} - 1\right)^2 \tag{9}$$

$$\Gamma_t^{K,j} \equiv \frac{\gamma^K}{2}\left(\frac{I_t^j}{K_{t-1}^j} - \delta\right)^2 p_t^I \bar{K}_{t-1}^J + \frac{\gamma^I}{2} p_t^I \left(\Delta I_t^j\right)^2 \tag{10}$$

Next, we discuss the two transmission channels of the pandemic on the production side: Labour input restrictions and liquidity constraints. Lockdown measures imply a downward shift in labour demand. In addition, the reduction in output induced by the pandemic and its associated demand and supply lockdown measures leads to falling investment via liquidity problems for firms.

2.2.1. Labour input constraints: Lockdown shocks

To prevent infections at the workplace, governments impose restrictions on labour input. Firms can only use $\bar{N}_t$ employees during the lockdown. Precautionary distancing measures at the workplace imply then $N_t \leq \bar{N}_t$. A binding constraint shifts down the labour demand schedule:

$$\frac{\partial Y_t^j}{\partial N_t^j} = W_t^r + \theta_t^N, \tag{11}$$

where $\theta_t^N \geq 0$ is the Lagrange multiplier of the labour input constraint.

2.2.2. Investment liquidity constraints

Below we consider an extended model in which a subset $0 \leq s_t^{li} < 1$ of intermediate goods firms faces temporary binding liquidity constraints of the form:[9]

$$L_t \leq \mu K_{t-1}, \tag{12}$$

where μ is the loan-to-value ratio. For these firms, adverse demand and supply shocks increase liquidity needs and trigger a credit tightening. A binding collateral constraint binding imposes a reduction in investment. Therefore, the investment rate for constrained firms follows:

$$\left(\frac{I_t^j}{K_{t-1}^j} - \delta\right) = \mathcal{H}\left(\frac{GOS_t^j}{K_{t-1}^j}\right) \equiv \zeta_1\left(\frac{GOS_t^j}{K_{t-1}^j} - \delta\right) - \zeta_2, \tag{13}$$

where parameters ζ_1 and ζ_2 govern the strength of the liquidity constraint. The share of constrained firms is endogenous and follows:

$$s_t^{li} = a_0 - a_1 GOS_t. \tag{14}$$

As shown in Annex A, this reduced-form equation with parameters a_0 and a_1 is consistent with a micro-founded liquidity constraint.

[9] By contrast, in the baseline model $s_t^{li} = 0\ \forall t$.

Covid Economics 30, 19 June 2020: 120-145

The remaining unconstrained firms decide investment according to a standard Q-equation $Q_t^j = 1 + \Gamma_t^{j,\prime}\left(\frac{I_t^j}{K_t^j}\right)$, where Q_t^j represents the discounted value of physical capital. The net investment rate is a function of Q_t^j, i.e. $\left(\frac{I_t^j}{K_{t-1}^j} - \delta\right) = \mathcal{F}(Q_t^j)$ for $j \in [s_t^{li}, 1]$. Thus, the aggregate net investment follows:

$$\left(\frac{I_t}{K_{t-1}} - \delta\right) = \int_0^{s_t^{li}} \mathcal{H}\left(\frac{GOS_t^j}{K_{t-1}^j}\right) dj + \int_{s_t^{li}}^1 \mathcal{F}(Q_t^j)\, dj\,. \tag{15}$$

2.3. FINAL GOOD FIRMS

Final good producers have access to a CES production technology $Y_t = \left[(s^d)^{1/\nu} O_t^{\frac{\nu}{\nu-1}} + (1 - s^d)^{1/\nu} M_t^{\frac{\nu}{\nu-1}}\right]^{\frac{\nu-1}{\nu}}$, with home bias $0.5 < s^d < 1$. $O_t = \left[\int_0^1 (Y_t^j)^{\frac{\varepsilon-1}{\varepsilon}} dj\right]^{\frac{\varepsilon}{\varepsilon-1}}$ is an aggregate of the local intermediates, where ε is the exogenous substitution elasticity between varieties. M_t denotes intermediate imports from the RoW. The final good is used for domestic private and government consumption, and investment.

2.4. WAGE SETTING

A trade union 'differentiates' homogenous labour hours provided by the two domestic households into imperfectly substitutable labour services. Both households work the same hours and receive the same wage. The labour input N_t in the production process of intermediate goods is a CES aggregate of these differentiated labour services. The union sets wage rates at a mark-up μ_t^W over the marginal rate of substitution between leisure and consumption. μ_t^W is inversely related to the degree of substitution between labour varieties. The mark-up is countercyclical because of nominal wage adjustment costs. Following Blanchard and Gali (2007), we allow for real wage inertia; the current period real wage rate is a weighted average of the desired net real wage and the past (net) real wage:

$$(1 - \tau_t^N) W_t^r = [(1 + \mu_t^W) mrs_t]^{1-\xi} [W_t^r (1 - \tau_{t-1}^N)]^{\xi}, \tag{16}$$

where mrs_t is a weighted average of the two households' marginal rates of substitution between consumption and leisure. The parameter ξ is an index of real wage rigidity.

2.5. PUBLIC POLICY

2.5.1. Monetary policy

EU monetary policy is subject to a ZLB constraint.[10] The notional interest rate follows a smooth Taylor rule with respect to inflation and the output gap:

[10] Monetary policy in the RoW does not hit the ZLB.

Covid Economics 30, 19 June 2020: 120-145

$$i_t = \max\left\{0, \rho_i i_{t-1} + (1-\rho_i)\left(\bar{r} + \pi^{tar} + \tau_\pi\left(\frac{\pi^C_{t,yoy}}{4} - \pi^{tar}\right) + \tau_y y_t^{gap}\right)\right\}. \quad (17)$$

The central bank has an inflation target π^{tar}, adjusts its policy rate relative to the steady-state value $\bar{r}$ when actual CPI inflation deviates from the target, where $\pi^C_{t,yoy} \equiv P^C_t / P^C_{t-4} - 1$ is year-on-year CPI inflation, or in case of a non-zero output gap (y_t^{gap}).[11]

2.5.2. Fiscal policy

We assume that the government keeps its expenditure (G_t) constant in real terms. Real government debt evolves as:

$$B^G_t = (1 + r^G_t)B^G_{t-1} + G_t + (uben_t)(1 - npart - N_t) + stw_t(N_t - N_0) + TR^L_t + TR^R_t - R^G_t, \quad (18)$$

where r^G_t denotes the government interest rate. R^G_t, government revenues, are the sum of consumption, labour, and profit taxes. Time-varying labour taxation stabilises the debt-to-GDP ratio:

$$\tau^N_t = \tau^N_{t-1} + d^G_t\left(\tau^B\left(\frac{B^G_t}{4Y_t} - \overline{btar}\right) + \tau^{def}\Delta B^G_t\right), \quad (19)$$

with $\overline{btar}$ being the target level of government debt-to-GDP. Parameters τ^{def} and τ^B control the feedback rule. d^G_t is a dummy that allows to turn off the debt rule temporarily.

2.6. MODEL CALIBRATION AND SOLUTION

In our model calibration, one period corresponds to one quarter. Tables 1 summarises the main parameter values. Table 2 features block-specific parameter values. Macroeconomic aggregates that characterise the steady state, like private and public consumption and investment, trade openness, and trade linkages match block-specific data from national accounts and the GTAP database (Narayanan and Walmsley, 2008).

Behavioural parameters that govern the dynamic adjustment to shocks are based on earlier estimates of QUEST model versions. In particular, the model estimations have identified high labour adjustment costs for the EU ($\gamma^N = 25$).[12] Annex B shows additional details on the convex adjustment costs related to price setting and capacity utilisation. Concerning financial market frictions, we set the share of the Ricardian household to 60% - close to the estimates in Ratto et al. (2009), Dolls et al. (2012) and Kaplan et al. (2014). We microfound the firm liquidity needs based on collateral constraints in Annex A. In our simulation, the endogenous share of constrained firms reaches around 30 per cent in 2020Q2 (including policy). This value is in line with recent estimates (OECD, 2020) and based on a collateral constraint parameter μ of 0.3.[13] The labour supply elasticity is set at 0.2 slightly below the estimate in Kollmann et al. (2016). Concerning adjustment costs on labour, goods, and capital, we broadly follow earlier QUEST-based estimates. The tax rule parameters assure a smooth transition to the long-run debt-to-GDP ratio. The latter reflects average pre-pandemic EU data. Taylor rule parameters are standard values in the literature.

[11] The output gap concepts comes from a production function framework. See Ratto et al. (2009).

[12] See, for example, in 't Veld et al. (2015) and Kollmann et al. (2016).

[13] EU corporate debt is around 30% of the private capital stock.

Covid Economics 30, 19 June 2020: 120-145

We solve the model nonlinearly under perfect foresight using a Newton-Raphson algorithm.

Table 1. **Selected parameter values**

Parameter	Value	Description
β	0.997	Discount factor
$1/\kappa$	0.2	Labour supply elasticity
γ^N	25	Head-count adjustment costs parameter
γ^P	20	Price adjustment costs parameter
$\gamma^{ucap,1}$	0.04	Linear capacity-utilisation adjustment cost
$\gamma^{ucap,2}$	0.1	Quadratic capacity-utilisation adjustment cost
γ^K	20	Capital adjustment cost
γ^I	75	Investment adjustment cost
ξ	0.8	Real wage inertia
ν	1.2	Elasticity of substitution in total trade
$1-1/\varepsilon$	0.12	Price mark-up
α	0.65	Cobb-Douglas labour share parameter
$\bar{G}$	0.17	Government expenditure (share in GDP)
μ^W	0.2	Steady-state wage mark-up
δ	0.015	Depreciation rate (quarterly)
ζ_1	1.1	Intensity of liquidity-constraints (firms) parameter 1
ζ_2	0.1	Intensity of liquidity-constraints (firms) parameter 2
a_1	42	Share of liquidity-constrained firms parameter
μ	0.3	LTV in affected sectors
$\bar{s}^{li}$	0	Steady-state share of liquidity-constrained firms
τ^b	0.05	Tax rule parameter on debt
τ^{def}	0.1	Tax rule parameter on deficit
ρ_i	0.8	Taylor rule persistence
τ_π	2	Reaction to inflation in Taylor rule
τ^y	0.1	Reaction to output gap in Taylor rule

Source: Commission services.

Table 2. **Region-specific parameter values**

Parameter	EA	RoW	Description
s^{lc}	0.4	0.5	Share of liquidity-constrained households
$1-npart$	0.71	0.76	Labour force to population
N	0.64	0.66	Steady-state employment to population
$benrr$	0.40	0.30	Benefit replacement rate
$\overline{TR}$	0.16	0.12	Transfer share (share in GDP)
s^d	0.22	0.06	Steady-state share of imports
$btar$	0.8	0.4	Baseline government debt-to-GDP ratio

Source: Commission services.

Covid Economics 30, 19 June 2020: 120-145

3. LOCKDOWN AND PANDEMIC IN A MACRO MODEL

We now discuss our assumption on exogenous shocks. For clarity, we assume that containment measures are active in March until June with a peak early in Q2. The pandemic shocks thus last two quarters, with a stronger effect in the second quarter. The shocks end in the third quarter. As discussed above, the pandemic shocks are the two shocks associated with supply and demand disruptions, i.e. restrictions on labour input and consumption (see Jonung and Roeger, 2006). In reality, the separation of demand and supply is difficult. For example, supply constraints are only one factor behind the closure of shops and factories. Moreover, we do not necessarily see a trade-off between economic costs and lockdown. Timely containment measures may also prevent disruptions by avoiding larger-scale outbreak at a later stage. Moreover, in multi-sector models with incomplete markets, supply shocks can have "Keynesian" features (Guerrieri et al., 2020).

All shocks are global, i.e. they are symmetric in both regions of the model. We calibrate the shock size as roughly consistent with the Spring Forecast of the European Commission (2020). The forecast employed detailed sectoral assumptions, e.g. particularly strong declines in air transport, accommodations, restaurants, tourism etc.

Two remarks are in order. First, we do not believe that the pandemic will end in 2020Q3. Yet, this assumption allows us to distinguish the direct pandemic shocks from their endogenous transmission in a transparent way. Second, we do not consider the exogenous shocks included in the simulations to be the only economic disruptions caused by the pandemic. Other channels, such as heightened uncertainty (e.g. Baker et al., 2020), financial risks such as cascading bankrupticies, or permanent changes in consumption patterns, complement the analysis presented in this paper.[14]

Regarding public policy, the simulations assume an inactive debt rule for 40 periods, i.e. $d_t^G = 0$ for $t = 2020Q1{:}\,2030Q1$ and 1 otherwise. This setting allows a clearer assessment of the budgetary effects of the pandemic and the economic policy response. In our simulations, the ZLB binds for two years. The simulations without discretionary policy intervention assume that governments only rely on automatic stabilisers (in particular unemployment benefits). All scenarios assume that government consumption and other transfers are constant in real terms (unless explicitly specified).

4. MACROECONOMIC TRANSMISSON OF THE PANDEMIC

This section looks at the transmission channels of the pandemic shock and quantifies its impact. The next section then adds the economic policy response to the analysis.

4.1. THE COVID-19 SHOCK ABSENT LIQUIDITY CONSTRAINTS

The COVID-19 crisis has a very large detrimental economic impact on the EU and the world economy as shown in Figure 1. In the basic model version without firm liquidity constraints, the economic impact closely follows the pandemic and required containment measures (shown with red dashed

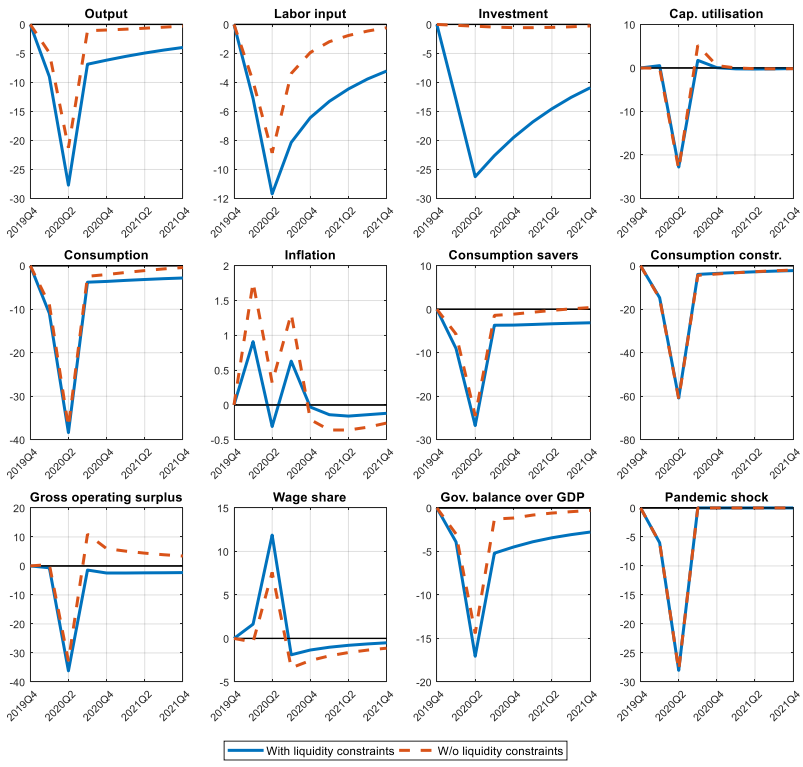
[14] As mentioned above, our related work in the European Commission's (2020) spring forecast addresses these two points. It considers additional channels, relaxes the assumption on duration, and considers other pandemic patterns such as second waves.

Covid Economics 30, 19 June 2020: 120-145

lines).[15] The result for this model is essentially a V-shaped recession in Q1 and Q2. There is a massive decline in consumption, where reduced labour income adds to the adverse impact of the demand shock. Some persistence in adjustment frictions in employment prevent an immediate adjustment. Higher capacity utilisation partly offsets a delayed response of the labour inputs, which, however, remains limited as we assume a partial short-run complementarity between capital and labour. The crisis also has a distributional dimension. Consumption of the constrained households depends more on labour income and falls more strongly than the consumption of the Ricardian household.

The baseline model version only generates a small decrease in investment. Investors foresee the temporary nature of the shock (as we abstract from uncertainty effects). Supply and demand shocks alone cannot generate a fall of investment. Yet, we find a strong decline in the gross operating surplus since labour costs remain high. This result indicates the relevance of liquidity constraints, as we discuss next.

Figure 1: **Simulations absent policy response**



Note: This figure expresses the wage share, quarterly inflation, and the government balance in percentage point deviation from steady state. All other variables are expressed in percent deviation from steady state. The pandemic shock is an illustrative index of the exogenous shock process.

Source: Commission services.

[15] We always maintain the assumptions that a share of households is liquidity constrained.

Covid Economics 30, 19 June 2020: 120-145

4.2. THE IMPACT OF LIQUIDITY CONSTRAINTS

We now show that the firm liquidity constraints lead to a deeper and more U-shaped recession. In contrast to the baseline setup, the firm liquidity channel in this model version amplifies supply and demand shocks and generates a sizable decline in investment as shown in Figure 1 (blue solid lines). The endogenous fall in the GOS increases the share of liquidity-constrained firms and generates a sizable contraction in private investment. Investment adjustment costs, a plausible empirical feature, generate additional persistence. The magnitude of the investment decline is roughly in line with the Commission's investment forecast (European Commission, 2020). In the baseline model supply constraints dominate and generate (quarterly) inflationary pressure. By contrast, the impact from liquidity constraints amplifies the deflationary demand effects leading to a more balanced picture. Once the lockdown can be lifted, higher capacity utilisation and recovering consumption lead to an increase prices.

The amplification from the occasionally binding liquidity constraints also renders the reduction in labour input more persistent even though the constraint binds only in the first and second quarter. This effect strongly reduces output growth in 2020 and 2021 as shown in Table 4. GDP growth falls by 13 per cent below the no-shock path, compared to -8 per cent in the absence of the liquidity constraints. In sum, the amplification leads to a more U-shaped pattern of output and motivates the focus on liquidity constraints when analysing the EU policy response in the next section.

5. ECONOMIC POLICY RESPONSE

This section analyses the economic policy response in the EU with a focus on (i) short-time work allowances and (ii) loan guarantees. Both measures target the distortion arising from firms' liquidity constraints. As shown above, given the sharp fall in the gross operating surplus, liquidity constraints substantially prolong the recession, if not addressed appropriately by economic policy. These measures, however, cannot prevent a temporary decline in production because of sickness, restrictions on the mobility of workers or supply chain interruptions.

5.1. OVERVIEW STABILISATION MEASURES

The fiscal measures announced in the Member States consist of stimulus measures with a direct impact on the budget, as well as liquidity measures without direct budgetary impact. Table 3 provides information about fiscal measures and their composition for the EU-27. Total liquidity support by EU Member States amounts to approximately 22% of GDP, mostly in form guarantees and tax delays. In addition, the stimulus measures amount to around 2.8% of GDP, mainly as STW allowances and transfers. We leave the analysis of the sizable supranational EU support for future work.

Table 3. **Overview of announced measures by EU-27 Member States**

	bln EUR	*% of GDP*
1. Measures with a direct budgetary impact	368	2.8
2. Liquidity measures without budgetary impact		
a. Tax delays	248	1.9
b. Public guarantees	2301	17.6
c. Others	334	2.6
Total liquidity support	**2882**	**22.1**

Source: Commission services. Cut-off date 29/04/2020.

Covid Economics 30, 19 June 2020: 120-145

5.2. SHORT-TERM WORK ALLOWANCES

Our modelling focuses on two differences of STW allowances compared to unemployment benefits: First, STW reduces employment adjustment costs. Unemployment benefits are paid to workers who have lost their job. By contrast, STW allowances are paid for workers staying in the firm. Since workers stay in the firm, firms avoid matching frictions and hiring or firing costs. We capture this improved allocative efficiency under STW by setting $\gamma^N = 0$ – a parsimonious way to capture the absence of matching frictions. Second, STW allowances are more generous than unemployment benefits. We capture this effect by including additional transfers to liquidity-constrained households, $TR^C_{2020Q3:2020Q4}$, which can be spent only after the pandemic.[16] For analytical purposes, we separate the two channels via two simulations.

5.2.1. Stabilisation gains under identical ex-ante costs

We first compare stabilisation gains of STW under identical ex-ante fiscal cost, i.e. the STW rate equals the benefit replacement rate. Assuming the same generosity of STW allowance and unemployment benefits, allows us to highlight the difference in allocative efficiency. Temporary unemployment entails layoff and search costs, which STW schemes avoid. To keep capture the absence of matching friction and employment adjustment costs, we eliminate labour adjustment costs from the model. Figure 2 then compares the dynamics under STW (red dotted lines) to those of the baseline model discussed in Section 4.2, where only unemployment benefits active (solid blue line).

The absence of employment adjustment frictions under STW allows firms to adjust their labour input more strongly during the peak of the shock (2020Q2). However, by the same logic, STW schemes support the exit from the pandemic shock by avoiding a costly and time-consuming hiring process and allowing firms to increase labour input more rapidly.

The interaction of STW with liquidity constraints. STW addresses two interacting frictions in the model: Employment and liquidity frictions. The cost-saving effect of STW improves the corporate liquidity position. By allowing for a stronger reduction of labour input, STW schemes help lower the wage bill without incurring firing costs. Importantly, the stabilisation of the gross operating surplus reduces the impact of firm liquidity constraints, as fewer firms become constrained. The smaller distortion from liquidity constraints implies that investment declines less, and labour input recovers faster and more strongly – especially after the pandemic. This, in turn, makes the recession less persistent. This result highlights the gains from allocative efficiency even under identical ex-ante fiscal costs.

STW implies a larger government deficit in 2020. The expenditure on STW allowances is higher than under unemployment benefits scheme since firms reduce labour input more strongly (even under same generosity). However, once the government can lift the lockdown restrictions, STW policy achieves higher revenues from labour, consumption, and profit taxation. In annual terms, the (total) government balance to GDP ratio falls by around 9 percentage points in 2020 and remains at -3 pps in 2021.

The fact that STW is a desirable policy depends on the significant labour adjustment costs in European countries. The relative gains compared to unemployment benefits could be smaller for economies (or sectors) with a higher degree of labour market churning.[17]

5.2.2. The generosity of STW allowances

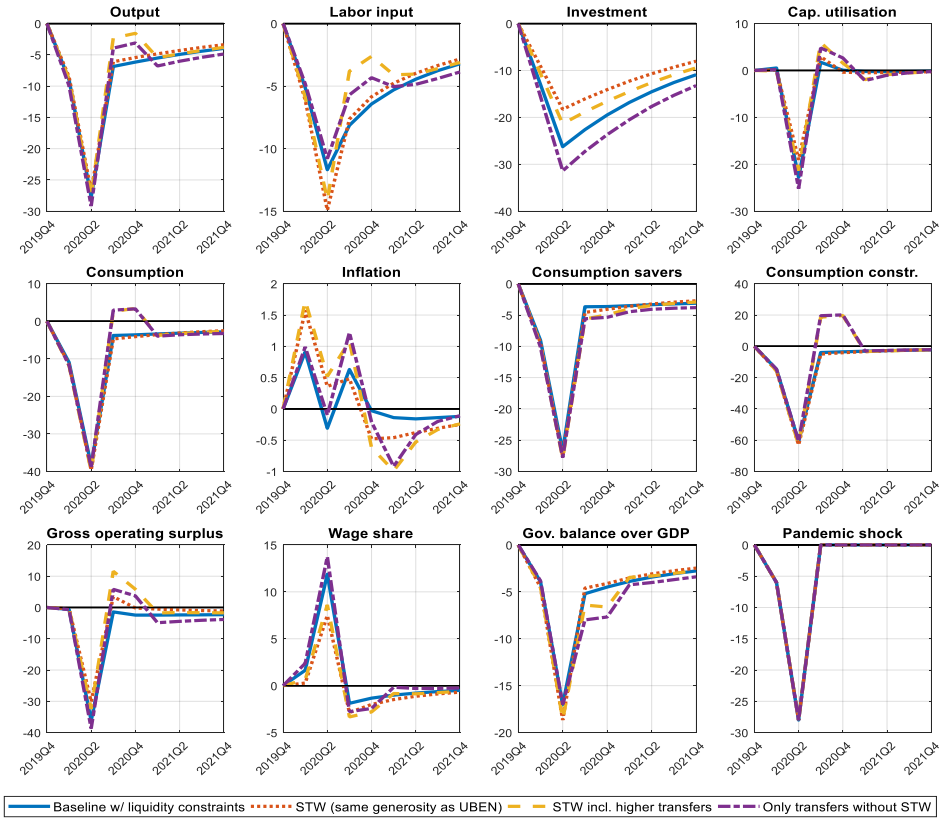
We now turn to the more realistic case, where the generosity under STW is higher than under unemployment benefits. Unlike in the previous simulation, STW allowances entail higher ex-ante

[16] The Annex shows results for transfers to both households (non-targeted).
[17] See, e.g., Davis et al. (2012) for a discussion on labour market flows in the US.

Covid Economics 30, 19 June 2020: 120-145

budgetary costs than unemployment benefits. A second simulation captures this aspect by adding transfers to constrained households of 2.8% of EU-27 GDP. This value corresponds to the policy measures credibly announced as of now (see Table 3). Since the containment measures partially curtail spending opportunities, households can spend transfers only in Q3 and Q4. Moreover, we assume that transfers target liquidity-constrained households. The Annex relaxes this assumption.

Figure 2: **Short-term work allowances**



Note: This figure expresses the wage share, quarterly inflation, and the government balance in percentage point deviation from steady state. All other variables are expressed in percent deviation from steady state. The pandemic shock is an illustrative index of the exogenous shock process.

The higher generosity implies sizable output gains. Constrained households have a high marginal propensity to consume, implying a higher multiplier than non-targeted transfers. Since the containment measures partially curtail spending opportunities, we assume that transfers are paid and spent in Q3 and Q4. A sizable rebound of consumption materialises in these periods. The simulation shows that the support for household consumption facilitates exit. In 2020, the level of consumption is 2.7% higher than without discretionary policy intervention. Higher transfers provide a strong boost to employment, which under STW can expand more rapidly. STW thereby also improves risk-sharing among households, with a more balanced distribution of consumption across households. In sum, STW and associated higher transfer allowances cushion the fall in real GDP by around 2.2 pps. in 2020.

Covid Economics 30, 19 June 2020: 120-145

The higher generosity of STW also implies ex-ante a stronger deterioration of the government balance.[18] However, the faster recovery implies increased tax revenues (from relative increases in consumption, labour, and profits) as well as lower unemployment benefits. In sum, STW allowances increase the deterioration in the deficit-to-GDP ratio by less than 1 pps. (on average in 2020 and 2021) compared to the baseline simulations with only unemployment benefits active.

Interestingly, the two aspects of STW analysed here, namely avoiding job match destruction and higher generosity, interact. The improved allocative efficiency under STW leads to stronger effects of transfers. Figure 2 shows the results by adding a simulation with "only transfers" where employment adjustment costs remain at the baseline values. In this case, additional transfers provide smaller stabilisation gains, because employment adjustment costs slow down the response of hours worked and wages, translating into a smaller increase in household income.

Finally, note that the effectiveness of transfers depends on the marginal propensity to consume of the receiving households. The Annex shows that targeted transfers are significantly more efficient by supporting households with a higher marginal propensity to consume in line with the findings provided by Bayer et al. (2020).

Covid Economics 30, 19 June 2020: 120-145

5.3. LIQUIDITY SUPPORT

Liquidity support in the form of lending guarantees amounts to a maximum of 22% of GDP. One important goal of these programs is to stabilise investment of liquidity-constrained firms.

It is challenging to operationalise the liquidity guarantees in a macro model. Since there is some heterogeneity of initial conditions and on how severely individual firms are affected by the shock, a fraction of firms will defer investment even with guarantees. There are also specific eligibility criteria, which exclude certain types of firms from the schemes or restrict the schemes to certain sectors. Also generally, an upper bound on the guarantee per firm is imposed. Modelling the take up rate would require more information about the distribution of the shock across firms and the constraints imposed by governments.

In the absence of detailed up-to-date information, we will assume that 50% of the liquidity-constrained firms are not revising their investment plans or are excluded from funding. This allows us to say something about the 'guarantee multiplier'. Under the assumption that firms keep dividend payouts stable and that investment does not affect gross operating surplus of the firm in the current period, the investment multiplier to a loan increase for liquidity-constrained firms is one.

To see this, consider a liquidity-constrained firm (superscript C). The budget constraint restricts restricts investment of this firm to the loan supply of the bank and current gross operating surplus minus debt service and dividend payments

$$I_t^C = \Delta L_t^C + GOS_t^C - rL_{t-1} - D_t^C \tag{20}$$

Keeping current period dividends and GOS constant, the increase in investment due to a (guarantee-secured) extension of the loan is given by

$$I_t^G - I_t^C = \Delta L_t^G - \Delta L_t^C, \tag{21}$$

where $\Delta L_t^G - \Delta L_t^C$ is the loan expansion fully guaranteed by the government and supescript G denotes variables following the provision of guarantees. This is likely to be an upper bound since some of the additional funds may be diverted to increase dividends. However, apart from firm specific preferences, diversion of funds is limited because guarantee schemes by EU governments generally impose temporary restrictions on dividend payments for firms, which receive funding under public loan

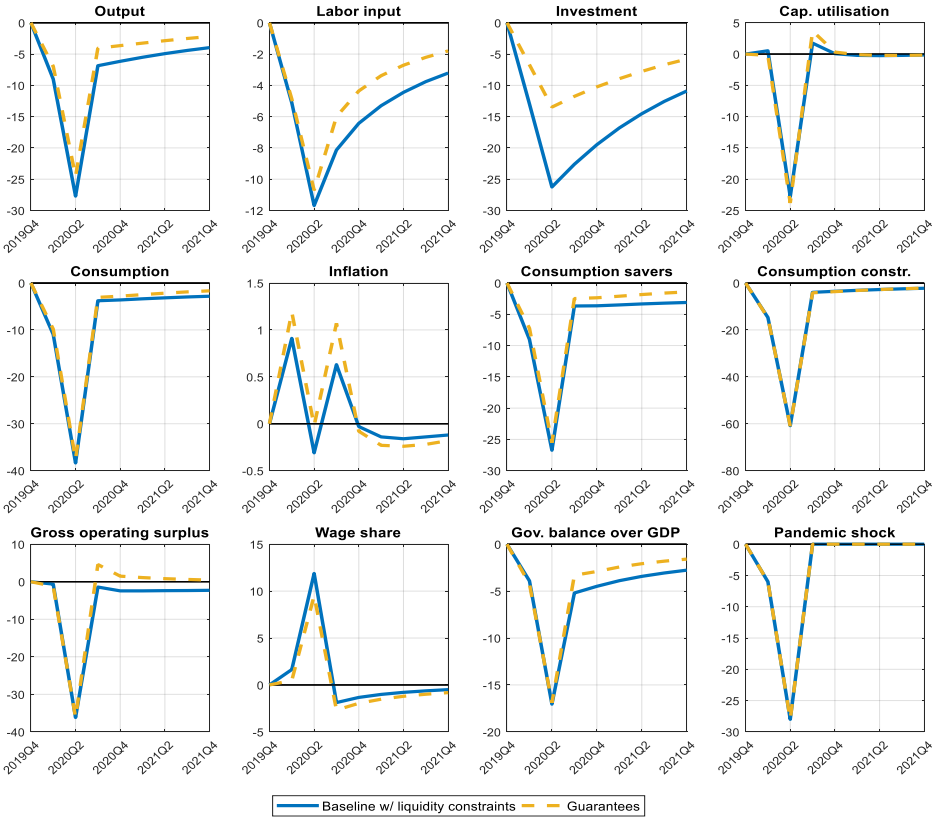
[18] We assume that the government covers the additional generosity. There are no additional costs for the firms.

guarantee schemes. Based on information about non-performing loans[19], we can assume that about ten per cent of these loans will default. This value gives a guarantee-multiplier of

$$m^G = \frac{Y_t^G - Y_t^C}{(\Delta L_t^G - \Delta L_t^C)} = m^I * loss \tag{22}$$

Thus, essential for the fiscal multiplier is the investment multiplier (m^I) and the loss rate of the guaranteed loan.[20] This multiplier is an upper bound since it ignores possible windfall gains to the banking sector. The banking sector might use the loan guarantees also for loans to unconstrained firms, thereby covering losses, which would otherwise be borne by the banking sector.

Figure 3 : **Lending guarantees**



Note: This figure expresses the wage share, quarterly inflation, and the government balance in percentage point deviation from steady state. All other variables are expressed in percent deviation from steady state. The pandemic shock is an illustrative index of the exogenous shock process

[19] Consolidated banking data from the ECB shows a ratio of lower non-performing loans to total loans. However, during peak crisis times around 2013, the share reach around eight percent. See also: https://www.ecb.europa.eu/press/pr/date/2019/html/ecb.pr191106_1~a993d312e7.en.html

[20] The loss rate is defined by how much of the value of the investment project associated with defaulting loans must be written off.

Covid Economics 30, 19 June 2020: 120-145

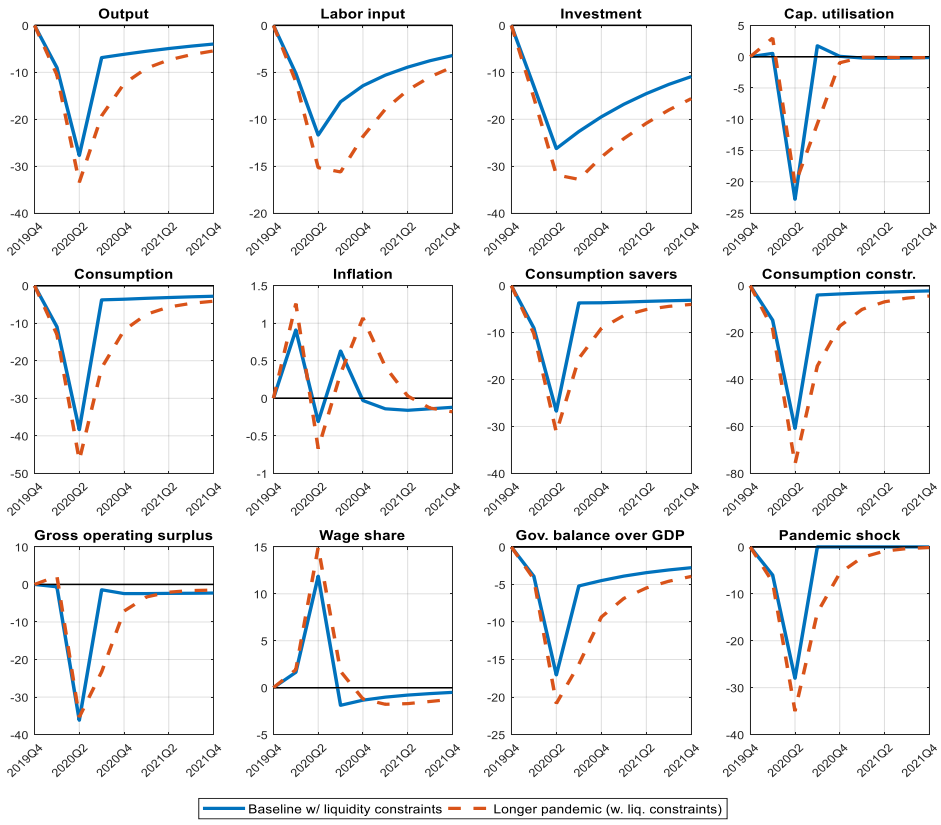
Covid Economics 30, 19 June 2020: 120-145

6. ROBUSTNESS

For analytical clarity, we have considered a very short pandemic. Yet, it is increasingly clear that the pandemic and the associated lockdown will continue for a longer time. Therefore, Figure 4 considers also a longer pandemic, extending partially also into 2021 (see bottom right figure). Dashed red lines show the simulations of the longer pandemic, while blue solid lines correspond to the main simultions (with firm liquidity constraints) as shown in Figure 1.

Figure 4 shows that the duration of the pandemic is crucial for the length of the economic downturn: the adverse impact on investment and output increases and becomes more persistent into the second year (see final columns in Table 4 below). Note that the longer pandemic scenario includes automatic stabilisers, but no STW and guarantees. The Commission's forecast (European Commission, 2020) has also considered other pandemic scenarios such as a second wave.

Figure 4: **Longer pandemic**



Note: This figure expresses the wage share, quarterly inflation, and the government balance in percentage point deviation from steady state. All other variables are expressed in percent deviation from steady state. The pandemic shock is an illustrative index of the exogenous shock process.

7. DISCUSSION

COVID-19 is a rare disaster with enormous economic costs. This paper has analysed the transmission channels of the pandemic using a TANK model. In the baseline model, demand and supply shocks lead to a V-shaped recession. However, an augmented model with firm liquidity constraints provides a different picture. In this framework, the strong decline in the gross operating surplus following the pandemic and its associated lockdown induces a strong reduction in investment due to a tightening of liquidity constraints. This effect amplifies the economic fallout and generates persistence. It also provides an important entry point for economic policy.

Table 4. **Overview of results**

Variable/Scenario	No liq. constr.		Liq. constraints		STW (without add. transfers)		STW		Guarantees		STW + Guarantees		Longer pandemic (absent policy)	
	2020	*2021*	*2020*	*2021*	*2020*	*2021*	*2020*	*2021*	*2020*	*2021*	*2020*	*2021*	*2020*	*2021*
GDP	-7.9	-0.9	-13.0	-4.8	-12.1	-3.7	-10.8	-4.2	-10.7	-3.1	-8.7	-2.5	-19.8	-7.9
Consumption	-12.9	-1.2	-14.1	-3.0	-15.0	-2.9	-11.4	-3.2	-13.4	-2.1	-10.8	-2.3	-23.3	-5.6
Investment	-0.5	-0.6	-20.2	-13.6	-14.3	-10.0	-16.9	-11.8	-10.7	-7.4	-8.6	-6.3	-27.0	-19.6
Labor input	-4.8	-0.9	-7.8	-4.1	-8.6	-3.7	-6.6	-3.7	-6.6	-2.6	-5.3	-2.3	-12.2	-6.4
GOS	-4.4	3.8	-10.1	-2.2	-6.8	-0.7	-3.9	-1.6	-8.1	0.6	-1.9	0.7	-15.9	-2.1
Wage share	0.4	0.4	2.6	2.6	0.7	-1.1	0.8	-0.8	1.4	-1.1	0.0	-1.0	4.4	-1.5
Gov. balance/GDP	-5.3	-0.8	-7.7	-3.2	-8.0	-2.9	-9.2	-3.2	-7.0	-2.1	-8.4	-2.1	-12.5	-5.2

Source: Commission services.

Our analysis then quantifies two central policy responses in the EU-27: Short-term work (STW) allowances and liquidity support. Both policies support the recovery after the pandemic shocks. Given that the demand effects of lockdowns cannot be stabilised at the time of implementation, policies should target a rapid exit, once the pandemic ceases and governments can lift the associated lockdown. Our paper shows that STW is a desirable policy, given the significant labour adjustment costs in European countries. Apart from reducing stress for employees associated job loss, STW allows more labour input flexibility and softens liquidity constraints of firms. In addition, liquidity guarantees target investment and employment and address the externalities arising from constrained firms. Together both policies reduce the output loss of COVID19 and its associated lockdown by about one fourth (see Table 4): Instead of dropping by 13.0 per cent in 2020, the simulated fall in real GDP reaches 8.7%. However, the policy measures cannot prevent a sharp temporary decline in production during the lockdown phase.

The focus of the paper has been on time-limited short-term support schemes. In particular, we have abstracted from the negative impact that these support policies may have on medium-run allocative efficiency through reducing exit-entry rates and labour market churning. We leave this important topic for future research.

Covid Economics 30, 19 June 2020: 120-145

REFERENCES

Acemoglu D., Chernozhukov V., Werning I., and Whinston M.D. (2020). A Multi-Risk SIR Model with Optimally Targeted Lockdown. NBER Working Paper No. 27102, May 2020

Ampudia, M., Georgarakos, D., Slacalek, J., Tristani, O., Vermeulen, P., and Violante, G. (2018). Monetary Policy and Household Inequality. ECB Working Paper, No. 2170.

Baldwin, R. and Weder di Mauro B. (2020), Mitigating the COVID Economic Crisis: Act Fast and Do Whatever It Takes, a VoxEU.org eBook, CEPR Press.

Baker, S., Bloom N., Davis S. J., Kost K., Sammon M. and Viratyosin T. (2020), The Unprecedented Stock Market Reaction to COVID-19, Covid Economics: Vetted and Real-Time Papers 1, 3 April.

Baqaee D. and Farhi E. (2020), Keynesian Production Networks with an Application to the Covid-19 Crisis, Technical Report 2020.

Bayer, C., Born, B., Luetticke, R., and Müller, G. (2020), The coronavirus stimulus package: how large is the transfer multiplier?, CEPR Discussion Paper DP14600.

Bénassy-Quéré, A., Marimon, R., Pisani-Ferry, J., Reichlin, L, Schoenmaker, D., Weder di Mauro, B. (2020), COVID-19: Europe needs a catastrophe relief plan, VoxEU 11/03/202, https://voxeu.org/article/covid-19-europe-needs-catastrophe-relief-plan.

Bigio, S., Zhang M., and Zilberman E. (2020), Transfers vs Credit Policy, Technical Report.

Blanchard, O. and Galí J. (2007). Real Wage Rigidities and the New Keynesian Model. Journal of Money, Credit and Banking 39 (S1), 35-65.

Brunnermeier, M., Landau J.-P., Pagano M., and Reis R. (2020), Throwing COVID Liquidity Life-Line.

Claeys, G. and Wolff G. (2020), COVID-19 Fiscal response: What are the options for the EU Council? https://www.bruegel.org/2020/03/esm-credit-lines-corona-bonds-euro-area-treasury-one-off-joint-expenditures-what-are-the-options-for-the-eu-council/.

Coenen, G., C. Erceg, C. Freedman, D. Furceri, M. Kumhof, R. Lalonde, D. Laxton, J. Linde, A. Mourougane, D. Muir, S. Mursula, C. de Resende, J. Roberts, W. Roeger, S. Snudden. M. Trabandt, J. in 't Veld (2012). Effects of Fiscal Stimulus in Structural Models. American Economic Journal: Macroeconomics, 4 (1): 22-68, available at: http://pubs.aeaweb.org/doi/pdfplus/10.1257/mac.4.1.22.

Davis, S, Faberman J. and Haltiwanger J. (2012), "Labor Market Flows in the Cross Section and over Time", Journal of Monetary Economics, 59 (1), 1–18.

Dolls, M., Fuest, C., and Peichl, A. (2012). Automatic Stabilisers and Economic Crisis: US vs. Europe. Journal of Public Economics 96: 279-294.

Eichenbaum, M., Rebelo S., and Trabandt, M. (2020). The Macroeconomics of Epidemics. NBER Working Papers 26882, National Bureau of Economic Research, Inc. https://ideas.repec.org/p/nbr/nberwo/26882.html.

European Commission (2020), European Economic Forecast, Institutional Paper 125, May. https://ec.europa.eu/info/sites/info/files/economy-finance/ip125_en.pdf.

Covid Economics 30, 19 June 2020: 120-145

Faria e Castro, M. (2020), Fiscal Policy during a Pandemic, Federal Reserve Bank of St. Louis Research Division Working Paper 2020-006D, https://research.stlouisfed.org/wp/more/2020-006.

Fornaro, L. and M. Wolf (2020), Covid-19 coronavirus and macroeconomic policy.

Glover, A., J. Heathcote, D. Krueger and J.-V. Rios-Rull (2020), "Health versus Wealth: On the Distributional Effects of Controlling a Pandemic", CEPR Discussion Paper 14606.

Gopinath G. (2020), Limiting the economic fallout of the coronavirus with large targeted policies, in R. Baldwin and B. Weder di Mauro (eds.), Mitigating the COVID economic crisis: Act fast and do whatever it takes, VoxEU.org eBook, CEPR Press.

Guerrieri, V., Lorenzoni, G., Straub, L. and I. Werning (2020), Macroeconomic Implications of COVID-19: Can Negative Supply Shocks Cause Demand Shortages?

Hagedorn, M. and K. Mitman (2020), Corona Policy According to HANK, CEPR Discussion Paper DP14694.

Jones, C., T. Philippon and V. Venkateswaran (2020), Optimal Mitigation Policies in a Pandemic: Social Distancing and Working from Home, Manuscript, NYU: 6 April.

Jonung, L. and W. Roeger (2006), The macroeconomic effects of a pandemic in Europe - A model-based assessment. ECFIN Economic Papers No 251.

Kaplan, G., Violante, G., and Weidner, J. (2014). The Wealthy Hand-to-Mouth. Brookings Papers on Economic Activity 45: 77-153.

Kollmann, R., Pataracchia, B., Raciborski, R., Ratto, M., Roeger, W., and Vogel, L. (2016). The Post-Crisis Slump in the Euro Area and the US: Evidence from an Estimated Three-Region DSGE Model. European Economic Review 88(C): 21-41.

In 't Veld, J., Pagano, A., Raciborski, R., Ratto, M., and Roeger, W. (2015). Imbalances and Rebalancing in an Estimated Structural Model for Spain. International Journal of Central Banking 11: 1-41.

Jermann, U., and Quadrini V. (2012). "Macroeconomic Effects of Financial Shocks." American Economic Review, 102 (1): 238-71.

Lane, P. (2020), 'The Monetary Policy Package: An Analytical Framework' at https://www.ecb.europa.eu/press/blog/date/2020/html/ecb.blog200313~9e783ea567.en.html.

Narayanan, B., and Walmsley, T. (2008). Global Trade, Assistance, and Production: The GTAP 7 Data Base. Center for Global Trade Analysis, Purdue University.

OECD (2020), "Corporate sector vulnerabilities during the Covid-19 outbreak: assessment and policy responses", Tackling Coronavirus Series.

Ratto, M., Roeger, W., and in 't Veld, J. (2009). QUEST III: An Estimated Open-Economy DSGE Model of the Euro Area with Fiscal and Monetary Policy. Economic Modelling 26: 222-233.

Schmitt-Grohé, S., and Uribe, M. (2003). Closing Small Open Economy Models. Journal of International Economics 61: 163-185.

Covid Economics 30, 19 June 2020: 120-145

Annex A DETAILS ON FIRM LIQUIDITY CONSTRAINTS

The main text has used the following equations describing the investment behaviour of constrained firms:

$$\mathcal{H}\left(\frac{GOS_t^j}{K_{t-1}^j}\right) \equiv \zeta_1\left(\frac{GOS_t^j}{K_{t-1}^j} - \delta\right) - \zeta_2,$$

$$s_t^{li} = a_0 - a_1 GOS_t$$

Liquidity-constrained investment

In this appendix, we show how a share of firms can become liquidity constrained if a shock hits in period t, which reduces the gross operating surplus of the firm. First, we look at an individual firm with budget constraint (dropping i indices to ease notation)

$$L_t = (1 + r_{t-1})L_{t-1} + D_t + I_t - GOS_t \qquad \text{(A.1)}$$

To facilitate our discussion, we make three assumptions: (i) the economy is initially on a balanced growth path, (ii) all firms are financially unconstrained before the shock occurs, and (iii) D_t does not change much across constrained and unconstrained regimes (see Jermann and Quadrini (2012) for a discussion). The firm faces an upper limit on loans which is determined by its capital stock

$$L_t \leq \mu K_t \qquad \text{(A.2)}$$

Prior to the unanticipated adverse shock, the firm is not constrained,

$$L_{t-1} < \mu K_{t-1} \qquad \text{(A.3)}$$

and becomes constrained after receiving a temporary negative GOS shock in period t, which increases borrowing to the collateral limit.

$$L_t^C = \mu K_t^C. \qquad \text{(A.4)}$$

Since the firm is unconstrained in $t-1$

$$K_{t-1} = \left(\frac{1}{\mu} + x\right) L_{t-1} > \frac{1}{\mu_1} L_{t-1} \qquad \text{(A.5)}$$

Since collateral constraint is binding in period t,

$$\Delta K_t^C = (K_t^C - K_{t-1}) = \frac{1}{\mu}\Delta L_t^C - xL_{t-1} < \frac{1}{\mu}\Delta L_t^C, \qquad \text{(A.6)}$$

where the inequality indicates that capital stock falls more than the loan since the firm is now facing a financial constraint. How much does investment decline relative to a situation where the firm is not hit by a negative GOS shock and a standard (unconstrained) Q-equation determines investment. Without adverse shock in t, balanced growth implies that the change in the capital stock (denoted by superscript B) would have exceeded ΔK_t^C:

$$\Delta K_t^B = \left(\frac{1}{\mu} + x\right)\Delta L_t^B > \Delta K_t^C. \qquad \text{(A.7)}$$

The difference between the change of capital with and without constraint $\Delta K_t^C - \Delta K_t^B = \frac{1}{\mu}\Delta L_t^C - xL_{t-1} - \left(\frac{1}{\mu} + x\right)\Delta L_t^B$. Thus,

$$\frac{1}{\mu+z}(\Delta L_t^C - \Delta L_t^B) = \Delta K_t^C - \Delta K_t^B < \frac{1}{\mu}(\Delta L_t^C - \Delta L_t^B) \qquad \text{(A.8)}$$

Since

$$\Delta K_t^C - \Delta K_t^B = I_t^C - I_t^B \qquad \text{(A.9)}$$

Covid Economics 30, 19 June 2020: 120-145

We get

$$(\mu + z)(I_t^C - I_t^B) = (\Delta L_t^C - \Delta L_t^B) \quad \text{(A.10)}$$

From the budget constraint of the firm (and neglecting differences in dividend payouts in both regimes) we obtain the following relationship between investment and GOS in the constrained and unconstrained regime

$$(\Delta L_t^C - \Delta L_t^B) = (\mu + z)(I_t^C - I_t^B) = (I_t^C - I_t^B) - (GOS_t^C - GOS_t^B) \quad \text{(A.11)}$$

$$(I_t^C - I_t^B) = \frac{1}{1-(\mu+z)}(GOS_t^C - GOS_t^B) \quad \text{(A.12)}$$

This result holds for the individual firm or for the case with all firms becoming liquidity constrained. The next section discusses the case when only a fraction of firms becomes constrained.

Aggregate relationship between GOS and Investment

In this section, we show how movement of (average) GOS affect the share of constrained firms by introducing a minimum amount of heterogeneity across firms. We first introduce some notation. There are i firms with $i \in [0,1]$. GOS has an aggregate and an idiosyncratic component governed by σ.

$$GOS_{it} = GOS_t + \sigma(i - 0.5) \quad \text{(A.13)}$$

Firm i will respond to a temporary decline of gross operating surplus by increasing borrowing because of a smoothness restriction on dividend payouts, and convex investment adjustment costs, which makes investment a function of the present discount value of profits. Investment of the unconstrained firm deviates marginally from the investment the firm would have undertaken in t without the COVID shock (see scenario 1). We denote this difference e_t.

$$I_t^U = I_t^B - e_t \quad \text{(A.14)}$$

$$L_{it}^C = (1 + r_{t-1})L_{t-1}^B + D_t^B + I_t^B - e_t - GOS_{it} = \mu K_t^B \quad \text{(A.15)}$$

We denote the marginal firm, which stays unconstrained with $\bar{\imath}$

$$\frac{\sigma}{\mu K_t^B}(\bar{\imath} - 0.5) = \frac{(1+r_{t-1})(\mu+x)K_t^B + D_t^B + I_t^B - e_t - GOS_t}{\mu K_t^B} \quad \text{(A.16)}$$

If GOS_t declines $\bar{\imath}$ increases, i.e. the profitability threshold increases for firms to remain unconstrained. Since $\bar{\imath}$ ranges between zero and one, it can also be interpreted as the share of constrained firms and the relationship between $\bar{\imath}$ and GOS^C can be approximated linearly (where we ignore the term e_t)

$$s_t^{li} = \bar{\imath} = a_0 - a_1 \frac{GOS_t}{K_t^B} \quad \text{(A.17)}$$

Equation (A.17) corresponds to the firm share equation presented in the main text. The parameter values can be determined using information about the share of liquidity constrained firms in the constrained and unconstrained regime. For the unconstrained regime, we assume a share equal to zero. By contrast, in the constrained regime, we use information about the share of output produced by firms directly affected from lockdown measures to set a lower bound of investment undertaken by constrained firms in the COVID regime. This share is set to 0.3. Therefore, we have two equations to determine the two parameter

$$0.3 = a_0 - a_1 \frac{GOS_t}{K_t^B} \quad \text{(A.18)}$$

and

$$0.0 = a_0 - a_1 \frac{GOS_t^B}{K_t^B} \quad \text{(A.19)}$$

Given the investment rule of constrained firm $i \in (0, \bar{\imath})$

$$\left(I_{it}^C - I_t^B\right) = \frac{1}{1-(\mu+z)}\left(GOS_{it}^C - GOS_t^B\right) \quad \text{(A.20)}$$

Covid Economics 30, 19 June 2020: 120-145

We obtain total investment of constrained firm by

$$I_t^C = \int_0^{\bar{\iota}} I_{it}^C \, di = \int_0^{\bar{\iota}} I_t^B \, di + \frac{1}{1-(\mu+z)} \int_0^{\bar{\iota}} \left(GOS_{it}^C - GOS_t^B\right) di \tag{A.21}$$

$$I_t^C = \bar{\iota} I_t^B + \bar{\iota} \frac{1}{1-(\mu+z)} \left((GOS_t - GOS_t^B) + \sigma 0.5(\bar{\iota}^2 - \bar{\iota}) \right)$$

$$I_t^C = \bar{\iota} (I_t^B + \frac{1}{1-(\mu+z)} \left(\left(GOS_t \ - GOS_t^B\right) + \sigma 0.5(\bar{\iota} - 1) \right)) \tag{A.22}$$

With $\sigma 0.5(\bar{\iota} - 1) < 0$

Firm i which is not constrained invests according to the Q equation

$$I_{it}^U = \phi(Q_{it} - 1)K_{it-1} + \delta K_{it-1} = \phi(Q_t + \varphi(i - 0.5) - 1)K_{t-1} + \delta K_{t-1} \tag{A.23}$$

$K_{it-1} = K_{t-1}$ because firms are identical in t-1

$$I_t^U = \int_{\bar{\iota}}^{1} I_{it}^U \, di = \int_{\bar{\iota}}^{1} (\phi(Q_t + \varphi(i - 0.5) - 1)K_{t-1} + \delta K_{t-1}) \, di$$

$$I_t^U = (1 - \bar{\iota})(\phi(Q_t - 1)K_{t-1} + \delta K_{t-1}) + \varphi 0.5\left((1 - \bar{\iota}^2) - (1 - \bar{\iota})\right)K_{t-1}$$

$$I_t^U = (1 - \bar{\iota})(\phi(Q_t - 1)K_{t-1} + \delta K_{t-1}) + \varphi 0.5(\bar{\iota} - \bar{\iota}^2)K_{t-1} \tag{A.24}$$

Total investment

$$\frac{I_t}{K_{t-1}} = \frac{I_t^U}{K_{t-1}} + \frac{I_t^C}{K_{t-1}}$$

$$= (1 - \bar{\iota})(\phi(Q_t - 1) + \delta) + \varphi 0.5(\bar{\iota} - \bar{\iota}^2)$$

$$+ \bar{\iota} \left(\frac{I_t^B}{K_{t-1}} + \frac{1}{1-(\mu+z)} \left(\left(\frac{GOS_t}{K_{t-1}} - \frac{GOS_t^B}{K_{t-1}} \right) + \frac{\sigma 0.5}{K_{t-1}} (\bar{\iota} - 1) \right) \right)$$

$$\frac{I_t}{K_{t-1}} - \delta = \frac{I_t^U}{K_{t-1}} - (1 - \bar{\iota})\delta + \frac{I_t^C}{K_{t-1}} - \bar{\iota}\delta =$$

$$(1 - \bar{\iota})\left(\phi(Q_t - 1)\right) + \varphi 0.5(\bar{\iota} - \bar{\iota}^2) + \bar{\iota} \left(\frac{I_t^B}{K_{t-1}} - \delta + \frac{1}{1-(\mu+z)} \left(\left(\frac{GOS_t}{K_{t-1}} - \frac{GOS_t^B}{K_{t-1}} \right) + \frac{\sigma 0.5}{K_{t-1}} (\bar{\iota} - 1) \right) \right) \tag{A.25}$$

We assume that $\varphi 0.5(\bar{\iota} - \bar{\iota}^2)$ is small since Q_{it} does not respond a lot to temporary shocks.

Covid Economics 30, 19 June 2020: 120-145

Annex B FUNCTIONAL FORMS

Price and capacity utilisation costs follow convex functions:

$$\Gamma_t^{P,j} \equiv 0.5\gamma_P(\pi_t^j)^2 P_{t-1}^j Y_t^j \text{ with } \pi_t^j \equiv P_t^j/P_{t-1}^j - 1\#$$

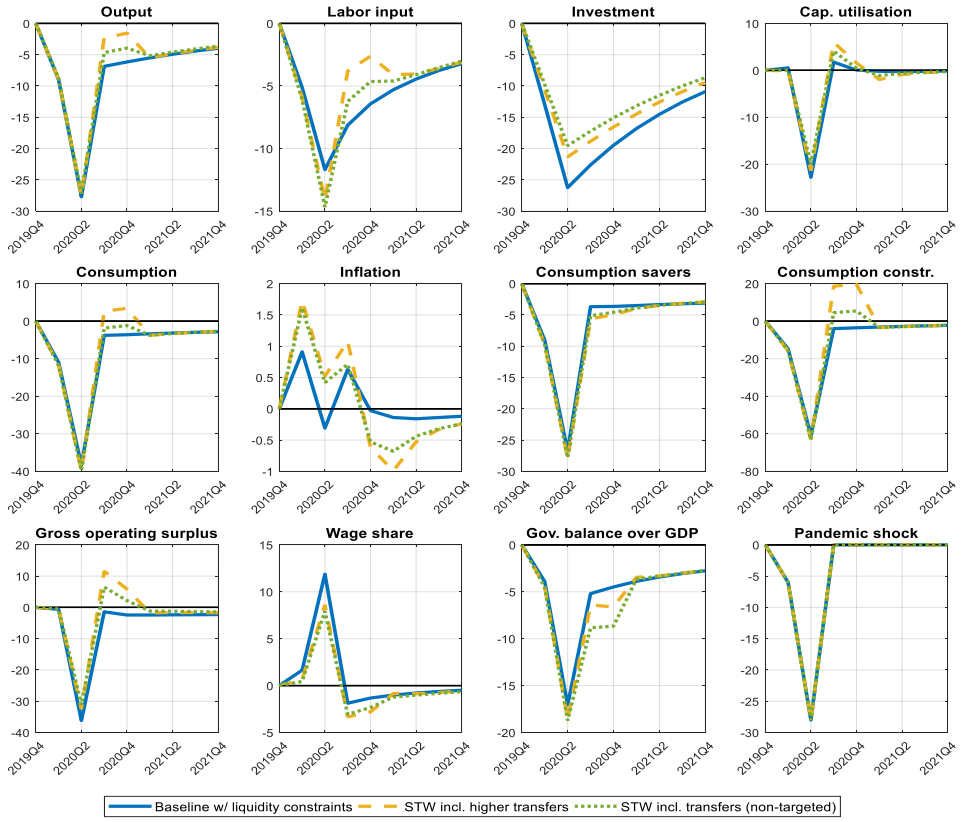
$$\Gamma_t^{ucap,j} \equiv \left(\gamma_{ucap,1}\,(u_t^j - 1) + \frac{\gamma_{ucap,2}}{2}(u_t^j - 1)^2\right)\frac{p_t^I}{p_t^j}K_t^j$$

Covid Economics 30, 19 June 2020: 120-145

CEPR PRESS

Annex C ADDITIONAL RESULTS

Figure 5: **Non-targeted transfers**



Note: This figure expresses the wage share, quarterly inflation, and the government balance in percentage point deviation from steady state. All other variables are expressed in percent deviation from steady state. The pandemic shock is an illustrative index of the exogenous shock process.

Covid Economics 30, 19 June 2020: 120-145