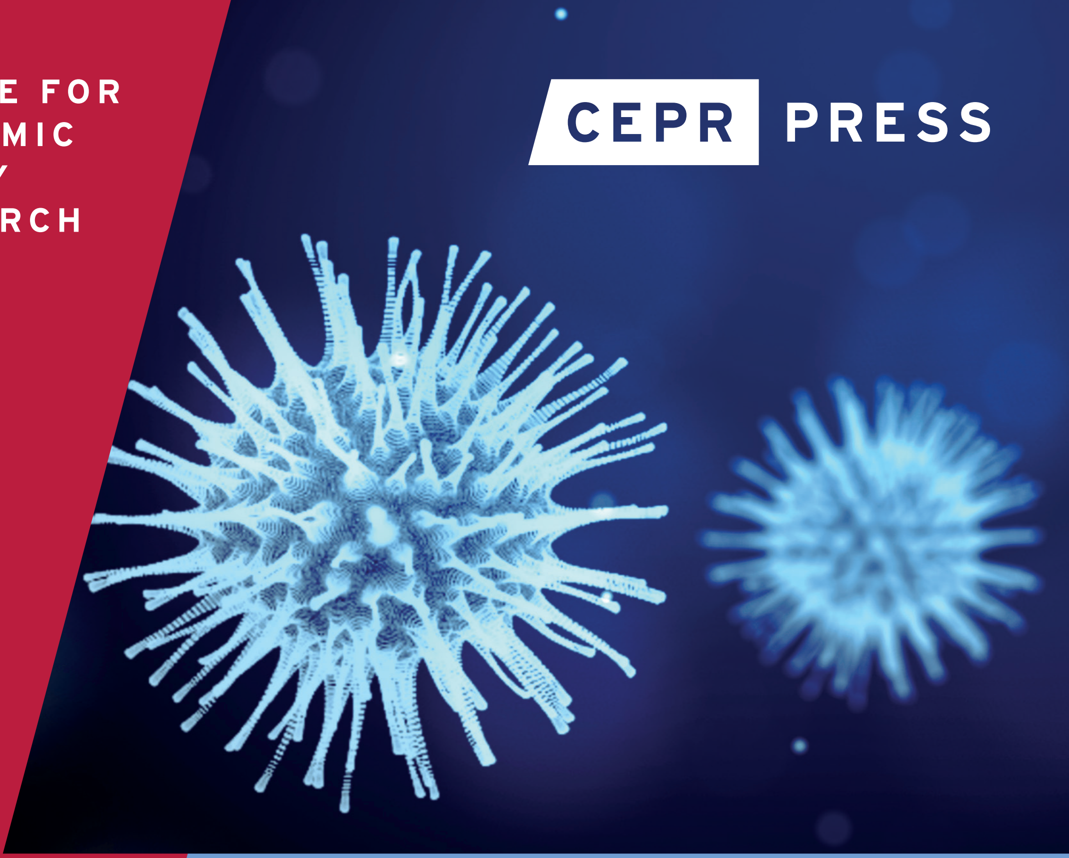
CENTRE FOR ECONOMIC POLICY RESEARCH

CEPR PRESS



COVID ECONOMICS

VETTED AND REAL-TIME PAPERS

ISSUE 66
28 JANUARY 2021

BINARY TESTING
Gabriel Ziegler

EARLY AND TIGHT LOCKDOWNS
Francesca Caselli, Francesco Grigoli and Damiano Sandri

THE WELLBEING CHALLENGE
Jacques Bughin, Michele Cincera, Dorota Reykowska, Marcin Zyszkiewicz and Rafal Ohme

LABOUR MARKETS IN EMERGING MARKET COUNTRIES
Michael A. Nelson

A NEW STRINGENCY INDEX FOR EUROPE
Daniel Gros, Alexandre Ounnas and Timothy Yu-Cheong Yeung

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 Journal, Applied Economics
American Economic Journal, Economic Policy
American Economic Journal, Macroeconomics
American Economic Journal, Microeconomics
American Economic Review
American Economic Review, Insights
American Journal of Health Economics
Canadian Journal of Economics
*Econometrica**
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 Health Economics
Journal of International Economics
*Journal of Labor Economics**
Journal of Monetary Economics
Journal of Public Economics
Journal of Public Finance and Public Choice
Journal of Political Economy
Journal of Population Economics
Quarterly Journal of Economics
*Review of Corporate Finance Studies**
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 66, 28 January 2021

Contents

Covid Economics Issue 66, 28 January 2021

Binary classification tests, imperfect standards, and ambiguous information[1]

Gabriel Ziegler[2]

Date submitted: 19 January 2020; Date accepted: 21 January 2021

New binary classification tests are often evaluated relative to a pre-established test. For example, rapid Antigen tests for the detection of SARS-CoV-2 are assessed relative to more established PCR tests. In this paper, I argue that the new test can be described as producing ambiguous information when the pre-established is imperfect. This allows for a phenomenon called dilation—an extreme form of non-informativeness. As an example, I present hypothetical test data satisfying the WHO's minimum quality requirement for rapid Antigen tests which leads to dilation. The ambiguity in the information arises from a missing data problem due to imperfection of the established test: the joint distribution of true infection and test results is not observed. Using results from Copula theory, I construct the (usually non-singleton) set of all these possible joint distributions, which allows me to assess the new test's informativeness. This analysis leads to a simple sufficient condition to make sure that a new test is not a dilation. I illustrate my approach with applications to data from three COVID-19 related tests. Two rapid Antigen tests satisfy my sufficient condition easily and are therefore informative. However, less accurate procedures, like chest CT scans, may exhibit dilation.

1 Thanks to Dan Sacks, Charles Manski, and Jörg Stoye for literature pointers. I thank Filip Obradovic for very valuable comments. Christopher Stapenhurst provided excellent research assistance. All errors are of course mine.

2 University of Edinburgh, School of Economics.

Covid Economics 66, 28 January 2021: 1-36

1 Introduction

An important aspect of evaluating a new diagnostic test is to assess its accuracy. Intuitively, a sensible binary test should have test results highly correlated with the underlying health condition. In other words, a positive test result should be likely if and only the tested person is indeed infected or sick.[1] However, establishing whether a person is truly infected is often costly or even impossible. Therefore, a new test is analyzed relative to an established test. An established test is perfect when a positive test result occurs if and only if the person is truly infected. The medical literature calls these perfect tests a "gold standard" (Watson et al., 2020). In these situations the joint distribution of the new test's outcomes and the underlying true health condition is the *same* as the joint distribution of test results from both tests. Thus, this observed joint distribution can be used to evaluate the new test's accuracy.

In practice, however, a perfect reference test does not exist. In such a case, the researcher would need the joint distribution of the health conditions and the outcomes of both tests.[2] This overall joint distribution is not observable (or maybe only if the researcher incurs a high costs for obtaining the data). This missing data problem leads to two distinct problems: (*i*) the marginal distribution of the underlying health condition is missing and (*ii*) the *correlation* between new test's outcome and health status is missing too.[3] The latter of these problems will introduce ambiguity in the information provided by the new test.

The first of these problems, missing data about the underlying health condition, is well-known. Recently, Manski and Molinari (2021) use methods known from the literature on partial identification to provide bounds on prevalence—the fraction of infected people in the population.[4] Measuring prevalence is different from the usual inference problem because the tested population might not be representative of the overall population. Here the data are observed selectively which corresponds to a *selection problem* as introduced by Manski (1989). Furthermore, Manski (2020) illustrates how this problem carries over to evaluating accuracy of new

[1] In the following, I will not differentiate between being infected and being sick.

[2] Depending on the question, it might suffice to consider the joint distribution of the new test's outcomes and the underlying true health condition. For example, the analysis Subsection 2.2 requires only this bivariate joint distribution. The trivariate distribution is needed to evaluate the informativeness of performing both tests as discussed in Subsection 2.3.

[3] In the introduction, I use the word correlation loosely and informal.

[4] Similar approaches were used by Stoye (2020) and Sacks et al. (2020).

Covid Economics 66, 28 January 2021: 1-36

Covid Economics 66, 28 January 2021: 1-36

tests in the context of COVID-19 Antibody tests maintaining the assumption of a perfect reference test.

The second problem of missing data about the correlation is different in nature and avoided when a perfect reference is available. Even if one would assume knowledge of prevalence, potentially multiple 'correlation structures' are consistent with the observed data. The reason for this multiplicity is well-known from copulas as studied in probability theory. Knowledge of prevalence provides the marginal distribution of the health condition, whereas the observed testing data provides (a bivariate) marginal distribution. In general, there are multiple (trivariate) joint distributions with these marginal distributions. Due to this multiplicity a simple, unambiguous interpretation of the new test is not possible. Without knowledge of prevalence, the problem identified before carries over and therefore exacerbates the overall multiplicity. However, as discussed in more detail later, the ambiguous information stems only from the missing data on correlation and therefore occurs whether or not the researcher has knowledge about prevalence.

In this paper, I provide a theoretic framework that combines insights from Manski and Molinari (2021) and Stoye (2020) about selective testing with the missing correlation data due to an imperfect reference test. Within this framework, it is possible to address informativeness of both tests. First, Proposition 1 shows that the established test's negative predictive value[5] is usually not given by a unique number, but it always informative nevertheless. This multiplicity arises because of problem (i) only. Then, I analyze the new test's informativeness for the test population only. The focus on the tested population simplifies the algebra and furthermore shuts down the ambiguity about prevalence (cf. problem (i)) and therefore allows me to study the essence of ambiguous information for the new test in separation (problem (ii) only). Finally, I study the implications on informativeness if both effects are present.

Studying the informativeness of tests has a long tradition in probability theory, statistics, economics, and philosophy. Blackwell (1951, 1953) introduces a notion of "(more) informative" for (what is now called Blackwell) experiments.[6] An experiment is a mapping from states of the world to a distribution over signals.

[5]A test's negative predictive values is the probability of being healthy conditional on obtaining a negative test result. Another important informativeness measure is the positive predictive value, which is the probability of being infected conditional on a positive test result. I will assume throughout that the established has a perfect predictive value in line with the application to SARS-CoV-2 testing.

[6]de Oliveira (2018) provides a more recent treatment.

In the current setting, an experiment is a function that associates a distribution over test results to each of the possible health conditions, i.e. for being infected and for being healthy. In such a setting, the value of information is defined as the amount a Bayesian decision maker is willing to pay for the experiment. Since every experiment is more informative than an uninformative experiment,[7] Blackwell's theorem shows that the value of information is (weakly) positive for every Bayesian decision maker.[8] Ideally a diagnostic test should satisfy Blackwell's definition of an experiment in order to ensure that it is always informative. However, this is typically only true for the established test in my framework.

The new test fails to be a Blackwell experiment because it does not map each state to a unique distribution over test results. Rather, due to the multiplicity of joint distributions, there is a *set* of distributions over test results for a given health condition.[9] Therefore, Blackwell's informativeness notion does not apply to the new test. Furthermore, the value of information needs to be adjusted because a Bayesian analysis does not readily apply with sets of probabilities. Such a situation is usually referred to as a situation of "ambiguity" and the literature has identified several extensions of Bayesian decision making to the realm of ambiguity.[10]

Instead of defining the value of information for a specific decision criterion in such a situation, I adopt a very weak notion of informativeness: the diagnostic test is informative if and only if it is not a *dilation*. Seidenfeld and Wasserman (1993) introduce the notation of dilation for situations with sets of probabilities. In the current context, a dilation occurs if, no matter what test result is obtained, the set of probabilities conditional on this information contains the original set of probabilities. Figure 1 illustrates an example of a dilation. Here, the set of probabilities indicating the infection likelihood before the test (black set) lies within both sets after the test result (blue for a positive result and red indicating the set after a negative result). Thus, in a sense, the decision maker is worse-off after taking the test than before taking the test *no matter* what the test result is. For this reason, Seidenfeld and Wasserman call a dilation a "counterintuitive phenomenon" and Gul and Pesendorfer (2018) refer to it as "all news is bad news".

[7] An experiment is uninformative if the mapping mentioned above is a constant function?

[8] More generally, Blackwell characterizes his notion of "more informative" with the requirement that every Bayesian decision maker has a higher value of information for the more informative experiment.

[9] Formally, the new test can be seen a correspondence or set-valued function.

[10] Machina and Siniscalchi (2014) provide a recent overview about this topic.

Covid Economics 66, 28 January 2021: 1-36

CEPR PRESS

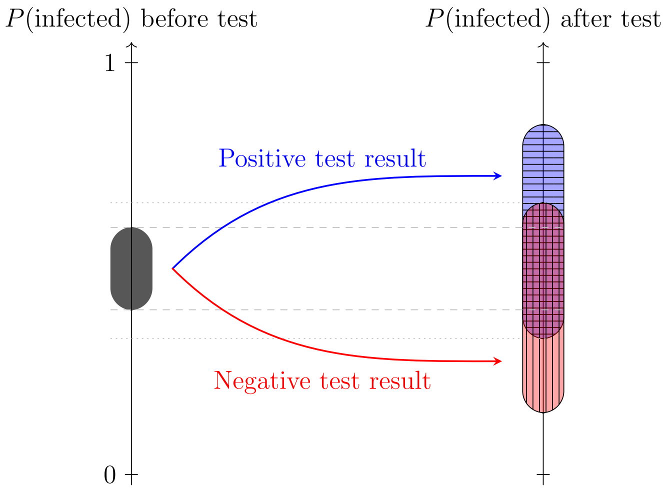


Figure 1: A diagnostic test as dilation

Table 1: Dilation Test Data

$z\backslash y$	$y=0$	$y=1$	Sum
$z=0$	39.5%	11.5%	51%
$z=1$	1.2%	47.8%	49%
Sum	40.7%	59.3%	

$y=1$ indicates a positive PCR-test result (i.e. the established test). $z=1$ denotes a positive Antigen test.

My framework allows to fully characterize when a new diagnostic test is a dilation (cf. Expression (3)). Since the definition of informativeness for tests is rather weak, any reasonable test should satisfy this criterion. The characterization provides a method to verify whether the new test is informative.

The WHO (2020) recommends a minimum standard of accuracy for rapid Antigen tests.[11] Usually a PCR test is the established test used to evaluate these Antigen tests (Esbin et al., 2020). Table 1 illustrates hypothetical test data, which fulfill the minimal requirements of the WHO. However, as the analysis will reveal, this test is actually a dilation and therefore not informative.[12] For minimum required accuracy standards, the dilation characterization provides an easy to

[11]For the informed reader, the WHO recommends a sensitivity of at least 80% and a specificity of at least 97%. These measures will be formally introduced and defined later.

[12]This statement depends on the accuracy of the established PCR test. A dilation occurs only if the PCR has sensitivity at the lower range identified by the literature.

Covid Economics 66, 28 January 2021: 1-36

verify sufficient condition to avoid dilation. The new test is informative (in the population of tested people) if

$$\boxed{\begin{array}{c}\text{Fraction of people with a positive new test}\\ \leq \\ \text{minimum sensitivity of new test} \times \text{sensitivity of established test.}\end{array}}$$

Besides theoretic applications of dilation, there is not much empirical evidence in the literature yet. Recently, economists started to investigate dilation and ambiguous information experimentally. The only experiment focusing on how decision makers react to dilation and relate the behavior to the value of information is conduced by Shishkin and Ortoleva (2020). To the best of my knowledge, the possible occurrence of dilation with diagnostic tests (or SARS-CoV-2 tests more specifically) is the first observation of this phenomenon 'in the field.'[13]

Of course, researchers studying diagnostic tests are well aware of the general issues addressed here. The problem of selection leading to unobserved prevalence is known as *Verification Bias*, whereas the problem arising from unobserved correlation due to an imperfect reference test is descriptively named *Imperfect Gold Standard Bias.* (Zhou et al., 2014, Chapters 10–11) This paper is not the first to document that either of of these problem leads to non-identified models; rather the novelty of this paper comes in the approach. Diagnostic test research seeks to avoid non-identified models by introducing additional assumptions and then address resulting biases relative to a baseline assumption. Proposed methods include simply imputing missing data or considering more sophisticated correction methods. Moreover, the two problems are often addressed separately. By contrast, my framework requires minimal assumptions and addresses both problems simultaneously.[14]

2 Main Analysis

I consider the following situation. Let $x = 1$ denote that a person is infected and $x = 0$ if the person is healthy. Initially, there is binary test available where $y = 1$

[13] Manski (2018) mentions that a dilation might occur in a different medical context, but does not address this issue further.

[14] Reitsma et al. (2009) provide a flowchart as guidance for applied researchers to address several problems arising when establishing accuracy of diagnostic tests. The two problems addressed here are in two distinct branches of the flowchart.

Covid Economics 66, 28 January 2021: 1-36

indicates a positive test result and $y = 0$ a negative result. Finally, a new test is introduced which again can be either positive ($z = 1$) or negative ($z = 0$).

Let $P(x, y, z)$ denote the population distribution under consideration with $p := P(x = 1)$ denoting prevalence. However, the population distribution is not directly observable. This is, of course, almost always the case, because a researcher usually only observes a sample from the population distribution. This leads to the usual inference problem. Throughout, I will abstract away from inference altogether. Instead, the data is given for people who were tested to obtain data on the new test. For this denote tested people with $t = 1$ and $t = 0$ otherwise. Then, the data are given by $P(y, z|t = 1)$ and I assume that $P(t = 1) > 0$.[15,16]

Furthermore, since the established test is well-known, precise information about the sensitivity and specificity of this test is available as well. The following assumption ensures that both of these measures are well defined.

Assumption 1 (Non-trivial prevalence)**.** *The population satisfies* $p \in (0, 1)$*.*

With this assumption, sensitivity and specificity for the initial test are respectively defined as:

$$P(y = 1|x = 1) = \frac{P(x = 1, y = 1)}{p} =: \sigma, \quad \text{and} \tag{1}$$

$$P(y = 0|x = 0) = \frac{P(x = 0, y = 0)}{1 - p}. \tag{2}$$

As discussed in Manski (2020), for decision making sensitivity and specificity are not the relevant measures. The relevant measures are *positive predictive value (PPV)* and *negative predictive value (NPV)*. For the established test these measures can be obtained from specificity and sensitivity via Bayes' rule if prevalence p and $P(y = 1)$ are known:

$$\text{PPV}_y := P(x = 1|y = 1) = \frac{p}{P(y = 1)} P(y = 1|x = 1) = \frac{p}{P(y = 1)} \sigma$$
$$\text{NPV}_y := P(x = 0|y = 0) = \frac{1 - p}{P(y = 0)} P(y = 0|x = 0).$$

[15] Furthermore, the following logical implications of (not) being tested hold: (i) $t = 0 \implies z = 0$ and (ii) $z = 1 \implies t = 1$. Note that $y = 1$ is possible even if not tested, because the participation pool concerns only the new test.

[16] Equivalently, the data is given by sensitivity and specificity of the new test *relative* to the established test with the additional information about how many established or new tests had a positive result.

Covid Economics 66, 28 January 2021: 1-36

Since the tested people are usually not representative of the overall population,[17] even for the established test these two measures are not point-identified. (Manski and Molinari, 2021; Manski, 2020; Stoye, 2020)

To simplify the analysis and in-line with the application to SARS-CoV-2 testing, I also consider the following three baseline assumptions.

Assumption 2 (No False-Positives for established test)**.** *The population satisfies* $P(x = 0, y = 1) = 0$.

Assumption 2 implies that the established test achieves a maximum specificity and PPV_y of 1.[18]

Additionally, I will assume *test-monotonicity* as in Manski and Molinari (2021), meaning conditional on being tested the probability of being infected is greater than if not being tested.[19]

Assumption 3 (Test-monotonicity)**.** *The population satisfies* $P(x = 1|t = 1) \geq P(x = 1|t = 0)$.

Lastly, I assume that the established test's sensitivity does depend on the underlying health status x, but not on whether the person is in the testing pool $t = 1$.[20]

Assumption 4 (Health-sufficiency)**.** *The population satisfies* $P(y = 1|x = 1, t = 1) = P(y = 1|x = 1) = \sigma$.

To reduce cumbersome lengthy notation in the following, I will use this simplified notation henceforth:

$$\gamma := P(y = 1|t = 1) \qquad \ldots \text{ test yield for established test}$$
$$\zeta := P(z = 1|t = 1) \qquad \ldots \text{ test yield for new test}$$
$$\tau := P(t = 1) \qquad \ldots \text{ measure of data representativeness.}$$

[17] For example, supposedly infected people may be oversampled in order to get meaningful results.

[18] This holds because $P(x = 0, y = 0) = P(x = 0, y = 0) + P(x = 0, y = 1) = P(x = 0) = 1 - p$ and $P(x = 1, y = 1) = P(x = 1, y = 1) + P(x = 0, y = 1) = P(y = 1)$.

[19] This might not be true, if there is voluntary enrollment into the testing pool. However, for establishing the accuracy of new tests this assumptions seems to be applicable often. See Footnote 17.

[20] Recall that the testing pool is obtained for the new test. This assumption might be violated, if, for example, the medical staff performing the established test for the testing pool is extra careful. In this case, the established test might be more sensitive for the testing pool.

Covid Economics 66, 28 January 2021: 1-36

Covid Economics 66, 28 January 2021: 1-36

To avoid trivial cases, assume that $\gamma, \zeta, \tau > 0$. Note that τ has a slightly different interpretation as in Manski and Molinari (2021) or Stoye (2020). Here, $\tau = 1$ means the data $P(y, z|t = 1)$ is perfectly representative of the overall population. In particular, such a parameter value implies no oversampling of infected participants.[21] In particular, even if the participation pool is small (as is often the case), this does not mean that τ should be close to zero.[22]

With this notation, we have $P(z = 1) = \tau\zeta$ since the new test is positive only if the person was tested. Furthermore, Assumption 2 combined with Assumption 4 gives $P(x = 1|t = 1) = \gamma/\sigma$. Then, the Law of Total Probability together with Assumption 3 provides sharp bounds[23] on prevalence $p \in [\tau\gamma/\sigma, \gamma/\sigma] =: [\underline{\chi}, \overline{\chi}]$ because

$$p := P(x = 1) = \underbrace{P(x = 1|t = 1)}_{=\frac{\gamma}{\sigma}}\tau + \underbrace{P(x = 1|t = 0)}_{\in\left[0, \frac{\gamma}{\sigma}\right] \text{ by Assumption 3}}(1 - \tau).$$

In turn, bounds on the established test's overall positivity rate are implied by sensitivity σ and Assumption 2: $P(y = 1) = p\sigma \in [\tau\gamma, \gamma]$.

Since we consider the non-trivial case of $p \in (0, 1)$, consistency of the data with the maintained assumptions requires the established test's sensitivity to be sufficiency high , i.e. $\gamma < \sigma \leq 1$. In turn, the assumptions imply $P(y = 1) \in (0, 1)$.

2.1 The established test

Assumption 2 implies a perfect positive predictive value for the established test ($\text{PPV}_y = 1$). However, the negative predictive value is only partially identified and Proposition 1 provides sharp bounds.

Proposition 1. *Under Assumption 1–Assumption 4, the established test's negative predictive value is sharply bounded as follows:*

$$\text{NPV}_y \in \left[\frac{1}{\sigma}\frac{\sigma - \gamma}{1 - \gamma}, \frac{1}{\sigma}\frac{\sigma - \tau\gamma}{1 - \tau\gamma}\right].$$

[21] See Footnote 17 for why such an assumption might be problematic.

[22] A small participation pool might worsen the statistical inference problem: suppose the participation pool is perfectly representative but small. In this case $\tau = 1$, but inference usually relies on some sort of central limit theorem which would not be appropriate in this scenario. However, recall that I abstract away from inference problems as mentioned above.

[23] A bound for a given set is called *sharp* if the bound itself is a member of this set.

Proof. Fix $\alpha = P(y=1) \in [\tau\gamma, \gamma]$ and define prevalence as a function of α by $P_\alpha(x=1) = \alpha/\sigma$. Then

$$\text{NPV}_y(\alpha) = \frac{1 - P_\alpha(x=1)}{1-\alpha} \underbrace{P(y=0|x=0)}_{=1 \text{ by Assumption 2}} = \frac{1}{\sigma}\frac{\sigma-\alpha}{1-\alpha}$$

Since $\sigma - \alpha \leq 1 - \alpha$ for all $\alpha \in [\tau\gamma, \gamma]$, $NPV_y(\cdot)$ is decreasing. Therefore, $\text{NPV}_y \in [\text{NPV}_y(\gamma), \text{NPV}_y(\tau\gamma)]$. ■

With this in hand, the established test's informativeness can be analyzed. Table 2 summarizes the prevalence before and after observing a test result from the established test, which are the relevant measures for defining informativeness (cf. Figure 1). Formally, the established test is a *dilation* if every possible prevalence $p \in [\underline{\chi}, \overline{\chi}]$ is a possible value of both $P(x=1|y=1) = \text{PPV}_y$ and $P(x=1|y=0) = 1 - \text{NPV}_y$. Obviously, a positive test result gives perfect knowledge due to the maintained assumptions. Thus, the established test cannot be a dilation. On the other hand, a negative result lowers the lower and upper bound of prevalence conditional on a negative result because $\tau\gamma \leq \gamma < \sigma$.[24] Furthermore, the interval width for any test result shrinks the set of possible values for prevalence conditional on either test result.[25] In this sense, the established test is not just informative (i.e. not a dilation), but also strictly shrinks the size of the set of possible prevalence values after a negative test result.

Table 2: Informativeness of the established test

$P(x=1\vert\cdot)$	Lower bound	Upper bound	Interval Width
Prior testing	$\tau\frac{\gamma}{\sigma}$	$\frac{\gamma}{\sigma}$	$\frac{\gamma}{\sigma}(1-\tau)$
Positive result ($y=1$)	1	1	0
Negative result ($y=0$)	$\frac{\tau\gamma}{\sigma}\frac{1-\sigma}{1-\tau\gamma}$	$\frac{\gamma}{\sigma}\frac{1-\sigma}{1-\gamma}$	$\frac{\gamma}{\sigma}\frac{1-\tau}{1-\tau\gamma}\frac{1-\sigma}{1-\gamma}$

Remark: The second row corresponds to PPV_y and the third row is given by $1 - \text{NPV}_y$.

It is well known that knowledge of prevalence is needed in order to apply Bayes' rule to obtain NPV. Since in most applications prevalence is not known, a common practice is to assume a given prevalence level. For example the United States Food and Drug Administration (FDA, 2020b) assumes a prevalence of 5% to cal-

[24] Of course, this has to hold since the Law of Total Probability holds pointwise.

[25] It is obvious for a positive test result. For a negative result, note that the width strictly increases if and only if $1 - \sigma > (1-\gamma)(1-\tau\gamma)$, which is equivalent to $\tau(1+\gamma) > \frac{\sigma}{\gamma} + 1 \geq 2$ leading to a contradiction.

Covid Economics 66, 28 January 2021: 1-36

CEPR PRESS

culate PPV and NPV. If such an assumption ($p = \chi$) is added to the maintained assumptions, then $P(y=1) = \chi\sigma$ and furthermore $P(y=1|t=0) = \frac{\chi\sigma-\gamma\tau}{1-\tau}$.[26] This additional assumption allows to exactly pin down the established test's NPV as $\frac{1-\chi}{1-\chi\sigma}$ and therefore $P(x=1|y=0) = \chi\frac{1-\sigma}{1-\chi\sigma}$. Thus, this additional assumption not only assumes away the ambiguity about prevalence, but also illustrates that the established test does not provide ambiguous information itself.[27] The apparent ambiguity reflected in the non-trivial interval for values of prevalence after a negative test result (cf. Table 2) or NPV (cf. Proposition 1) is only a manifestation of the ambiguity about prevalence, but it is not due to the test itself.

2.2 The new test

Next, the new test's informativeness is analyzed. First, I will discuss informativeness only based on the tested population. For this subpopulation the prevalence is given by $\overline{\chi} = \gamma/\sigma$ and therefore the ambiguity about prevalence is muted. Subsection 2.4 extends the analysis then to the informativeness of the new test for the overall population. For the test population, the relevant measures are again positive-predictive value (PPV) and negative-predicative value (NPV), but now they are also conditional on being tested:

$$\mathrm{PPV}_z := P(x=1|z=1,t=1) = \frac{P(x=1,z=1|t=1)}{P(z=1|t=1)} = \frac{P(x=1,z=1|t=1)}{\zeta}$$
$$\mathrm{NPV}_z := P(x=0|z=0,t=1) = \frac{P(x=0,z=0|t=1)}{P(z=0|t=1)} = \frac{P(x=0,z=0|t=1)}{1-\zeta}.$$

To obtain these measures, the distribution $P(x,z|t=1)$ is needed. For a fixed τ, I use a result from Joe (1997) that provides the set of all possible joint distributions $P(x,y,z)$ compatible with the data $P(x,y|t=1)$ (cf. Appendix A). Setting $\tau = 1$ in this construction gives the possible distributions $P(x,y,z|t=1)$. Finally, $P(x,z|t=1)$ is obtained by marginalization.

[26] Alternatively, one could drop the assumption that $P(t=1) = \tau$ is known exactly. In this case (and allowing $P(y=1|t=0) \in [0,\gamma]$ as in the general case) the assumed prevalence bounds τ. Calculations show that $\tau \in \left[0, \frac{\chi\sigma}{\gamma}\right]$. Since the lower bound is always $\tau_{\min} = 0$, we do not find this case very interesting.

[27] Technically, the established test is an experiment á la Blackwell (1951), where sensitivity and specificity can be seen as functions mapping (health) states to distributions over signals (i.e. test results). As mentioned in the introduction, this implies that the established test's value of information is (weakly) positive under these assumptions.

Covid Economics 66, 28 January 2021: 1-36

Covid Economics 66, 28 January 2021: 1-36

To simplify the algebraic expressions it will be useful to differentiate between four cases defined in Table 3. Fixing the established test's sensitivity σ, the test data $P(x, y|t = 1)$ immediately reveals the case the test belongs to. Figure 2 illustrates this for three real tests considered later (StQ, BiN, CT) and three hypothetical tests (including the dilation test from Table 1). When $\sigma \to 1$, then all but the informative case (I) cease to be relevant. For SARS-CoV-2 detecting Antigen test the WHO recommends a minimum specificity close to one. Tests close to the (top-right) frontier in Figure 2 satisfy this criterion.[28] Thus, for most applications, either the *confirmatory* (if $\sigma < 1$) or the *informative* case (if $\sigma \approx 1$) will be the relevant ones.

Table 3: Cases relating the two tests

Case name	Parameter restriction
Confirmatory (C)	$P(y = 0, z = 0\|t = 1) \geq \max\{\overline{\chi}(1-\sigma), 1-\overline{\chi}\}$
Informative (I)	$1-\overline{\chi} > P(y = 0, z = 0\|t = 1) \geq \overline{\chi}(1-\sigma)$
Uninformative (U)	$\overline{\chi}(1-\sigma) > P(y = 0, z = 0\|t = 1) \geq 1-\overline{\chi}$
Contradictory (X)	$\min\{\overline{\chi}(1-\sigma), 1-\overline{\chi}\} > P(y = 0, z = 0\|t = 1)$

Recall $\overline{\chi} = \gamma/\sigma$ is the upper bound on prevalence and $\gamma = P(y = 1|t = 1)$ is established test's yield.

In contrast to the established test, the new test's PPV could be less than one and is, in general, only set-identified. The reason for set-identification is that $P(x, z|t = 1)$ is not directly observed. As explained above, there are multiple distributions $P(x, z|t = 1)$ consistent with the data and each distribution leads to a potentially different PPV. Proposition 2 establishes the sharp identified set for values of PPV.

Proposition 2 (PPV)**.** *Under Assumption 1, Assumption 2, and Assumption 4, the new test's positive predictive value is*

$$\text{PPV}_z \in \begin{cases} [P(y=1|z=1,t=1), 1] & \textit{in case (C)} \\ \left[P(y=1|z=1,t=1), P(y=1|z=1,t=1) + \overline{\chi}\frac{1-\sigma}{\zeta}\right] & \textit{in case (I)} \\ \left[1 - \frac{1-\overline{\chi}}{\zeta}, 1\right] & \textit{in case (U)} \\ \left[1 - \frac{1-\overline{\chi}}{\zeta}, P(y=1|z=1,t=1) + \overline{\chi}\frac{1-\sigma}{\zeta}\right] & \textit{in case (X).} \end{cases}$$

[28] CT, Uni, and Anti do not satisfy the WHO minimum requirement of a 97% minimum specificity.

CEPR PRESS

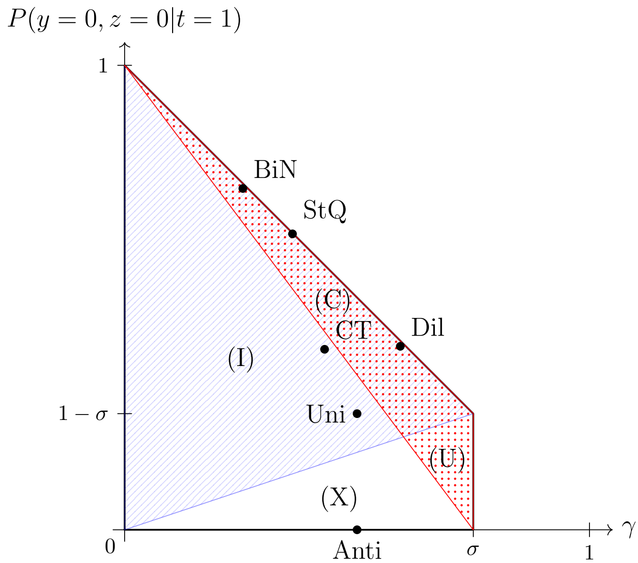


Figure 2: Illustrating the four cases with $\sigma = 0.75$:

Uni is a test corresponding to a uniform distribution $P(y, z|t = 1) = 1/4$. *Anti* always produces the opposite result of the established test $P(y = 1, z = 0|t = 1) = P(y = 0, z = 1|t = 1) = 1/2$. *StQ*, *BiN*, and *CT* are real tests studied later in Section 3. *Dil* is the dilation test given by Table 1.

Proof. Conditional on $t = 1$ is the same as if $\tau = 1$. Thus, From Table 20 and Table 21, we obtain respectively:[29]

$$\underline{P}(x = 1, z = 1|t = 1) := P(y = 1, z = 1|t = 1) + \max\left\{0, \underbrace{P(y = 0, z = 1|t = 1) - (1 - \overline{\chi})}_{=\overline{\chi}-\gamma-P(y=0,z=0|t=1)}\right\}$$

$$\begin{aligned}\overline{P}(x = 1, z = 1|t = 1) &:= P(y = 1, z = 1|t = 1) + \min\{\overline{\chi}, 1 - P(y = 0, z = 0|t = 1)\} - \gamma \\ &= \min\{\overline{\chi}, 1 - P(y = 0, z = 0|t = 1)\} - P(y = 1, z = 0|t = 1) \\ &= \min\left\{\underbrace{\overline{\chi} - P(y = 1, z = 0|t = 1)}_{=\overline{\chi}-\gamma+P(y=1,z=1|t=1)}, \zeta\right\}.\end{aligned}$$

Now, note that $\overline{\chi} - \gamma = \overline{\chi}(1 - \sigma)$ and divide by ζ to obtain PPV. ■

[29] Here and in the following I will use $\underline{P}$ to denote lower bound distributions and $\overline{P}$ for upper bounds.

Covid Economics 66, 28 January 2021: 1-36

Covid Economics 66, 28 January 2021: 1-36

To avoid partially identified predictive values, these measures for the new tests are often reported *as if* the reference test is perfect. In this case, the data $P(y, z)$ alone delivers a unique predictive value:

Corollary 1 (Perfect Gold Standard - PPV)**.** *Suppose Assumption 1, Assumption 2, and Assumption 4 hold. If $\sigma = 1$, i.e. the established test has perfect sensitivity, then $PPV_z = P(y = 1|z = 1, t = 1)$.*

Proof. If $\sigma = 1$, then $\overline{\chi} = \gamma = P(y = 1|t = 1)$. Thus, the relevant case is (I).[30] Therefore the lower bound is $P(y = 1|z = 1, t = 1)$ and the upper bound is $P(y = 1|z = 1, t = 1) + \underbrace{\frac{\gamma}{\zeta}\frac{1-\sigma}{\sigma}}_{=0}$. ■

We saw before that the established test always achieves a maximal PPV of one and therefore provides a lot of information in case it delivers a positive result. How informative is a positive result of the new test? To answer this question, note that since we condition on being tested, there is no prior uncertainty as the prevalence in the testing pool is given by $\overline{\chi}$. Even without this prior ambiguity there remains ambiguity in the test result. For example, in the confirmatory case (C) the interval's width of possible values for PPV_z is $1 - P(y = 1|z = 1, t = 1)$, which is usually small—but non-zero—in applications. Thus, the information obtained from a new test is ambiguous at least after a positive test result.[31]

In contrast to the established test as discussed in Subsection 2.1, the ambiguity arising from the new test allows for the occurrence of dilation. In the current setting a dilation occurs if $\overline{\chi} = P(x = 1|t = 1)$ is contained in the intersection of the two sets with possible values for $P(x = 1|y = i, t = 1)$ for each test result $i \in \{0, 1\}$. Is it possible that after a positive test result the set of possible values for $P(x = 1|y = 1, t = 1)$ contain $\overline{\chi} = P(x = 1|t = 1)$? Corollary 2 provides a full characterization. The corresponding case after a negative test result will be discussed afterwards.

Corollary 2 (Bounds Increase - PPV)**.** *Suppose Assumption 1, Assumption 2, and Assumption 4 hold. The new test's possible values for PPV contain $\overline{\chi}$ =*

[30] Technically, one could be on the border of case (C) or (U), but by continuity the resulting bounds do not change.

[31] This observation alone implies the test is *not* an experiment à la Blackwell (1951). Similarly to deriving PPV, it is possible to derive the new test's sensitivity. This will be a set in general too. Thus, there is a correspondence from (health) states to distributions over signals (test results).

Covid Economics 66, 28 January 2021: 1-36

$P(x = 1|t = 1)$ *if and only if*

$$\sigma \leq \min\left\{\frac{\gamma(1-\zeta)}{P(y=1,z=0|t=1)}, \frac{\gamma\zeta}{P(y=1,z=1|t=1)}\right\},$$

where the first entry corresponds to an increase in the upper bound, and the second condition ensures the lower bound decreases.

Proof. For the upper bounds, note that a strict decrease can only happen if and only if cases (I) or (X) occur (i.e. $1-\overline{\chi} > P(y=0,z=0|t=1)$) and $P(y = 1|z=1,t=1) + \overline{\chi}\frac{1-\sigma}{\zeta} < \overline{\chi}$. The first is equivalent to $\sigma > \frac{\gamma}{1-P(y=0,z=0|t=1)}$ and the second to $\sigma > \frac{1-\zeta}{P(z=0|y=1,t=1)} = \frac{\gamma}{P(y=1|z=0,t=1)}$. Since $P(y=0,z=0|t=1) \leq P(y = 0|z=0,t=1)$ we also have $P(y=1|z=0,t=1) \leq 1-P(y=0,z=0|t=1)$. Thus, a strict decrease happens if and only if $\sigma > \frac{\gamma}{P(y=1|z=0,t=1)}$.

For the lower bound, note that $1-\frac{1-\overline{\chi}}{\zeta} \leq \overline{\chi}$ always holds. For the other cases (C and I, i.e. $P(y=0,z=0|t=1) \geq \overline{\chi}(1-\sigma)$), a strict increase is $P(y=1|z=1,t=1) > \chi$, which is equivalent to $\sigma > \frac{\gamma}{P(y=1|z=1,t=1)}$. Whereas the case condition is equal to $\sigma \geq \frac{\gamma}{1-P(y=0,z=1|t=1)}$. Similar to above, $P(y=1|z = 1,t=1) \leq 1-P(y=0,z=1|t=1)$ and therefore a strict increase happens if and only if $\sigma > \frac{\gamma}{P(y=1|z=1,t=1)}$. ∎

The inequality of Corollary 2 becomes non-trivial in case the testing data does not correspond to an independent distribution, which will be the case for most applications. In these cases, a dilation cannot occur when the non-trivial inequality of Corollary 2 is violated.

It remains to analyze the information contained in a negative test result. Proposition 3 establishes sharp bounds for the negative predictive value of the new test. This is the relevant measure for analyzing how informative a negative test result is.

Proposition 3 (NPV)**.** *Under Assumption 1, Assumption 2, and Assumption 4, the new test's negative predictive value is*

$$\text{NPV}_z \in \begin{cases} \left[P(y=0|z=0,t=1) - \overline{\chi}\frac{1-\sigma}{1-\zeta}, \frac{1-\overline{\chi}}{1-\zeta}\right] & \textit{in case (C)} \\ \left[P(y=0|z=0,t=1) - \overline{\chi}\frac{1-\sigma}{1-\zeta}, P(y=0|z=0,t=1)\right] & \textit{in case (I)} \\ \left[0, \frac{1-\overline{\chi}}{1-\zeta}\right] & \textit{in case (U)} \\ [0, P(y=0|z=0,t=1)] & \textit{in case (X).} \end{cases}$$

Proof. Conditional on $t = 1$ is the same as if $\tau = 1$. Thus, From Table 20 and Table 21, we obtain respectively:

$$\begin{aligned}\underline{P}(x=0,z=0|t=1) &= \max\{0, 1-\overline{\chi} - P(y=0,z=1|t=1)\} \\ &= \max\left\{0, \underbrace{\gamma - \overline{\chi}}_{=-\overline{\chi}(1-\sigma)} + P(y=0,z=0|t=1)\right\} \\ \overline{P}(x=0,z=0|t=1) &= \min\{1-\overline{\chi}, P(y=0,z=0|t=1)\}\,.\end{aligned}$$

Division by $P(z=0|t=1) = 1-\zeta$ gives NPV. ■

The uninformative (U) and contradictory (X) case seem problematic in light of Proposition 3. In both cases, the lower bound is zero and also the width of the interval is rather large. This is another indication that any reasonable test should not fall in either of these two cases. However, even for the other cases—and like for PPV—the NPV is generally only set-identified. Therefore a negative test result also produces ambiguous information.

Avoiding this ambiguity can be achieved with a perfect reference test. Corollary 3 verifies that if the reference test is perfect, Proposition 3 reduces to the expression used in many applications and can be calculated directly from the data $P(y, z)$.

Corollary 3 (Perfect Gold Standard - NPV)**.** *Suppose Assumption 1, Assumption 2, and Assumption 4 hold. If $\sigma = 1$, i.e. the established test has perfect sensitivity, then* $\text{NPV}_z = P(y=0|z=0,t=1)$.

Proof. If $\sigma = 1$, then $\frac{\gamma}{1-\zeta}\frac{1-\sigma}{\sigma} = 0$ and as in the proof of Corollary 1 the relevant case is (I). ■

If there is no perfect reference test available, the negative new test's result leads to ambiguity. Similar to the case of a positive test result, this ambiguity allows for the occurrence of dilation. Using Proposition 3, Corollary 4 provides a characterization of when the set of possible values of $P(x=1|z=0,t=1) = 1-\text{NPV}_z$ contains the prior information $P(x=1|t=1) = \overline{\chi}$.

Corollary 4 (Bounds Increase - NPV)**.** *Suppose Assumption 1, Assumption 2, and Assumption 4 hold. The set of possible values for* $P(x=1|z=0,t=1)$

Covid Economics 66, 28 January 2021: 1-36

includes the prior information $\overline{\chi} = P(x = 1|t = 1)$ *if and only if*

$$\sigma \leq \min\left\{\frac{\gamma\zeta}{P(y = 1, z = 1|t = 1)}, \frac{\gamma(1-\zeta)}{P(y = 1, z = 0|t = 1)}\right\},$$

where the first entry corresponds to an increase in the upper bound, and the second condition ensures the lower bound decreases.

Proof. For the upper bound (of $1 - \mathrm{NPV}_z$) a strict decrease can only happen in cases (C) and (I), i.e. $P(y = 0, z = 0|t = 1) \geq \overline{\chi}(1 - \sigma)$. In these cases, a strict decrease is equivalent to $(1 - \zeta) - [P(y = 0, z = 0|t = 1) - \chi(1 - \sigma)] < \chi(1 - \zeta)$ or $1 - \zeta + \overline{\chi}(\zeta - \sigma) < P(y = 0, z = 0|t = 1) = 1 - \zeta - P(y = 1, z = 0|t = 1)$. Rearranging gives,

$$\sigma > \frac{\zeta}{P(z = 1|y = 1, t = 1)} = \frac{\gamma\zeta}{P(y = 1, z = 1|t = 1)}.$$

As in the proof of Corollary 2 the condition for being in case (C) or (I) is implied by this condition.

For the lower bound, a decrease occurs if

$$\chi(1-\zeta) \geq \\ (1-\zeta) - \min\{1 - \chi, P(y = 0, z = 0|t = 1)\} = \max\{\chi - \zeta, P(y = 1, z = 0|t = 1)\}.$$

First, $\chi - \zeta \leq \chi(1 - \zeta)$ always holds as $\chi \leq 1$. Second, rearranging $P(y = 1, z = 0|t = 1) \leq (1 - \zeta)\frac{\gamma}{\sigma}$ provides the condition. ■

Corollary 2 combined with Corollary 4 provides an exact characterization for when the new test is a dilation. In fact, as the conditions are the same a dilation occurs if and only if

$$\sigma \leq \min\left\{\frac{\gamma\zeta}{P(y = 1, z = 1|t = 1)}, \frac{\gamma(1-\zeta)}{P(y = 1, z = 0|t = 1)}\right\}. \tag{3}$$

When evaluating a new test's accuracy it is important to make sure the data violates Expression (3). Otherwise, the test is uninformative in an extreme sense. In typical applications, the data often satisfies $P(y = 1, z = 0|t = 1) \leq \gamma(1-\zeta)$.[32] test In these cases, a dilation can only occur if $\sigma \leq \frac{\gamma\zeta}{P(y=1,z=1|t=1)}$.

[32] Even data that regards a test as inadequate as in Cassaniti et al. (2020) satisfies this inequality. I thank Filip Obradovic for making me aware of this report.

Covid Economics 66, 28 January 2021: 1-36

The WHO (2020) recommends minimum quality requirements using only information directly provided by the data $P(y, z|t = 1)$. In light of this analysis, an evaluation should also take σ, the established test's sensitivity, into account and with this also make sure that the test is not a dilation. $\sigma \leq \frac{\gamma\zeta}{P(y=1,z=1|t=1)}$ combined with a given minimum standard provides an easy-to-verify sufficient condition to avoid dilation.

For this, let $\underline{\Sigma}$ be a minimum (apparent) sensitivity threshold below which a test is deemed not reliable and denote the new test's apparent sensitivity with $\Sigma = P(z = 1|y = 1, t = 1)$, so that a test is reliable if $\Sigma > \underline{\Sigma}$.[33] Then, the application relevant case from Expression (3) to avoid a dilation can be expressed as $\sigma > \zeta/\Sigma$ or equivalently as $\Sigma > \zeta/\sigma$. If $\underline{\Sigma} \geq \zeta/\sigma$, then any test meeting the minimum requirement cannot be a dilation. Thus, it suffices to make sure the new test's yield is not too high:[34]

$$\zeta := P(z = 1|t = 1) \leq \sigma \times \underline{\Sigma}. \tag{4}$$

If the established test is highly specific, i.e. $\sigma \approx 1$, then Expression (4) is satisfied unless the new test's yield is extremely high.

For SARS-CoV-2 Antigen tests the WHO recommendation is $\underline{\Sigma} = 0.8$ and if the PCR test is not highly specific then Expression (4) might be violated. For example, the dilation test of Table 1 has $\zeta = 0.49$ and if the PCR has sensitivity of $\sigma = 0.6$ then not only Expression (4) is violated but the test is a dilation. More specifically, Expression (3) can be used to find the exact threshold sensitivity σ^* below which a given test turns into a dilation. For the dilation test this value is $\sigma^* = 60.79\%$.

2.3 Informativeness of Additional Testing

Whereas the new test produces ambiguous information, the established test is always informative. Therefore, practitioners might want to perform an additional established test depending on whether a person obtains a negative or positive result from the new test. For example, if an Antigen test is used to detect SARS-

[33] Usually, the minimum requirements include also a threshold for specificity, but this does not matter here.

[34] It is worth recalling that this is a sufficient condition when, additionally, $\frac{\gamma\zeta}{P(y=1,z=1|t=1)} \leq \frac{\gamma(1-\zeta)}{P(y=1,z=0|t=1)}$ and in many applications this inequality becomes irrelevant for Expression (3) because the right-hand side is greater than one.

Covid Economics 66, 28 January 2021: 1-36

CoV-2 and the result is positive, a common practice is verifying the result by means of a PCR test. Since PCR tests are the reference test for evaluating the accuracy of Antigen tests, the current framework can be used to shed light on how informative this additional test is.

Proposition 4 (Combined testing)**.** *Under Assumption 1, Assumption 2, and Assumption 4,*

$$P(x=1|y=1,z=1,t=1)=P(x=1|y=1,z=0,t=1)=1,$$

$$P(x=0|y=0,z=0,t=1)\in\begin{cases}\left[1-\frac{\overline{\chi}(1-\sigma)}{P(y=0,z=0|t=1)},1\right] & \textit{in case (C)}\\ \left[1-\frac{\overline{\chi}(1-\sigma)}{P(y=0,z=0|t=1)},\frac{1-\overline{\chi}}{P(y=0,z=0|t=1)}\right] & \textit{in case (I)}\\ [0,1] & \textit{in case (U)}\\ \left[0,\frac{1-\overline{\chi}}{P(y=0,z=0|t=1)}\right] & \textit{in case (X)},\end{cases}$$

and

$$P(x=0|y=0,z=1,t=1)\in\begin{cases}[0,1] & \textit{in case (C)}\\ \left[0,\frac{1-\overline{\chi}}{P(y=0,z=1|t=1)}\right] & \textit{in case (I)}\\ \left[1-\frac{\overline{\chi}(1-\sigma)}{P(y=0,z=1|t=1)},1\right] & \textit{in case (U)}\\ \left[1-\frac{\overline{\chi}(1-\sigma)}{P(y=0,z=1|t=1)},\frac{1-\overline{\chi}}{P(y=0,z=1|t=1)}\right] & \textit{in case (X)}.\end{cases}$$

Proof. If $y=1$ the the PPV has to be one independent of the new test's result because of Assumption 2. For NPV, again start from Table 20 and Table 21 with $\tau=1$. First, the case of both tests matching, i.e. $y=0=z$:

$$\begin{aligned}\underline{P}(x=0,y=0,z=0|t=1)&=\max\{0,\gamma-\overline{\chi}+P(y=0,z=0|t=1)\}\\&=\max\{0,P(y=0,z=0|t=1)-\overline{\chi}(1-\sigma)\}\\&\text{and}\\ \overline{P}(x=0,y=0,z=0|t=1)&=\min\{1-\overline{\chi},P(y=0,z=0|t=1)\}.\end{aligned}$$

Now, divide by $P(y=0,z=0|t=1)$ to obtain $P(x=0|y=0,z=0,t=1)$.

Covid Economics 66, 28 January 2021: 1-36

In case of differing test results, the relevant probability are:

$$\begin{aligned}\underline{P}(x=0,y=0,z=1|t=1) &= \min\left\{1-\frac{\gamma}{\sigma}, P(y=0,z=1|t=1)\right\}\\&= \min\left\{1-\frac{\gamma}{\sigma}, 1-\gamma-P(y=0,z=0|t=1)\right\}\\&= 1-\gamma-\max\left\{\overline{\chi}(1-\sigma), P(y=0,z=0|t=1)\right\}\\&\text{and}\\\overline{P}(x=0,y=0,z=1|t=1) &= \max\left\{1-\overline{\chi}-P(y=0,z=0|t=1), 0\right\}\\&= \max\left\{P(y=0,z=1|t=1)-\overline{\chi}(1-\sigma), 0\right\}.\end{aligned}$$

Note that $\underline{P}(x=0,y=0,z=1|t=1) \geq \overline{P}(x=0,y=0,z=1|t=1)$ in this case. Division by $P(y=0,z=1|t=1)$ gives $P(x=0|y=0,z=1,t=1)$. ■

Proposition 4 once more reveals that tests in the category (U) and (X) should be avoided. Even if the two test results match and are both negative, the possibility of zero (negative) predictive value cannot be ruled out. Proposition 4 also makes clear the naming convention of the cases defined in Table 3. A confirmatory test (C) provides accurate information when both test produce a negative result, but is completely uninformative if and only if the new test has a positive result. An informative test (I), however, always provides some information in the sense of producing not completely trivial bounds. A contradictory test (X) provides information, but leans against the result of the established test. The uninformative test (U) provides no information at all even when both tests agree on a negative result. Of course, performing the additional test is always informative in the sense of not being a dilation. The established test does not produce false-negatives and therefore a positive result from the established test is always a perfect predictor of being infected regardless of the new test.

2.4 Predictive values for the overall population

Proposition 1 bounds the established test's NPV for the overall population, not only for the tested population. The new test, on the other hand, was analyzed for the testing pool only so far. The full characterization in Appendix A allows to extend the analysis of the new test to make an evaluation for the overall population. Since this involves more cumbersome notation, I only illustrate the resulting bounds for the NPV $= P(x=0|z=0)$. The analysis of PPV would proceed in a similar matter.

Covid Economics 66, 28 January 2021: 1-36

Proposition 5 (Unconditional NPV)**.** *Under Assumption 1–Assumption 4, the new test's (unconditional) negative predictive value is sharply bounded by*

$$
\begin{cases}
\left[\dfrac{1-\overline{\chi}-\tau P(y=0,z=1|t=1)}{1-\tau\zeta}, \dfrac{1-\underline{\chi}}{1-\tau\zeta}\right] & \text{in case } (C^*)\\
\left[\dfrac{1-\overline{\chi}-\tau P(y=0,z=1|t=1)}{1-\tau\zeta}, 1-\dfrac{P(y=1,z=0|t=1)}{1-\tau\zeta}\right] & \text{in case } (I^*)\\
\left[0, \dfrac{1-\underline{\chi}}{1-\tau\zeta}\right] & \text{in case } (U^*)\\
\left[0, 1-\dfrac{P(y=1,z=0|t=1)}{1-\tau\zeta}\right] & \text{in case } (X^*),
\end{cases}
$$

where

$$
\begin{aligned}
(C^*) \quad \ldots \quad & P(y=0,z=0|t=1) \geq \max\left\{1-\gamma-\frac{1-\overline{\chi}}{\tau}, 1-\overline{\chi}\right\}\\
(I^*) \quad \ldots \quad & 1-\overline{\chi} > P(y=0,z=0|t=1) \geq 1-\gamma-\frac{1-\overline{\chi}}{\tau}\\
(U^*) \quad \ldots \quad & 1-\gamma-\frac{1-\overline{\chi}}{\tau} > P(y=0,z=0|t=1) \geq 1-\overline{\chi}\\
(X^*) \quad \ldots \quad & \min\left\{1-\gamma-\frac{1-\overline{\chi}}{\tau}, 1-\overline{\chi}\right\} > P(y=0,z=0|t=1).
\end{aligned}
$$

Proof. From Table 20 and Table 21, we obtain respectively:

$$
\begin{aligned}
\underline{P}(x=0,z=0) &= \max\{0, 1-\overline{\chi}-\tau P(y=0,z=1|t=1)\}\\
\overline{P}(x=0,z=0) &= 1-\tau\max\{\overline{\chi}, 1-P(y=0,z=0|t=1)\}\\
&= 1-\tau\max\{\overline{\chi}, \zeta+P(y=1,z=0|t=1)\}.
\end{aligned}
$$

Now, the result follows from dividing by $P(z=0)=1-\tau\zeta$. ∎

If instead of predictive values the interest lies in the new test's sensitivity or specificity in the whole population another complication arises. Conditional on the testing pool, both of these measures can be derived as in Subsection 2.2. For example, for sensitivity one could use the proof of Proposition 2 and divide by $P(x=1|t=1)=\overline{\chi}$ instead of $P(z=1|t=1)=\zeta$. The bounds for sensitivity are again determined by considering the extremes of Table 20 and Table 21. For the unconditional sensitivity, however, the the numerator and the denominator are both set identified because $P(x=1)\in[\underline{\chi},\overline{\chi}]$. Therefore, the lower bound might not be attained at either of the extreme distributions. This makes solving for

Covid Economics 66, 28 January 2021: 1-36

a closed-form expression for sensitivity intractable. Nonetheless, the bounds can easily be obtained computationally by considering a fixed $\Gamma := P(y = 1) \in [\tau\gamma, \gamma]$ with corresponding $p = \Gamma/\sigma$. For this Γ, sharp bounds of sensitivity, say $[L_\Gamma, H_\Gamma]$, can be obtained by using Table 18 and Table 19. To find the overall bounds for sensitivity, two (non-linear) optimization problems across all values of Γ need to be performed to give $[\min_\Gamma L_\Gamma, \max_\Gamma H_\Gamma]$.

3 Applications

In this section, the theoretic framework will be illustrated with several applications. First, I analyze the (hypothetical) dilation test presented in the introduction. Then, I examine two real SARS-CoV-2 detecting tests. Finally, I show that CT-scanning procedures to detect COVID-19 are prone to being dilations.

3.1 A Dilation Test

As argued before the hypothetical test data in Table 1 corresponds to a dilation. Suppose the test data is derived for an Antigen test to detect SARS-CoV-2 and the reference test is a PCR test.[35]. The test satisfies the WHO's (2020) minimum requirements with apparent sensitivity ($\Sigma = 80.6\%$) and specificity (97.1%) above the specified thresholds of 80% and 97%, respectively.[36] For such a setting the current framework is applicable. Especially, Assumption 2 seems to be warranted because a PCR test is highly specific. However, it is known that a PCR test might lack high sensitivity. Alcoba-Florez et al. (2020) report sensitivity for several PCR tests with point estimates ranging from $\sigma = 60.2\%$ to $\sigma = 97.9\%$.[37] All of the 95% confidence intervals exclude perfect sensitivity, $\sigma = 1$.

Using the results from Subsection 2.2, Table 4 summarizes some key statistics for the dilation test. When the PCR sensitivity is close to one, the new (hypothetical) test produces relative accurate measurements with PPV close to one and NPV above 75%. However, if the PCR test lacks high sensitivity then we cannot be sure of the dilation test's quality. In the worst-case for PCR sensitivity ($\sigma = 0.6$), the new test is indeed a dilation: Before a test result was obtained the prevalence (in

[35] Recall that for SARS-CoV-2 detection a PCR test is the established test used to evaluate other tests. (Esbin et al., 2020)

[36] These numbers are calculated as if the reference test is perfect. This is similar to Corollary 1 and Corollary 3.

[37] Alcoba-Florez et al. differentiate values based on the targeted gene. The range reported here is across all genes and tests.

Covid Economics 66, 28 January 2021: 1-36

the testing pool) is 98.8%, after obtaining *either* dilation test's result the possible probability of being infected is at least the interval $[97.7\%, 100\%]$. In fact, potentially even more puzzling is that the lowest value after a negative test is strictly higher than after a positive result. Using Expression (3), $\sigma^* = 60.8\%$ represents the cutoff PCR sensitivity below which a dilation occurs.

Table 4: Dilation Test Statistics

σ	0.6	0.85	0.98	1
$\overline{\chi} = P(x = 1 \mid t = 1)$	98.8%	69.8%	60.5%	59.3%
PPV_z	$[97.6\%, 100\%]$	$[97.6\%, 100\%]$	$[97.6\%, 100\%]$	97.6%
$1 - \text{NPV}_z$	$[97.7\%, 100\%]$	$[40.7\%, 43.1\%]$	$[22.6\%, 24.9\%]$	22.6%
NPV_z	$[0\%, 2.29\%]$	$[56.9\%, 59.3\%]$	$[75.1\%, 77.43\%]$	77.43%
Dilation Threshold		$\sigma^* = 60.8\%$		

3.2 Standard Q COVID-19 Rapid Antigen Test

Next, consider the Standard Q (*StQ*) COVID-19 Rapid Antigen Test of SD Biosensor/Roche for detection of SARS-CoV-2 as analyzed by Kaiser et al. (2020). They use results of PCR tests as comparison (see Footnote 35). The testing data is summarized in Table 5. When $\sigma = 1$, then StQ's PPV and NPV are obtained

Table 5: StQ Test results from Kaiser et al. (2020, p. 3)

$z \backslash y$	$y = 0$	$y = 1$	Sum
$z = 0$	63.71%	3.97%	67.67%
$z = 1$	0.19%	32.14%	32.33%
Sum	63.89%	36.11%	

with Corollary 1 and Corollary 3 which yields 99.42% and 94.13%, respectively. These are the reported values of Kaiser et al. (2020). However, as explained above PCR are not perfectly sensitive.[38] Thus, to evaluate the StQ test the current framework is applicable.

Focusing first on the testing pool only, Table 6 summarizes PPV and NPV for different values of PCR sensitivity (σ) using Proposition 2 and Proposition 3. Even if the PCR test lacks high sensitivity, StQ has a close to perfect positive

[38] Kaiser et al. (2020) use PCR tests targeting E genes, which tend to have higher sensitivity in the analysis of Alcoba-Florez et al. (2020). The lowest reported sensitivity for a PCR test targeting E genes is 65.33%.

predicative value ($\text{PPV}_z \approx 1$). However, the values for NPV drop significantly as σ decreases. In the worst case, a negative StQ result becomes close to a fair coin flip. However, the test is very informative overall as can be seen by the low dilation threshold $\sigma^* = 36.3\%$.

Table 6: Accuracy of StQ

σ	0.6	0.85	0.98
PPV_z	$[99.4\%, 100\%]$	$[99.4\%, 100\%]$	$[99.4\%, 100\%]$
NPV_z	$[58.6\%, 58.9\%]$	$[84.7\%, 85\%]$	$[93.1\%, 94.1\%]$
Dilation Threshold		$\sigma^* = 36.3\%$	

Kaiser et al. (2020, p. 1) state "study participants were representative of the usual population seeking testing in our center (main testing center in Geneva). The majority were presenting with symptoms compatible with a SARS-CoV2 infection and a minority were asymptomatic but with a known positive contact or were asymptomatic healthcare workers." The current framework allows to use the obtained testing data to evaluate StQ's quality for the overall population (of Geneva) as analyzed in Subsection 2.4. Furthermore, this explanation supports Assumption 3.

Table 7 shows bounds on prevalence using the baseline analysis in Section 2. For low values of τ, i.e. the testing pool was highly non-representative of the overall population, the width of the intervals is rather wide. However, even the lowest number is close to 2% indicating a thorough spread of the virus in Geneva at the time of testing.[39] When testing becomes representative ($\tau \to 1$) the prevalence converges to the prevalence in the testing pool.

Table 7: Bounds on prevalence $p \in [\underline{\chi}, \overline{\chi}]$

$\tau \diagdown \sigma$	0.6	0.85	0.98
$1/20$	$[3.01\%, 60.2\%]$	$[2.12\%, 42.5\%]$	$[1.84\%, 36.8\%]$
$1/10$	$[6.02\%, 60.2\%]$	$[4.25\%, 42.5\%]$	$[3.68\%, 36.8\%]$
$1/2$	$[30.1\%, 60.2\%]$	$[21.2\%, 42.5\%]$	$[18.4\%, 36.8\%]$
$19/20$	$[57.2\%, 60.2\%]$	$[40.4\%, 42.5\%]$	$[35.7\%, 36.8\%]$
1	60.2%	42.5%	36.8%

[39] Note that this is a one time analysis. It does not answer the question of how many people were cumulatively infected by SARS-CoV-2 up to the time of testing.

Covid Economics 66, 28 January 2021: 1-36

At this time, if a Genevese obtains a negative PCR result, what is the probability of her being infected? If testing is not competently representative, a unique number cannot be given. However, Proposition 1 provides sharp bounds for this case and the results are shown in Table 8. How do these PCR results compare to

Table 8: Bounds on NPV_y

$\tau \diagdown \sigma$	0.6	0.85	0.98
$^1/_{20}$	[62.3%, 98.8%]	[90.0%, 99.7%]	[98.9%, 100%]
$^1/_{10}$	[62.3%, 97.5%]	[90.0%, 99.3%]	[98.9%, 99.9%]
$^1/_2$	[62.3%, 85.3%]	[90.0%, 96.1%]	[98.9%, 99.6%]
$^{19}/_{20}$	[62.3%, 65.2%]	[90.0%, 90.8%]	[98.9%, 98.9%]
1	62.3%	90.0%	98.9%

results from StQ? Table 9 provides the numbers using Proposition 5. The lower bounds are significantly lower than for the PCR test. This makes the width of the interval also significantly wider. The widening is a reflection of the combination of the two missing data problems inherit in the testing procedure without a perfect reference test: (i) unknown overall prevalence (which also affects PCR's NPV) and (ii) missing correlation data (which does not affect the PCR's NPV).

Table 9: Bounds on STQ's NPV_z for overall population

$\tau \diagdown \sigma$	0.6	0.85	0.98
$^1/_{20}$	[40.5%, 96.0%]	[58.5%, 96.0%]	[64.2%, 96.0%]
$^1/_{10}$	[41.1%, 95.9%]	[59.4%, 95.9%]	[65.3%, 95.9%]
$^1/_2$	[47.4%, 83.4%]	[68.5%, 94.0%]	[75.2%, 95.2%]
$^{19}/_{20}$	[57.2%, 61.8%]	[82.8%, 86.1%]	[90.9%, 93.8%]

3.3 BiaxNOW Covid-19 Ag Home Test

The BiaxNOW (*BiN*) Covid-19 Ag Home Test of Abbott is one of the first rapid Antigen tests for use at home which is able to detect the SARS-CoV-2 virus that was emergency approved by the FDA (2020a). BiN's clinical performance for approval by the FDA was conducted with a PCR test as a reference. The data are shown in Table 10 and Table 11 shows the implied accuracy measures.

Relative to StQ, BiN has significantly lower PPV and also rules out perfect PPV for lower values of PCR specificity. On the other hand, NPV is uniformly

Covid Economics 66, 28 January 2021: 1-36

Table 10: BiN Test results from FDA (2020a, p. 20)

$z \backslash y$	$y = 0$	$y = 1$	Sum
$z = 0$	73.48%	3.91%	77.39%
$z = 1$	1.09%	21.52%	22.61%
Sum	74.57%	25.43%	

Table 11: Accuracy of BiN

σ	0.6	0.85	0.98
PPV_z	$[95.2\%, 100\%]$	$[95.2\%, 100\%]$	$[95.2\%, 97.5\%]$
NPV_z	$[73.0\%, 74.4\%]$	$[89.1\%, 90.6\%]$	$[94.3\%, 94.9\%]$
Dilation Threshold		$\sigma^* = 26.7\%$	

greater for BiN compared to StQ. Even for the worst-case PCR sensitivity, BiN's possible NPV values are reasonably high. Furthermore, the dilation threshold is extremely low at $\sigma^* = 26.7\%$.

The FDA (2020a) also provides additional data about BiN results by including the cycle threshold obtained by the PCR test.[40] Table 12 shows this data. This

Table 12: BiN Test results from FDA (2020a, p. 21) with Cycle Threshold Count

$z \backslash y$	$Ct < 33$	$Ct \geq 33$	Sum
$z = 0$	16.2%	6.94%	23.12%
$z = 1$	9.83%	67.05%	76.89%
Sum	26.01%	73.99%	

is additional data a PCR test produces, which can be used to refine bounds on predictive values of the Antigen test. However, an extension of the current setting is needed, because such additional information is not accounted for in the current setting with binary tests. Subsection 4.2 discusses a possible extension of the current setting to allow for this additional data.

[40]Public Health England (2020) explains: "Cycle threshold (Ct) is a semi-quantitative value that can broadly categorise the concentration of viral genetic material in a patient sample following testing by RT PCR as low, medium or high — that is, it tells us approximately how much viral genetic material is in the sample. A low Ct indicates a high concentration of viral genetic material, which is typically associated with high risk of infectivity. A high Ct indicates a low concentration of viral genetic material which is typically associated with a lower risk of infectivity."

Covid Economics 66, 28 January 2021: 1-36

3.4 CT scan to detect COVID-19

Ai et al. (2020) and Gietema et al. (2020) propose using chest CT scans for early identifying COVID-19 in patients. In Gietema et al.'s study, all COVID-19 symptomatic patients of a single Dutch emergency department have a chest CT scan and a PCR test for detecting SARS-CoV-2. Their study design exactly fits the framework of the current paper: (i) non-representative sampling of the testing pool and (ii) missing correlation data due to use of an imperfect reference test (with perfect specificity). The testing data are reproduced in Table 13. Compared

Table 13: CT scan data from Gietema et al. (2020, Table 2)

$z \backslash y$	$y = 0$	$y = 1$	Sum
$z = 0$	38.86%	4.66%	43.52%
$z = 1$	18.13%	38.34%	56.48%
Sum	56.99%	43.01%	

to the previously studies Antigen tests, the data for CT scans seems less aligned with the PCR test results. This is an indication that such a CT test is less informative: the dilation threshold of $\sigma^* = 63.35\%$ is higher than for the Antigen tests. Thus, this testing procedure is completely uninformative if $\sigma = 60\%$—the lowest sensitivity of a PCR test reported by Alcoba-Florez et al. (2020). In this case, Table 14 shows (sharp bounds on) population prevalence, PCR NPVs, and CT scan NPVs. The PCR's (assumed) low sensitivity means that its NPV might

Table 14: Implications for overall population from Gietema et al. (2020).

$\tau\downarrow$	$p := P(x = 1)$	NPV_y	NPV_z
$^1/_{20}$	$[3.58\%, 71.7\%]$	$[49.7\%, 98.5\%]$	$[23.4\%, 65.1\%]$
$^1/_{10}$	$[7.17\%, 71.7\%]$	$[49.7\%, 97.0\%]$	$[23.4\%, 65.1\%]$
$^1/_2$	$[35.8\%, 71.7\%]$	$[49.7\%, 81.7\%]$	$[23.4\%, 65.1\%]$
$^{19}/_{20}$	$[68.1\%, 71.7\%]$	$[49.7\%, 54.0\%]$	$[23.4\%, 65.1\%]$

be quite low, but it as at least close to 50%, irrespective of τ. On the other hand, the CT scan has both a sizable interval and a low lower bound of possible NPVs. Since $\sigma = 0.6$ is below the dilation threshold, the CT scan is completely uninformative for the tested population (equivalently for $\tau = 1$). Furthermore, this remains true for the overall population if $\tau \geq 1/2$ as shown in Table 14. For exam-

Covid Economics 66, 28 January 2021: 1-36

Covid Economics 66, 28 January 2021: 1-36

ple, for a non-COVID-indicative CT scan the set of possible infection probabilities $P(x = 1|z = 0)$ increases relative to the prior information p.[41]

Even more striking is the data of Ai et al. (2020) shown in Table 15.[42] Ai et al. also use Chest CT scans to test for COVID-19. Their data is obtained in Wuhan, China and like the study of Gietema et al. (2020) a PCR test is used as a reference. The data reveals a very low apparent specificity but a high apparent sensitivity of $\Sigma = 96.51\%$. Expression (4) indicates that a high yield of the new test, ζ, might be problematic. Here, this yield is very high with $\zeta = 87.57\%$. The problem becomes even more apparent by looking at the dilation threshold, which is high with a value of $\sigma^* = 90.74$. This implies that even if the PCR test is quite sensitive, the CT scan is completely uninformative for the tested people in Wuhan.

Table 15: CT scan data from Ai et al. (2020, Table 2)

$z\backslash y$	$y = 0$	$y = 1$	Sum
$z = 0$	10.36%	2.07%	12.43%
$z = 1$	30.37%	57.20%	87.57%
Sum	40.73%	59.27%	

4 Discussion and Extension

4.1 Imperfect Specificity of the Established Test and the Use of Additional Data

Assumption 2 might be too strong for some applications. Although, this assumption simplifies the algebraic expression sometimes significantly, it is not a crucial assumption conceptually. The crucial characterization of joint distributions in Appendix A can easily be extended to allow for false-positives of the established test.

Evaluations of a new test sometimes have more data available than just $P(y, z|t = 1)$. For example, blood samples from before the existence of a virus can serve as true-negative samples. On the other hand, specific blood samples could be analyzed with more sophisticated (and usually much more expensive) methods than just using an established test as reference. These methods would lead to samples

[41] Recall that $P(x = 1|z = 0) = 1 - \text{NPV}_z$.
[42] I thank Filip Obradovic for providing this reference.

with true positives (or at least with very high probability).[43] Either of these methods would be provide additional data and therefore would also reduce the missing data problem. In general, this supplementary knowledge leads to narrower bounds, but unless these extra methods are performed for the whole tested population, the missing correlation issues remains. Of course, these methods cannot be applied for the untested population. Therefore the the missing data on prevalence cannot be avoided with these extraneous data.

4.2 Beyond Binary Tests

The current framework only allows for binary outcomes for both tests and also for the underlying health state. This seems to be the most common situation studied in the literature on diagnostic testing. (Zhou et al., 2014) Often tests provide ternary results (with the additional result of ‘inconclusive’ or ‘invalid’), or allow for even more detailed information, like the Cycle Threshold Count of a PCR test as mentioned in Subsection 3.3. In such situations, the theoretic analysis does not provide the appropriate machinery. However, the crucial application to characterize the set of all joint distribution is a result in copula theory (Joe, 1997, Theorem 3.10), which does not rely on any dimension being binary. Indeed, the result even works for continuous outcomes on each dimension.

Similarly, one could use other results in Joe (1997) to characterize the set of possible joint distributions if multiple tests are conducted simultaneously as studied in Zhou et al. (2014, Chapter 9). In this case, and like in the characterization of Appendix A, the testing data are higher-dimensional marginal distribution of an overall joint distribution with an additional dimension (the health state). When considering such an extension, a caution has to be taken because sometimes sharp bounds on the set of possible higher-dimensional distributions may not be known.

4.3 COVID-19 related testing

Since testing has a potentially big impact on the economy, an accurate description of the available testing technology is crucial. From the microeconomic perspective, the testing technology affects how test should be optimally allocated (see Ely et al. (2020), Lipnowski and Ravid (2020)) and also how much people engage in social distancing as studied by Acemoglu et al. (2020). But also the macroeconomy

[43]For example, Olbrich et al. (2020) combine these methods to evaluate SARS-CoV-2 antibody tests.

is highly affected by testing strategies and an optimal choice might reduce the economic costs of pandemics considerably. (Alvarez et al., 2020; Eichenbaum et al., 2020) Although, these studies establish the importance of testing and also address varying testing technologies, all of them assume that a test corresponds to an experiment à la Blackwell (1951) and therefore is always informative (sometimes the assumption is even that the test itself provides perfect information).

This paper demonstrates that the assumption of unambiguous information in test results is only applicable if a perfect reference is available when evaluating new tests. In particular, new Antigen test for detection of SARS-CoV-2 are evaluated relative to an imperfect PCR test and therefore—as shown in this paper—these Antigen test produce ambiguous information. An optimal testing procedure should take this ambiguity into account. Similarly, practitioner guides (like Galeotti et al. (2020); Watson et al. (2020)) might want to consider addressing uncertainty in test results in more detail.

A Characterization of the set of joint distributions $P(x, y, z)$

Consider a fixed $\Gamma := P(y = 1) \in [\tau\gamma, \gamma]$ with corresponding prevalence $p := P(x = 1) = \Gamma/\sigma$.

Lemma 1. *Suppose Assumption 1 and Assumption 2 hold. For a given* Γ*,* $P(x, y)$ *is given by Table 16.*

Table 16: Joint distribution of $P(x, y)$

$P(x\backslash y)$	$y = 0$	$y = 1$	
$x = 0$	$1 - {}^{\Gamma}/_{\sigma}$	0	$1 - {}^{\Gamma}/_{\sigma}$
$x = 1$	$\Gamma\frac{1-\sigma}{\sigma}$	Γ	${}^{\Gamma}/_{\sigma}$
	$1 - \Gamma$	Γ	

Proof. 1. $P(x = 1, y = 1) = P(y = 1) - \underbrace{P(x = 0, y = 1)}_{=0 \text{ by Assumption 2}} = \Gamma$

2. $P(x = 1, y = 0) = P(x = 1) - P(x = 1, y = 1) = \Gamma(1 - \sigma)/\sigma.$

3. $P(x = 0, y = 0) = P(y = 0) - P(x = 1, y = 0) = 1 - \Gamma/\sigma.$

■

Using the law of total probability and rearranging gives $P(y = 1|t = 0) = \frac{\Gamma - \tau\gamma}{1-\tau}$. Furthermore, by the nature of testing $P(y = 0, z = 1|t = 0) = P(y = 1, z = 1|t = 0) = 0$. Therefore, $P(y, z)$ is given by Table 17.

Table 17: Joint distribution of $P(y, z)$

$P(z\backslash y)$	$y = 0$	$y = 1$	
$z = 0$	$1 - \Gamma - P(y = 0, z = 1\|t = 1)\tau$	$\Gamma - P(y = 1, z = 1\|t = 1)\tau$	$1 - \tau\zeta$
$z = 1$	$P(y = 0, z = 1\|t = 1)\tau$	$P(y = 1, z = 1\|t = 1)\tau$	$\tau\zeta$
	$1 - \Gamma$	Γ	

By Joe (1997, Theorem 3.10) the set of all distributions $P(x, y, z)$ with marginals given by $P(x, y)$ and $P(y, z)$ is bounded by two extreme distributions:[44] $\underline{F}_\Gamma \leq F \leq \overline{F}_\Gamma$, where F is the CDF corresponding to $P(x, y, z)$, $\underline{F}_\Gamma$ is given by Table 18, and $\overline{F}_\Gamma$ is given by Table 19.

Covid Economics 66, 28 January 2021: 1-36

Table 18: CDF $\underline{F}_\Gamma$

	$x=1$		$x=0$	
	$z=1$	$z=0$	$z=1$	$z=0$
$y=1$	1	$1-\tau\zeta$	$1-\frac{\Gamma}{\sigma}$	$\max\{0, 1-\Gamma/\sigma - P(y=0,z=1\|t=1)\tau\}$
$y=0$	$1-\Gamma$	$1-\Gamma-P(y=0,z=1\|t=1)\tau$	$1-\frac{\Gamma}{\sigma}$	$\max\{0, 1-\Gamma/\sigma - P(y=0,z=1\|t=1)\tau\}$

Table 19: CDF $\overline{F}_\Gamma$

	$x=1$		$x=0$	
	$z=1$	$z=0$	$z=1$	$z=0$
$y=1$	1	$1-\tau\zeta$	$1-\frac{\Gamma}{\sigma}$	$\min\left\{1-\frac{\Gamma}{\sigma}, 1-\Gamma-P(y=0,z=1\|t=1)\tau\right\}$
$y=0$	$1-\Gamma$	$1-\Gamma-P(y=0,z=1\|t=1)\tau$	$1-\frac{\Gamma}{\sigma}$	$\min\left\{1-\frac{\Gamma}{\sigma}, 1-\Gamma-P(y=0,z=1\|t=1)\tau\right\}$

Since $\underline{F}_\Gamma$ and $\overline{F}_\Gamma$ are both nonincreasing in Γ, sharp bounds for the CDF F across all $\Gamma := P(y=1) \in [\tau\gamma, \gamma]$ are $\underline{F} := \underline{F}_\gamma \leq F \leq \overline{F}_{\tau\gamma} =: \overline{F}$. For the lower, we have $1 - \Gamma/\sigma - P(y=0, z=1|t=1) = (1-\gamma)(1-\tau) + \tau$.

For the upper bound, note that $1 - \tau\gamma - P(y=0, z=1|t=1)\tau = 1 - \tau(1 - P(y=0, z=0|t=1))$. The corresponding probability mass functions are given by Table 20 and Table 21.

Table 20: Lower bound PMF with $\mathcal{P}_{01} := P(y=0, z=1|t=1)$

	$x=1$		$x=0$	
	$z=1$	$z=0$	$z=1$	$z=0$
$y=1$	$P(y=1,z=1\|t=1)\tau$	$\gamma - P(y=1,z=1\|t=1)\tau$	0	0
$y=0$	$\max\left\{0, \frac{\gamma}{\sigma} + \mathcal{P}_{01}\tau - 1\right\}$	$\min\left\{\gamma\frac{1-\sigma}{\sigma}, 1-\gamma-\mathcal{P}_{01}\tau\right\}$	$\min\left\{1-\frac{\gamma}{\sigma}, \mathcal{P}_{01}\tau\right\}$	$\max\left\{0, 1-\frac{\gamma}{\sigma} - \mathcal{P}_{01}\tau\right\}$

Table 21: Upper bound PMF with $\mathcal{P}_{00} := P(y=0, z=0|t=1)$

	$x=1$		$x=0$	
	$z=1$	$z=0$	$z=1$	$z=0$
$y=1$	$P(y=1,z=1\|t=1)\tau$	$P(y=1,z=0\|t=1)\tau$	0	0
$y=0$	$\tau\left[\min\left\{\frac{\gamma}{\sigma}, 1-\mathcal{P}_{00}\right\} - \gamma\right]$	$\tau\max\left\{\mathcal{P}_{00} + \frac{\gamma}{\sigma} - 1, 0\right\}$	$\tau\max\left\{1-\frac{\gamma}{\sigma} - \mathcal{P}_{00}, 0\right\}$	$1 - \tau\max\left\{\frac{\gamma}{\sigma}, 1-\mathcal{P}_{00}\right\}$

References

Acemoglu, D., A. Makhdoumi, A. Malekian, and A. Ozdaglar (2020): "Testing, Voluntary Social Distancing and the Spread of an Infection," Tech. Rep. w27483, National Bureau of Economic Research, https://www.nber.org/papers/w27483.

[44]The set of all these distributions, often called Fréchet class, is a convex set. Hence, here it suffices to consider the extreme points only.

Ai, T., Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, and L. Xia (2020): "Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases," *Radiology*, 296, E32–E40, https://pubs.rsna.org/doi/10.1148/radiol.2020200642.

Alcoba-Florez, J., H. Gil-Campesino, D. G.-M. de Artola, R. González-Montelongo, A. Valenzuela-Fernández, L. Ciuffreda, and C. Flores (2020): "Sensitivity of Different RT-qPCR Solutions for SARS-CoV-2 Detection," *International Journal of Infectious Diseases*, 99, 190–192, http://www.sciencedirect.com/science/article/pii/S1201971220306032.

Alvarez, F. E., D. Argente, and F. Lippi (2020): "A Simple Planning Problem for COVID-19 Lockdown," *Covid Economics*, 14, 1–32, https://cepr.org/sites/default/files/CovidEconomics14.pdf.

Blackwell, D. (1951): "Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, The Regents of the University of California, https://projecteuclid.org/euclid.bsmsp/1200500222.

——— (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272, https://projecteuclid.org/euclid.aoms/1177729032.

Cassaniti, I., F. Novazzi, F. Giardina, F. Salinaro, M. Sachs, S. Perlini, R. Bruno, F. Mojoli, and F. Baldanti (2020): "Performance of VivaDiag COVID-19 IgM/IgG Rapid Test Is Inadequate for Diagnosis of COVID-19 in Acute Patients Referring to Emergency Room Department," *Journal of Medical Virology*, 92, 1724–1727, https://onlinelibrary.wiley.com/doi/abs/10.1002/jmv.25800.

de Oliveira, H. (2018): "Blackwell's Informativeness Theorem Using Diagrams," *Games and Economic Behavior*, 109, 126–131, http://www.sciencedirect.com/science/article/pii/S0899825617302270.

Eichenbaum, M. S., S. Rebelo, and M. Trabandt (2020): "The Macroeconomics of Testing and Quarantining," Tech. Rep. w27104, National Bureau of Economic Research, https://www.nber.org/papers/w27104.

Covid Economics 66, 28 January 2021: 1-36

Ely, J., A. Galeotti, O. Jann, and J. Steiner (2020): "Optimal Test Allocation," Tech. rep., http://home.cerge-ei.cz/steiner/allocation.pdf.

Esbin, M. N., O. N. Whitney, S. Chong, A. Maurer, X. Darzacq, and R. Tjian (2020): "Overcoming the Bottleneck to Widespread Testing: A Rapid Review of Nucleic Acid Testing Approaches for COVID-19 Detection," *RNA*, 26, 771–783, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7297120/.

FDA (2020a): "BinaxNOW COVID-19 Ag Card Home Test: Healthcare Provider Instructions for Use," *FDA*, https://www.fda.gov/media/144574/download.

——— (2020b): "EUA Authorized Serology Test Performance," *FDA*, https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices/eua-authorized-serology-test-performance.

Galeotti, A., J. Steiner, and P. Surico (2020): "Merit of Test: Perspective of Information Economics," *Health Policy and Technology*, 9, 575–577, http://www.sciencedirect.com/science/article/pii/S2211883720300848.

Gietema, H. A., N. Zelis, J. M. Nobel, L. J. G. Lambriks, L. B. van Alphen, A. M. L. O. Lashof, J. E. Wildberger, I. C. Nelissen, and P. M. Stassen (2020): "CT in Relation to RT-PCR in Diagnosing COVID-19 in The Netherlands: A Prospective Study," *PLOS ONE*, 15, e0235844, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0235844.

Gul, F. and W. Pesendorfer (2018): "Evaluating Ambiguous Random Variables and Updating by Proxy," https://www.princeton.edu/~fgul/updating.pdf.

Joe, H. (1997): *Multivariate Models and Multivariate Dependence Concepts*, CRC Press.

Kaiser, L., I. Eckerle, M. Schibler, and A. Berger (2020): "Validation Report: SARS-CoV-2 Antigen Rapid Diagnostic Test," Tech. rep., https://www.hug.ch/sites/interhug/files/structures/laboratoire_de_virologie/documents/Centre_maladies_virales_infectieuses/ofsp_rdt_report_gcevd_27.10.2020.pdf.

Covid Economics 66, 28 January 2021: 1-36

Lipnowski, E. and D. Ravid (2020): "Pooled Testing for Quarantine Decisions," SSRN Scholarly Paper ID 3633360, Social Science Research Network, Rochester, NY, https://papers.ssrn.com/abstract=3633360.

Machina, M. J. and M. Siniscalchi (2014): "Chapter 13 - Ambiguity and Ambiguity Aversion," in *Handbook of the Economics of Risk and Uncertainty*, ed. by M. Machina and K. Viscusi, North-Holland, vol. 1 of *Handbook of the Economics of Risk and Uncertainty*, 729–807, http://www.sciencedirect.com/science/article/pii/B9780444536853000131.

Manski, C. F. (1989): "Anatomy of the Selection Problem," *The Journal of Human Resources*, 24, 343–360, https://www.jstor.org/stable/145818.

——— (2018): "Credible Ecological Inference for Medical Decisions with Personalized Risk Assessment," *Quantitative Economics*, 9, 541–569, https://onlinelibrary.wiley.com/doi/abs/10.3982/QE778.

——— (2020): "Bounding the Accuracy of Diagnostic Tests, with Application to COVID-19 Antibody Tests," .

Manski, C. F. and F. Molinari (2021): "Estimating the COVID-19 Infection Rate: Anatomy of an Inference Problem," *Journal of Econometrics*, 220, 181–192, http://www.sciencedirect.com/science/article/pii/S0304407620301676.

Olbrich, L., N. Castelletti, Y. Schälte, M. Garí, P. Pütz, A. Bakuli, M. Pritsch, I. Kroidl, E. Saathoff, J. M. Guggenbuehl Noller, V. Fingerle, R. Le Gleu, L. Gilberg, I. Brand, P. Falk, A. Markgraf, F. Deák, F. Riess, M. Diefenbach, T. Eser, F. Weinauer, S. Martin, E.-M. Quenzel, M. Becker, J. Durner, P. Girl, K. Müller, K. Radon, C. Fuchs, R. Wölfel, J. Hasenauer, M. Hoelscher, and A. Wieser (2020): "A Serology Strategy for Epidemiological Studies Based on the Comparison of the Performance of Seven Different Test Systems - The Representative COVID-19 Cohort Munich," http://www.klinikum.uni-muenchen.de/Abteilung-fuer-Infektions-und-Tropenmedizin/download/de/KoCo19/KoCo19_Olbrichetal_Serostrategy_SARS-CoV-2.pdf.

Covid Economics 66, 28 January 2021: 1-36

Public Health England (2020): "Understanding Cycle Threshold (Ct) in SARS-CoV-2 RT-PCR," Tech. rep., Department of Health and Social Care, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/926410/Understanding_Cycle_Threshold__Ct__in_SARS-CoV-2_RT-PCR_.pdf.

Reitsma, J. B., A. W. S. Rutjes, K. S. Khan, A. Coomarasamy, and P. M. Bossuyt (2009): "A Review of Solutions for Diagnostic Accuracy Studies with an Imperfect or Missing Reference Standard," *Journal of Clinical Epidemiology*, 62, 797–806, http://www.sciencedirect.com/science/article/pii/S0895435609000638.

Sacks, D. W., N. Menachemi, P. Embi, and C. Wing (2020): "What Can We Learn about SARS-CoV-2 Prevalence from Testing and Hospital Data?" *arXiv:2008.00298 [econ, stat]*, http://arxiv.org/abs/2008.00298.

Seidenfeld, T. and L. Wasserman (1993): "Dilation for Sets of Probabilities," *Annals of Statistics*, 21, 1139–1154, https://projecteuclid.org/euclid.aos/1176349254.

Shishkin, D. and P. Ortoleva (2020): "Ambiguous Information and Dilation: An Experiment," http://denisshishkin.com/papers/ambiguous_information.pdf.

Stoye, J. (2020): "Bounding Disease Prevalence by Bounding Selectivity and Accuracy of Tests: The Case of COVID-19," *arXiv:2008.06178 [econ, stat]*, http://arxiv.org/abs/2008.06178.

Watson, J., P. F. Whiting, and J. E. Brush (2020): "Interpreting a Covid-19 Test Result," *BMJ*, 369, m1808, https://www.bmj.com/content/369/bmj.m1808.

WHO (2020): "Antigen-Detection in the Diagnosis of SARS-CoV-2 Infection Using Rapid Immunoassays," https://www.who.int/publications-detail-redirect/antigen-detection-in-the-diagnosis-of-sars-cov-2infection-using-rapid-immunoassa

Zhou, X.-H., N. A. Obuchowski, and D. K. McClish (2014): *Statistical Methods in Diagnostic Medicine*, John Wiley & Sons.

Covid Economics 66, 28 January 2021: 1-36

Protecting lives and livelihoods with early and tight lockdowns[1]

Francesca Caselli,[2] Francesco Grigoli[3] and Damiano Sandri[4]

Date submitted: 20 January 2021; Date accepted: 20 January 2021

Using high-frequency proxies for economic activity over a large sample of countries, we show that the economic crisis during the first seven months of the COVID-19 pandemic was only partly due to government lockdowns. Economic activity also contracted severely because of voluntary social distancing in response to higher infections. Furthermore, we show that lockdowns substantially reduced COVID-19 cases, especially if they were introduced early in a country's epidemic. This implies that, despite involving short-term economic costs, lockdowns may pave the way to a faster recovery by containing the spread of the virus and reducing voluntary social distancing. Finally, we document that lockdowns entail decreasing marginal economic costs but increasing marginal benefits in reducing infections. This suggests that tight short-lived lockdowns are preferable to mild prolonged measures.

Covid Economics 66, 28 January 2021: 37-57

1 The views expressed in this working paper are those of the authors and do not necessarily represent those of the IMF, its Executive Board, or its management. Working papers describe research in progress by the authors and are published to elicit comments and to encourage debate. We thank Jörg Decressin, Gabriel Di Bella, Gian Maria Milesi Ferretti, Gita Gopinath, Yuriy Gorodnichenko, Òscar Jordà, Malhar Nabar, Mikkel Plagborg-Møller, and Antonio Spilimbergo for insightful comments.
2 Economist, IMF.
3 Economist, IMF.
4 Deputy Division Chief, IMF and CEPR Research Fellow.

CEPR PRESS

Covid Economics 66, 28 January 2021: 37-57

1 Introduction

The COVID-19 pandemic has raised unprecedented health challenges on a global scale. To contain the spread of the virus in the spring of 2020, most countries have resorted to stringent lockdown measures, for example closing schools and business activities and sometimes even preventing people from leaving their homes except for essential reasons. The resurgence of COVID-19 cases in the fall of 2020 rekindled the debate about the desirability of lockdown measures. The discussion is often based on the notion that lockdowns entail a trade-off before protecting lives and supporting the economy. In this paper, we revisit this prevailing narrative by examining the economic and epidemiological developments across a large set of countries during the first seven months of the COVID-19 pandemic.

We begin the analysis by examining which factors drove the economic contraction over a large panel of advanced, emerging, and low-income countries. We document that lockdowns contributed substantially to the drop in economic activity while they were in place. Nonetheless, they were not the only factor to wreak havoc on the global economy. Voluntary social distancing also took a severe tool on economic activity, as people isolated themselves in fear of contracting the virus when infections increased.

We reach these conclusions by examining two high-frequency proxies for economic activity, namely mobility indicators provided by Google and job openings advertised on the website Indeed.[1] As shown in Figure 1, the collapse in mobility over the first six months of 2020 correlates well with the decline in real GDP growth (panel 1a). Similarly, job postings display a tight negative correlation with unemployment rates over the same period (panel 1b). These correlations indicate that mobility and job postings serve as good high-frequency proxies of economic activity.

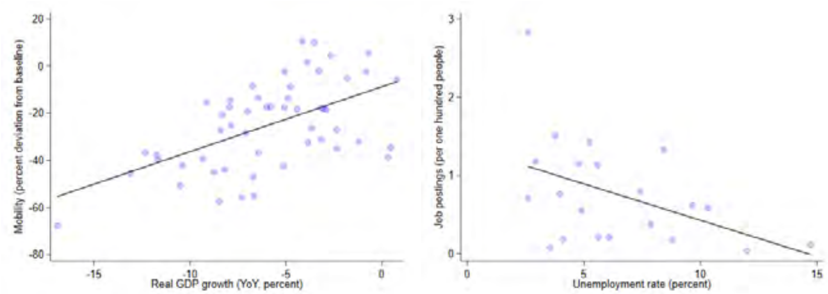
Identifying the causal impact of lockdowns is a challenging task primarily because government measures were imposed in response to epidemiological developments, which in turn affect the economy. To alleviate this concern, the econometric specifications examine the effects of lockdowns while controlling for the stage of the epidemic. Specifically, we use local projections to regress the level of mobility index and number of job postings over the stringency of lockdowns and the number of COVID-19 infections. The regressions are estimated using national data for more than 100 countries. By controlling for COVID-19 infections, the regression framework can also shed light on the extent of voluntary social distancing for a given level of lockdowns. The response of mobility and job postings to rising infections should indeed capture how people change behavior when health risks become more severe.

We validate the results based on national data by replicating the analysis using subnational data for a smaller set of countries that allow us to strengthen identification. More precisely, we re-estimate the regressions using only data from regions less affected by COVID-19 in countries that adopted national lockdowns. The identification assumption is based on the observation that national

[1] Google Community Mobility Reports provides information on daily attendance rates at various locations relative to pre-crisis levels. Data are available at the national level for a large set of advanced, emerging, and developing economies. For various countries, mobility information is also available at the sub-national level. Indeed provides information about daily job postings in 22 countries. See Appendix A for more details about data sources and sample coverage.

Figure 1: High-Frequency Proxies of Economic Activity for the First Semester of 2020

(a) Mobility and real GDP growth (b) Job postings and unemployment rate



Notes: Mobility and job postings are computed as the daily average over the first semester of 2020. Real GDP growth for the first semester of 2020 is computed with respect to the first semester of 2019. The unemployment rate is computed as the average of the monthly unemployment rate over the first semester of 2020.

lockdowns were often imposed in response to localized outbreaks and were thus largely exogenous to the conditions prevailing in regions with low infections.

Our results show that both lockdowns and voluntary social distancing in response to rising COVID-19 infections can have strong detrimental effects on the economy. Indeed, lockdowns and voluntary social distancing played a comparable roles in driving the drop in mobility across our full set of countries. Similar results are obtained using job postings. The analysis also reveals significant heterogeneity across countries. The contribution of voluntary distancing was stronger in advanced economies, where people can work from home more easily and sustain periods of temporary unemployment because of personal savings and government benefits. Lockdowns played instead a much stronger role in low-income countries where people do not have the financial means to temporarily refrain from economic activities.

Second, we proceed with the analysis by assessing the effectiveness of lockdowns in containing infections. Using a similar empirical framework to the one employed for the analysis of mobility and job postings, the paper also documents that lockdowns can substantially reduce infections. The results are again robust to using subnational data to strengthen identification. The effects of lockdowns on COVID-19 cases tend to materialize a few weeks after the introduction of lockdowns, consistent with the incubation period of the virus and testing times. This underscores the importance of rapid intervention. Indeed, the analysis shows that lockdowns are particularly effective in curbing infections if they are introduced at an early stage of a country's epidemic.

Finally, the paper examines whether lockdowns involve non-linear effects on mobility and infections. We find evidence that more stringent lockdowns have decreasing marginal costs in restricting mobility and thus they likely entail progressively smaller damages to the economy. On the contrary, lockdowns display increasing marginal benefits in reducing infections. In fact, low-intensity lockdowns do not appear to curb the number of infections. This implies that, to reduce infections by a certain amount at the lowest short-run economic cost, more stringent shorter-lived lockdowns could

Covid Economics 66, 28 January 2021: 37-57

be preferable to mild prolonged measures.

The fact that lockdowns can reduce infections but impose short-term economic costs while they are in place is often used to argue that lockdowns involve a trade-off between saving lives and protecting livelihoods. However, the findings in the paper that infections also severely depress economic activity through voluntary social distancing calls for a re-assessment of this narrative. By bringing infections under control, stringent and early lockdowns may pave the way to a faster economic recovery as people feel more comfortable to resume normal activities. In other words, the short-term economic costs of lockdowns could be compensated through higher future economic activity, possibly leading to a positive overall effect on the economy.[2] This remains a crucial area for future research as the pandemic progresses across different waves, making it possible to better assess the ultimate consequences of lockdowns.

The paper is organized as follows. After discussing the related literature, section 2 presents an assessment of the economic impact of lockdowns and voluntary social distancing relying on high-frequency proxies of economic activity. Section 3 examines the effect of lockdowns on COVID-19 infections. Section 4 explores the non-linear effects of lockdowns on mobility and infections. Section 5 concludes.

Related Literature. The literature provides conflicting evidence about the importance of lockdowns and voluntary social distancing in driving the economic contraction. Some papers find that lockdowns have a severe impact on the economy. Using customized survey data, Coibion et al. (2020) document that lockdowns accounted for much of the decline in employment and consumer spending in the US during the first months of the country's epidemic. Beland et al. (2020) and Gupta et al. (2020) use data from the US Current Population Survey and also find that stay-at-home orders led to large increases in unemployment. Analyzing transaction level data from bank accounts, Baker et al. (2020) find that consumer spending dropped twice as much in US states that issued shelter-in-place orders. Evidence about the severe impact of lockdowns extends to studies beyond the US. For example, Carvalho et al. (2020) exploit high-frequency transaction data in Spain to show that expenditures fell sharply in conjunction with the national lockdown. Similarly, Chronopoulos et al. (2020) use transaction level data showing that consumer spending declined in line with lockdown measures in the UK.

Other papers argue instead that voluntary social distancing was the key driver of the economic contraction. Combining high-frequency data from payroll and financial firms in the US, Chetty et al. (2020) find that spending and employment fell before state-at-home orders and that re-openings had modest effects on economic activity. Goolsbee and Syverson (2020) analyze customers' visits to businesses located nearby but that faced different lockdown restrictions because belonging to different counties. They conclude that the drop in economic activity was mostly due to people voluntarily reducing visits in line with rising COVID-19 deaths. Baek et al. (2020), Bartik et al. (2020), Forsythe et al. (2020) and Rojas et al. (2020) also find that lockdown restrictions had a modest impact on the US labor market. Chen et al. (2020) document that lockdowns in Europe did

[2] Correia et al. (2020) argue that lockdowns during the 1918 Flu Pandemic were associated with better medium-term economic outcomes.

Covid Economics 66, 28 January 2021: 37-57

CEPR PRESS

not have systematic effects on electricity consumption and Maloney and Taskin (2020) find that in most countries the decline in mobility was related to rising infections rather than to lockdowns. The importance of voluntary social distancing is also attested by the economic contractions in countries that did not adopt stringent lockdowns, such as South Korea and Sweden (Andersen et al., 2020; Aum et al., 2020; Born et al., 2020).

Compared to these two strands of the literature, a key contribution of our paper is to jointly assess the relevance of lockdowns and voluntary social distancing using a common empirical framework over a very large set of countries. This provides a full picture of the nature of the economic crisis across the globe and ensures that the results are not driven by specific country features.

The paper is also related to a growing body of evidence on the effectiveness of lockdowns in reducing infections (Chernozhukov et al., 2020; Dave et al., 2020; Friedson et al., 2020; Glaeser et al., 2020, Fang et al., 2020, Imai et al., 2020, Jinjarak et al., 2020; Yilmazkuday, 2020). In particular, consistent with our results, Demirgüç-Kunt et al. (2020) find that lockdowns are particulary effective if they are introduced early on. The literature also documents the importance of face masks and testing to contain the virus (Chernozhukov et al., 2020; Gapen et al., 2020).

The paper is organized as follows. Section 2 presents an assessment of the economic impact of lockdowns and voluntary social distancing relying on high-frequency proxies of economic activity. Section 3 examines the effect of lockdowns on COVID-19 infections. Section 4 explores the non-linear effects of lockdowns on mobility and infections. Section 5 concludes.

2 Lockdowns and Voluntary Social Distancing

In this section we examine the economic impact of lockdowns and voluntary social distancing using high-frequency data. Specifically, we rely on two types of data to proxy for economic activity, both of which are available at daily frequency. First, we use mobility data provided by Google, which reports the attendance rate at various locations relative to pre-crisis levels.[3] These data have the key advantages of covering a large set of countries and being available also at the subnational level. Second, we corroborate the analysis of mobility using job posting data reported by Indeed, an online job search engine. Data from Indeed are available for fewer countries but capture labor market conditions more directly.

2.1 Impact on Mobility

Assessing the impact of lockdowns on mobility is a challenging task since the decision to deploy lockdowns is not random. Cross-country identification is precluded by omitted variable concerns because the introduction of lockdowns can reflect time-invariant country characteristics that also affect economic outcomes. For example, countries with higher social capital may not require stringent lockdowns—as people take greater precautions against infecting others—and could also better with-

[3]Data are based on cell phones' locations for people that own smart phones and accept to share location data with Google. A drawback of this data is that, since this category of people may have characteristics that differ from the broader population (e.g., relative to income level, age, and access to internet, among others), the mobility indices may not be fully representative of the entire country, especially in poorer countries where fewer people have smart phones.

Covid Economics 66, 28 January 2021: 37-57

stand the economic impact of the crisis. When using time variation in the data, the main challenge is that the adoption of lockdowns depends on the stage of the epidemic. For example, governments are more likely to impose lockdowns when health risks become more acute. At that time, people tend to voluntarily reduce social interactions because they fear being infected or infecting others. This may generate a spurious correlation between the introduction of lockdowns and the reduction in mobility.

To alleviate endogeneity concerns, the analysis relies on panel regressions that control for country fixed-effects and the stage of the country's epidemic. More specifically, we assess the dynamic response of mobility to lockdowns using the following local projection regressions (Jordà, 2005):

$$mob_{i,t+h} = \alpha_i^h + \tau_t^h + \sum_{p=0}^{P} \beta_p^h ln\Delta cases_{i,t-p} + \sum_{p=0}^{P} \delta_p^h lock_{i,t-p} + \sum_{p=1}^{P} \rho_p^h mob_{i,t-p} + \varepsilon_{i,t+h} \qquad (1)$$

The variable $mob_{i,t+h}$ denotes the level of mobility for country i at time $t+h$, with h being the horizon;[4] $ln\Delta cases_{i,t-p}$ is the log of daily COVID-19 cases, which is used to track the stage of the pandemic, with p being the lag length; and $lock_{i,t-p}$ is an index measuring the stringency of lockdowns.[5] The specification also features lags of the dependent variable to account for pre-existing trends, and country and time fixed effects to control for country characteristics and global factors. The estimation includes a week worth of lags.[6] Standard errors are clustered at the country level. The sample of analysis includes 128 countries between early February and mid-July, 2020.

Our identification assumption is that by controlling for the stage of the pandemic (proxied by daily cases) and country fixed effects, the coefficient δ_0^h should isolate the impact of lockdowns. At the same time, for a given level of lockdown stringency, the coefficient β_0^h should reveal the extent of voluntary social distancing, capturing the responsiveness of mobility to rising infections. Finally, to control for the persistence of the stringency index and of the number of COVID-19 cases, we include lags of both variables.

To address endogeneity concerns further, we validate our findings using an alternative identification strategy that takes advantage of the sub-national disaggregation of the Google mobility data. This is based on the observation that various countries imposed lockdowns on a *national* scale in reaction to *localized* outbreaks. For example, in Italy—one of the first countries severely hit by the pandemic after China—the government imposed a national lockdown in early March even though most of the infections were concentrated in Lombardy. In these countries, the adoption of national lockdowns was largely exogenous to the conditions prevailing in those regions that had few COVID-19 infections. This provides an opportunity to considerably strengthen identification by analyzing

[4]The mobility index used in the analysis is constructed as the average of the mobility indices for groceries and pharmacies, parks, retails and recreation, transit stations, and workplaces. In the case of China, the mobility index is based on data from Baidu.

[5]We employ the lockdown stringency index provided by the University of Oxford's Coronavirus Government Response Tracker. This index is a simple average of nine sub-indicators capturing school closures, workplace closures, cancellations of public events, gatherings restrictions, public transportation closures, stay-at-home requirements, restrictions on internal movement, controls on international traveling, and public information campaigns. Since we want to measure the impact of *actual* restrictions, we re-construct the index excluding public information campaigns as they aim to promote voluntary social distancing. The results, however, are similar when public information campaigns are included in the index.

[6]A richer lag structure does not affect the results.

Covid Economics 66, 28 January 2021: 37-57

the effects of national lockdowns on the mobility in regions less affected by COVID-19.

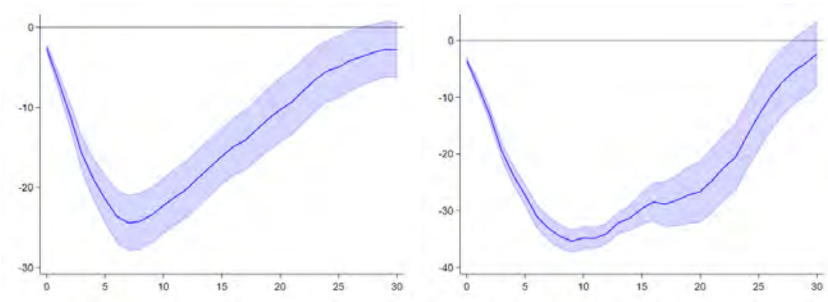
Formally, we re-estimate equation (1) using data for 422 subnational regions in 15 G20 countries that adopted national lockdowns. For each country, we exclude the region with the largest number of COVID-19 cases and any region that had more than 20 percent of the country's total cases. The regression thus analyzes the mobility response in those regions less affected by the virus for which the national lockdown was an exogenous event triggered by conditions elsewhere in the country.

Figure 2 shows the impact on mobility arising from a full lockdown that includes all measures used by governments during the pandemic. Panels 2a and 2b display the results from the national and subnational regressions, respectively. We see that in both cases a full lockdown leads to a very significant decline in mobility. When using national level data, the impact reaches about 25 percent after a week and then mobility starts to resume gradually as the lockdown tightening dissipates.[7] The estimates based on subnational data corroborate the negative effect of lockdowns on mobility. The shape of the mobility response is remarkably similar to the one obtained with national data. The impact is modestly larger and more persistent, possibly reflecting differences in the sample coverage.

Figure 2: Impact of a Full Lockdown on Mobility
(Percent)

(a) National data

(b) Subnational data



Notes: The x-axes denote the number of days, the lines denote the point estimates, and the shaded areas correspond to 90 percent confidence intervals computed with standard errors clustered at the country level.

As discussed, lockdowns are not the only contributing factor to the decline in mobility during the pandemic. People also voluntarily reduce exposure to each other as infections increase and they fear becoming sick. Aum et al. (2020), Goolsbee and Syverson (2020), and Maloney and Taskin (2020) document indeed that mobility has been tightly correlated to the spread of COVID-19 even after controlling for government lockdowns, especially in advanced economies. In line with this literature, the regression framework provides estimates that can shed light on the strength of

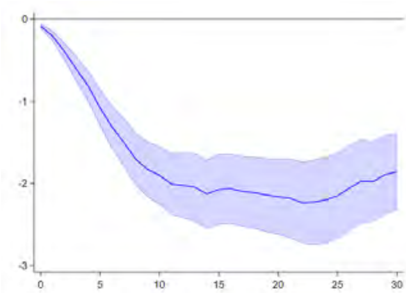
[7]Results are robust to controlling for COVID-19 deaths instead of cases; using sub-indicators of mobility provided by Google; controlling for testing, contact tracing, and public information campaigns; and testing for possible cross-country heterogeneity in the mobility response depending on population density and indicators of governance and social capital.

Covid Economics 66, 28 January 2021: 37-57

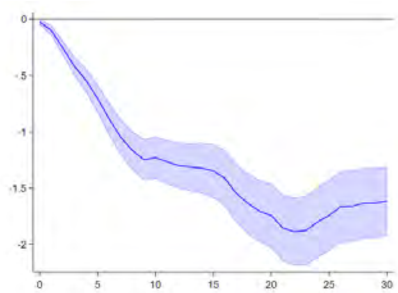
voluntary social distancing by capturing the response of mobility to rising COVID-19 infections for a given lockdown stringency.[8] Figure 3 presents the estimates of the strength of voluntary social distancing by capturing the response of mobility to rising COVID-19 infections for a given lockdown stringency. Using national data, panel 3a shows that an increase in COVID-19 cases has a considerable negative effect on mobility. A doubling of daily COVID-19 cases leads to a contraction in mobility by about 2 percent.[9] Panel 3b shows the impact of COVID-19 on mobility using subnational data. The results are in line with the ones obtained at the national level: a doubling of COVID-19 cases leads to a contraction in mobility of 1.7 percent after 30 days.

Figure 3: Impact of Voluntary Social Distancing on Mobility
(Impact of a doubling in daily COVID-19 cases, percent)

(a) National data



(b) Subnational data



Notes: The x-axes denote the number of days, the lines denote the point estimates, and the shaded areas correspond to 90 percent confidence intervals computed with standard errors clustered at the country level.

The national and sub-national results thus convey a consistent message. Both lockdowns and voluntary social distancing in response to rising infections severely reduce mobility. To gain further insights into the relative importance of these two factors, we calculate the contributions of lockdowns and voluntary social distancing in driving the decline in mobility during the first three months of each country's epidemic. The effects of lockdowns and voluntary distancing are likely to differ across countries depending on the stage of economic development. For example, in more advanced countries people can more easily opt for voluntary social distancing thanks to the prevalence of teleworking, the presence of contactless delivery services, the amount of personal savings to sustain periods of temporary unemployment, etc. To capture some of these nuances, we amend the specification in equation (1) allowing the impact of lockdowns and rising COVID-19 cases to vary between advanced,

[8]Besides reacting to the spread of COVID-19, people may opt to voluntarily self distance also in response to other factors, such as public health announcements, news about celebrities being infected, or even the adoption of government lockdowns. As such, the analysis may underestimate the true extent of voluntary social distancing. Also, as shown by Adda (2016), higher mobility and economic activity might lead to faster spread of viral diseases, generating some reverse causality between the outcome variables and COVID-19 infections. The dynamic structure of the estimation should alleviate this endogeneity concern.

[9]The results are robust to controlling for COVID-19 deaths instead of cases.

Covid Economics 66, 28 January 2021: 37-57

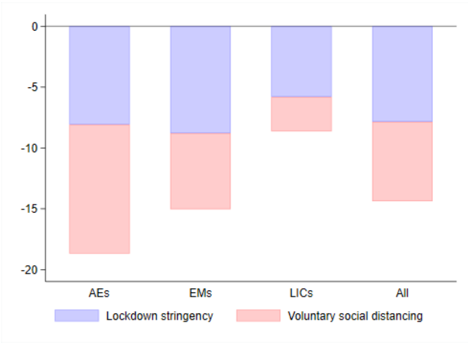
emerging, and low-income countries:

$$
\begin{aligned}
mob_{i,t+h} = {} & \alpha_i^h + \tau_t^h + \sum_{p=0}^{P} \beta_p^h ln\Delta cases_{i,t-p} + \sum_{p=0}^{P} \delta_p^h lock_{i,t-p} \\
& + AE_i \times \left(\sum_{p=0}^{P} \beta_p^{h,AE} ln\delta cases_{i,t-p} + \sum_{p=0}^{P} \delta_p^{h,AE} lock_{i,t-p} \right) \\
& + EM_i \times \left(\sum_{p=0}^{P} \beta_p^{h,EM} ln\Delta cases_{i,t-p} + \sum_{p=0} P\delta_p^{h,EM} lock_{i,t-p} \right) \\
& + \sum_{p=1}^{P} \rho_p^h mob_{i,t-p} + \varepsilon_{i,t+h} \qquad (2)
\end{aligned}
$$

The variables AE_i and EM_i are dummies that denote advanced economies and emerging markets, respectively, with low-income countries being the omitted category. Thus, the impact of lockdowns on mobility at the horizon h for advanced economies can be obtained as $\delta_0^h + \delta_0^{h,AE}$, for emerging markets as $\delta_0^h + \delta_0^{h,EM}$, and for low-income countries as δ_0^h.

We then compute the contributions of lockdowns and voluntary social distancing to the decline in mobility during the first three months of each country's epidemic. We do so by multiply the average impact of lockdowns and COVID-19 cases during the 30-day local projection horizon by the average stringency of lockdowns and number of COVID-19 cases during the first three months of each country's epidemic.

Figure 4: Contributions to the Mobility Decline
(Percent)



Notes: The bars denote the cross-country averages of the contributions of lockdowns and voluntary social distancing, computed using the coefficients on lockdowns and the log of daily COVID-19 cases multiplied by the average of the corresponding variables for each country group during the first three months of each country's epidemic.

Figure 4 illustrates the contributions of lockdowns and voluntary social distancing in reducing mobility across country groups. Both lockdowns and voluntary social distancing had a large impact on mobility, playing a roughly similar role across the full set of countries. The contribution of voluntary social distancing was significantly stronger in advanced economies, likely because people can work from home more easily and can even afford to stop working temporarily by relying on personal savings and social security benefits. On the contrary, voluntary social distancing was quite limited in low-income countries where the drop in mobility was mostly due to lockdowns.

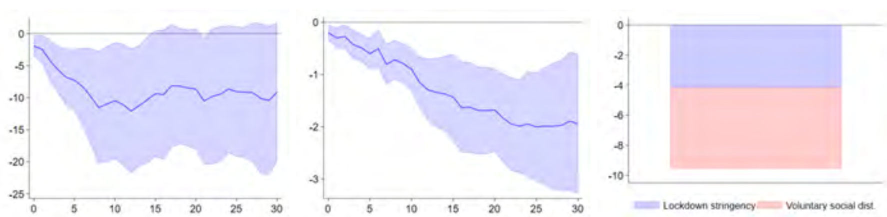
2.2 Impact on Job Postings

In the previous section, we found that both lockdowns and voluntary social distancing played a very substantial role in reducing mobility. We now show that similar results are obtained when analyzing job postings data provided by Indeed. We re-estimate the panel regression in equation (1) substituting the level of mobility with the log of the number of job postings. The sample includes daily data for 22 countries from January 1 to June 28, 2020. In line with the analysis of mobility, the specification includes seven lags of the dependent and independent variables, and country and time fixed effects to control for time invariant country characteristics and global factors.

Figure 5 shows that both lockdowns and voluntary social distancing have negative and significant effects on job postings. In panel 5a, a full lockdown is associated with a decline in job postings of about 12 percent two weeks after the introduction of the lockdown. In panel 5b, a doubling COVID-19 cases leads to a 2 percent decline in job postings after 30 days. Using these estimates, we can compute the contributions of lockdowns and voluntary social distancing in reducing job postings during the first three months of each country's epidemic. Panel 5c shows that both lockdowns and voluntary social distancing were important factors behind the drop in job postings. The contribution of voluntary social distancing was relatively stronger. This is consistent with the results based on mobility data since the Indeed sample includes primarily advanced economies.

Figure 5: Impact of Lockdowns and Voluntary Social Distancing on Job Postings
(Percent)

(a) Impact of a full lockdown on job postings | (b) Impact of a doubling in daily COVID-19 cases on job postings | (c) Contributions to the job posting decline



Notes: The x-axes in panels 5a and 5b denote the number of days, the lines denote the point estimates, and the shaded areas correspond to 90 percent confidence intervals computed with standard errors clustered at the country level. The bars in panel 5c denote the cross-country averages of the contributions of lockdowns and voluntary social distancing, computed using the coefficients on lockdowns and the log of daily COVID-19 cases multiplied by the average of the corresponding variables during the first three months of each country's epidemic.

Covid Economics 66, 28 January 2021: 37-57

Covid Economics 66, 28 January 2021: 37-57

3 Lockdowns and COVID-19 Infections

After having examined the economic effects of lockdowns, we now turn to the question of whether these tools can succeed in their intended goal of curbing infections. To address this issue, we estimate the following local projections:

$$lncases_{i,t+h} - lncases_{i,t-1} = \alpha_i^h + \tau_t^h + \sum_{p=0}^{P} \beta_p^h X_{i,t-p} + \sum_{p=0}^{P} \delta_p^h lock_{i,t-p} + \sum_{p=1}^{P} \rho_p^h \Delta lncases_{i,t-p}$$
$$+trend_i^h + trend_i^{2,h} + \varepsilon_{i,t+h} \qquad (3)$$

where $X_{i,t-p}$ is a vector of controls including the average temperature and humidity in the country (Adda, 2016, for instance, finds that higher temperatures reduce the spread of influenza and other viral diseases), as well as indicators for whether widespread testing and contact tracing policies are in place; and $trend_i^h$ and $trend_i^{2,h}$ are country-specific linear and quadratic trends. The sample includes 89 countries based on data availability.

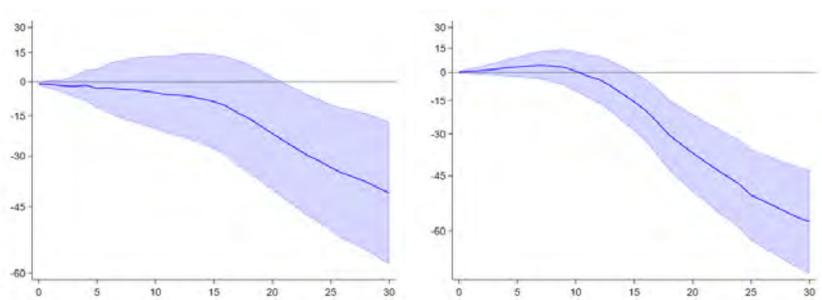
As done for the analysis of mobility, to improve the identification we re-estimate equation 3 using sub-national data for 339 units in 15 G20 countries. The sample excludes sub-national units with the largest number of cases per country and those that had more than 20 percent of the country's total COVID-19 cases. It thus focuses on regions with fewer cases for which the adoption of national lockdowns was largely an exogenous event. The sub-national regressions exclude the controls $X_{i,t-p}$ since they are not available at the sub-national level.

Figure 6 presents the results of the impact of lockdowns on COVID-19 infections. Using national level data, panel 6a shows that a full lockdown leads to a large reduction in cumulated infections, equal to about 40 percent after 30 days. The results based on sub-national data in panel 6b point to an even larger effect, reducing infections by about 58 percent after 30 days.

Figure 6: Impact of a Full Lockdown on COVID-19 Infections
(Percent)

(a) National data

(b) Subnational data



Notes: The x-axes denote the number of days, the lines denote the point estimates, and the shaded areas correspond to 90 percent confidence intervals computed with standard errors clustered at the country level.

CEPR PRESS

Figure 6 also shows that the effects of lockdowns on confirmed COVID-19 cases tend to materialize with a delay of at least two weeks. This is consistent with the incubation period of the virus and the time required for testing. Acknowledging this delayed effect is crucial to guide people's expectations about the effectiveness of lockdowns. Furthermore, it points to the need to adopt lockdowns before infection rates increase too rapidly.

Indeed, lockdowns appear to be particularly effective in curbing infections if they are adopted early in a country's epidemic. This can be seen by comparing the epidemiological outcomes of countries that adopted lockdown measures at different times. We differentiate countries between early and late adopters using two alternative criteria. First, we consider the number of days that passed from the first COVID019 case to when lockdown measures reached their maximum stringency. As shown in panel 7a of Figure 7, there is a considerable cross-country heterogeneity. Half of the countries reached their maximum lockdown stringency within a month from the beginning of the epidemic but some waited up to four months. Second, we differentiate countries based on the number of weekly cases at the time in which the maximum lockdown stringency was reached. Panel 7b shows that virtually all countries reached the maximum stringency before daily cases reached 0.1 cases per thousand people.

We then compares the epidemiological outcomes of early and late lockdown adopters 90 days after the first COVID-19 case, splitting the country sample with respect to the median of the distributions in panels 7a and 7b. Panel 7c shows the evolution of infections since the first COVID-19 case, differentiating countries by the number of days passed from the first case to the time that authorities adopted the most stringent lockdown measures. Countries that imposed lockdowns faster experienced better epidemiological outcomes. The differences are even more striking if the sample is split with respect to the number of COVID-19 cases at the time of lockdowns as in panel 7d. Countries that adopted lockdowns when COVID-19 cases were still low witnessed considerably fewer infections during the first three months of the epidemic relative to countries that introduced lockdowns when cases were already high.

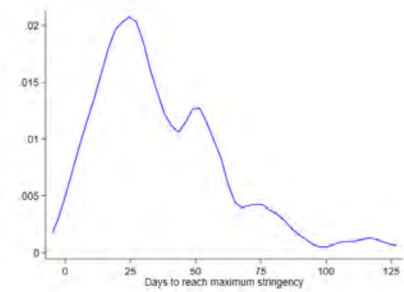
4 Nonlinear Effects of Lockdowns

So far, we have used a lockdown stringency index that combines a broad range of underlying measures. These includes for example travel restrictions, school and workplace closures, and stay-at-home orders, among others. Disentangling the effects of these measures is an arduous task because they are highly correlated, as countries often introduced them in rapid succession to contain infections. Furthermore, countries have generally followed a similar sequence, from restrictions on international travel to stay-at-home orders as illustrated in Figure 8. A regression specification that features all the lockdown measures as independent variables would thus capture the marginal effect of each measure conditional on those that have been adopted beforehand. This underestimates the importance of measures that are adopted at a later stage. For example, stay-at-home orders are generally found to have a modest impact on mobility because various other measures are already in

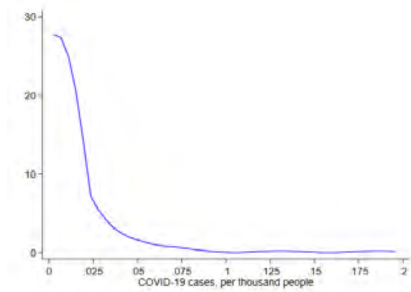
Covid Economics 66, 28 January 2021: 37-57

Figure 7: The Importance of Speed and Timing of Lockdowns

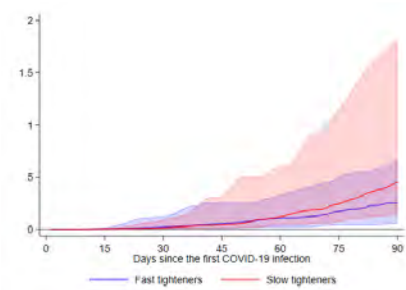
(a) Cross-country distribution of the speed of lockdowns
(Density)



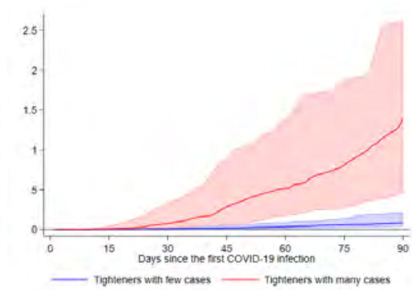
(b) Cross-country distribution of the timing of lockdowns
(Density)



(c) Infections since the start of the countries' epidemics for fast and slow tighteners
(Per thousand people)



(d) Infections since the start of the countries' epidemics for tighteners with few and many COVID-19 cases
(Per thousand people)

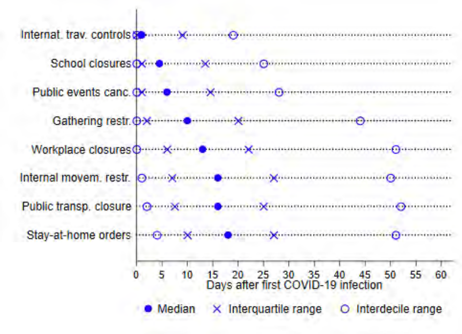


Notes: In panels 7c and 7d, the lines denote the medians and the shaded areas correspond to the interquartile ranges. In panel 7c, countries are split based on the cross-country median value of the distribution in panel 7a; in panel 7d, countries are split based on the cross-country median value of the distribution in panel 7b.

Covid Economics 66, 28 January 2021: 37-57

place.[10]

Figure 8: Sequencing of Lockdown Measures



Notes: The blue dots denote the cross-country median number of days since the first COVID-19 case and the day in which each lockdown measure was introduced, the blue crosses denote the interquartile ranges, and the empty circles denote the interdecile ranges.

An analytically sounder approach is to examine whether adding multiple lockdown measures continues to have similar economic and epidemiological effects. This can inform policymakers on whether it is best to rely on a protracted mild lockdowns or to opt for stringent temporary measures. To examine nonlinearities in the effects of lockdowns on mobility, we expand equation (1) by adding a quadratic term for the lockdown stringency:

$$mob_{i,t+h} = \alpha_i^h + \tau_t^h + \sum_{p=0}^{P} \beta_p^h ln\Delta cases_{i,t-p} + \sum_{p=0}^{P} \delta_p^h lock_{i,t-p} + \sum_{p=0}^{P} \omega_p^h lock_{i,t-p}^2 + \sum_{p=1}^{P} \rho_p^h mob_{i,t-p} + \varepsilon_{i,t+h} \quad (4)$$

We do the same for infections modifying equation (3) as follows:

$$lncases_{i,t+h} - lncases_{i,t-1} = \alpha_i^h + \tau_t^h + \sum_{p=0}^{P} \beta_p^h X_{i,t-p} + \sum_{p=0}^{P} \delta_p^h lock_{i,t-p} + \sum_{p=0}^{P} \omega_p^h lock_{i,t-p}^2 + \sum_{p=1}^{P} \rho_p^h \Delta lncases_{i,t-p} + trend_i^h + trend_i^{2,h} + \varepsilon_{i,t+h} \quad (5)$$

[10]For example, replacing the lockdown stringency index in equation (1) with the (rescaled) indices for each individual lockdown measure would produce results for which measures that are introduced later (e.g., stay-at-home orders or transportation restrictions) display a smaller impact on mobility, while the measures that are introduced first (e.g., international movement restrictions or school closures) are associated with a larger impact. In the case of infections, while the point estimates are negative, the confidence intervals include the zero for most of the measures. Results are available upon request. Another approach could be to allow for interaction terms across all measures to better capture the impact on mobility of a given measure conditional on the others being in place or not. However, the regression becomes cumbersome and the results are inconclusive.

Covid Economics 66, 28 January 2021: 37-57

Covid Economics 66, 28 January 2021: 37-57

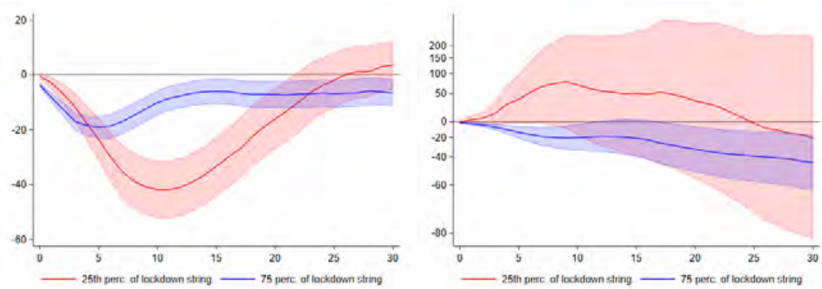
The results in panel 9a of Figure 9 show that lockdowns have decreasing marginal effects on mobility. Introducing additional measures when the lockdown stringency index is already elevated has a weaker impact on mobility compared to introducing them when the lockdown stringency is low. For example, stay-at-home orders may have only a modest negative impact on economic activity if governments have already imposed workplace closures. Formally, these findings reflect that the quadratic term in equation (4) is positive and statistically significant at various horizons.

Whereas lockdowns have decreasing marginal effects on mobility, panel 9b shows that they have increasing marginal effects on infections. Lockdown measures are effective in reducing COVID-19 cases only if they are sufficiently stringent. A possible interpretation is that preventing only a few instances of personal contacts—such as by closing schools alone—is not enough to significantly reduce community spread. More stringent measures—such as workplace closures or stay-at-home orders—are needed to effectively bring the virus under control. The quadratic term in equation (5) is negative and statistically significant at various horizons.

Taken together, these results suggest that to achieve a given reduction in infections, policymakers may want to opt for stringent lockdowns over a shorter period rather than resort to prolonged mild lockdowns. Tighter lockdowns appear indeed to entail only modest additional economic costs while leading to a considerably stronger decline in infections.

Figure 9: Nonlinear Effects of Lockdowns
(Percent)

(a) Impact on mobility (b) Impact on infections



Notes: The x-axes denote the number of days, the lines denote the point estimates, and the shaded areas correspond to 90 percent confidence intervals computed with standard errors clustered at the country level.

5 Conclusions

This paper documents that lockdowns and voluntary social distancing have both played a crucial role in reducing economic activity during the first phase of the COVID-19 pandemic. Consistent results are obtained from examining mobility and job posting data and from employing identification strategies based on national and sub-national data. Therefore, despite lockdowns entail significant economic costs while they are in place, letting infections grow uncontrolled can also have dire economic consequences because people voluntarily refrain from economic activities if they fear contracting the virus.

We also find that lockdowns are powerful instruments to reduce infections, especially if they are introduced early in a country's epidemic. Furthermore, the analysis reveals that lockdowns impose decreasing marginal costs on economic activity as they become more stringent but they involve increasing marginal benefits in reducing infections. Therefore, policymakers should lean towards adopting tight lockdowns rapidly when infections increase rather than rely on protracted mild measures.

The effectiveness of lockdowns in reducing infections coupled with the finding that rising infections can considerably harm economic activity provide an important new perspective on the overall costs of lockdowns. The prevailing narrative often portrays lockdowns as involving a trade-off between saving lives and supporting the economy. This characterization neglects that, despite imposing short-term economic costs, lockdowns may lead to a faster economic recovery by containing the virus and reducing voluntary social distancing. More research is warranted as the pandemic progresses to provide a fuller assessment of the overall economic effects of lockdowns.

Meanwhile, policymakers should also look for alternative ways to contain infections that may entail even lower short-run economic costs. These include expanding contact tracing, promoting the use of face masks, and encouraging working from home. As the understanding of the virus transmission improves, countries may also be able to use targeted lockdown measures more effectively, for example by limiting large indoor gatherings and better protecting vulnerable people. These remain important areas for future research.

Covid Economics 66, 28 January 2021: 37-57

References

Adda, Jérôme (Feb. 2016). " Economic Activity and the Spread of Viral Diseases: Evidence from High Frequency Data". *The Quarterly Journal of Economics* 131(2), pp. 891–941.

Andersen, Asger Lau, Emil Toft Hansen, Niels Johannesen, and Adam Sheridan (2020). "Pandemic, Shutdown and Consumer Spending: Lessons from Scandinavian Policy Responses to COVID-19". arXiv preprint arXiv:2005.04630.

Aum, Sangmin, Sang Yoon Tim Lee, and Yongseok Shin (2020). "COVID-19 Doesn't Need Lockdowns to Destroy Jobs: The Effect of Local Outbreaks in Korea". National Bureau of Economic Research Working Paper No. 27264.

Baek, ChaeWon, Peter B McCrory, Todd Messer, and Preston Mui (2020). "Unemployment Effects of Stay-at-Home Orders: Evidence from High Frequency Claims Data". *Institute for Research on Labor and Employment Working Paper* (101-20).

Baker, Scott R, Robert A Farrokhnia, Steffen Meyer, Michaela Pagel, and Constantine Yannelis (2020). "How Does Household Spending Respond to an Epidemic? Consumption During the 2020 COVID-19 Pandemic". Covid Economics 18.

Bartik, Alexander W, Marianne Bertrand, Feng Lin, Jesse Rothstein, and Matt Unrath (2020). "Measuring the Labor Market at the Onset of the COVID-19 Crisis". National Bureau of Economic Research Working Paper No. 27613.

Beland, Louis-Philippe, Abel Brodeur, and Taylor Wright (2020). "COVID-19, Stay-at-Home Orders and Employment: Evidence from CPS Data". IZA Discussion Paper No. 13282.

Born, Benjamin, Alexander Dietrich, and Gernot J Müller (2020). "Do Lockdowns Work? A Counterfactual for Sweden". CEPR Discussion Paper No. DP14744.

Carvalho, Vasco M, Stephen Hansen, Alvaro Ortiz, Juan Ramon Garcia, Tomasa Rodrigo, Sevi Rodriguez Mora, and Pep Ruiz de Aguirre (2020). "Tracking the COVID-19 Crisis with High-Resolution Transaction Data". CEPR Discussion Paper No. DP14642.

Chen, Sophia, Deniz Igan, Nicola Pierri, and Andrea F Presbitero (2020). "Tracking the Economic Impact of COVID-19 and Mitigation Policies in Europe and the United States". Covid Economics 36.

Chernozhukov, Victor, Hiroyuki Kasaha, and Schrimpf Paul (2020). "Causal Impact of Masks, Policies, Behavior on Early Covid-19 Pandemic in the U.S." Centre for Economic Policy Research, COVID Economics Vetted and Real-Time Papers, Issue 35.

Chetty, Raj, John N Friedman, Nathaniel Hendren, Michael Stepner, et al. (2020). "How Did COVID-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data". National Bureau of Economic Research Working Paper No. 27431.

Chronopoulos, Dimitris K, Marcel Lukas, and John OS Wilson (2020). "Consumer Spending Responses to the COVID-19 Pandemic: An Assessment of Great Britain". Covid Economics 34.

Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber (2020). "The Cost of the COVID-19 Crisis: Lockdowns, Macroeconomic Expectations, and Consumer Spending". Covid Economics 20.

Correia, Sergio, Stephan Luck, and Emil Verner (2020). "Pandemics Depress the Economy, Public Health Interventions Do Not: Evidence from the 1918 Flu". Available at SSRN 3561560.

Dave, Dhaval, Andrew I Friedson, Kyutaro Matsuzawa, and Joseph J Sabia (2020). "When Do Shelter-in-Place Orders Fight COVID-19 Best? Policy Heterogeneity Across States and Adoption Time". *Economic Inquiry*.

Demirgüç-Kunt, Asli, Michael Lokshin, and Ivan Torre (2020). "The Sooner, the Better : The Early Economic Impact of Non-Pharmaceutical Interventions during the COVID-19 Pandemic". *World Bank Policy Research Working Paper No. 9257.*

Fang, Hanming, Long Wang, and Yang Yang (2020). "Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-ncov) in China".

Forsythe, Eliza, Lisa B Kahn, Fabian Lange, and David Wiczer (2020). "Labor Demand in the Time of COVID-19: Evidence from Vacancy Postings and UI Claims". *Journal of Public Economics*, p. 104238.

Friedson, Andrew I, Drew McNichols, Joseph J Sabia, and Dhaval Dave (2020). "Did California's Shelter-in-Place Order Work? Early Coronavirus-Related Public Health Effects". National Bureau of Economic Research Working Paper No. 26992.

Gapen, Michael, Jonathan Millar, Blerina Uruçi, and Pooja Sriram (2020). "Assessing the Effectiveness of Alternative Measures to Slow the Spread of COVID-19 in the United States". *Covid Economics*, p. 46.

Glaeser, Edward L, Caitlin S Gorback, and Stephen J Redding (2020). "How Much does COVID-19 Increase with Mobility? Evidence from New York and Four Other US Cities". National Bureau of Economic Research Working Paper No. 27519.

Goolsbee, Austan and Chad Syverson (2020). "Fear, Lockdown, and Diversion: Comparing Drivers of Pandemic Economic Decline". National Bureau of Economic Research Working Paper No. 27432.

Gupta, Sumedha, Laura Montenovo, Thuy D Nguyen, Felipe Lozano Rojas, Ian M Schmutte, Kosali I Simon, Bruce A Weinberg, and Coady Wing (2020). "Effects of Social Distancing Policy on Labor Market Outcomes". National Bureau of Economic Research Working Paper No. 27280.

Imai, Natsuko, Katy AM Gaythorpe, Sam Abbott, Sangeeta Bhatia, Sabine van Elsland, Kiesha Prem, Yang Liu, and Neil M Ferguson (2020). "Adoption and Impact of Non-Pharmaceutical Interventions for COVID-19". *Wellcome Open Research* 5.

Jinjarak, Yothin, Rashad Ahmed, Sameer Nair-Desai, Weining Xin, and Joshua Aizenman (2020). "Accounting for Global COVID-19 Diffusion Patterns, January-April 2020". National Bureau of Economic Research Working Paper No. 27185.

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

Maloney, William and Temel Taskin (2020). "Determinants of Social Distancing and Economic Activity During COVID-19: A Global View". The World Bank, Policy Research Working Paper 9242.

Covid Economics 66, 28 January 2021: 37-57

Rojas, Felipe Lozano, Xuan Jiang, Laura Montenovo, Kosali I Simon, Bruce A Weinberg, and Coady Wing (2020). "Is the Cure Worse than the Problem Itself? Immediate Labor Market Effects of COVID-19 Case Rates and School Closures in the U.S." *NBER Working Paper 27127*.

Yilmazkuday, Hakan (2020). "COVID-19 Deaths and Inter-County Travel: Daily Evidence from the US". Available at SSRN 3568838.

Appendix A. Data Sources and Country Coverage

Table A.1 lists the data sources used in the analysis. The country coverage for the different sections of the analysis is reported in Table A.2, with the selection of countries being driven by data availability. For the analysis relying on high-frequency indicators, the sample includes 22 countries when job postings are used and 128 countries when mobility is used. When we employ sub-national data on mobility, the sample consists of 422 units for 15 G20 countries. Finally, the analysis of infections is based on a sample of 89 countries for which information on temperature, humidity, testing, and contact tracing is available. At the sub-national level, the sample consists of 373 units for G20 15 countries.

Table A.1: Data Sources

Indicator	Source
Contact tracing	Oxford COVID-19 Government Response Tracker
COVID-19 cases	Oxford COVID-19 Government Response Tracker
Humidity	Air Quality Open Data Platform
Lockdown stringency index	Oxford COVID-19 Government Response Tracker
Mobility	Google Community Mobility Reports, Baidu for China
Stock of job postings	Indeed
Temperature	Air Quality Open Data Platform
Testing	Oxford COVID-19 Government Response Tracker

Covid Economics 66, 28 January 2021: 37-57

Table A.2: Country Coverage

Country	Samples	Country	Samples	Country	Samples
Afghanistan	*Mn, In*	Iraq	*Mn, In*	Guatemala	*Mn, In*
Algeria	*In*	Ireland	*Mn, In, Jp*	Guinea	*In*
Angola	*Mn*	Israel	*Mn, In*	Haiti	*Mn*
Argentina	*Mn, Ms, In, Is*	Italy	*Mn, Ms, In, Is, Jp*	Honduras	*Mn*
Aruba	*Mn*	Jamaica	*Mn*	Hong Kong SAR	*Mn, In, Jp*
Australia	*Mn, Ms, In, Is, Jp*	Japan	*Mn, Ms, In, Is, Jp*	Hungary	*Mn, In*
Austria	*Mn, In, Jp*	Jordan	*Mn, In*	Iceland	*In*
Bahrain	*Mn, In*	Kazakhstan	*Mn, In*	India	*Mn, Ms, In, Is*
Bangladesh	*Mn, In*	Kenya	*Mn*	Indonesia	*Mn, Ms, In, Is*
Barbados	*Mn*	Korea	*Mn, In*	Iran	*In*
Belarus	*Mn*	Kosovo	*In*	Puerto Rico	*Mn*
Belgium	*Mn, In, Jp*	Kuwait	*Mn, In*	Qatar	*Mn*
Belize	*Mn*	Kyrgyz Republic	*Mn, In*	Romania	*Mn, In*
Benin	*Mn*	Lao P.D.R.	*Mn, In*	Russia	*Mn, In*
Bolivia	*Mn, In*	Latvia	*Mn*	Rwanda	*Mn*
Bosnia and Herzegovina	*Mn, In*	Lebanon	*Mn*	Saudi Arabia	*Mn, Ms, In, Is*
Botswana	*Mn*	Libya	*Mn*	Senegal	*Mn*
Brazil	*Mn, Ms, In, Is, Jp*	Lithuania	*Mn, In*	Serbia	*Mn, In*
Bulgaria	*Mn, In*	Luxembourg	*Mn*	Singapore	*Mn, In, Jp*
Burkina Faso	*Mn*	Macao SAR	*In*	Slovak Republic	*Mn, In*
Cambodia	*Mn*	Malaysia	*Mn, In*	Slovenia	*Mn*
Cameroon	*Mn*	Mali	*Mn, In*	South Africa	*Mn, Ms, In, Is*
Canada	*Mn, Ms, In, Is, Jp*	Mauritius	*Mn*	Spain	*Mn, In, Jp*
Chile	*Mn, In*	Mexico	*Mn, Ms, In, Is, Jp*	Sri Lanka	*Mn, In*
China	*Mn, Ms, In, Is*	Moldova	*Mn*	Sweden	*Mn, In, Jp*
Colombia	*Mn, In*	Mongolia	*Mn, In*	Switzerland	*Mn, In, Jp*
Costa Rica	*Mn, In*	Morocco	*Mn*	Taiwan Province of China	*Mn*
Croatia	*Mn, In*	Mozambique	*Mn*	Tajikistan	*Mn, In*
Czech Republic	*Mn, In*	Myanmar	*Mn, In*	Tanzania	*Mn*
Côte d'Ivoire	*Mn, In*	Namibia	*Mn*	Thailand	*Mn, In*
Cyprus	*In*	Nepal	*Mn, In*	Togo	*Mn*
Denmark	*Mn, In*	Netherlands	*Mn, In, Jp*	Trinidad and Tobago	*Mn*
Dominican Republic	*Mn*	New Zealand	*Mn, In, Jp*	Turkey	*Mn, In*
Ecuador	*Mn, In*	Nicaragua	*Mn*	Uganda	*Mn, In*
Egypt	*Mn*	Niger	*Mn*	Ukraine	*Mn, In*
El Salvador	*Mn, In*	Nigeria	*Mn*	United Arab Emirates	*Mn, In, Jp*
Estonia	*Mn, In*	Norway	*Mn, In*	United Kingdom	*Mn, Ms, In, Is, Jp*
Ethiopia	*In*	Oman	*Mn*	United States	*Mn, In, Jp*
Fiji	*Mn*	Pakistan	*Mn, In*	Uruguay	*Mn*
Finland	*Mn, In*	Panama	*Mn*	Uzbekistan	*In*
France	*Mn, Ms, In, Is, Jp*	Papua New Guinea	*Mn*	Venezuela	*Mn*
Gabon	*Mn*	Paraguay	*Mn*	Vietnam	*Mn, In*
Georgia	*Mn, In*	Peru	*Mn, In*	Yemen	*Mn*
Germany	*Mn, Ms, In, Is, Jp*	Philippines	*Mn, In*	Zambia	*Mn*
Ghana	*Mn, In*	Poland	*Mn, In, Jp*	Zimbabwe	*Mn*
Greece	*Mn, In*	Portugal	*Mn, In*		

Notes: *Mn* = national-level regressions of mobility; *Ms* = subnational-level regressions of mobility; *In* = national-level regressions of infections; *Is* = subnational-level regressions of infections; *Jp* = job postings.

Covid Economics 66, 28 January 2021: 37-57

Covid Economics **Issue 66, 28 January 2021**

The great employee divide: Clustering employee 'well-being' challenge during Covid-19

Jacques Bughin,[1] Michele Cincera,[2] Dorota Reykowska,[3] Marcin Zyszkiewicz[4] and Rafal Ohme[5]

Date submitted: 17 January 2021; Date accepted: 18 January 2021

The Covid-19 pandemic has triggered unprecedented levels of disruption and stress for workers. Still, little is relatively known about the state of mind of the workforce, even if its well-being is increasingly recognized as a driver of productivity. This paper encompasses multiple forms of stress – health, economic, social, and psychological – faced by the workforce, and demonstrates that not only have workers been facing large levels of stress during the Covid-19 pandemic beyond health issues, but that stress is not uniformly distributed among workers. While it is known that Covid-19 has been building a divide between remote and on-site workers, we uncover a much larger divide than the ones induced by work location alone, with the divide being due to different perceptions of mix and level of worries. Human resources practices may have to be much more personalized and include all forms of stress to diagnose the level of workers' state of fragility if they wish to create a much more resilient and productive workforce.

Covid Economics 66, 28 January 2021: 58-88

1 Université libre de Bruxelles, Solvay Brussels School of Economics and Management, iCite and ECARES; MachaonAdvisory.
2 Université libre de Bruxelles, Solvay Brussels School of Economics and Management, iCite and ECARES.
3 Neurohm.
4 Neurohm.
5 Neurohm and WSB University.

1. Introduction

According to the Worldometers database[1], the Covid-19 pandemic has affected more than 60 million people worldwide by early December 2020, resulting in more than 10 million hospitalizations, and 1.5 million official deaths. This figure matches the worse cases of annual flu and is already twice as large as the global pandemic of H1N1 by 2009 (Bughin, 2020).

While waiting for an effective vaccine, half of the governments on the planet have taken radical measures of quarantines. In effect, they also have shut down a material part of economies, through significant reduction of the face-to-face business interactions, and accelerating the shift towards home working.

The debate has been growing over the financial risk of shutting down businesses, especially small businesses (e.g. Bartik et al., 2020). The International Labour Organization (ILO) recently warned that more than 400 million enterprises were facing high risks of serious disruption worldwide, due to Covid-19.[2] Still, the debate has not focused *inside firms*, on *the perception* of the working population.[3] This is rather surprising, as most economies' sources of added value are still driven by the labor force, and the workforce is not doing that well (Pfefferbaum et al., 2020). Recent US research has for instance publicized a 3 times surge in burn-out among the full US working population, of which 75% of the surge can be traced to the Covid-19 pandemic[4]. Academic studies confirm this surge in stress, in most exposed occupations such as medical workers, teachers, or security forces (Sokal et al., 2020; or Dinibutun, 2020).

Further, there is a clear link being made in the literature between workforce well-being and labor productivity boost. A recent study by DiMaria et al. (2020) pervasively shows

[1] Coronavirus Update (Live): 59,027,330 Cases and 1,394,240 Deaths from Covid-19 Virus Pandemic. Worldometer (worldometers.info).

[2] See Covid-19: Stimulating the economy and employment: ILO: As job losses escalate, nearly half of global workforce at risk of losing livelihoods. These enterprises are operating in the hardest-hit economic sectors, including some 232 million in wholesale and retail, 111 million in manufacturing, 51 million in accommodation and food services, and 42 million in real estate and other business activities.

[3] Exception include VanderWeele (2019); Carnevale and Hatak (2020).

[4] FlexJobs, Mental Health America Survey: Mental Health in the Workplace

Covid Economics 66, 28 January 2021: 58-88

that European countries lagging in workers' wellbeing may gain up to 4% of productive efficiency for each extra point increase in subjective wellbeing.[5]

Many firms have been working hard to put into place health prevention measures against Covid-19, including the option of home-working. This strategy is not only guided by government directives but may be optimal to the extent that workers are afraid to come back to work, even if remote working limits productivity gains (Bughin and Cincera, 2020; Rahman, 2020). In Silicon Valley, normally a location that champions pushing for back-to-work, *70% of the tech* professionals have expressed fears of returning to work on-site[6], the so-called FOG (fear of going back to work).[7] Another September 2020 survey commissioned by the work platform Envoy, found that about 3 out of 4 US employees remained worried about going back to work on-site.[8]

The fact that about 40% of workers, and most notably, managers, were able to work from home, without health fear, "in the comfort of their home", while some workers had no choice but to work on-site. This has led to the debate about the divide created by the Covid-19 pandemic (see Dingel and Neiman, 2020; Sostero et al., 2020).

But the divide issue is not exclusively about health (and FOG). Divides may have different flavors. Divide may emerge from the difference of impact of Covid-19 on job preservation and finance or divide may arise in terms of ability to protect close ties. Regarding the former, ILO recently warned in its *ILO Monitor third edition: Covid-19* that a drop in working hours in the current (second) quarter of 2020 would be in the range of a 10.5% deterioration, equivalent to 305 million full-time jobs at risk, and will

Covid Economics 66, 28 January 2021: 58-88

[5] France is set to gain the most among the EU countries analysed.
[6] See https://spectrum.ieee.org/view-from-the-valley/at-work/tech-careers/coronavirus-is-triggering-fear-of-going-to-work.
[7] As result of Covid-19, for instance, ILO has developed multiple advices as to how maximize the returns to teleworking technologies, see:
https://www.ilo.org/global/about-the-ilo/newsroom/news/WCMS_739879/lang--en/index.htm
[8] https://envoy.com/content/new-survey-highlights-employees-fears-about-returning-to-work/

especially affect workers in consumer-facing industries, with lower skills, and in countries, where most employment is self-employment.[9,10]

Regarding the latter, on top of mental health issues, the largest source of stress for medical workers is usually their family, rather than self-worry (see Vagni et al., 2020).

To our knowledge, this article is the first one to look at a comprehensive *fragility* of the workforce, through a broad look of risks the workforce perceives from the Covid-19 crisis. Risks assessed include job and financial risk (micro-economic risk), basic needs provisioning risk (supply chain), violence and psychological risks (social), country finance (macroeconomic risk), on top of health risk. The analysis covers 5 countries in continental Europe (Italy, Spain, France, Sweden, and Germany) so that one can sort out, country effects from common risk effects.

The first insight is that health risk (including about self and third parties like close family) is important but accounts for just above 40% of all risks expressed by the workers. Clearly, there is more than health that stresses the workplace. Second, and as expected, job and financial preservation risk is clearly important for a set of the working population and in all cases, is indeed a more important matter than for the non-working population, e.g. the retired, or the unemployed (who are already without a job, or voluntary unemployed).

Third, fragility is not evenly distributed. Resorting to clustering analysis, we uncover five major segments of the working population concerning the amount and profile of risk perception encountered. 45% of the working population has a large breadth of worries, and in our wording, is rather fragile. There is also large polarization as another, smaller cluster group than the worried workforce one, composed of 17% of the European

[9] See Covid-19: Stimulating the economy and employment (ILO): As job losses escalate, nearly half of global workforce at risk of losing livelihoods. These enterprises are operating in the hardest-hit economic sectors, including some 232 million in wholesale and retail, 111 million in manufacturing, 51 million in accommodation and food services, and 42 million in real estate and other business activities.
[10] This is only a first-order effect, as damages may drag alone. Ten years after the sub-prime 2008 crisis and the Lehman Brothers bankruptcy, about 60% of countries still have an output trajectory below pre-crisis levels, according to research by the IMF. Likewise, consumption might be pressed downwards during and post covid-time, leading to a demand shortfall, and a risk of recession fuelling a new wave of unemployment (Barro et al., 2020)

population seems to have limited risk perception, except *outside* of work (and related to social violence).

This distribution of risk can be traced to a set of employees' features, e.g. the segment that is more worried about job and finance has a higher portion of workers with lower education, less income generation, and are more in the midlife career. Those only worried about social violence, tend to be more of higher education. But also, macro-elements shape (non-) fragility, i.e. the segment whose worries are essentially more health than wealth-related is hopefully trusting the health system better than the other segments. On the negative side, those whose main worry is job preservation and finance stabilization are the least to trust the governmental actions so far in handling the Covid-19 crisis.

Last, but not least, the results are based on a representative sample survey, conducted online, but relying on respondents' statements. In order, to limit any bias, we use response time online, to adjust survey answers, based on the neuroeconomics principle that response time is an indicator of attitude strength (see Fazio et al., 1989). As we correct for this response time, we essentially make answers re-centered towards a neutral response. Thus, our statistically significant results are reinforced by this procedure.

The paper reads as follows. The next section discusses the methodology and sampling. Section 3 discusses the clustering analysis and implications. Section 4 concludes.

2. Background and sampling

The background of this research is a part of an extensive multinational Covid-19 Fever project aimed at understanding people's attitudes, emotions, and behaviors connected with the pandemic. The full list of questions is described in Appendix 1 to this paper.

The focus of the research is on understanding people's perception of disruption and stress brought by the pandemic, as it is well-known that risk perception may support larger protective behavior against the virus exposure (Wise et al., 2020; Harper et al.,

Covid Economics 66, 28 January 2021: 58-88

2020). The general point is that *individual* behavior is badly needed to limit the *social* diffusion of a fast reproducing Covid-19 (see Viceconte and Petrosillo, 2020).[11]

Using the same data set of this article, a companion paper (Bughin et al., 2020a) supports the link between risk and protection, but emphasize that the intensity and type of protection, as well as the intensity of the link with risk, is not homogeneously distributed in the population, casting doubt on « one size fits all » analysis.

For this paper, we also look at the heterogeneity of behaviors during the first wave of the Covid-19, at its peak of April 2020 in European countries, but focuses on *the workforce* population. The workforce population is typically 50% of all citizens, and in aggregate, labor is one of the largest drivers of productivity growth for our economies. Yet, the focus on how Covid-19 affects this specific population has been rare to date. Risks we look at are furthermore not only health risks, but other risks specific to work, such as fear of job and financial stability among others.

2.1. Data sampling and scope

We focus on Europe. Five countries are being analyzed: France, Germany, Italy, Spain, Sweden. Those countries are the largest of Europe and/or are representative of different socio-economic models (Esping-Andersen, 1999), as well as have been chosen because they stand for different archetypes of policy responses to the Covid-19 crisis.

The data collection was performed online[12], based on country representative samples for age (above 18 years old) and gender, and recruited via a panel agency in April 2020, with a total sample of more than 5,000 answers, or a minimum of 1,000 per country. Considering employees only, the total sample is just above 2,780 employees across 5

[11] With a Ro of 2-3, the Covid-19 reproduction rate implies a fast rate of diffusion. This is similar to the 2002 SRAS, with Ro was in the range of 2.2 to 3.6, but say much lower than the MERS-Cov, which broke out in the Middle East by 2012, and with a confined Ro, at less than 0.5 in Saudi Arabia and Middle East. Ebola by 2014, is said to have a reproduction rate, Ro between 1.5 to 2.3. The 1918 Spanish influenza Ro was estimated imprecisely between 1.8 to 4 (this was the case for Covid too, as the Ro range published varied between 1.5 to 6.5, with a mean of Ro = 3.3).

[12] We would like to thank Neurohm and Syno for collecting the data in all the countries.

countries (Table 1a), or 55% of the sample, a ratio in line with workers participation in the 18+ population of those countries.

Table 1a. Number of (employee) respondents and demographic split by country

	Total	Gender		Age			Total employees
	N	Females	Males	18-35	36-49	50+	N
FRANCE	1,024	51%	49%	29%	28%	43%	639
GERMANY	1,017	49%	51%	27%	24%	50%	535
ITALY	1,021	51%	49%	26%	30%	44%	507
SPAIN	1,019	50%	50%	32%	32%	36%	635
SWEDEN	1,006	51%	49%	30%	20%	49%	466

Table 1b provides high-level demographics of employees in the sample. In terms of largest frequency, the sample is also well representative of Western Europe - that is, it contains more male (53%) workers, whose age is between 36-49 years range (37%), who are relatively well educated (35%), have one child, and a monthly income above 2,000 Euro per month (71%). Note that by the time the sample got collected (by April and May 2020), about 26% of employees reported knowing at least someone being infected by the Covid-19. The sample is balanced in terms of traditional left/right political orientations.

Respondents got email invites and were informed about the study scope. No personal data were collected. The task of the respondents was to evaluate if they agree with the statements presented on the screen.[13] To avoid people being « forced » to respond, or respond with answers that are not reflective of actual behavior, each question was structured to respond, on a 3 point scale (yes, hard to tell, no) with hard to tell allowing not to force an answer.

A caveat of surveys is the uncertainty of the fit between what people report and their actual attitudes/behaviors. This is critical in a study like this one, as results may lead to managerial human resources or broader public policy implications. We thus apply response time measurement, and adjust data, in line with Fazio et al. (1989) who find a high correlation between report and actual behavior among people with fast reaction time when expressing their opinions. iCode Smart test was used to collect the data

[13] See Appendix 1.

Covid Economics 66, 28 January 2021: 58-88

(Ohme et al., 2020), with response time (RT) collected for each answer. RT given with a latency lower than 500 milliseconds (ms) (suspected to be given randomly) or higher than 10,000 ms (suspected to have been given after distraction) were eliminated. In total, this amounts to only 0.52% of dubious responses.[14]

Table 1b. Employees high-level demographics and Covid exposure

Features	Types	Percent	Features	Types	Percent
Gender	Female	47%	Location	<100,000 inhab.	56%
	Male	53%		>100,000 inhab.	44%
			Income	<20,000€	29%
Age	<18	0%		>20,000€	71%
	18-25	7%		Don't want to answer	7%
	26-35	23%	Infected	Yes	26%
	36-49	37%		No	68%
	50-64	31%		Don't know	6%
	>64	2%		Don't want to answer	1%
Education	Primary schools	2%	Political orientation	Left	23%
	Middle school	8%		Right	26%
	Vocational	28%		Other	21%
	High school	26%		Don't associate with politics	21%
	Bachelor or higher	35%		Don't want to answer	8%
Kids	0 children	50%			
	1 child	25%			
	2 children	19%			
	3 children	4%			
	>3 children	1%			

To account for individual differences in reaction speed, we standardize reaction time data measured in milliseconds, with STDRT being the z-score of log(RT), with mean = 0 and standard deviation = 1. We then build the variable, RTC, that takes into account both the explicit answer as well as the reaction time (RT) needed to produce the answer,

[14] Furthermore, to ensure high quality of data and eliminate test biases a calibration phase and control screen have been added. Calibration preceded the test phase and consisted of 3 steps:

a. Familiarization with the scale. The task of the respondents was to press certain answer options – this task made sure respondents are aware of the position of the buttons on the screen.
b. Familiarization with the purpose of the task. A few statements were presented describing the test and the task. After each screen respondents had to press a button. This part served as a motoric warm up.
c. Increasing the focus on the task. During the study a screen appeared asking the respondent to indicate the statement that was presented last. The aim of this task was to make sure respondents focus their attention on the presented statements. Such screen was presented twice.

The control screen was introduced to eliminate the effect of the position of the mouse on the screen. It was presented before each statement, forcing a standardized position of the mouse (the distance to the yes and no answers was always the same).

Covid Economics 66, 28 January 2021: 58-88

that is RTC' - 1/2 = (1 - a) x (Y - N)/4 (0 < RTC'< 1) where (1 - a) = max(SDRT, 2)/2 and Y-N is the difference between the portion of reported Yes and of reported N. Thus 0 < a < 1 acts as a factor that reduces the difference in responses, in the function of answer reaction time, which we call the confidence index. When RTC' converges to 50%, this implies either that everyone's answer oscillates around "Hard to tell", or simply because all the answers are not at all credible because of unusual reaction time. The more extreme RTC' value is, the stronger the survey answer is taking a firm position on the statement qualification asked in the survey, thus RTC = 0 is a strong and dominant NO, and RTC = 1 is an overwhelming YES. We notice here-after in Table 2 that the confidence index is not immaterial, and we thus use the adjusted responses as a more reliable dataset for our analyses in this paper.

2.3. High-level data statistics

2.3.1. Breadth of Worries

Remember that we look at four types of worries mostly, health (henceforth, H), economic (E), social (S), and psychological (P). Table 2 provides the RTC' value as well as the confidence index of answers, associated with each risk measures perception, ranked from the largest to the lowest, for the total sample, and from 16 constructs allocated to H, E, S, and P.

First, if one sums up all the RTC' values, the total goes to 9.1 out of 16, or a value of 56.8%. Clearly, a majority of worries prevails in the employee population, during the first wave of the Covid-19 pandemics. As our sample selects only employees, we can also compare the extent of worries to the one of non-employees, e.g. retirees or working-age people not working. There, the total for retirees is 7.8 (or 15% lower than employees), while it is 8.7 (6% lower than employees) for the other non-working population. Otherwise stated, the employee population expresses a broader risk than non-employees.

Looking at the different drivers, there is no surprise that half of the gap is linked to economic consideration, e.g. the largest difference between retirees and employees is by far job preservation risk. This worry for employees should be even more prevalent as the countries we cover, except for Sweden, had forced full blanket lockdown, with large pressure of economic activity (Coibion et al., 2020). But other (and expected) differences

Covid Economics 66, 28 January 2021: 58-88

still exist between employees and retirees. Not surprisingly, retirees are more worried about their health, and less so about social risk (as they tend to be more standalone).

Taking the average of the 4 constructs by type of risk, for the employee only, which is our focus here, we have that H = 61.2%, S = 59.5%, P = 56.0%, and E = 50.7%. All constructs are above 50%, meaning that each is majorly present in the employees' population.

H has the largest value but stands for only 40% of the total worries. H includes the two highest ranks in Table 2. Interestingly, worry about self is only average in the ranking, and the main risk is linked to people with a high risk of fatalities, like the older family members. This is consistent with other literature findings, e.g. Dryhurst et al. (2020).

Covid Economics 66, 28 January 2021: 58-88

Table 2. European employees worry during wave 1 of the Covid-19 pandemic

RTC'	Confidence	Statement
0.7	0.43	I am worried about the health of my older family members (H)
0.67	0.41	I am worried about the health of people in my country (H)
0.63	0.49	COVID-19 increases domestic violence (S)
0.62	0.61	The COVID-19 outbreak will make society more unequal (S)
0.60	0.53	I am worried that our country will run out of money (E)
0.60	0.49	I am worried about not being able to meet with my family (P)
0.59	0.52	COVID-19 will increase divorce rates(S)
0.57	0.52	I am anxious about not being able to meet with friends(P)
0.54	0.58	Living in isolation negatively impacts my wellbeing (P)
0.54	0.31	I am worried about my own health(H)
0.54	0.35	I am worried about the health of my children(H)
0.54	0.63	Being together all the time increases family tensions (S)
0.53	0.59	I worry how living in isolation will affect me (P)
0.52	0.35	I am worried about my financial situation (E)
0.47	0.36	I am worried about my job situation (E)
0.44	0.59	I am worried that of not enough necessities in the stores (E)

Economic elements, E, has the lowest risk value. Supply chain risk (as measured by necessary goods availability) is a minor risk, but the macro-economic risk of a country running out of money is a larger risk than personal risk, as we have noticed also for health. Psychological and social risks are clearly important too. Domestic violence and divorce rates are clearly signaled as a risk among the employee's sample.[15]

3.2.2. Contextual drivers of worries.

We also have collected responses linked to various attitudes and beliefs that may affect employee risk expression.

As we have a large list of statements (see Appendix 1), we first have applied Principal Categorical Component Analysis (CATPCA) to reduce the information. CATPCA was performed using Varimax rotation with Kaiser Normalization, to maximize the sum of the variance of the factor coefficients.

Ten factors were derived, which stands for 19.6% of the total variance. Table 3a provides the ten factors and associated dimensions, in order of how they emerge from the data rotation. Table 3b reports the RTC' and the confidence values, ranked from highest to lowest importance of the Factors.

Three factors (1, 4, 6) are linked to *third party trust*. The first is linked essentially to governmental institutions, the second is linked to healthcare, while the last relates to how people are reacting around the Covid-19 crisis. Factors 2 and 5 relate to *precautionary measures*, with Factor 2 encompassing the most important NPIs in terms of controlling the disease (Bo et al, 2020).

Factors 8, 9, 10 are all linked to the perception of a *lasting danger* linked to the virus.

Factor 10 relates to the duration of the crisis, Factors 8 and 9 relate to the vulnerability to the virus and the prioritization bias towards health versus wealth. Finally, Factors 3 and 7 are more social care about self and family.

[15] Note evidently that the risks measured concerns employees at their broad life - not only at work. But it can be correlated with work situation. For instance, domestic violence or divorce rate may arise from the fact that lockdown made people work at home, through teleworking, and that there is no escape from a close group.

Table 3a. PCA Factors from European employees' statements linked to the Covid-19 pandemic

Factor	Dimensions	Dimension loading
1. Trust in institutions	I am satisfied with how my government is handling this crisis	0.912
	The government is doing a good job dealing with COVID-19	0.908
	The government discloses real numbers of coronavirus infections and deaths	0.702
	[PRESIDENT] is doing a good job dealing with COVID-19	0.608
	Media provide reliable information about the pandemic	0.519
2. NPI compliance	I comply with the recommendations for physical distancing	0.683
	I comply with the restrictions to stay home	0.619
	I wash hands for 20 seconds when necessary	0.600
	I am grateful to our healthcare professionals	0.594
	I actively encourage others to follow the restrictions and guidelines	0.549
3. Social Fabric/citizenship	Since COVID-19 I exercise at home more	0.607
	Since COVID-19 I eat healthier	0.603
	I'm worried about my children's education	0.561
	I would like to help people who are more vulnerable to COVID-19	0.534
	COVID-19 will bring countries closer	0.482
	I worry that there will be an increase in break-ins and thefts	0.435
4. Trust in Healthcare	In case of coronavirus infection, I will get appropriate medical help	0.711
	I am grateful to our essential workers	0.599
	I am satisfied with how our healthcare system is handling this crisis	0.590
	I believe we will beat COVID-19 soon	0.304
5. Extra caution	I disinfect groceries before putting them away	0.902
	I disinfect mail and deliveries before opening them	0.899
6. Trust in people	COVID-19 reveals the worse in people	-0.788
	COVID-19 reveals the best in people	0.775
	People will stop following the restrictions soon	-0.571
7. Lifestyle maintenance	Since COVID-19 I exercise less	-0.745
	Since COVID-19 I eat unhealthier	-0.740
8. Vulnerability	My chance of getting COVID-19 is high	0.839
	Coronavirus is dangerous for my health	0.663
9. Covid a top priority	Media exaggerate the situation with COVID-19	-0.845
	Slowing the spread of COVID-19 is more important than the economy	0.434
	When a COVID-19 vaccine is available, I'd like to be vaccinated	0.375
10. Duration of Covid	The restrictions caused by COVID-19 will be over in a month	-0.807
	The restrictions caused by COVID-19 will continue at least until the fall	0.795

Notes: Variable Principal Normalization. Rotation Method: Varimax with Kaiser Normalization.

Regrouping by themes, Table 3a highlights good NPI compliance (average = 62%), even after correcting for likely over-statement in answers. In effect, the confidence level is the lowest of all themes. This overstatement might originate from appearing to obey the public mandate of quarantines and social distancing measures to limit the diffusion of the pandemics. Still, the RTC' < 100%, that is, "true" NPI compliance, is not complete, as

Covid Economics 66, 28 January 2021: 58-88

found in many studies (Zickfeld et al., 2020). A third-party trust is relatively well acknowledged (58%), yet people feel majorly vulnerable (56%). Lifestyle impact is felt more minor (48%). Expectations linked to the duration of the crisis is that it may be more short-term than long-term so that it seems that most European employees were not necessarily expecting the current second wave.

Table 3b. How employees perceive and act upon the Covid-19 pandemic

RTC'	Confidence	Factors
0.72	0.44	NPI Compliance
0.66	0.52	Trust in healthcare
0.57	0.51	Lifestyle maintenance
0.56	0.47	Vulnerability
0.55	0.54	Trust in people
0.54	0.50	Trust in Government
0.53	0.58	Covid top priority
0.52	0.52	Social fabric
0.44	0.49	NPI Extra caution
0.41	0.67	Crisis duration

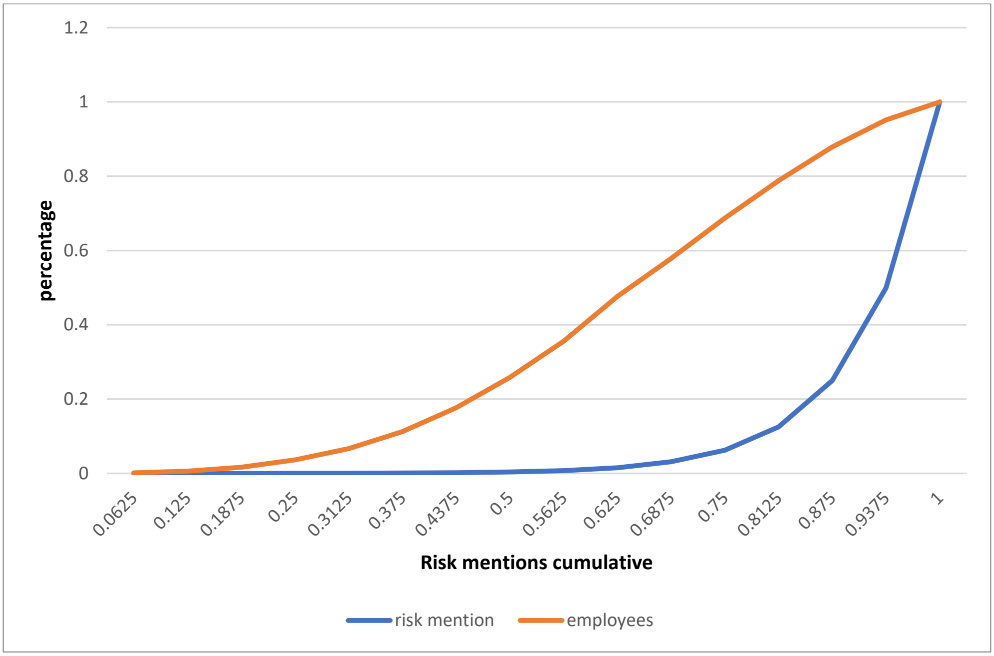
Trust is especially larger towards healthcare than for the government. It is nevertheless important that public authorities are trusted in their way of managing the crisis so that citizens actually adopt recommended protective actions (Li et al., 2018). Finally, vulnerability is more than less perceived by the employees' population and in general, there is a majority to think that the crisis will last until this time (as it did indeed).

3. Clustering results

One caveat of Table 2 about risk expressed by employees is that it only shows averages, but the key insight is not that people are worried about a pandemic, but more that the risk distribution is very wide. Figure 1 for instance plots the distribution of risk expression among the employees, and clearly the distribution is not uniform, building up a major divide; from Figure 1, one among others derives that 20% of employees express less than 50% of the type of the H, S, P, E risks surveyed, and 20% of others mention at least 81% of them during wave 1 of the Covid-19 pandemic. Furthermore, we find that 18% of employees make up 90% of all risks mentions, or more concentrated than a typical Pareto distribution.

Covid Economics 66, 28 January 2021: 58-88

Figure 1. Distribution of Covid-19 related risk expression among employees



3.1. Method

We resort to *clustering analysis* around the 16 elements that feature the four H, S, P, E risk domains. We have sued K-means clustering intending to partition the population into cohesive and stable segments, and in the hope to identify high risk and low-risk segments, as per Figure 1 above.[16]

The K-means technique minimizes the sum of square distances within each possible risk cluster to its centroid. Several analyses with different solutions of clusters number were conducted. The 5-cluster solution appears to be the most informative.

[16] To the best of our knowledge, the only study that segments risk attitudes is the one by Bodrud-Doza et al. (2020), in a study for Bangladesh. The authors demonstrate four homogenous groups linked to risk attitudes towards Covid-19, linked to health risks, socio-economic issues, and mental health problems. The study however only covers 340 people online, and given the country current digital development, is non- representative of the population. Finally, drivers of cluster belonging are not tested, which we do in our current study. We remind as well that we focus on employees only, where job risk may be acute, and for a large sample around Western Europe; among final innovations, we also have adjusted response rate for their confidence, based on large difference versus a base line of response time.

Table 4. K-means cluster size of European employees for different risks associated with the Covid-19 pandemic

Cluster	Total	Germany	Spain	France	Italy	Sweden
1	30.4%	19.8%	46.5%	32.1%	31.2%	17.8%
2	15.6%	10.7%	18.6%	13.8%	23.1%	11.8%
3	15.4%	21.1%	9.4%	13.9%	16.8%	17.4%
4	21.2%	23.0%	20.8%	25.8%	15.8%	19.5%
5	17.3%	25.4%	4.7%	14.4%	13.2%	33.5%
Total	100%	100%	100%	100%	100%	100%

Table 4 shows the size of the segments in aggregate, then, split by country. We see that the size of a segment is country dependent. Cluster 5 is the dominant one for Sweden and Germany, but the smallest one for Italy and Spain. As seen later, the 5th segment is composed of the least worried employees in contrast to the first segment. That Italy and Spain have such a large worried workforce can be traced to the fact that Italy and Spain have suffered relatively high contamination, healthcare under-capacity, and a largely enforced lockdown. This contrasts with Sweden, where no lockdown was applied, or with Germany where the healthcare capacity is rather large, and contamination was less spread than in the South of Europe.

3.2. Clusters details

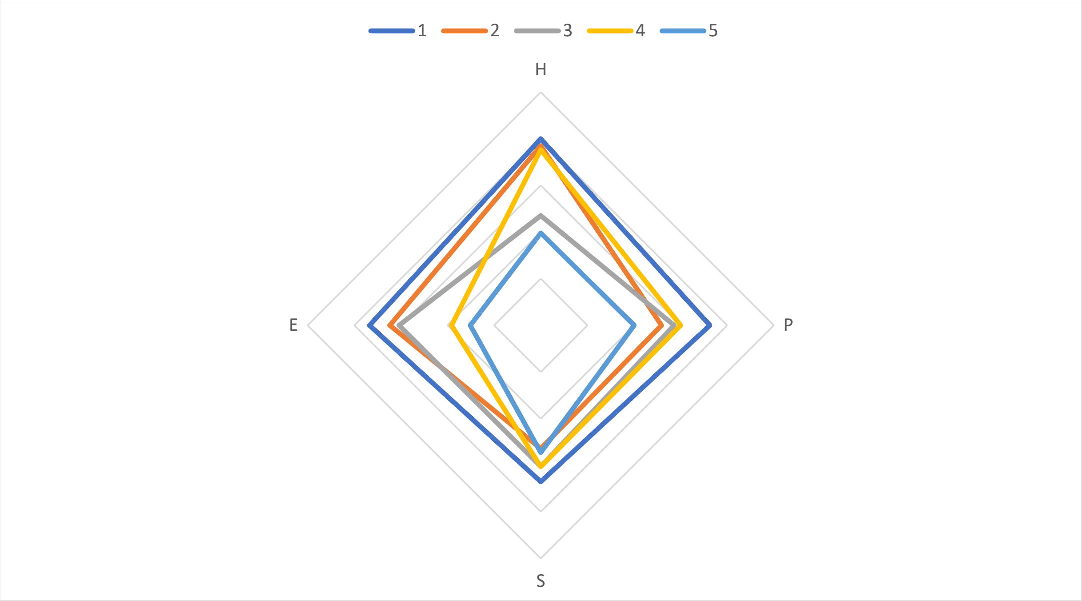
Table 5 provides the RTC' values by segment across the 16 risks analyzed, while Figure 2 aggregates the risk profile along with the four domains H, P, S, E.

Segments vary both in level and mix of risks expressed. As said, the risk expression level is the largest for the first segment. Risk mention then decreases along for each other segment.

The 5th Segment is the only segment with an average RTC' < 50%, Segment 5 has the lowest risk perception across all types of risk domains (See Figure 2), except that Segment 5 exhibits similar social risk perception as the average of other segments. Its main worry is social and linked to home violence and divorce rate.

Covid Economics 66, 28 January 2021: 58-88

Figure 2. **Covid-19 Risk profile radar by segment, European population**



Among the four other segments, Segment 1 has the largest risk perception across all dimensions. Segment 2 suffers less from the lack of social contacts than other segments, but this Segment expresses large concern across all other types of risk. Segment 3 perceives lower health problems than other segments, and finance is its key concern. Segment 4 has relatively low economic risk perception but is especially health concerned.

Those risk profile differences are striking. We see that the 3rd Segment has an opposite concern to the 4th Segment when it comes to the health-wealth trade-off. The 1st Segment is rather fragile, as the breadth of risk mention (12.9) is three times larger than the 5th Segment (4.3).

Figure 3 reports the distribution of risk per segment, normalized to the most risk-prone segment (Segment 1). The probability to mention more than 50% (= 8) of all worries, is just 15% for an employee in the 5th Segment but raises to 99% for the 5th Segment.

Covid Economics 66, 28 January 2021: 58-88

Table 5. Risk expressions linked to Covid-19 pandemic by European employee segments

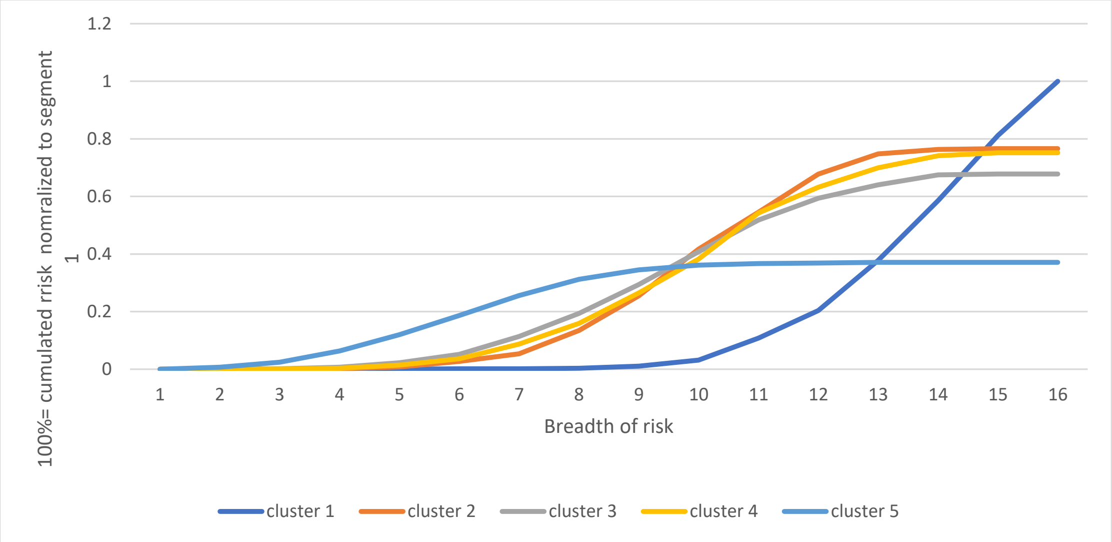
Segment	1	2	3	4	5
I am worried about my financial situation	80%	77%	74%	28%	22%
I am worried about my job situation	79%	73%	66 %	22%	20%
I am worried that our country will run out of money	73%	66%	62%	61%	46%
I am worried that there will not be enough basic necessities in the stores	62%	43%	42%	42%	33%
I am worried about my health	84%	82%	30%	74%	22%
I am worried about the health of my children	77%	71%	37%	69%	30%
I am worried about the health of my older family members	79%	78%	63%	79%	57%
I am worried about the health of people in my country	80%	77%	59%	80%	49%
I am anxious about not being able to meet with friends	71%	50%	59%	63%	41%
I am worried about not being able to meet with my family	77%	60%	58%	69%	40%
I worry how living in isolation will affect me	73%	50%	53%	52%	37%
Living in isolation negatively impacts my wellbeing	69%	48%	58%	56%	41%
The COVID-19 outbreak will make society more unequal	68%	60%	63%	64%	56%
Being together all the time increases family tensions	65%	47%	54%	53%	47%
COVID-19 increases domestic violence	67%	51%	65%	65%	61%
COVID-19 will increase divorce rates	68%	53%	59%	61%	54%
Total risk mentions out of 16	12.91	8.79	7.78	8.62	4.26

In fact, we can compute that about 85% of the 18% of employees accountable in Figure 1, for 90% of total volume stress expression belongs to the 1st Segment. This is an odd ratio of 85%/30% = 2.83, (where 30% is the share of employees in the 1st Segment). In contrast, the odd ratio is only 3%/17% = 17.6% (where 17% is the share of employees in the 5th Segment), or just above 1 chance of 6, for the less risk-prone 5th Segment.

Using further Figure 3 the probability to express more than 8 worries, over the 16 possibilities is just 15% for the 5th Segment, but still 70% for Segments 2, 3, and 4 and 99% for the 5th Segment.

Based on those distribution profiles, we can compute that about 85% of the 1st Segment makes up for the 18% of employees accountable in Figure 1, for 90% of total volume stress expression. This is an odd ratio of 85/30, or close to 3 for the Segment (where 30% is the share of employees in the 1st Segment in total).

Figure 3. Distribution of risk mentions by population segments



In contrast, the odd ratio is 3%/17%; or just above 1 chance of 6, for the less risk-prone 5th Segment. The segmentation allows thus to provide some significant information as to the skewed distribution of risk expression among employees, especially the most fragile, as the latter has an 85% probability to be linked to the 1st Segment.

As a further cross-check to Table 5, Table 6 also correlates the compliance to key non-pharmaceutical interventions (NPI) and risk mentions; as higher risk perception would lead to more extensive use of NPIs (see Bughin et al., 2020a; Harper, 2020; or Hammond, 2020 among others). This is indeed what we observe as a simple indicative log-log regression running from employee risk expression to her NPI compliance has a largely positive, highly significant elasticity (2.9, $p < 0.01$, $R^2 = 0.71$).[17]

Especially, the 5th Segment is less compliant to any measure as it suffers from the lowest level of risk. Segments 1 and 2 actually prefer to be quarantined or prefer the least interactions possible as they bear the largest burden of risk.

[17] Equation controls for employee socio-demographics from Table 2, and country dummies. A log-log specification is used as per the prevailing distribution of risk mention, and because typical risk aversion is said to be exponential.

The 2nd Segment is the most compliant to quarantine as it also suffers relatively less from being alone. The 3rd Segment also has a relatively low health concern, and more job preservation issue, so that it complies more with social distancing than quarantine.

Table 6. NPI compliance by European employees, per risk segments

Segment	1	2	3	4	5
NPI	77%	78%	70%	76%	68%
I comply with the recommendations for physical distancing	77%	78%	72%	78%	72%
I comply with the restrictions to stay home	79%	81%	68%	76%	64%
I wash hands for 20 seconds when necessary	76%	75%	69%	75%	69%

3.2. Cluster belonging

The above demonstrates that a factor such as NPI compliance is a good marker as to where an employee lies in terms of a risk-perceptual segment. Here, we formally test factors as markers of segmentation. We include socio-economic drivers, as they are well known to impact attitudes and risk expression (e.g. Dryhurst et al., 2020; Papageorge et al., 2020), as well as condition the ability to work remotely (Sostero et al., 2020).[18]

The detail of each logit regression per segment is presented in Appendix 2. Table 7 synthesizes the results, presenting only markers that are statistically significant at the 10% statistical threshold, and omitting country effect and constant. For simplicity, we also have regrouped factors into 4 major categories (trust, NPI compliance, vulnerability, and lifestyle). A negative sign means a lower impact on the probability to belong to a segment.

Regarding socio-demographics, neither the kid's family composition nor gender has any impact. Income, type of location (countryside or not), as well as education, play a role, as expected.

[18] In this sample,we neither have information on the rank of the employees, nor her work status (part versus full time, and home or site working). In other work (Bughin and Cincera, 2020), we test this specifically in the context of the French market. On-site workers are indeed significantly more health-stressed related (a FOG effect). Higher rank employees are less prone to risk perception, but a part of this is linked to their higher propensity to work from home. In general, the work location effect exists, but is a minor driver of the full risk perception.

Low income (less than 2,000 Euro per month) reduces the likelihood to belong to the 5th Segment. One reason, already highlighted in the introduction of this study, is that lower-income is often associated with essential work, exposing people more to health risks and vulnerability to the virus. Leaving in the countryside (in places with less than 100,000 inhabitants) makes an employee less likely to belong to the 3rd Segment. Education achievement plays on the likelihood to belong to various segments.

Table 7. Probit estimates of risk segment belonging

Segment	1		2		3		4		5	
	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.
1.Trust in institutions	-0.2	0.082	0.15	0.09			-0.24	0.08	0.45	0.1
4. Trust in Healthcare			-0.4	0.13	-0.33	0.13	0.67	0.13		
6. Trust in people	-0.42	0.123	0.24	0.13					0.26	0.15
2. NPI compliance	0.7	0.123	0.64	0.16	-0.46	0.12	0.51	0.13	-0.63	0.14
5. Extra caution	0.15	0.058	0.28	0.06	-0.27	0.07	-0.17	0.06	-0.16	0.08
8. Vulnerability	0.72	0.123	0.51	0.09	-0.72	0.08	0.37	0.07	-0.91	0.09
9. Covid top priority	0.46	0.067	-0.5	0.13						
10. Covid duration										
Social Fabric/citizenship							-0.31	0.11	-1.77	0.16
Lifestyle maintenance				0.07	0.16	0.07	-0.15	0.07	-0.48	0.09
Primary school	-0.74	0.442	0.85	0.38						
Middle school			0.38	0.21						
Vocational									-0.26	0.16
High school			0.24	0.14						
< 100,000 habitants					-0.29	0.12				
< 2,000 euros/month									-0.53	0.26

Of interest are the markers of trust, NPI, vulnerability, and lifestyle. As expected, vulnerability perception is a significant discriminant across all segments, as it drives health risk.

NPI compliance (and extra caution) are behaviors emerging out of health risk, but we see that they play a role on top of vulnerability. One reason is that NPI has been imposed as a government mandate response, and thus, NPI here also captures the compliance to authoritative measures.

Covid Economics 66, 28 January 2021: 58-88

Trust matters for each segment too, and especially the mix determines what segment an employee will belong to. Segments 1 and 3 are especially less inclined to accept their government actions to fight the Covid-19 pandemic. The 1st Segment is also a segment that is more trusting its peers than institutions for example.

Using the exponential of point estimates of Table 7, we can compute the marginal probability impact for the four categories of markers in Table 8. It becomes apparent that markers can truly discriminate among segments. Consider one employee among many with low institutions trust, which further complies to NPI, and is feeling vulnerable to the virus, belongs to the 1st Segment. The true mirror opposite belongs to the 5th Segment. Those two segments are also the most and least fragile among employees. One who is not scared too much about the virus trusts the healthcare system and complies with NPI while keeping its healthy lifestyle habit belongs to the 4th Segment.

Table 8. How markers determine risk segments associated with Covid-19

Segment	1	2	3	4	5
Trust	-18%	4%	-9%	25%	29%
NPI	59%	61%	-30%	26%	-31%
Virus dangerouness	55%	9%	-17%	15%	-20%
Lifestyle	0%	48%	66%	53%	-7%
Total	96%	122%	10%	118%	-29%

What is also crucial about those segments is that the 5th Segment may be less fragile, but may cause a risk to other segments, given low NPI compliance. Likewise, the 4th Segment may be ok with the Covid related health situation but is more stressed about jobs. The 1st Segment, and to a lesser extent the 2nd Segment, are rather stressed, and the physical and psychological health and, to a lesser extent, wealth are three considerations that employers should consider keeping those segments productive.

3.3. Country specificities

A final note concerns differences among countries, as, among others, it has been seen that the most / read least fragile segments size happens in South/ read Northern, Europe.

Covid Economics 66, 28 January 2021: 58-88

One reason for this is likely because of contrasts in the type of lockdown imposed, in the urgency of the sanitary crisis, as well as incapacity of the healthcare system.

In general, there could also be some key country nuance within segments. Table 9 illustrates the markers' impact on the most fragile employee segment in the five respective countries. As for the total, higher NPI compliance, higher Covid-19 vulnerability perception, strong social fabric, or lifestyle maintenance, are common to all countries.

Still, Spain discriminates in terms of trust in people versus the government, Sweden in terms of NPI compliance and social fabric, while German employees in the 1st Segment are biased towards more health than wealth in terms of social priority. Again, those can be traced to culture and background. Sweden has a large social culture versus the other countries (Esping-Andersen, 1999), and has not imposed lockdown. Thus, NPI compliance by Swedish employees is likely to be a more clear-cut discriminant behavior than in countries where NPI has been imposed.

Table 9. Marginal probability to belong to most fragile segment (Segment 1)

	Sweden	Germany	France	Spain	Italy
Trust in government				-33%	
Trust in healthcare					
Trust in people				95%	
NPI compliance	369%	191%	208%	132%	252%
Extra caution					
Vulnerability	247%	300%	185%	155%	278%
Top priority	0%	252%	0%	87%	136%
Duration					
Social fabric	753%	278%	397%	480%	272%
Lifestyle maintenance	116%	99%	107%	116%	108%

Note: Only statistically significant coefficients at 10% included.

Country differences thus prevail, but in general, a large set of common drivers allows to segment the workforce fragility and state of mind, across different countries.

4. Discussion and conclusion

For the workforce of an economy, the total number of physical contacts at work may be as important as the number of contacts at home. For the society at large, employees thus stand for a non-negligible channel of large contagion hazard and risk for absenteeism for companies. Further, the risk is not only health-based, and is much broader, including psychological stress, or job preservation worries, that, if not accounted for, may adversely affect productivity. The later worries can remain even if people telework, in which case, other stress may emerge, like at-home violence, or more.

We confirm in this research that the type of stress affecting the workforce is rather broad and that it goes beyond the pure physical health effect of the pandemic of Covid-19. Among 16 indicators of stress, the average worker reports to be affected by more than 9. We further show that the fragility of the workforce is not evenly distributed, with close to 20% of employees bearing 90% of the breadth of risk mentions by workers.

Using clustering techniques, we find five clear-cut segments that can be identified through a set of key markers. Those markers give not only an indication of fragility but how Human resources should engage in the appropriate selective dialogue with the various workers. For the human resources of companies, this is a potentially powerful tool to better engage with the workforce, improve their well-being during this pandemic. This is not only a question of corporate responsibility. This is one that may help keep high productivity and resilience for companies.

Covid Economics 66, 28 January 2021: 58-88

References

Bartik A.W., M. Bertrand, Z. Cullen, E.L. Glaeser, M. Luca, C. Stanton (2020). The impact of COVID-19 on small business outcomes and expectations. Proceedings of the National Academy of Sciences, 117(30):17656-17666.

Barro, R., J. Ursua, J. Weng (2020). Coronavirus meets the Great Influenza Pandemic. Voxeu.org, 20 March 2020.

Bo, Y., C. Guo, C. Lin, Y. Zeng, H.B. Li, Y.S. Zhang, S.Y. Wong (2020). Effectiveness of non-pharmaceutical interventions on COVID-19 transmission in 190 countries from 23 January to 13 April 2020. International Journal of Infectious Diseases.

Bodrud-Doza, M., M. Shammi, L. Bahlman, A.R.M. Islam, M. Rahman (2020). Psychosocial and socio-economic crisis in Bangladesh due to COVID-19 pandemic: a perception-based assessment. Frontiers in public health, 8, 341.

Bughin, J. (2020). Ten moments of truth in the Covid-19 Crisis Policy Punchline, Princeton.

Bughin, J., M. Cincera (2020). F.O.G. and teleworking: Some labor economics of Covid-19, Working Papers ECARES 2020-21. Université libre de Bruxelles.

Bughin, J., M. Cincera, R. Ohme, D. Reykowska, M. Żyszkiewicz (2020a). Perceptive risk clusters of European citizens and NPI compliance in face of the Covid-19 pandemics. Covid Economics 63: 126-158.

Bughin, J., M. Cincera, R. Ohme, D. Reykowska, M. Żyszkiewicz (2020b). The Worried, the Reckless, and the Carefree at Covid time: A cluster analysis of socio-economic and risks perception factors in France, iCite WP2020 - 038. Université libre de Bruxelles.

Carnevale, J.B., I. Hatak (2020). Employee adjustment and well-being in the era of COVID-19: Implications for human resource management. Journal of Business Research.

Coibion, O., Y. Gorodnichenko, M. Weber (2020). The cost of the Covid-19 crisis: Lockdowns, macroeconomic expectations, and consumer spending. Covid Economics 20: 1-51.

Dinibutun, S. R. (2020). Factors associated with burnout among physicians: An evaluation during a period of COVID-19 pandemic. *Journal of Healthcare Leadership*, *12*, 85.

Dingel, J and B Neiman (2020) "How Many Jobs Can be Done at Home?," Covid Economics: Vetted and Real-Time Papers 1, 3 April

DiMaria, C.-H., C. Peroni, F. Sarracino (2020). Happiness matters: Productivity gains from subjective well-being. Journal of Happiness Studies, 21(1): 139-160.

Dryhurst, S., C.R. Schneider, J. Kerr, A. Freeman, G. Recchia, A.M. Van Der Bles, D. Spiegelhalter, S. van der Linden (2020). Risk perceptions of COVID-19 around the world. Journal of Risk Research: 1-13.

Esping-Andersen, G (1999). Social foundations of postindustrial economies, Oxford University Press.

Covid Economics 66, 28 January 2021: 58-88

Fazio, R. H (1989). The role of attitude accessibility in the attitude to-behavior process. The Journal of Consumer Research, 16(3), 280–288.

Harper C.A., L. Satchell, D. Fido, R. Latzman (2020). Functional fear predicts public health compliance in the COVID-19 pandemic. Int. J. Ment. Health Addict (27): 1–14. 10.1017/dmp.2020.338

Li, Y.L., W.Z. Wang, J. Wang (2018). Government intervention behavior and optimization strategy of major epidemic control: Based on game theory and China's H7N9 prevention and control practice in 2013. J. Hunan Agri. Uni. (19) : 61–66.

Ohme, R., M. Matukin, P. Wicher (2020). Merging Explicit Declarations With Implicit Response Time to Better Predict Behavior. In Chkoniya, V., Madsen, A. O., & Bukhrashvili, P. (Ed.), Anthropological Approaches to Understanding Consumption Patterns and Consumer Behavior (pp. 427-448). IGI Global.

Papageorge, N. M. Zahn, M. Belot, E. van den Broek-Altenburg, S. Choi, J. Jamison, E. Tripodi (2020). Socio-demographic factors associated with self-protecting behavior during the Covid-19 pandemic, Covid Economics, 40, July.

Pfefferbaum, B., & North, C. S. (2020). Mental health and the Covid-19 pandemic. *New England Journal of Medicine*.

Rahman, A. (2020). Why can't everybody work remotely? Blame the robots, Covid Economics, 36.

Sokal, L., Trudel, L. E., & Babb, J. (2020). Canadian teachers' attitudes toward change, efficacy, and burnout during the COVID-19 pandemic. *International Journal of Educational Research Open*, 100016.

Sostero,M. S. Milasi , J. Hurley, E. Fernandez-Macías and M. Bisello (2020), Teleworkability and the COVID-19 crisis: a new digital divide? JRC Working Papers Series on Labour, Education and Technology 2020/05 A Joint European Commission–Eurofound Report

Vagni, M., Maiorano, T., Giostra, V., & Pajardi, D. (2020). Hardiness, stress and secondary trauma in Italian healthcare and emergency workers during the COVID-19 pandemic. *Sustainability*, *12*(14), 5592.

VanderWeele Tyler J. (2020). Challenges estimating total lives lost in COVID-19 decisions: consideration of mortality related to unemployment, social isolation, and depression. Jama 324(5): 445-446.

Viceconte, G., N. Petrosillo (2020). Covid-19 Ro: Magic number or conundrum, Infectious Disease Reports, (10)1. doi: 10.4081/idr.2020.8516

Wise T., T.D. Zbozinek, G. Michelini, C.C. Hagan, D. Mobbs (2020). Changes in risk perception and protective behavior during the first week of the COVID-19 pandemic in the United States. PsyarXiv (10).31234/osf.io/dz428

Zickfeld, J. H., T.W. Schubert, A.K. Herting, J. Grahe, K. Faasse (2020). Correlates of health-protective behavior during the initial days of the COVID-19 outbreak in Norway. Frontiers in psychology, 11.

APPENDIX 1. Tested statements

BEHAVIOR	
1.	I actively encourage others to follow the restrictions and guidelines
2.	I comply with the recommendations for physical distancing
3.	I comply with the restrictions to stay home
4.	I disinfect groceries before putting them away
5.	I disinfect mail and deliveries before opening them
6.	I wash hands for 20 seconds when necessary
7.	I would like to help people who are more vulnerable to COVID-19
8.	Since COVID-19 I eat healthier
9.	Since COVID-19 I eat unhealthier
10.	Since COVID-19 I exercise less
11.	Since COVID-19 I exercise at home more
12.	When a COVID-19 vaccine is available, I'd like to be vaccinated
EMOTIONS	
13.	I'm worried about my financial situation
14.	I'm worried about my job situation
15.	I'm worried that our country will run out of money
16.	I'm worried that there will not be enough basic necessities in the stores
17.	I am worried about my own health
18.	I am worried about the health of my children
19.	I am worried about the health of my older family members
20.	I am worried about the health of people in my country
21.	I worry that there will be an increase in break-ins and thefts
22.	I'm worried about my children's education
23.	I am anxious about not being able to meet with friends
24.	I am worried about not being able to meet with my family
25.	I worry how living in isolation will affect me
26.	Living in isolation negatively impacts my wellbeing
OPINIONS	
27.	The COVID-19 outbreak will make society more unequal
28.	Being together all the time increases family tensions
29.	COVID-19 increases domestic violence
30.	COVID-19 will increase divorce rates
31.	COVID-19 will bring countries closer
32.	I am grateful to our essential workers
33.	I am grateful to our healthcare professionals
34.	My chance of getting COVID-19 is high
35.	Slowing the spread of COVID-19 is more important than the economy
36.	Coronavirus is dangerous for my health
37.	Media exaggerate the situation with COVID-19
38.	Media provide reliable information about the pandemic
39.	[The President] is doing a good job dealing with COVID-19
40.	I am satisfied with how my government is handling this crisis
41.	The government is doing a good job dealing with COVID-19
42.	I am satisfied with how our healthcare system is handling this crisis
43.	In the case of coronavirus infection, I will get appropriate medical help
44.	The government discloses real numbers of coronavirus infections and deaths
45.	COVID-19 reveals the best in people
46.	COVID-19 reveals the worse in people
47.	I believe we will beat COVID-19 soon
48.	People will stop following the restrictions soon
49.	The restrictions caused by COVID-19 will continue at least until the fall
50.	The restrictions caused by COVID-19 will continue for about a month

Covid Economics 66, 28 January 2021: 58-88

Covid Economics 66, 28 January 2021: 58-88

APPENDIX 2. Probit estimates

CLUSTER 1 [K-Means 5 clusters for risk perception RTC][a]	B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
							Lower Bound	Upper Bound
Intercept	-2.478	0.203	149.006	1	0.000			
[Gender - Male=1,00]	-0.223	0.107	4.364	1	0.037	0.800	0.649	0.986
[Gender - Male=2,00]	0[b]			0				
[Kids - 0 children=1,00]	-0.293	0.112	6.823	1	0.009	0.746	0.599	0.929
[Kids - 0 children=2,00]	0[b]			0				
[Income - <20000€=1,00]	0.303	0.117	6.664	1	0.010	1.354	1.076	1.704
[Income - <20000€=2,00]	0[b]			0				
[Quarantine - yes=1,00]	-0.366	0.176	4.339	1	0.037	0.693	0.491	0.979
[Quarantine - yes=2,00]	0[b]			0				
Factor02_RTC - Compliance	0.517	0.154	11.321	1	0.001	1.677	1.241	2.266
Factor03_RTC - Social citizenship	1.428	0.126	129.322	1	0.000	4.171	3.261	5.335
Factor05_RTC - Extra caution	0.166	0.060	7.560	1	0.006	1.180	1.049	1.328
Factor06_RTC - Bad in people	0.585	0.120	23.933	1	0.000	1.795	1.420	2.269
Factor07_RTC - Lifestyle impact	0.641	0.070	83.711	1	0.000	1.899	1.655	2.179
Factor08_RTC - Percived vulnerability	0.570	0.083	46.882	1	0.000	1.768	1.502	2.081
Factor09_RTC - Fighting Covid top priority	0.594	0.125	22.384	1	0.000	1.811	1.416	2.315
Factor10_RTC - Predictions	0.298	0.121	6.087	1	0.014	1.347	1.063	1.708
GAP_INF [mean Std-RT from 16 risk perception attributes]	-4.019	0.292	188.861	1	0.000	0.018	0.010	0.032

a. The reference category is other clusters.
b. This parameter is set to zero because it is redundant.

CEPR PRESS

Covid Economics 66, 28 January 2021: 58-88

APPENDIX 2. Probit estimates

CLUSTER 2 [K-Means 5 clusters for risk perception RTC][a]	B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
							Lower Bound	Upper Bound
Intercept	-2.170	0.185	137.545	1	0.000			
[country_DE=1,00]	-0.462	0.171	7.261	1	0.007	0.630	0.450	0.882
[country_DE=2,00]	0[b]			0				
[country_IT=1,00]	0.474	0.137	12.067	1	0.001	1.607	1.230	2.100
[country_IT=2,00]	0[b]			0				
[Edu - Bachelor or higher=1,00]	-0.294	0.121	5.907	1	0.015	0.745	0.588	0.945
[Edu - Bachelor or higher=2,00]	0[b]			0				
[Infected - don't want to answer=1,00]	-20.349	0.000		1		1.455E-09	1.455E-09	1.455E-09
[Infected - don't want to answer=2,00]	0[b]			0				
Factor02_RTC - Compliance	0.628	0.151	17.220	1	0.000	1.874	1.393	2.522
Factor04_RTC - Trust in healthcare	-0.426	0.121	12.470	1	0.000	0.653	0.516	0.827
Factor05_RTC - Extra caution	0.327	0.060	29.954	1	0.000	1.387	1.234	1.560
Factor06_RTC - Bad in people	-0.270	0.119	5.147	1	0.023	0.763	0.604	0.964
Factor07_RTC - Lifestyle impact	-0.330	0.076	19.008	1	0.000	0.719	0.620	0.834
Factor08_RTC - Percived vulnerability	0.585	0.091	41.352	1	0.000	1.794	1.502	2.144
Factor09_RTC - Fighting Covid top priority	-0.449	0.129	12.041	1	0.001	0.638	0.495	0.823
GAP_INF [mean Std-RT from 16 risk perception attributes]	1.869	0.273	46.747	1	0.000	6.484	3.794	11.081

a. The reference category is other clusters.
b. This parameter is set to zero because it is redundant.

Covid Economics 66, 28 January 2021: 58-88

APPENDIX 2. Probit estimates

CLUSTER 3 [K-Means 5 clusters for risk perception RTC][a]	B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
							Lower Bound	Upper Bound
Intercept	-1.414	0.157	80.696	1	0.000			
[Gender - Female=1,00]	0.277	0.116	5.672	1	0.017	1.319	1.050	1.656
[Gender - Female=2,00]	0[b]			0				
[Age - 50-64=1,00]	-0.313	0.133	5.505	1	0.019	0.731	0.563	0.950
[Age - 50-64=2,00]	0[b]			0				
[Kids - 0 children=1,00]	0.458	0.118	15.050	1	0.000	1.581	1.254	1.993
[Kids - 0 children=2,00]	0[b]			0				
[Town - >100000 inhab.=1,00]	0.275	0.116	5.665	1	0.017	1.317	1.050	1.652
[Town - >100000 inhab.=2,00]	0[b]			0				
[Income - <20000€=1,00]	0.270	0.127	4.568	1	0.033	1.311	1.023	1.679
[Income - <20000€=2,00]	0[b]			0				
Factor02_RTC - Compliance	-0.426	0.118	13.018	1	0.000	0.653	0.518	0.823
Factor04_RTC - Trust in healthcare	-0.544	0.119	21.037	1	0.000	0.580	0.460	0.732
Factor05_RTC - Extra caution	-0.246	0.066	14.087	1	0.000	0.782	0.688	0.889
Factor07_RTC - Lifestyle impact	0.182	0.073	6.233	1	0.013	1.200	1.040	1.385
Factor08_RTC - Percived vulnerability	-0.759	0.079	91.549	1	0.000	0.468	0.401	0.547
GAP_INF [mean Std-RT from 16 risk perception attributes]	1.617	0.272	35.204	1	0.000	5.037	2.953	8.592

a. The reference category is other clusters.
b. This parameter is set to zero because it is redundant.

CEPR PRESS

Covid Economics 66, 28 January 2021: 58-88

APPENDIX 2. Probit estimates

CLUSTER 4 [K-Means 5 clusters for risk perception RTC][a]	B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
							Lower Bound	Upper Bound
Intercept	-2.262	0.168	181.827	1	0.000			
[country_IT=1,00]	-0.504	0.142	12.616	1	0.000	0.604	0.458	0.798
[country_IT=2,00]	0[b]			0				
[country_SE=1,00]	-0.401	0.141	8.102	1	0.004	0.670	0.508	0.883
[country_SE=2,00]	0[b]			0				
[Kids - 2 children=1,00]	0.254	0.119	4.580	1	0.032	1.289	1.022	1.627
[Kids - 2 children=2,00]	0[b]			0				
[Town - <100000 inhab.=1,00]	0.221	0.099	4.921	1	0.027	1.247	1.026	1.515
[Town - <100000 inhab.=2,00]	0[b]			0				
[Income - <20000€=1,00]	-0.341	0.115	8.840	1	0.003	0.711	0.568	0.890
[Income - <20000€=2,00]	0[b]			0				
Factor02_RTC - Compliance	0.684	0.130	27.704	1	0.000	1.981	1.536	2.555
Factor03_RTC - Social citizenship	-0.410	0.107	14.754	1	0.000	0.664	0.539	0.818
Factor04_RTC - Trust in healthcare	0.363	0.113	10.322	1	0.001	1.437	1.152	1.793
Factor05_RTC - Extra caution	-0.189	0.058	10.460	1	0.001	0.828	0.738	0.928
Factor07_RTC - Lifestyle impact	-0.146	0.063	5.344	1	0.021	0.864	0.763	0.978
Factor08_RTC - Percived vulnerability	0.399	0.074	29.431	1	0.000	1.491	1.291	1.722
GAP_INF [mean Std-RT from 16 risk perception attributes]	0.499	0.237	4.410	1	0.036	1.646	1.034	2.622

a. The reference category is other clusters.
b. This parameter is set to zero because it is redundant.

Covid Economics 66, 28 January 2021: 58-88

APPENDIX 2. Probit estimates

CLUSTER 5 [K-Means 5 clusters for risk perception RTC][a]	B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
							Lower Bound	Upper Bound
Intercept	-0.912	0.144	40.127	1	0.000			
[country_ES=1,00]	-0.622	0.222	7.876	1	0.005	0.537	0.347	0.829
[country_ES=2,00]	0[b]			0				
[country_SE=1,00]	0.537	0.147	13.251	1	0.000	1.710	1.281	2.283
[country_SE=2,00]	0[b]			0				
[Gender - Female=1,00]	-0.458	0.126	13.175	1	0.000	0.633	0.494	0.810
[Gender - Female=2,00]	0[b]			0				
[Income - <20000€=1,00]	-0.391	0.149	6.880	1	0.009	0.676	0.505	0.906
[Income - <20000€=2,00]	0[b]			0				
Factor01_RTC - Trust in Government	0.394	0.087	20.280	1	0.000	1.483	1.249	1.760
Factor02_RTC - Compliance	-0.555	0.127	19.069	1	0.000	0.574	0.448	0.737
Factor03_RTC - Social citizenship	-1.822	0.144	161.056	1	0.000	0.162	0.122	0.214
Factor06_RTC - Bad in people	-0.326	0.134	5.901	1	0.015	0.722	0.555	0.939
Factor07_RTC - Lifestyle impact	-0.456	0.084	29.378	1	0.000	0.634	0.537	0.747
Factor08_RTC - Percived vulnerability	-0.865	0.085	103.806	1	0.000	0.421	0.356	0.497
GAP_INF [mean Std-RT from 16 risk perception attributes]	1.119	0.305	13.458	1	0.000	3.063	1.684	5.570

a. The reference category is other clusters.
b. This parameter is set to zero because it is redundant.

Covid Economics Issue 66, 28 January 2021

COVID-19 closure and containment policies: A first look at the labour market effects in emerging nations

Michael A. Nelson[1]

Date submitted: 21 January 2021; Date accepted: 22 January 2021

Covid Economics 66, 28 January 2021: 89-114

Evidence is provided as to how government containment and closure policies in response to the COVID-19 pandemic in affected firm-level employment and hours worked and the differential employment impacts of such policies between men and women. The analysis uses data from the World Bank's Enterprise Analysis Unit survey of business enterprises owners and top managers located in 20 emerging nations about the impact that COVID-19 had on their business operations. Several principal conclusions are drawn from the analysis. First, containment and closure policies, viewed as a whole, impacted negatively permanent jobs and total hours worked at the firm level, but not temporary employment. Second, school and workplace closing policies increased the likelihood that firms reduced permanent employment, but the impact did not fall disproportionately on women. Third, public transport closings negatively impacted the employment prospects of all employment categories except temporary employment. Further, women were disproportionately affected by such polices. Fourth, policies directed at closure of public events had large negative effects across all employment categories, including temporary employment. Fifth, restrictions on internal movement negatively affected both permanent and temporary employment, but there is only weak evidence that such policies affect women disproportionally. Finally, at least for the set of emerging economies studied in this analysis, there is no evidence that international travel controls affected the likelihood that firms would reduce their total work hours, their levels of permanent and temporary employment, nor their reliance on women in their workforce.

1 Professor Emeritus, Department of Economics, University of Akron.

Covid Economics 66, 28 January 2021: 89-114

Introduction

Governments around the globe have put in place a whole range of containment and closure policies in response to the COVID-19 pandemic. These policies have varied considerably across governments, both in terms of their level of stringency and with respect to the length of time that authorities have allowed them to stay in place. At the same time, the merits of these policies have been widely debated in terms of their efficacy in constraining the spread of the disease ("flattening the curve"), their potential conflict with notions of individual liberty and freedom, and with regard to the collateral damage they inflict on the economy after they are put in place. In this paper the focus is placed on the latter, with the overriding goal of understanding better the economic consequences of containment/closure policies as it pertains to employment in private labor markets.

The specific research questions addressed in the empirical analysis below are twofold:

- Out of the menu of containment/closure policies that have been put in place, which ones have had the most effect on employment and hours worked?
- Has the effect of these policies on the workforce affected men and women differently?

In addressing these questions use is made of a COVID-19 related survey of business enterprises conducted by the World Bank's Enterprise Analysis Unit. In this survey, firms operating in 20 emerging nations were surveyed in the late spring and early summer of 2020 about the impact that pandemic had on their business operations.

The availability of unprecedented amounts of real time data tracking the global impact of the pandemic along with government responses to constrain its spread has already resulted in an impressive volume of literature in analyzing the economics of the disease.[2] While some of these papers have focused on the pandemic and labor market outcomes, the present paper is distinct in several important ways. First, it employs a data set on firm-level response to COVID-19 consisting of firms of all sizes. Second, the analysis is focused on emerging economies in contrast to most of the other papers that has addressed this general topic. Third, it assesses the comparative effect of a broad range of containment/closure policies. Fourth, the analysis distinguishes between impacts of these policies on permanent and temporary employment. Finally, an assessment is also made of the impact of policy responses to the pandemic on employment by gender.

Several principal conclusions are drawn from the analysis. First, containment and closure policies, viewed as a whole, impacted negatively permanent jobs and total hours worked at the firm level with little apparent impact on temporary employment. Second, school and workplace closing policies increased the likelihood that firms reduced permanent employment, but the impact did not fall disproportionately on women. The effect on women from such policies may have been mitigated to the extent they are considered "essential workers" and that many already have relatively high work-from-home occupations. Third, public transport closings negatively impacted the employment prospects of all employment categories considered in this analysis with the exception of temporary employment.

[2] Indeed, the pandemic has spawned a *Covid Economics* journal that features real-time papers on the topic. Since its launch in March 2020, over 60 issues of the journal have been published in that year alone.

Further, women were disproportionately affected by such polices. Fourth, policies directed at closure of public events have potentially large negative effects across all employment categories, including temporary employment. Fifth, restrictions on internal movement negatively affected both permanent and temporary employment, but there is only weak evidence that such policies affect women disproportionally. Finally, at least for the set of emerging economies studied in this analysis, there is no evidence that international travel controls affected the likelihood that firms reduced their total work hours, their levels of permanent and temporary employment, nor their reliance on women in their workforce.

The policy relevance of this research is clear as these closure and containment strategies will be relevant in confronting future waves of the COVID-19 and its variants that are expected to persist into the latter part of 2021. Furthermore, COVID-19 represents at least the sixth global pandemic since the Influenza Pandemic of 1918 and experts predict more frequent and deadlier pandemics going forward (IPBES, 2020). Understanding the economic costs of these closure/containment strategies will be important to policy makers as they select among the menu of options. It will also afford them the opportunity to modify these strategies or put in place support mechanisms to mitigate these costs.

Literature Review

While the global economy has recovered some since the beginning of the pandemic, overall employment remains below the levels experienced at the beginning of 2020. For example, in the US total employment was down nearly six percent in November compared to February of that year.[3] It has been observed that this downturn has been different from most past recessions where the manufacturing sector typically takes the brunt of economic collapse. Instead, in 2020 the sectors hit hardest included hospitality and travel, retail, culture and recreation, and the garment industry - the latter mainly attributed to export restrictions (Karabarbounis and Trachter, 2020). Women typically make up a relatively large share of employment in these industries (OECD, 2020; Schalatek, 2020; Alon, et. al, 2020) and hence their employment in these sectors has been disproportionately negatively affected by the pandemic.[4]

In regard to the employment effects of closure and containment measures, Gottlieb, et al., (2020) develop a measure of the ability of a country's workforce to "Work from Home (WFH)." They use this to simulate the GDP and employment effects of four lockdown scenarios ranging from complete lockdown (required workplace closing and can only work from home), to lockdown strategies that exempt the agricultural sector or other sectors that are deemed to be "essential," to a "soft" lockdown reflecting government easing of shutdown policies. They find the workers in low-income countries have less ability to work from home thereby affecting employment in these countries more, although the losses are mitigated to the extent the lockdown policies exclude the agriculture sector.

[3] Bureau of Labor Statistics, Current Population Survey, https://www.bls.gov/cps/home.htm. (accessed January 2021).

[4] While not directly relevant for the purposes of this paper, it has also been noted that women comprise nearly 70% of the health care workforce globally (e.g., nurses and midwives, long-term health care workforce) exposing them to greater risk of COVID-19 virus infection (OECD, 2020).

Covid Economics 66, 28 January 2021: 89-114

Using survey data of firms from 10 emerging economies conducted in the late spring and early summer of 2020, Beck, et al. (2020) study the impact of the pandemic on firm payroll and investment. Neither the impacts of specific closure/containment policies, nor possible differential consequences by gender, are considered in their analysis. They found that firms principally reacted to the pandemic by reducing investment rather than payroll, an outcome they attribute to the importance that firms placed on maintaining strong long-term relationships with labor.

Hupkau and Petrongolo (2020) employ household survey data from the UK to analyze how lockdowns affected employment, hours worked, and earnings in that country over the first half of 2020. They found that negative employment and hours worked outcomes are mitigated through previous WFH experience. Further, the likelihood of job loss is about the same for men and women, while women experience lower hours and earnings losses. This despite evidence pointing to women taking on most of the increased childcare hours resulting from the lockdown and school closings.

An analysis by the OECD (2020) also shows that employed women are likely relative to their male counterparts to spend more time on childcare and also have greater adult care responsibilities taking care of ill or elderly relatives during the pandemic. They further point out that outsourced home production such as childcare services often become more problematic during lockdowns and school closures and that "[m]uch of this additional burden is likely to fall on women." (p. 5).

Alon, et al. (2020), using survey data on American time use from 2017 and 2018, conclude that women will suffer more from typical lockdown measures because fewer females are employed in jobs that are highly telecommutable jobs and they are also less likely to work in critical occupations that may be exempt from lockdown measures. Further, the burden of childcare tends to fall disproportionately on women thereby likely to impact them more when school closure containment measures are imposed. They also emphasize that the childcare situation is exacerbated to the extent that daycare centers are subject to lockdown measures.

Brussevich, et al. (2020) construct a WFH measure for OECD countries and then show that the probability of remaining employed is directly associated with greater WFH capabilities using survey data from the US and Peru. Industries hardest hit by the pandemic such as retail, accommodation, and food services are least likely to work from home, hence these sectors are predicted to be among the hardest hit in terms of layoffs and furloughs. Further, since women tend to be heavily represented in these sectors, they are likely to be disproportionally affected by the pandemic. In contrast to Alon, et al. (2020), they conclude that this will be offset to some extent because more women than men are considered to be "essential workers".

Finally, the impact of regulatory restrictions on the movement of people across international borders has been consider by Benz et al. (2020). They focus on services-trade costs rather than employment impacts and conclude that such costs increase by an average of 12% when countries close borders to people, but not to freight.

To summarize, while the labor market effects surrounding COVID-19 has been addressed by several recent papers, more needs to be done, especially as it pertains to the employment impacts of specific policy measures designed to control the spread of the virus. As the World Bank notes, "many government measures to curtail contagion will have a direct impact on the private sector."[5] The goal of this paper to understand better these impacts and thereby build on the existing literature summarized above. This is accomplished through the use of a stratified random sample of firms and owner/manager assessment of the impacts of the pandemic on their enterprise over the first few months of the pandemic. In contrast to many other studies, firms of all sizes are considered in the analysis and the focus centers on emerging nations. Relative to the extant literature, a comparative assessment of a broad range of closure/containment policies is accomplished. Further, the data set employed in this analysis permits a separate assessment of containment/closure polices on temporary and permanent employment as well as the impact of policy responses to the pandemic by gender.

The model used to conduct this analysis is presented next.

Model

To assess the employment impact of COVID-19 containment and closure policies at the firm level, the following model is employed:

$$(\text{Employment Impact})_{ij} = \beta_0 + \beta_1[\text{COVID Policy}]_j + \beta_2\,[\text{Industry } (\mathbf{I})]_{ij} + \beta_3\,[\text{Firm Size } (\mathbf{S})]_{ij} + \alpha_i + \mu_{ij} \quad \text{...(1)}$$

where,

Employment Impact = One of five measures of employment change of firm i in country j during the early phases of the COVID-19 pandemic,

COVID Policy = a specific category of the government's containment and closure polices,

Industry (I) = *manufacturing*, *retail*, *other*,

Firm Size (S) = *small*, *medium*, *large*,

with α_j representing country-level fixed effects and μ_{ij} is the random disturbance term.

Based on available data, five alternative categories of firm-level employment change [Employment Impact] are considered in the analysis below. All are defined as a binary variable taking on a value of one if employment (or total hours worked) declined relative to pre-pandemic levels, and zero if the firm experienced zero or positive employment change over that time period. The five employment impact categories considered alternatively in the models presented below are as follows:

- Total Hours Worked
- Permanent, Full-time Workers
- Temporary Workers
- Female Permanent, Full-time Workers
- Female Share of all Permanent Full-time Workers

Further details on each of the five categories will be discussed in the data section below.

[5] https://www.enterprisesurveys.org/en/covid-19

COVID Policy is an index of the stringency policy in a country in one of eight areas (e.g., school closings, restrictions on internal movement). An overall stringency index capturing all containment and closure policies is also considered. These will be discussed further in the data section that follows. The firm size categorical variables (*small*, *medium*, *large*) are based on the number of permanent workers of the firm using the World Enterprise Survey definitions of firm size categories. Finally, a fixed-effect model specification is chosen to account for the other country-specific factors that might affect firm-level employment changes not directly accounted for elsewhere in the model.

Data

Data for model estimation are drawn from two sources. Firm-level data on employment and other enterprise characteristics are taken from the COVID-19 Follow-up Enterprise Surveys conducted through the World Bank.[6] These surveys were conducted during the months of May through July 2020 as a follow up to the regular business surveys they have periodically undertaken for many, primarily-emerging, nations since the early 2000s.[7] In the regular surveys business owners and top managers are asked about the characteristics, constraints and climate of their business operations in the country they are located. The Follow-up Surveys used in the present analysis focus on the impact of COVID-19 on the business operations of the respondents. At the time of this writing these surveys were conducted in 24 countries, although due to data limitations only 20 could be included in the analysis.[8] In all, the sample of usable data was approximately 7,000 firms, with the number of individual country surveys ranging from a low of 47 to a high of 780 with a mean of 146.

Firms operating in the manufacturing and service sectors are the primary focus of the survey, with services firms broadly defined to include construction, retail, wholesale, hotels, restaurants, transport, storage, communications, and IT. A limitation of the survey data is that enterprises operating in the informal sector are excluded from the analysis. Government employees are also not included. Beyond that, the sample is designed to be representative sample of firms operating in the private sector for that country.

Total hours worked and employment changes are based on the month the Follow-Up Survey was undertaken. Survey dates vary by country, with the earliest being May 2020 and the latest being August of that year. The time period covered to calculate the change is constrained by the specific questions asked in the survey. For the five measures included in the analysis, the periods considered are as follows:

- Total Hours Worked: The establishments total hours worked per week in the month preceding the survey relative to the same month in 2019.

[6] https://www.enterprisesurveys.org/en/covid-19.
[7] For further details, see: https://www.enterprisesurveys.org/en/survey-datasets.
[8] These countries are Albania, Belarus, Bulgaria, Chad, Cyprus, El Salvador, Greece, Guatemala, Guinea, Honduras, Jordan, Mongolia, Morocco, Nicaragua, Niger, Poland, Slovenia, Togo, Zambia, Zimbabwe.

Covid Economics 66, 28 January 2021: 89-114

- Permanent, Full-time Workers: The number of workers in the month preceding the survey compared to the end of December 2019.
- Temporary Workers: Response to the question: "Since the outbreak of COVID-19, has the total number of this establishments temporary workers increased, remained the same, or decreased?"
- Female Permanent, Full-time Workers: The number of female workers in the month preceding the survey compared to the end of December 2019.
- Female Share of all Permanent Full-time Workers: The share of female workers in the month preceding the survey compared to the end of December 2019.

Summarized in Table 1 are the details on how each of the five employment categories changed over the time period considered based on the Enterprise Survey data set used in this analysis. A majority of the firms in the sample (53.7%) reported that total hours worked in their establishment declined relative to the preceding year. As to the other four measures that deal directly with employment, more than half or the firms reported no change in employment levels since the outbreak of COVID-19. A sizable minority of firms, however, did experience employment declines while, not surprisingly, relatively few reported that employment in their establishment actually increased. Interestingly, the percentage of firms that reported the proportion of women (full-time) workers in their labor force increased (16.2%) was about the same percentage that reported a decrease (17.4%).

Table 1

Total Hours and Employment Changes in the Early COVID-19 Pandemic

(Percentage of Firms in Sample)

	Total Hours Worked	Permanent Full-time Workers	Temporary Workers	Female Employment	Proportion of Female Employment
Negative	53.7	29.7	29.1	20.4	16.2
Zero	43.4	63.8	67.2	75.3	66.5
Positive	2.9	6.5	3.8	4.3	17.4

Notes: See text for details.

The second source of data used in this analysis is drawn from the Oxford Covid-19 Government Response Tracker dataset.[9] Based at the University of Oxford, a wide range of measures that governments around the world are using to combat the spread of the virus and to deal with the related fallout on society and the economy are consistently tracked daily. In all, 17 factors are considered, addressing a government's response to the pandemic in three general areas: containment and closure policies, economic policies, and health system policies. For each indicator, efforts are made to account

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

Covid Economics 66, 28 January 2021: 89-114

for the degree of policy stringency and comprehensives of geographic area covered by the policy.[10] Of primary interest here is understanding better the employment impacts surrounding containment and closure policies at the firm level, including how it pertains to gender. These factors are grouped into eight policy areas and are summarized in the left-most column of Table 2. For each policy area, an index is calculated on an ordinal scale that goes from zero (no restrictions) to a high of 2 to 4 (depending on the policy) indicating maximum restrictions. Further, the analysis will also consider an overall *Stringency Index* which considers all eight containment/closure factors and policies pertaining to public information campaigns and weighs them equally in the construction of the index. It is scaled from zero to 100 with higher values indicating a stronger government policy response to the pandemic.

Table 2

Government COVID-19 Containment and Closure Policies

Policy	Ordinal Scale	Description of Low/High Scale Range	Sample Mean (Std. Dev.)
School Closing	0 to 3	0= no measures, 3 = required closing at all levels	2.28 (1.00)
Workplace Closing	0 to 3	0= no measures, 3 = required closing at all but essential workplaces	1.47 (0.92)
Public Events Cancellations	0 to 2	0= no measures, 2 = required cancelling	1.44 (0.77)
Restrictions on Gatherings	0 to 4	0= no restrictions, 4 = restrictions on gatherings < 10 people	2.30 (1.28)
Close Public Transport	0 to 2	0= no measures, 2 = require closing	0.62 (0.74)
Stay at Home Requirements	0 to 3	0= no measures, 3 = require not leaving home with minimal exceptions	0.93 (0.99)
Restrictions on Internal Movement	0 to 2	0= no measures, 2 = internal movement restrictions in place	1.11 (0.93)
International Travel Controls	0 to 4	0= no restrictions, 4 = bans on all regions or total border closure	3.16 (1.06)
Stringency Index	0 to 100	Overall index addressing government response in all eight areas listed above along with public information campaigns. Higher values imply stronger government response.	59.20 (25.15)

Notes: See text and sources below for further details.
Sources: https://github.com/OxCGRT/covid-policy-tracker/blob/master/documentation/codebook.md, https://github.com/OxCGRT/covid-policy-tracker/blob/master/documentation/index_methodology.md

[10] For further details, see: https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker and https://github.com/OxCGRT/covid-policy-tacker/blob/master/documentation/codebook.md.

Covid Economics 66, 28 January 2021: 89-114

Since the containment and closure polices were subject to daily changes in any country over the period analyzed, the approach taken here is to average the value of each indicator index over the preceding month leading up to the firm's current employment estimates. For example, the Enterprise Surveys' COVID-19 Follow Up Survey for Albania was conducted in June 2020. Survey respondents in this country were asked to use May 2020 as "the month preceding the survey" in answering the questions regarding employment and total work hours. Hence, the value of the "school closing" policy index used in this analysis for Albania was based on the average value of that index over the month of May 2020.[11] Likewise, a portion of the surveys in Poland were conducted in August. For those firms, the policy indices in Table 2 were based on averages for the month of July.[12]

The firm size variables are calculated using the number of permanent, full-time workers employed by the firm prior to the start of the pandemic (December 2019). Following Enterprise Survey conventions, *Small Firms* are defined as having 19 or fewer employees, *Medium Size* firms employ over 19 but less than 100 workers, and *Large Firms* are designated by having more than 100 employees prior to the start of 2020. In the sample, nearly half of the survey respondents (48%) were derived from small firms, one-third from medium-size firms and the remainder were drawn from large firms. In the Enterprise Survey COVID-19 Follow-up firms are classified into one of three areas - *Manufacturing*, *Retail*, and *Other Services*. Included in the latter are construction, hotels, restaurants, transport, storage, communications, and IT.

Results

Since the dependent variable is binary, estimation is carried out using the logit estimator with robust standard errors. Results for the key policy variables of interest for all five employment impact categories are summarized in Table 3. For example, in the upper left-hand corner in Panel A of the Table is the parameter estimate (0.01) using equation (1) for the *Stringency Index* policy variable when the binary dependent variable is based on whether or not the firm experienced a decline in Total Hours Worked in the months following the COVID-19 outbreak. The positive sign on the parameter estimate indicates that more aggressive overall containment/closure policies is associated with a greater probability that a firm experienced a decline in total hours worked, other factors being equal. Further, the finding is statistically significant at better than the 10 percent level.

To gain further insight into the significance of the COVID policy parameter estimates reported in the top panel of Table 3 the implied marginal effects stemming from more aggressive policy responses to the pandemic are summarized in Panel B in the lower half of the table. The calculations presented here are based on the effects from going from no containment/closure policy (0 on the ordinal scale) to the most aggressive policy (2, 3, or 4 depending on the specific policy category – see Table 2), holding all other variables in the model at their sample mean values. In the case of the overall *Stringency Index* the calculation is based on going from 10 (lowest value for any country/date in the data set) to 100 on the ordinal scale for that index. The upper-left corner of Panel B in the table shows that such an increase in the *Stringency Index* is associated with a 32% increase in the probability that a firm's total hours will

[11] Averages for each country were rounded to the nearest half point, although the results were substantially the same if the averages were not rounded.
[12] Not surprisingly, the correlation between the average index value for each of the policy measures listed in Table 2 was not trivial, ranging from a high of 0.87 to a low of 0.31.

Covid Economics 66, 28 January 2021: 89-114

Covid Economics 66, 28 January 2021: 89-114

Table 3
Firm-level Response to Government COVID-19 Containment and Closure Policies: Summary Impact on Various Dimensions to Employment
(Dependent variable: *Hours/Employment/Share Declined*)

COVID-19 Policy	Total Hours Worked	Permanent Employment	Temporary Employment	Female Employment	Female Labor Share
Panel A: Logit Model Parameter Estimates for COVID-19 Policy Variable					
Stringency Index	**0.01**** **(1.9)**	**0.03**** **(3.5)**	0.00 (0.6)	**0.02**** **(2.5)**	**0.02*** **(1.9)**
School Closing	-0.15 (0.6)	**0.42**** **(2.1)**	0.24 (1.5)	0.25 (1.2)	0.06 (0.3)
Workplace Closing	**0.25**** **(2.5)**	**0.25**** **(2.6)**	-0.01 (0.2)	0.10 (0.9)	0.04 (0.4)
Public Transport Closing	**0.81**** **(3.0)**	**0.93**** **(2.8)**	-0.04 (0.1)	**0.96**** **(2.7)**	**1.10**** **(2.6)**
Stay-at-Home Mandates	**0.18*** **(1.9)**	**0.25**** **(2.4)**	-0.04 (0.4)	**0.21*** **(1.8)**	0.13 (1.0)
Cancellation of Public Events	**1.40**** **(2.7)**	**1.34**** **(2.3)**	**1.67**** **(2.9)**	**1.42**** **(2.1)**	**1.66**** **(2.4)**
Restrictions on Gatherings	**0.53*** **(2.0)**	0.46 (1.6)	0.65 (1.4)	-0.43 (1.3)	-0.40 (1.1)
Restrictions on Internal Movement	0.34· (0.8)	**1.64**** **(3.2)**	**1.84**** **(2.9)**	0.53 (0.8)	1.52· (1.6)
International Travel Controls	-0.11 (1.5)	0.05 (0.6)	0.23 (1.1)	0.12 (1.3)	0.15 (1.5)
Panel B: Marginal Effects of Policy Change					
Stringency Index	32%	48%	7%	30%	22%
School Closing	-11%	22%	12%	10%	2%
Workplace Closing	19%	15%	-1%	5%	2%
Public Transport Closing	36%	40%	-1%	35%	36%
Stay-at-Home Mandates	13%	16%	-2%	10%	5%
Cancellation of Public Events	59%	41%	43%	31%	29%
Restrictions on Gatherings	48%	34%	41%	-28%	-22%
Restrictions on Internal Movement	16%	58%	58%	16%	38%
International Travel Controls	-11%	4%	14%	8%	7%

Table 3 - Continued
Firm-level Response to Government COVID-19 Containment and Closure Policies: Summary Impact on Various Dimensions to Employment
(Dependent variable: *Hours/Employment/Share Declined*)

Notes: Results presented in this table are for the key policy response variables only for each model estimated. Complete model results for each employment category can be found in Tables A1 – A5 in the Appendix. All models are fit using the logit estimator for a binary response by maximum likelihood. Marginal effects of policy change refer to the change in the probability of a decline in total work hours, employment or female labor share reflecting the relevant policy change index going from the lowest to highest value in the data set, when all other variables in the model are evaluated at their sample mean. Absolute value of robust z-statistics reported in parentheses.
* denotes statistical significance at the 10% level, and ** denotes significance at the 5% level (or better).

decline as a result of the implementation of these policies, when all other variables in the model are evaluated at their sample mean values.

To conserve space, complete results each variable in all models estimated in this analysis can be found in the Appendix. For example, equation (1) results for the complete model using the *Stringency Index* policy measure and the Total Hours Worked employment impact indicator can be found in left-most results column of Panel A of Table A1. This is followed by the results for the other closure and stay-at-home policies variables in the remainder of Panel A. The other containment/closure policies for this employment impact measure are summarized in Panel B. Similarly, complete model results for the other four employment impact categories can be found in the Appendix as Table A2 (Permanent Employment) through Table A5 (Female Labor Share).

A summary of key findings regarding the impact of closure/containment policies the employment categories considered in this analysis follows:

Total Hours Worked. As expected, many of the specific COVID policy variables have parameter estimates with statistically significant positive signs indicating that these containment/closure measures increase the probability that a firm will reduce total work hours once these measures are put in place (see Table 3, Panel A). The evidence is strongest when it comes to measures directed at *Workplace Closing*, *Public Transport Closing*, *Restrictions on Gatherings,* and the *Cancellation of Public Events*. Going from open public transit to required closing of the system (or prohibit most citizens from using it) is associated with a 36% increase the probability that a firm reduced total hours worked, other things equal. Required cancellation of public events and restrictions on public gatherings are associated with a 59% and 48% likelihood of reduced working hours, respectively. The marginal effects on *Workplace Closing* are more modest (19%). Interestingly, *School Closing* measures appear to have little impact on total hours worked, nor do policies restricting movements either internally or internationally.

Regarding the other variables in the model (see Table A1), there is evidence that firms in the retail services sector (*Retail*) are *less* likely to reduce work hours relative to firms operating in the (omitted category) manufacturing sector. Further, both medium and large size firms are also less likely to reduce total hours worked relative the smaller firms (excluded group) over the time period analyzed.

Permanent Employment. Nearly all the closure/containment measures considered in the analysis are associated with increased probability of permanent employee reductions based on results summarized

Covid Economics 66, 28 January 2021: 89-114

in the top panel of Table 3. The evidence is weakest when it comes to *International Travel Controls*, both in terms of statistical significance and estimated marginal effects of a policy change in that area. As to the other categories, *Workplace Closing*, *Public Transport Closing* and *Cancellation of Public Events* is shown to have a sizable influence on the probability of worker cuts. In contrast to the hours worked category, there is strong evidence that *School Closings*, along with *Restrictions on Internal Movement*, enhances the probability that a firm will reduce permanent employment in the face of these policy measures. Results for the industry control variables are in line with what was found earlier, however, both medium and large size firms are more likely to reduce permanent employment (Table A1). The latter results are at odds with what was found for Total Hours Worked and is suggestive that the reduced workforce may be expected to work longer hours.

Temporary Employment. Viewed as a whole, and judged by the composite *Stringency Index*, containment/closure measures did not appear to have much effect on temporary employment. This composite policy variable failed reach conventional levels of statistical significance, nor did it for most of the specific COVID policy areas considered in the analysis. Exceptions were measures related to the *Cancellation of Public Events* and *Restrictions on Internal Movement*. Both showed statistical significance at better than the 5-percent level, with large estimated marginal effects. Required cancellation of public events, for example, increased the probability by 43% that an establishment would decrease temporary workers following the outbreak of COVID-19, perhaps reflecting greater use of temporary workers surrounding such activities. Putting in place restrictions on internal movements between cities/region also substantially increased (58%) that firms will reduce the employment of temporary workers. Results for the model control variables (Table A3) are similar to the Permanent Employment category.

Female Employment. The estimated probabilities of a decline in female full-time employment at the firm level are substantial for several policy measure categories. For example, required *Public Transport Closing* (or prohibiting most citizens from using) increases the probability of a reduction in full-time female staff by 35%, all other factors held constant. Full *Cancellation of Public Events* yields similar probabilities, while the most aggressive *Stay-at-Home Mandates* is associated with a more modest 10 percent increase in the likelihood of reductions of female employment. Interestingly, there is only weak evidence that *School Closing* mandates are associated with reductions in female employment, at least permanent, full time employees addressed in the Enterprise Survey. Viewed from the perspective of overall containment/closure policies, an increase in the *Stringency Index* from 0 to 100 is linked to a 30% increase in the probability that female labor will be let go as a result of the implementation of these policies, other things equal.

Regarding the other variables in the model (see Table A4), there is evidence that firms in the retail services sector (*Retail*) are *less* likely to reduce work hours relative to firms operating in the non-manufacturing sector, but the evidence is not statistically strong. As with the other employment categories considered above, both medium and large size firms are more likely to reduce female employment relative the smaller firms following the COVID-19 outbreak.

Female Labor Share. The rightmost column of Table 3 presents the model results when employment change is based on Female Labor Share of firm permanent employment rather than absolute levels. This

Covid Economics 66, 28 January 2021: 89-114

has the advantage of controlling for a firm's pre-pandemic reliance on female labor, as a firm which places relatively little reliance on women in its workforce might reasonably be expected to reduce their female labor force by relatively small amounts given how their workforce was structured in the first place. With this specification of the dependent variable the issue becomes identifying the factors are relevant in a explaining why a firm would place greater or less relative reliance on female labor in the pandemic period after control/containment policies are put in place.

Results presented in Table 3 show that greater overall containment/closure polices as measured by the *Stringency Index* increases the probability that female labor share declined. Moreover, the estimated marginal effects are substantial – a 22% increase if the *Index* rises from the lowest value in the data set to 100, a level that several countries reached over the time period analyzed. As before, there is also strong evidence that restrictions on the usage of public transport and policies related to the cancellation of public events have a negative effect on female labor share. On the other hand, there is no evidence that either school or workplace closing measures had a differential effect between men and women. Regarding other variables in the model (Table A5), the sign on the parameter estimate on the retail industry variable remains negative and is statistically significant at the ten percent level across all models estimated.[13] Also, larger firms are more likely make relatively less use of female workers.

Firm Closures Due To COVID-19: For the countries included in this analysis approximately 15% of the firms from the baseline survey that were to be included in the Enterprise Surveys COVID-19 Follow-up were deemed to be "permanently closed". Unfortunately, for most of them it was not possible to determine when they closed hence these firms were excluded from the analysis presented above. For some, perhaps most, closure may very well have happened at a date well before 2020 and the start of the pandemic. These firms were originally surveyed as far back as 2016 and many could not be reached for the follow up survey.

One question that was asked when representatives of closed firms could be located was the following: *Did this establishment close temporarily (suspended services or production) due to the COVID-19 outbreak?* Nearly 50 percent (46%) answered this question in the affirmative. To assess the possible impact that containment/closure polices on how firms responded to this question, equation (1) was estimated using the response to this question as the binary dependent variable. The results, not included here to conserve space but available upon request, revealed that neither the overall *Stringency Index*, nor any of the specific policy area variables considered above, were statistically significant in the model estimations.

Discussion and Conclusions

Viewed as a whole it is not surprising that closure/containment polices cost jobs and reduced workhours. It is somewhat surprising that, at least for the emerging economies considered here, the

[13] About 200 observations in the data set had a decline in female labor share that was less than one percent. To see if similar results obtain if the analysis is restricted to more substantial share declines, in preliminary analysis the binary dependent variable was redefined such that it took on a value of one with employment share declines of at least one percent, zero otherwise. The results (not reported to conserve space) were similar to what is reported in Table 3 with statistically stronger evidence regarding the specific COVID-19 policies that impacted female employment shares.

Covid Economics 66, 28 January 2021: 89-114

evidence indicates that permanent jobs for the most part, rather the temporary employment, were most negatively impacted by these policies. The Enterprise Survey does not make a distinction between full and part-time employment, nor hours worked, when asking the question about the impact of COVID-19 on temporary employment, so possible impacts of government policies on this category of employment may be masked by the nature of the data set. The present findings might suggest, for example, that firms found it easier to use temporary workers with more flexible workhours when confronted with any downturn in business they faced since the start of the pandemic.

School and workplace closing policies are associated with a greater likelihood that a firm will reduce permanent employment, although the evidence is weak when it comes for female employment and labor share. These findings are noteworthy as they do not align with earlier work which has suggested that, due to the specific industries most affected by the pandemic, and because women to take on a disproportionate share of childcare duties within the household, that women would bear the brunt of job losses associated with these policies. Possible explanations are that there is some evidence that women are more likely than men to be employed in high work-from-home occupations and disproportionately represented in occupations where the workers are considered to be "essential" and potentially exempt from work closing policies (Alon, et al. (2020), Brussevich, et al.(2020), OECD (2020)).

There is strong evidence that closing public transport in emerging economies had a negative effect on all employment categories considered with the exception of temporary employment. This likely reflects the importance of public transit in the work commute and the fact that telecommunicating is less of a viable alternative in emerging nations (OECD, 2020). Further, Moreno-Monroy (2016) note that access to public transportation encourages low-income workers to switch from home-based informal jobs into the formal labor market, typically concentrated in urban centers. To the extent that this holds, public transit closings can be expected to do just the opposite resulting in employment losses by firms in the formal sector. Moreover, the decline in female labor share also points to women being disproportionately affected by such policies.

Interestingly, of all eight COVID-19 policy areas considered, policies on the closure of public events has statistically significant and large negative effects on all five employment categories considered. Film screenings, theatre performances and screenings, sporting events, concerts, are considered public events based on the OxCGRT Coding Interpretation Guide.[14] Given that many of these events are likely to employ temporary workers it is perhaps not surprising that this would be one policy area that would have a negative effect on this employment category.

Restrictions on internal movement range from recommendations not to travel between regions/cities to stricter ones involving curfew requirements or stay-in-place mandates. Results show that stronger policies have a negative effect on both permanent and total employment, with most of the impact on male workers given the statistically insignificant result for both the female employment and labor share employment categories.

Finally, there is no strong evidence that international travel controls had an effect on any of the five employment categories considered.

[14] https://github.com/OxCGRT/covid-policy-tracker/blob/master/documentation/interpretation_guide.md

Covid Economics 66, 28 January 2021: 89-114

In closing, it is important to point out that the informal labor market and firms with 100% government/state ownership were excluded from this analysis. Given that such activities are difficult to measure, unregistered (informal) business enterprises were not surveyed by the World Bank. The omission of this sector is not inconsequential, especially in emerging economies (Goel and Nelson, 2016), and given that women are more heavily represented than men in the informal economy.[15] Nevertheless, the present analysis does offer valuable perspective on the impacts of the pandemic, at least as the first wave is sweeping the globe, from the perspective of business owners and top managers of firms of all sizes operating in the formal sector.

References

Alon, T., Doepke, M., Olmstead-Rumsey, J., Tertilt, M., 2020. The impact of COVID-19 on gender equality, Covid Economics 4: 62-85.

Beck, T., Flynn, B., Homanen, M., 2020. COVID-19 in emerging markets: Firm-survey evidence, Covid Economics 38: 37-67.

Benz, S., Gonzales, F., Mourougane, A., 2020. The impact of COVID-19 international travel restrictions on services-trade costs: some illustrative scenarios, Covid economics 45, 65-76.

Brussevich, M., Dabla-Norris, E., Khalhid, S., 2020. Who will Bear the Brunt of Lockdown Policies? Evidence from Tele-workability Measures Across Countries, IMF Working Paper WP/20/88.

Goel, R., Nelson, M., 2016. Shining a light on the shadows: Identifying robust determinants of the shadow economy, Economic Modelling 58: 351-364.

Gottlieb, C., Grobovšek, J., Poschke, M. Saltiel, F., 2020. Lockdown accounting, Covid Economics 31, 103-129.

Hupkau, C., Petrongolo, B., 2020. Work, care and gender during the covid-19 crisis, Covid Economics 54, 1-28.

IPBES. 2020. Workshop Report on Biodiversity and Pandemics of the Intergovernmental Platform on Biodiversity and Ecosystem Services. Daszak, P., das Neves, C., Amuasi, J., Hayman, D., Kuiken, T., Roche, B., Zambrana-Torrelio, C., Buss, P., Dundarova, H., Feferholtz, Y., Foldvari, G., Igbinosa, E., Junglen, S., Liu, Q., Suzan, G., Uhart, M., Wannous, C., Woolaston, K., Mosig Reidl, P., O'Brien, K., Pascual, U., Stoett, P., Li, H., Ngo, H. T., IPBES secretariat, Bonn, Germany, DOI:10.5281/zenodo.4147317. http://www.ipbes.net/pandemics

[15] https://interactive.unwomen.org/multimedia/infographic/changingworldofwork/en/index.html

Karabarbounis, M., Trachter, N., 2020. How COVID-19 is affecting main street. Federal Reserve Bank of Richmond. https://www.richmondfed.org/publications/research/coronavirus/economic_impact_covid-19_04-17-20

Moreno-Monroy, A,. 2016. Access to public transport and labor informality. IZA World of Labor 2016: 274. doi: 10.15185/izawol.274

OECD, 2020. Women at the core of the fight against COVID-19 crisis. https://read.oecd-ilibrary.org/view/?ref=127_127000-awfnqj80me&title=Women-at-the-core-of-the-fight-against-COVID-19-crisis. Accessed 12.21.2020

Schalatek, L., 2020. The invisible coronavirus makes systemic gender inequalities and injustices visible, Heinrich-Böll-Stiftung: Washington, DC. https://us.boell.org/en/2020/04/30/invisible-coronavirus-makes-systemic-gender-inequalities-and-injustices-visible#top Accessed 1.6.2021.

Covid Economics 66, 28 January 2021: 89-114

Table A1
Firm-level Response to Government COVID-19 Containment and Closure Policies: Impact on Total Hours Worked

Dependent variable: ***Total Hours Worked Declined***

Panel A: Overall Policies; School, Work, Public Transport Closings; Stay at Home

	Stringency Index	**School Closing**	**Workplace Closing**	**Public Transport Closing**	**Stay-at-Home Mandates**
COVID-19 Policy	**0.01*** **(1.9)**	**-0.15** **(0.6)**	**0.25**** **(2.5)**	**0.81**** **(3.0)**	**0.18*** **(1.9)**
Retail	-0.14* (1.9)	-0.14* (1.9)	-0.14* (1.9)	-0.15** (2.0)	-0.14* (2.0)
Other Services	0.05 (0.8)	0.05 (0.8)	0.05 (0.8)	0.05 (0.8)	0.05 (0.8)
Medium Size Firm	-0.21** (3.4)	-0.21** (3.4)	-0.21** (3.5)	-0.21** (3.5)	-0.21** (3.5)
Large Size Firm	-0.34** (4.6)	-0.35** (4.7)	-0.35** (4.6)	-0.35** (4.7)	-0.35** (4.6)
Country Fixed Effects	yes	yes	yes	yes	yes
Marginal Effects of Policy Change	32%	-11%	19%	36%	13%
Number of Obvs.	7,007	7,007	7,007	7,007	7,007
Log likelihood chi-square	962.3**	961.0**	959.5**	970.7**	963.6**
Percentage correctly predicted	66.2	66.1	66.3	66.1	66.2
Pseudo R-sq.	0.13	0.13	0.13	0.13	0.13

Notes: See next page.

Table A1 - Continued

Firm-level Response to Government COVID-19 Containment and Closure Policies: Impact on Total Hours Worked

Dependent variable: ***Total Hours Worked Declined***

Panel B: Public Events, Gatherings, Domestic and International Movement

	Cancel Public Events	**Restrictions on Gatherings**	**Restrictions on Internal Movement**	**International Travel Controls**
COVID-19 Policy	**1.40**** **(2.7)**	**0.53*** **(2.0)**	**0.34** **(0.8)**	**-0.11** **(1.5)**
Retail	-0.15** (2.0)	-0.14* (1.9)	-0.14* (1.9)	-0.15** (2.0)
Other Services	0.05 (0.8)	0.05 (0.9)	0.05 (0.8)	0.05 (0.8)
Medium Size Firm	-0.20** (3.4)	-0.21** (3.4)	-0.21** (3.4)	-0.21** (3.5)
Large Size Firm	-0.34** (4.6)	-0.35** (4.6)	-0.35** (4.6)	-0.35** (4.7)
Country Fixed Effects	yes	yes	yes	yes
Marginal Effects of Policy Change	59%	48%	16%	-11%
Number of Obvs.	7,007	7,007	7,007	7,007
Log likelihood chi-square	967.6**	963.7**	962.1**	962.9**
Percentage correctly predicted	66.1	66.1	66.1	66.1
Pseudo R-sq.	0.13	0.13	0.13	0.13

Notes: All models included a constant term (not reported) and are fit using the logit estimator for a binary response by maximum likelihood. Cutoff is 0.5 percentage for correctly predicted. Marginal effects of policy change refer to the change in the probability of a firm-level decline in total hours worked reflecting the relevant policy change index going from the lowest to highest value in the data set, when all other variables in the model are evaluated at their sample mean. Absolute value of robust z-statistics reported in parentheses.
* denotes statistical significance at the 10% level, and ** denotes significance at the 5% level (or better).

Covid Economics 66, 28 January 2021: 89-114

CEPR PRESS

Covid Economics 66, 28 January 2021: 89-114

Table A2

Firm-level Response to Government COVID-19 Containment and Closure Policies: Impact on Full-time Permanent Employment

Dependent variable: ***Permanent Employment Declined***

Panel A: Overall Policies; School, Work, Public Transport Closings; Stay at Home

	Stringency Index	School Closing	Workplace Closing	Public Transport Closing	Stay-at-Home Mandates
COVID-19 Policy	**0.03**** **(3.5)**	**0.42**** **(2.1)**	**0.25**** **(2.6)**	**0.93**** **(2.8)**	**0.25**** **(2.4)**
Retail	-0.18** (2.3)	-0.18** (2.3)	-0.18** (2.3)	-0.18** (2.3)	-0.18** (2.3)
Other Services	-0.04 (0.7)	-0.04 (0.7)	-0.04 (0.7)	-0.04 (0.6)	-0.04 (0.7)
Medium Size Firm	0.22** (3.5)	0.21** (3.4)	0.22** (3.5)	0.22** (3.4)	0.22** (3.4)
Large Size Firm	0.49** (6.5)	0.48** (6.4)	0.49** (6.5)	0.48** (6.4)	0.48** (6.5)
Country Fixed Effects	yes	yes	yes	yes	yes
Marginal Effects of Policy Change	48%	22%	15%	40%	16%
Number of Obvs.	6,998	6,998	6,998	6,998	6,998
Log likelihood chi-square	488.4**	483.6**	486.1**	483.4**	484.6**
Percentage correctly predicted	70.6	70.5	70.8	70.5	70.5
Pseudo R-sq.	0.06	0.06	0.06	0.06	0.06

Notes: See next page.

Covid Economics 66, 28 January 2021: 89-114

Table A2 - Continued

Firm-level Response to Government COVID-19 Containment and Closure Policies: Impact on Full-time Permanent Employment

Dependent variable: ***Permanent Employment Declined***

Panel B: Public Events, Gatherings, Domestic and International Movement

	Cancel Public Events	**Restrictions on Gatherings**	**Restrictions on Internal Movement**	**International Travel Controls**
COVID-19 Policy	**1.34**** **(2.3)**	**0.46** **(1.6)**	**1.64**** **(3.2)**	**0.05** **(0.6)**
Retail	-0.19** (2.4)	-0.18** (2.3)	-0.18** (2.3)	-0.18** (2.3)
Other Services	-0.04 (0.6)	-0.04 (0.6)	-0.04 (0.6)	-0.04 (0.6)
Medium Size Firm	0.22** (3.5)	0.22** (3.5)	0.22** (3.5)	0.22** (3.5)
Large Size Firm	0.48** (6.5)	0.48** (6.4)	0.49** (6.5)	0.48** (6.4)
Country Fixed Effects	yes	yes	yes	yes
Marginal Effects of Policy Change	41%	34%	58%	4%
Number of Obvs.	6,998	6,998	6,998	6,998
Log likelihood chi-square	484.6**	481.8**	490.7**	480.1**
Percentage correctly predicted	70.5	70.5	70.5	70.5
Pseudo R-sq.	0.06	0.06	0.06	0.06

Notes: All models included a constant term (not reported) and are fit using the logit estimator for a binary response by maximum likelihood. Cutoff is 0.5 percentage for correctly predicted. Marginal effects of policy change refer to the change in the probability of a firm-level decline in permanent full-time employment reflecting the relevant policy change index going from the lowest to highest value in the data set, when all other variables in the model are evaluated at their sample mean. Absolute value of robust z-statistics reported in parentheses.
* denotes statistical significance at the 10% level, and ** denotes significance at the 5% level (or better).

Table A3

Firm-level Response to Government COVID-19 Containment and Closure Policies: Impact on Temporary Workers

Dependent variable: ***Temporary Employment Declined***

Panel A: Overall Policies; School, Work, Public Transport Closings; Stay at Home

	Stringency Index	**School Closing**	**Workplace Closing**	**Public Transport Closing**	**Stay-at-Home Mandates**
COVID-19 Policy	**0.00 (0.6)**	**0.24 (1.5)**	**-0.01 (0.2)**	**-0.04 (0.1)**	**-0.04 (0.4)**
Retail	-0.16* (1.9)	-0.16* (1.9)	-0.16* (1.9)	-0.16* (1.9)	-0.16* (1.9)
Other Services	0.02 (0.2)	0.02 (0.2)	0.02 (0.2)	0.02 (0.3)	0.02 (0.3)
Medium Size Firm	0.20** (3.0)	0.20** (3.0)	0.20** (3.0)	0.20** (3.0)	0.20** (3.0)
Large Size Firm	0.28** (3.3)	0.28** (3.3)	0.28** (3.3)	0.28** (3.3)	0.28** (3.3)
Country Fixed Effects	yes	yes	yes	yes	yes
Marginal Effects of Policy Change	7%	12%	-1%	-1%	-2%
Number of Obvs.	6,653	6,653	6,653	6,653	6,653
Log likelihood chi-square	779.6**	779.6**	778.4**	778.6**	778.6**
Percentage correctly predicted	73.3	73.3	73.3	73.3	73.3
Pseudo R-sq.	0.14	0.14	0.14	0.14	0.14

Notes: See next page.

Table A3 - Continued

Firm-level Response to Government COVID-19 Containment and Closure Policies: Impact on Temporary Workers

Dependent variable: ***Temporary Employment Declined***

Panel B: Public Events, Gatherings, Domestic and International Movement

	Cancel Public Events	**Restrictions on Gatherings**	**Restrictions on Internal Movement**	**International Travel Controls**
COVID-19 Policy	**1.67**** **(2.9)**	**0.65** **(1.4)**	**1.84**** **(2.9)**	**0.23** **(1.1)**
Retail	-0.16* (2.0)	-0.16* (1.9)	-0.16* (1.9)	-0.16* (1.9)
Other Services	0.02 (0.2)	0.02 (0.3)	0.02 (0.3)	0.02 (0.3)
Medium Size Firm	0.20** (3.0)	0.20** (3.0)	0.20** (3.0)	0.20** (3.0)
Large Size Firm	0.29** (3.4)	0.28** (3.3)	0.29** (3.4)	0.28** (3.3)
Country Fixed Effects	yes	yes	yes	yes
Marginal Effects of Policy Change	43%	41%	58%	14%
Number of Obvs.	6,653	6,653	6,653	6,653
Log likelihood chi-square	785.0**	776.6**	787.3**	777.9**
Percentage correctly predicted	73.3	73.3	73.3	73.3
Pseudo R-sq.	0.14	0.14	0.14	0.14

Notes: All models included a constant term (not reported) and are fit using the logit estimator for a binary response by maximum likelihood. Cutoff is 0.5 percentage for correctly predicted. Marginal effects of policy change refer to the change in the probability of a decline in firm-level temporary employment reflecting the relevant policy change going from the lowest to highest value in the data set, when all other variables in the model are evaluated at their sample mean. Absolute value of robust z-statistics reported in parentheses.
* denotes statistical significance at the 10% level, and ** denotes significance at the 5% level (or better).

Covid Economics 66, 28 January 2021: 89-114

Table A4

Firm-level Response to Government COVID-19 Containment and Closure Policies: Impact on Female Employment

Dependent variable: ***Female Employment Declined***

Panel A: Overall Policies; School, Work, Public Transport Closings; Stay at Home

	Stringency Index	School Closing	Workplace Closing	Public Transport Closing	Stay-at-Home Mandates
COVID-19 Policy	**0.02**** **(2.5)**	**0.25 (1.2)**	**0.10 (0.9)**	**0.96**** **(2.7)**	**0.21*** **(1.8)**
Retail	-0.14 (1.6)	-0.14 (1.6)	-0.14 (1.6)	-0.14 (1.6)	-0.14 (1.6)
Other Services	-0.11 (1.6)	-0.11 (1.6)	-0.11 (1.6)	-0.11 (1.6)	-0.11 (1.6)
Medium Size Firm	0.38** (5.3)	0.38** (5.3)	0.38** (5.3)	0.38** (5.3)	0.38** (5.3)
Large Size Firm	0.70** (8.3)	0.69** (8.3)	0.70** (8.3)	0.69** (8.3)	0.70** (8.3)
Country Fixed Effects	yes	yes	yes	yes	yes
Marginal Effects of Policy Change	30%	10%	5%	35%	10%
Number of Obvs.	6,807	6,807	6,807	6,807	6,807
Log likelihood chi-square	386.3**	380.3**	381.2**	385.1**	383.0**
Percentage correctly predicted	79.5	79.6	79.5	79.5	79.6
Pseudo R-sq.	0.06	0.06	0.06	0.06	0.06

Notes: See next page.

Covid Economics 66, 28 January 2021: 89-114

Covid Economics 66, 28 January 2021: 89-114

Table A4- Continued

Firm-level Response to Government COVID-19 Containment and Closure Policies: Impact on Female Employment

Dependent variable: ***Female Employment Declined***

Panel B: Public Events, Gatherings, Domestic and International Movement

	Cancel Public Events	Restrictions on Gatherings	Restrictions on Internal Movement	International Travel Controls
COVID-19 Policy	**1.42**** **(2.1)**	**-0.43** **(1.3)**	**0.53** **(0.8)**	**0.12** **(1.3)**
Retail	-0.14 (1.6)	-0.14 (1.6)	-0.14 (1.6)	-0.14 (1.6)
Other Services	-0.11 (1.5)	-0.11 (1.6)	-0.11 (1.5)	-0.11 (1.6)
Medium Size Firm	0.39** (5.4)	0.38** (5.3)	0.38** (5.3)	0.38** (5.3)
Large Size Firm	0.70** (8.3)	0.69** (8.3)	0.70** (8.3)	0.70** (8.3)
Country Fixed Effects	yes	yes	yes	yes
Marginal Effects of Policy Change	31%	-28%	16%	8%
Number of Obvs.	6,807	6,807	6,807	6,807
Log likelihood chi-square	383.9**	379.0**	380.2**	380.7**
Percentage correctly predicted	79.6	79.6	79.6	79.5
Pseudo R-sq.	0.06	0.06	0.06	0.06

Notes: All models included a constant term (not reported) and are fit using the logit estimator for a binary response by maximum likelihood. Cutoff is 0.5 percentage for correctly predicted. Marginal effects of policy change refer to the change in the probability of a decline in female employment reflecting the relevant policy change index going from the lowest to highest value in the data set, when all other variables in the model are evaluated at their sample mean. Absolute value of robust z-statistics reported in parentheses.
* denotes statistical significance at the 10% level, and ** denotes significance at the 5% level (or better).

Table A5

Firm-level Response to Government COVID-19 Containment and Closure Policies: Impact on Female Labor Share

Dependent variable: ***Female Labor Share Declined***

Panel A: Overall Policies; School, Work, Public Transport Closings; Stay at Home

	Stringency Index	**School Closing**	**Workplace Closing**	**Public Transport Closing**	**Stay-at-Home Mandates**
COVID-19 Policy	**0.02*** **(1.9)**	**0.06** **(0.3)**	**0.04** **(0.4)**	**1.10**** **(2.6)**	**0.13** **(1.0)**
Retail	-0.17* (1.8)	-0.17* (1.8)	-0.17* (1.8)	-0.17* (1.8)	-0.17* (1.8)
Other Services	-0.05 (0.7)	-0.05 (0.7)	-0.05 (0.7)	-0.05 (0.7)	-0.05 (0.7)
Medium Size Firm	0.33** (4.3)	0.33** (4.3)	0.33** (4.3)	0.33** (4.3)	0.33** (4.3)
Large Size Firm	0.46** (5.0)	0.46** (5.0)	0.46** (5.0)	0.46** (5.0)	0.46** (5.0)
Country Fixed Effects	yes	yes	yes	yes	yes
Marginal Effects of Policy Change	22%	2%	2%	36%	5%
Number of Obvs.	6,736	6,736	6,736	6,736	6,736
Log likelihood chi-square	207.4**	203.0**	204.1**	207.1**	205.2**
Percentage correctly predicted	83.9	83.9	83.9	83.9	83.9
Pseudo R-sq.	0.04	0.04	0.04	0.04	0.04

Notes: See next page.

Covid Economics 66, 28 January 2021: 89-114

Table A5 - Continued

Firm-level Response to Government COVID-19 Containment and Closure Policies: Impact on Female Labor Share

Dependent variable: *Female Labor Share Declined*

Panel B: Public Events, Gatherings, Domestic and International Movement

	Cancel Public Events	**Restrictions on Gatherings**	**Restrictions on Internal Movement**	**International Travel Controls**
COVID-19 Policy	**1.66** **(2.4)**	**-0.40** **(1.1)**	**1.52·** **(1.6)**	**0.15** **(1.5)**
Retail	-0.17* (1.8)	-0.17* (1.8)	-0.17* (1.7)	-0.17* (1.7)
Other Services	-0.05 (0.7)	-0.05 (0.7)	-0.05 (0.7)	-0.05 (0.7)
Medium Size Firm	0.33** (4.3)	0.33** (4.3)	0.33** (4.3)	0.33** (4.3)
Large Size Firm	0.46** (5.0)	0.46** (5.0)	0.46** (5.0)	0.46** (5.0)
Country Fixed Effects	yes	yes	yes	yes
Marginal Effects of Policy Change	29%	-22%	38%	7%
Number of Obvs.	6,736	6,736	6,736	6,736
Log likelihood chi-square	208.9**	202.8**	204.5**	205.1**
Percentage correctly predicted	83.9	83.9	83.9	83.9
Pseudo R-sq.	0.04	0.04	0.04	0.04

Notes: All models included a constant term (not reported) and are fit using the logit estimator for a binary response by maximum likelihood. Cutoff is 0.5 percentage for correctly predicted. Marginal effects of policy change refer to the change in the probability of a decline in firm-level female employment share reflecting the relevant policy change index going from the lowest to highest value in the data set, when all other variables in the model are evaluated at their sample mean. Absolute value of robust z-statistics reported in parentheses.
* denotes statistical significance at the 10% level, and ** denotes significance at the 5% level (or better).

Covid Economics Issue 66, 28 January 2021

A new Covid policy stringency index for Europe[1]

Daniel Gros,[2] Alexandre Ounnas[3] and
Timothy Yu-Cheong Yeung[4]

Date submitted: 22 January 2021; Date accepted: 23 January 2021

We present a new policy stringency index for Europe, based on a compilation of country response measures to Covid-19 provided by the European Centre for Disease Prevention and Control (ECDC). This new index is available for dozens of different types of mandatory social distancing measures most frequently applied (e.g. school closures, face covering, closure of restaurants and other sectors of the economy, stay at home orders, etc.) and takes into account that many measures are graduated. An aggregate index is also provided. First tests indicate that (changes in) this policy index are highly correlated with contemporaneous and future economic activity. An increase in the overall restrictiveness indicator of one standard deviation is associated with a fall in GDP of about 3 percentage points. Increases in this indicator are usually followed by a fall in infections. The aggregate 'CEPS-PERISCOPE index' is highly correlated (correlation coefficient 80-90 % in levels and changes) with the Oxford government response tracker in both level and monthly changes. However, the correlation is much smaller for individual elements, such as school closures, prohibitions on mass gatherings, etc. The underlying data is available for researchers to use for further empirical work.

Covid Economics 66, 28 January 2021: 115-137

1 This research has received funding from the Horizon 2020 research and innovation programme of the EU under grant agreement No. 101016233, H2020-SC1-PHE CORONAVIRUS-2020-2-RTD, PERISCOPE (Pan European Response to the Impacts of Covid-19 and future Pandemics and Epidemics). Our indices are available here.
2 Distinguished Fellow, Centre for European Policy Studies.
3 Research Fellow, Centre for European Policy Studies.
4 Research Fellow, Centre for European Policy Studies.

CEPR PRESS

1 Introduction

Measuring complex phenomena in a numerical form for empirical research is a recurring issue in social sciences (Marozzi, 2016; Diamantopoulos and Winklhofer, 2001). This challenge is particularly acute if one considers the governmental response to the spread of Covid-19. Governments almost everywhere used many so-called Non-pharmacological interventions (NPI) to stop, or at least slow down contagion. Typical measures were ordering the wearing of face masks in public spaces, the closure of schools, restaurants and other other places, sometimes culminating in a lockdown under which the population was ordered to stay at home. Theses measures have had an immediate and dramatic impact on the economy. However, little is known so far about which specific measures had the strongest impact on the economy and to what extent the deep recession was mainly caused by governmental NPI's or spontaneous reactions from the fear of the disease. Given that so many countries have adopted similar NPIs, research would benefit from cross-country studies. However, comparing the experience of many countries requires an index that is at least roughly comparable and numerical in order to be useful for standard statistical approaches.

This contribution explains the construction of such an index for 30 European countries. The proposed new Covid-19 policy stringency index is based on a compilation of country response measures to Covid-19 provided by the European Centre for Disease Prevention and Control (hereafter, ECDC)[2] which we transform into a quantitative measure. Until now the main data source for comparative empirical research on the impact of governmental social distancing measures had to rely on the Oxford Stringency Index (Hale, Angrist, Cameron-Blake, Hallas, Kira, Majumdar, Petherick, Phillips, Tatlow and Webster, 2020), which has been widely employed by researchers across disciplines (Edejer et al., 2020; Yan et al., 2020).[3] Our new index provides a new source for European countries and adds several elements not covered by the Oxford Index.

The ECDC documents 58 categories of measures for 31 European countries (the EEA plus the UK), covering 1,274 measures. For each measure entry, a start date and an end date (if the policy ended) are provided. We are thus able to build a panel dataset

[2]Data can be found here: https://www.ecdc.europa.eu/en/publications-data/download-data-response-measures-covid-19

[3]A search on Google Scholar on December 15, 2020 yielded close to 2000 entries when employing the search term "oxford stringency index covid"

Covid Economics 66, 28 January 2021: 115-137

CEPR PRESS

of all these measures since the start of the pandemic. The compilation of measures provided by the ECDC has a number of important advantages: it provides information on some types of measures not considered by the Oxford Stringency, for example, the restrictions of certain indoor activities. It also provides important detail on measures regarding school closures, distinguishing between different types of schools, from nursery to higher education. Stay at home regulations (lockdown) measures are also measured in a graduated way, distinguishing orders for all areas from those for specific areas only. It also provides a distinction between measures which are mandatory, or voluntary, for example government advice to stay at home or mandatory stay at home orders.[4]

To build a stringency index, we selected 40 measures and group them into eight broader categories. Table 1 provides a detailed list of the measures categorised by the ECDC upon which we group them into broader categories. The selection of measures among the 58 measures is constrained by the fact that some measures are in fact duplicatory in nature. For example, "Adaptation of Workplaces" and "Workplace Closure" have been adopted along with "Teleworking" in some countries but adding them together in a composite index may risk a double-count of the restrictions on working routine. Some countries did not have a specific measure on Workplace Closure but somehow integrated with Teleworking. Another example is the restriction of numbers of people at indoor gatherings. We leave it out because "Closure: Indoor Activities" in a broad sense measures the same restriction and is more granular.

2 Construction of a European Covid-19 Policy Stringency Index

Most measures within a category are mutually exclusive and can be ranked in an ordinal (and put into a cardinal scale when building the index). For example, in the category of "Closure Daycare," a partial closure is less stringent than a total closure, and they are mutually exclusive. The maximum unweighted score of this category is thus two, which corresponds to the maximum stringency level. The same logic applies to "Clousure: Public Area," "Mass Gathering Ban," and "Teleworking." The category "Stay Home (Population)" is slightly different as there are four mutually exclusive measures. The

[4]For a large compilation of Covid related measures see also Cheng et al. (2020)

Covid Economics 66, 28 January 2021: 115-137

Table 1: Categorization of Measures

Category	Measure	Cardinal Scale
Closure: Daycare	Partial Closure Daycare	1
	Total Closure Daycare	2
Closure: School	Partial Closure Primary Schools	1
	Total Closure Primary Schools	2
	Partial Closure Secondary Schools	1
	Total Closure Secondary Schools	2
	Partial Closure Higher Education	1
	Total Closure Higher Education	2
Closure: Public Area	Partial Closure Public Area	1
	Total Closure Public Area	2
Closure: Indoor Activities	Partial Closure Public Transport	1
	Total Closure Public Transport	2
	Partial Closure Entertainment Venues	1
	Total Closure Entertainment Venues	2
	Partial Closure Sports Centres and Gyms	1
	Total Closure Sports Centres and Gyms	2
	Partial Closure Hotels	1
	Total Closure Hotels	2
	Partial Closure Non-essential Shops	1
	Total Closure Non-essential Shops	2
	Partial Closure Worship Places	1
	Total Closure Worship Places	2
	Partial Closure Private Gatherings	1
	Total Closure Private Gatherings	2
	Partial Closure Restaurants and Cafes	1
	Total Closure Restaurants and Cafes	2
Mass Gathering Ban	Partial Ban of Mass Gathering	1
	Total Ban of Mass Gathering	2
Facial Covering	Partial Masking Closed Spaces	1
	Total Masking Closed Spaces	2
	Partial Masking All Spaces	2
	Total Masking All Spaces	5
Stay Home (Population)	Stay Home (advice, specific areas)	1
	Stay Home (advice, all areas)	2
	Stay Home (enforced, specific areas)	3
	Stay Home (enforced, all areas)	4
Teleworking	Relaxed Teleworking Recommendation	1
	Teleworking Recommendation	2

Covid Economics 66, 28 January 2021: 115-137

maximum unweighted score of this category is four. For "Closure: School," we have three levels of education and the measures of each level are mutually exclusive. The maximum unweighted score of this category is thus six. The category "Closure: Indoor Activities" contains eight different indoor activities, and thus the maximum unweighted score is 16. Their relative importance in the aggregate index will be adjusted by a weighting scheme to be explained soon after.

"Facial covering", or masks represent a more complicated category. We have included all mandatory orders of facial covering in the ECDC classification scheme while leaving out advisory orders that were voluntary in nature. "Facial covering" is mainly divided into two types, namely, masking in open spaces and masking in closed spaces. The two types were however not mutually exclusive. Some countries, for example, Luxembourg, had overlapping measures (Partial Masking in Closed Spaces and Partial Masking in All Spaces) since 22 June 2020 according to the ECDC dataset (not yet end by the time of writing). However, the combined effect of these two parallel measures is arguably less stringent than a complete mandatory masking order in all spaces. Therefore, we allow five levels of stringency in the "Facial covering" category by recognising possibilities of combinations of parallel measures. Referring to Table 1, "Total Masking All Spaces" is the most stringent level and the unweighted score will be five even if there were some other parallel and less stringent measures. If there was no "Total Masking All Spaces," we add up the stringency levels of any parallel measures. Taking Luxembourg again as the example, the stringency level since 22 June 2020 is three. Precisely, we score this category as the following:

$$s_{face,t} = min\{5, \sum_{n} x_{nt}\}, \tag{1}$$

for $n \in \{\text{PartialClosed}, \text{TotalClosed}, \text{PartialOpen}, \text{TotalOpen}\}$.

It is clear that the different categories are not perfectly comparable. An index value of two in "Facial Covering" does not imply the same level of stringency of an index value of two in "Teleworking." It is thus important to keep indices of categories separate. However, it is helpful to have a composite stringency index. Summation does not exactly reflect the relative stringency levels of categories but is straightforward and transparent. Despite that, the distribution of weight among categories is arguably reasonable. The most heavily weighted categories are "Closure: Indoor Activities" and "Stay Home (Population)" that

Covid Economics 66, 28 January 2021: 115-137

Covid Economics 66, 28 January 2021: 115-137

their effectiveness and cost have often been discussed by researchers and governments. We provide a simple test of the usefulness of the aggregation below. One could consider the construction of the aggregate index as filling in a list of check boxes on a score sheet. One unavoidable limitation of this index (as with the Oxford response tracker) is that in many countries key measures are the competence of sub-national entities (regions, sometimes cities, Laender in Germany, etc.) the national values for the index can thus only present a rough average of the different regional measures. This level of aggregation is sufficient for estimating impacts on the economy since high frequency economic data is available only at the national level.

Sub-national Covid policy indicators are being constructed especially for the US and Canada (Adeel et al., 2020; Hale, Atav, Hallas, Kira, Phillips, Petherick and Pott, 2020), but do not exist yet systematically for European countries. [5]

The next step is to assign weights, listed in Table 2, to the eight categories so that the addictive aggregate score is bounded between 0 and 100.

Table 2: Weights of Categories in the Composite Index

Category	**Weight**
Closure: Daycare	10%
Closure: School	10%
Closure: Indoor Activities	20%
Closure: Public Area	10%
Mass Gathering Ban	10%
Facial Covering	10%
Stay Home (Population)	20%
Teleworking	10%
Total	100%

Given that the ECDC provides the day of entry into force (and its repeal), it would be possible to construct a daily index. We limit ourselves here to either weekly and monthly frequency. For the monthly index, we multiply each index value by the fraction of the month for which the measure in question has been in force. For the weekly index, we assume the measure takes effect starting from the week after the first day of implementation.[6] For the tests using actual economic data the weekly or monthly frequency

[5]For more detail about Covid measures at the regional level see (Cheng et al., 2020).

[6]This different treatment is due to the existence of different numbering schemes of weeks that makes exact matching of days difficult. The proposed method eliminates any unrealistic results for which researchers observed effects taking place before the measures had been implemented.

CEPR PRESS

is adequate. We proceed as follows. We compute the unweighted score for each category, multiply them by their corresponding weight, and then sum up all scores of eight categories.

The unweighted score of category i (except "Facial Covering") of month t is computed according to the following equation:

$$s_{it} = \sum_{n=1}^{N_i} \rho_{nt} v_n, \tag{2}$$

where N_i is the number of measures within the category i, ρ_{it} the fraction of days in month t that measure i is in effect, and v_n the associated cardinal value of the respective measure. Finally, ignoring the comparability issue between categories, we sum up scores of eight categories taking into account of their corresponding weight:

$$S_t = \sum_{i=1}^{8} s_{it} w_i \tag{3}$$

The constructed index is available for download and researchers are free to use it for their research work.[7]

3 Summary Statistics

This section shows some basic summary statistics and illustrative graphs, aiming to describe the trend of Covid restriction stringency for the first 11 months of 2020.[8] The ECDC dataset contains information on 31 European countries and we compute an index value for each month since January 2020 until November 2020. We begin by Table 3 that shows the mean value of the aggregate Index along with its standard deviation, maximum and minimum values of each country in the sample. On average, Iceland is the most relaxed country in terms of Covid restriction stringency, followed by Estonia and Hungary. At the other end, Ireland is the most stringent countries. Slovenia recorded the most stringent policy mix in April 2020.

[7]Data can be found: https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnx0aW1vdGh5Mjc5OXxneDozNzRlYjU3YWYxNmFmZTMy.

[8]By the time of writing, the ECDC dataset was updated up to 17 December 2020.

Covid Economics 66, 28 January 2021: 115-137

Table 3: Aggregate Index

	Aggregate Stringency Index, Monthly, Jan-Nov 2020				
Country	**Mean**	**Sd**	**Max**	**Min**	**Min** (Mar.-Nov.)
Austria	37	26	85	0	26
Belgium	39	23	69	0	30
Bulgaria	26	18	58	0	17
Croatia	32	24	85	0	27
Cyprus	28	25	81	0	14
Czechia	37	27	79	0	24
Denmark	24	16	56	0	21
Estonia	21	14	48	0	13
Finland	22	16	48	0	16
France	35	28	81	0	18
Germany	23	17	57	0	15
Greece	37	26	80	0	26
Hungary	21	25	69	0	2
Iceland	14	11	34	0	5
Ireland	42	29	79	0	22
Italy	38	27	86	0	26
Latvia	31	20	66	0	26
Liechtenstein	23	16	56	0	14
Lithuania	29	22	72	0	17
Luxembourg	35	22	73	0	29
Malta	38	20	59	0	27
Netherlands	32	21	67	0	25
Norway	28	16	54	0	25
Poland	38	22	78	0	31
Portugal	36	21	72	0	35
Romania	39	27	90	0	30
Slovakia	24	17	55	0	16
Slovenia	37	28	94	0	22
Spain	38	23	76	0	24
Sweden	23	13	37	0	15
Switzerland	28	15	50	0	26
United Kingdom	39	24	68	0	29
Overall	31	22	94	0	2

We construct an index for all 32 countries present in the ECDC dataset, which includes Iceland, Liechtenstein, Norway, Switzerland and the United Kingdom. When we move to compare with the Oxford data, we limit to a smaller sample that contains 29 countries (Liechtenstein and Malta are absent). In our regression exercise, we focus on EU27 countries.

Figure 1 illustrates the aggregate index of each country averaged over March-November 2020 on a map. Southern Europe, such as Spain, Italy and Greece, as well as the UK, were stringent, while Germany, the Scandinavian and Baltic states have been, on average, less stringent. The former group of countries comprises those with the sharpest fall in GDP. The next section explores more systematically the impact of restrictions on the economy.

Figure 2 illustrates the trends of stringency for all 32 countries covered by the ECDC

Covid Economics 66, 28 January 2021: 115-137

(EU-27 plus Iceland, Liechtenstein, Norway, Switzerland, and the UK). Most governments were quite quick to respond to the pandemic and the policy stringency level peaked locally in many cases in April. While some, e.g. Croatia, Hungary, Latvia, etc., relaxed their restrictions relatively quickly towards the summer, other countries, such as Belgium, Denmark, Luxembourg, Portugal, etc., kept most of their restrictions in the summer and adjusted slowly over the period. Following the start of the pandemic's second wave in September/October, the index started to increase again in many countries.

Finally, the evolution of the index for Sweden is consistent with what has been reported (in the media) for this country. The index increased up to April but even at its peak it remains much lower level than the values reached by most other countries. It then remained constant for the rest of the period. However, Sweden is not the country with the lowest average degree of restrictiveness. Two smaller neighbours, Estonia and Finland show lower values and even Germany is close.

Covid Economics 66, 28 January 2021: 115-137

4 A first use of the index

In this section we provide two examples of how our CEPS-PERISCOPE index could be used in Covid research. We first test for the impact of NPIs on economic growth, finding that increases in the restrictiveness have a negative impact on growth which is both statistically and economically significant. We then test for the impact of restrictions on the course of infection, finding again a strong association between our index and subsequent falls in infections. These two findings suggest that our index reflects adequately important policy measures, which have had a measurable impact on the economy and the course of the disease.

4.1 The economic impact of social distancing measures

A key issue for policymakers is the economic cost of different social distancing measures. Our CEPS-PERISCOPE index allows one to directly estimate the economic costs of the different classes of measures followed by the ECDC. One can consider the measures taken to slow down the spread of the virus as exogenous to the initial state of the economy. The impact of these measures on the economy could thus be estimated directly using ordinary least squares. We thus perform a set of linear regressions in which we explain

Covid Economics 66, 28 January 2021: 115-137

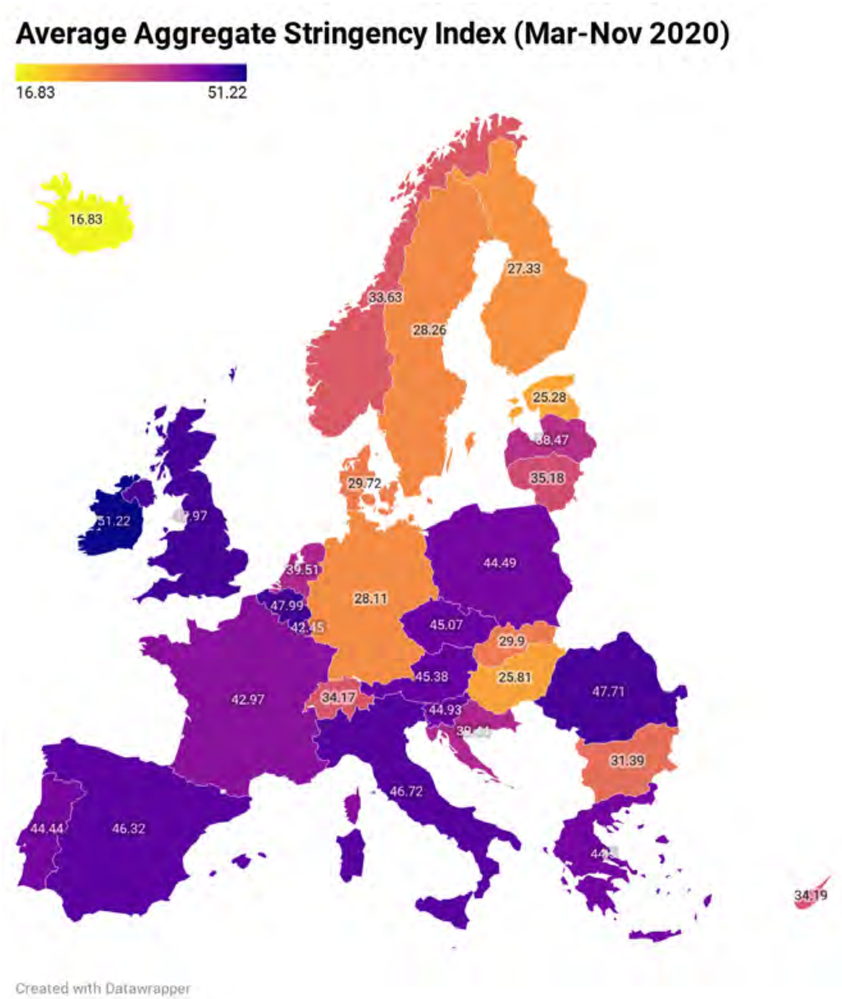


Figure 1: Average Aggregate CEPS-PERISCOPE Stringency Index by Country, Mar-Nov 2020

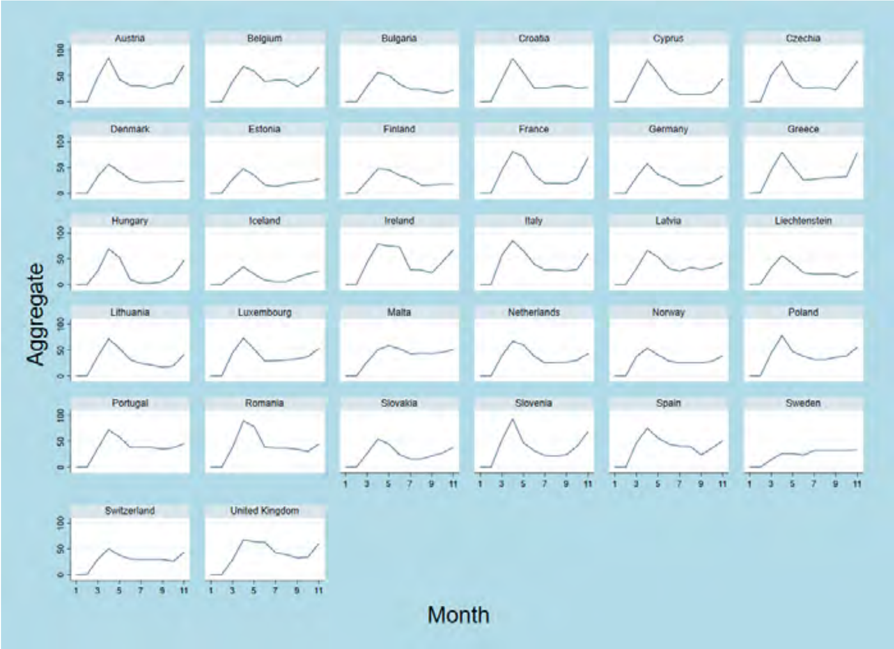


Figure 2: Aggregate CEPS-PERISCOPE Stringency Index by Country, Jan-Nov 2020

an indicator of GDP growth with current and lagged values of changes in the CEPS-PERISCOPE indicator. The dependent variable is the Economic Sentiment indicator[9] (ESI) produced by the European Commission because this is a widely used indicator with a good track record (Gayer et al., 2018), and it is available on a monthly basis. Given that the ESI serves as a proxy for contemporaneous GDP growth (not its level), we first-difference the indices which have been previously normalised to take values between 0 and 100. Hence, we regress the ESI for country i and period t, on the contemporaneous change of index j, $\Delta I_{j,t}$, and two lags. We also include a set of country fixed effects, d_i:

$$ESI_{i,t} = \beta_1 \Delta I_{j,t} + \beta_2 \Delta I_{j,t-1} + \beta_3 \Delta I_{j,t-2} + d_i + \varepsilon_{i,t} \tag{4}$$

We perform a total of nine regressions using data for the period March 2020 to October 2020. Given the relatively simple specification (4) and the small number of periods, the results displayed in Table 4 should be seen as exploratory.

Table 4 shows that almost all the indices have negative and highly significant effects on

[9]A description of the ESI can be found through the following link: https://ec.europa.eu/eurostat/web/products-datasets/-/teibs010.

Covid Economics 66, 28 January 2021: 115-137

the ESI. This confirms that more restrictive policies ($\Delta I_j > 0$) are negatively correlated with current and future economic developments. The estimated coefficients on lag values are generally greater than contemporaneous coefficients which could suggest that stricter policies have larger effects on economic activity with a certain delay. Furthermore, the indices perform quite well to explain the variance of the ESI over the sample period. The R^2 is greater than 50% except for Facial Covering and Stay Home indices. School Closures and the aggregate index, both explain around 70% of the ESI's variance.

The point estimate of the coefficients for the aggregate index can be used to provide an order of magnitude of the impact of increasing (or reducing) average restrictions on the economy. The sum of the coefficients for the three lags considered (see last row of Table 4) is close to 0.4. This would imply that an increase of the overall indicator by one standard deviation or 22 points would lead to a fall in the ESI of about 8.8 points. A standard result in the literature (Gayer et al., 2018) is that each point increase in ESI corresponds to an increase in GDP growth (measured on a year-on-year basis) by one third of a percentage point. It follows that an increase in the restrictiveness indicator of one standard deviation should lead to a fall in GDP of close to 3 percentage points. The total increase in the indicator to the peak of about 80, reached for some countries in April of 2020 could thus explain a fall in GDP of about 12% - close to the value observed in some cases.

4.2 The effectiveness of social distancing measures in controlling infections

A second use of the index would be to test the effectiveness of the NPIs. Identifying the impact of social distancing measures encoded by the ECDC on the course of the disease, as measured for example by the time path of infections, is difficult because governments tend to impose measures when infections are high and rising. Nevertheless, there is already a considerable literature studying the effects of NPIs (usually face masks and lockdown) on infections and deaths, which reports in general a significant impact of these measures on the spread of the virus (Amuedo-Dorantes et al., 2020; Chernozhukov et al., 2021; Karaivanov et al., 2020; Mitze et al., 2020). Our intention is not to present totally novel results. In this work we attempt to illustrate how our index could be used to estimate the effectiveness of NPIs.

Covid Economics 66, 28 January 2021: 115-137

CEPR PRESS

Table 4

Index	Economic Sentiment Indicator $\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	R^2
Closure: Daycare	-0.048^{***} (0.0135)	-0.081^{***} (0.0113)	-0.106^{***} (0.0138)	0.582
Closure: School	-0.096^{***} (0.0153)	-0.082^{**} (0.0152)	-0.150^{***} (0.0137)	0.714
Closure: Public Area	-0.055^{**} (0.0245)	-0.106^{***} (0.0226)	-0.122^{***} (0.0254)	0.543
Closure: Indoor	-0.039^{*} (0.0202)	-0.114^{***} (0.0194)	-0.146^{***} (0.0180)	0.594
Mass Gathering Ban	-0.063^{***} (0.0185)	-0.125^{***} (0.0160)	-0.108^{***} (0.0179)	0.630
Masks	-0.105^{***} (0.0207)	-0.088^{***} (0.0236)	-0.079^{**} (0.0317)	0.424
Stay Home	-0.034^{**} (0.0138)	-0.077^{***} (0.0146)	-0.109^{***} (0.0132)	0.481
Teleworking	-0.044 (0.0460)	-0.160^{***} (0.0305)	-0.134^{***} (0.0161)	0.643
Aggregate	-0.085^{***} (0.0203)	-0.112^{***} (0.0219)	-0.171^{***} (0.0205)	0.638

Each row corresponds to a specific regression. The first three columns display the estimated coefficients for each subcategories (current value, first lag and second lag respectively). Standard errors are given in parenthesis and are clustered at the country level. Stars indicate statistical significance at 10%, 5% and 1% levels. The fourth column shows the R^2 from the regression.

We use an approach which has become standard in econometric analysis of interdependent economic variables, namely a Vector Auto-regressive Model (VAR).[10] Given that the number of infections (and other variables describing the course of the pandemic) are available at a higher frequency. We thus use weekly data at this point[11]. We concentrate on nine countries, including the four largest EU member states.

Our illustrative example uses a simple Structural VAR (SVAR), which is identified by the assumption that the impact of NPIs on infections is not contemporaneous, but occurs with a lag of at least one week. Table 5 displays the estimation results for nine member

[10]For details of the exact model used for our estimations see the appendix.

[11]Daily data would also be available, but it appears that most measures are taken towards the end of the week. The daily data for the restrictiveness indicator thus shows jumps mostly towards the end of the week, which suggests that the use of weekly data is appropriate.

Covid Economics 66, 28 January 2021: 115-137

Covid Economics 66, 28 January 2021: 115-137

states. To keep the specification as simple as possible, we estimate a SVAR with only one lag of dependent variables (SVAR(1)) for all countries. The dependent variables is the log difference of the number of infections and the Aggregate Index in levels.

The results for infections show that the coefficients on the lagged values of the Aggregate Index are negative and statistically significant at a 5% level for all countries except Sweden. This tends to confirm that stricter restrictive measures captured by an increase in the Aggregate Index have an effect in reducing the number of infections. Moreover, our specification assumes that the Aggregate Index reacts contemporaneously with fluctuations in the number of infections, which can give indications on how fast national authorities reacts to surge in infections. The estimated coefficients are positive for most countries but are not significant, except for Sweden at the 10% level.[12] On the other hand, the coefficients on the lag values of infections are positive and statistically significant. These results seem to indicate that national authorities react to infections with a lag of one week. Exceptions are again the Nordic countries and France.[13]

The VAR specification jointly models the dynamics of the number of infections and the Aggregate Index. It is then possible to study how the system evolves following a shock to one of the variables (impulse response functions or irf). For our purpose, we are mainly interested in the response of infections to a positive shock to the Aggregate Index (higher restrictions). These irf and the cumulative irf are depicted in Figures 3 and 4. The irf for the Aggregate Index can be found in the Appendix. The irf in Figure 3 usually display a hump-shaped pattern which implies that the restrictions reached their maximum effects (in absolute values) 2-3 weeks after their implementations (exceptions are France and Finland). This maximum effect lies around 0.1 for most countries and implies a reduction of around 10% in the number of cases. This reduction is statistically significant as indicated by the 95% confidence bounds. For Sweden, infections increase after a shock but this result is not significant. On the other hand, the cumulative irf in Figure 4 suggest that a one standard deviation increase of the index leads to a reduction in infections of about one half three months after the the shock. Sweden constitutes the only exception to this general finding. The confidence bounds for each country are large but this is not really surprising given the simple specification used to generate these

[12] Note that the signs of estimated coefficients for Sweden usually differ from other countries, which is consistent with the evidence that this country pursued a very different strategy to fight the virus.

[13] Fana et al. (2020) indicates a tamer response by the French national authorities to the spread of Covid.

Table 5: SVAR(1) estimation results

Dep. var		AT	BE	FI	FR	DE	IT	NL	SP	SW
		Member States								
$\Delta \ln(\text{Inf}_t)$										
	$\Delta \ln(\text{Inf}_{t-1})$	0.211^{*}	0.471^{***}	-0.144	-0.123	0.484^{***}	0.392^{***}	0.547^{***}	0.320^{**}	0.437^{***}
		(0.106)	(0.105)	(0.130)	(0.083)	(0.114)	(0.092)	(0.097)	(0.107)	(0.119)
	Index_{t-1}	-0.010^{***}	-0.008^{**}	-0.014^{**}	-0.011^{***}	-0.010^{***}	-0.008^{***}	-0.007^{***}	-0.008^{**}	0.008
		(0.002)	(0.004)	(0.005)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.006)
	Constant	0.501^{***}	0.369^{***}	0.450^{***}	0.547^{***}	0.301^{***}	0.385^{***}	0.326^{***}	0.365^{**}	-0.211
		(0.124)	(0.132)	(0.165)	(0.129)	(0.080)	(0.079)	(0.107)	(0.142)	(0.267)
	R^2	0.576	0.649	0.155	0.348	0.711	0.749	0.678	0.440	0.262
Index										
	$\Delta \ln(\text{Inf}_t)$	2.860	-3.769	1.711	4.629	2.899	2.359	0.323	-2.154	3.158^{*}
		(5.834)	(6.376)	(2.521)	(5.831)	(4.817)	(9.851)	(6.468)	(6.367)	(1.773)
	$\Delta \ln(\text{Inf}_{t-1})$	13.43^{***}	11.28^{**}	1.857	9.682^{*}	11.72^{**}	18.33^{**}	11.76^{**}	13.67^{***}	-3.415^{**}
		(4.024)	(5.113)	(2.050)	(4.928)	(3.398)	(6.774)	(5.231)	(4.652)	(1.516)
	Index_{t-1}	1.058^{***}	0.865^{***}	0.889^{***}	0.908^{***}	1.086^{***}	1.040^{***}	0.961^{***}	0.759^{***}	0.578^{***}
		(0.101)	(0.111)	(0.089)	(0.081)	(0.085)	(0.120)	(0.103)	(0.120)	(0.068)
	Constant	-2.843	6.994	3.192	0.712	-2.717	-2.385	1.775	11.22^{*}	13.39^{***}
		(5.339)	(5.681)	(2.810)	(5.630)	(2.786)	(6.138)	(4.746)	(6.045)	(2.074)
	R^2	0.827	0.719	0.775	0.760	0.861	0.792	0.758	0.572	0.751
Sample		weeks 12-52	weeks 12-52	weeks 12-52	weeks 12-52	weeks 12-52	weeks 12-52	weeks 12-52	weeks 12-52	weeks 12-52

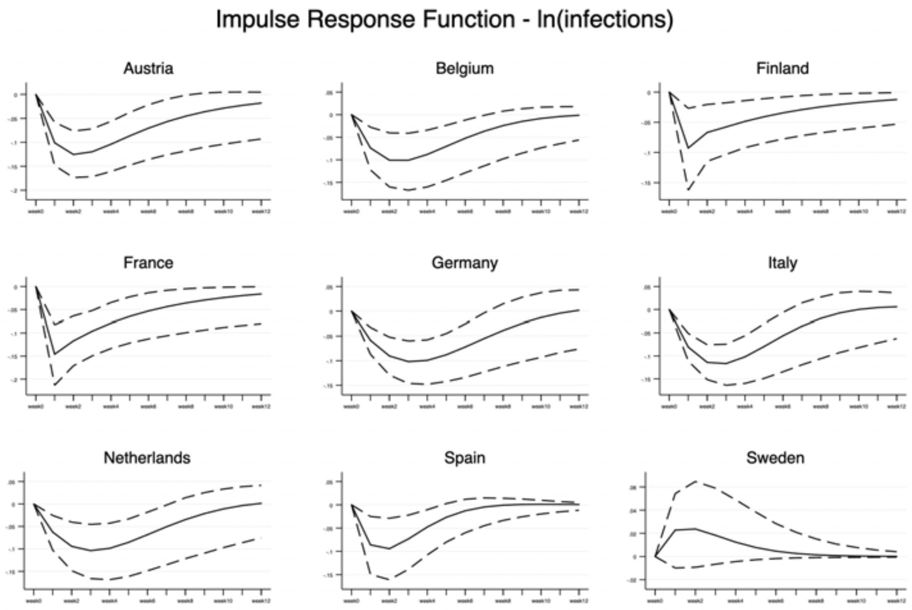
.

results. Nevertheless, the similarity in the results across countries suggests that higher restrictions captured by an increase in our index are effective in reducing the number of infections.

5 Comparison with Oxford Stringency Index

We provide two illustrative comparisons with the Oxford stringency index. Figure 5 shows a scatter plot of monthly values at level for the 29 European countries of our CEPS-PERISCOPE Stringency Index on the vertical axis against the corresponding monthly values for the Oxford Stringency Index on the horizontal. The two indices are likely to refer to the same underlying phenomenon since their correlation coefficient is almost 80 percent. However, there are some observations in February for which our CEPS-PERISCOPE index indicates no restrictions, whereas the Oxford index shows some positive values when some border restrictions were in place in some European countries, which are taken into account by the Oxford Stringency Index but not the ECDC dateset. Figure 6 plots the changes of the Oxford Stringency Index against the changes of CEPS-PERISCOPE Stringency Index. They are also strongly positively correlated. It is apparent that the two indices

Covid Economics 66, 28 January 2021: 115-137



.

Figure 3

also move together over time. Most points are clustered around the 45 degrees diagonal. The correlation coefficient at is 89 percent at change.

We also present a heat-map of the correlation coefficients of the individual elements of our index and five indices produced by Oxford Coronavirus Government Response Tracker (listed as the last five) in Figure 7. Our Aggregate Index is positively correlated with the five different indices produced by Oxford, which are themselves highly correlated. The positive correlations between individual elements are weaker, showing adequate variations of policy mix.

Covid Economics 66, 28 January 2021: 115-137

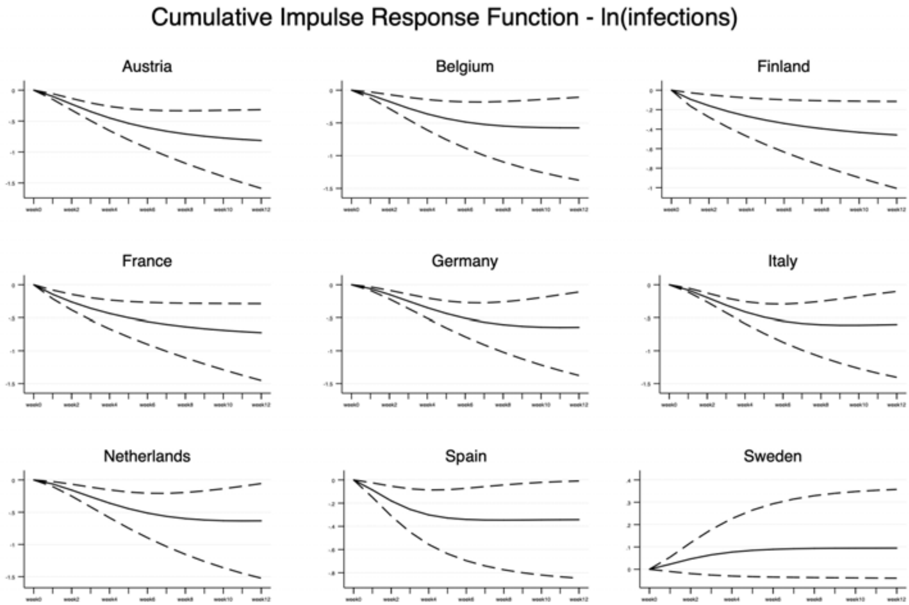


Figure 4

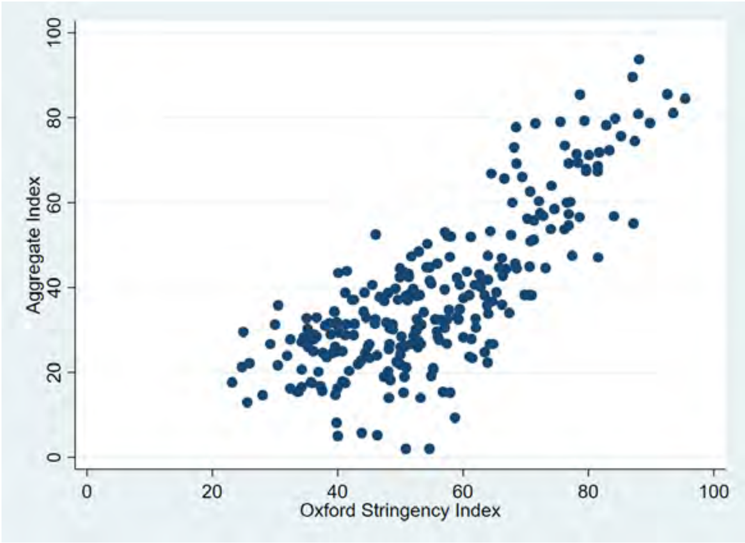


Figure 5: Scatter plot of CEPS-PERISCOPE Aggregate Index and Oxford Stringency Index, March-Nov 2020

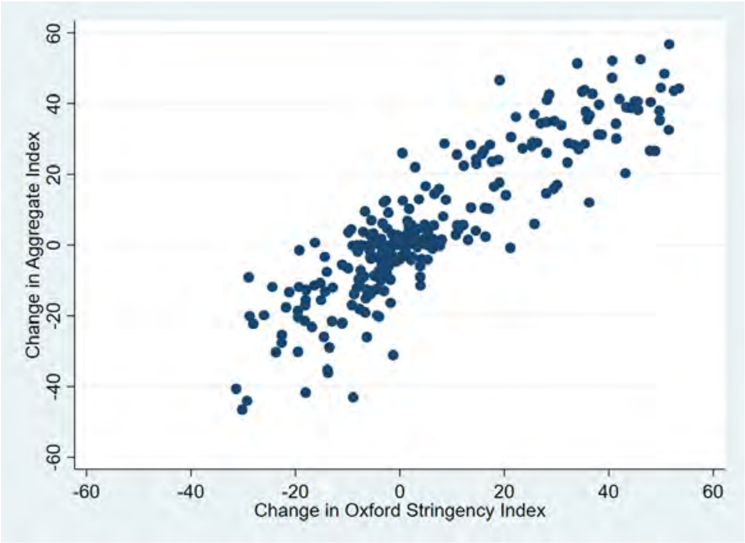


Figure 6: Scatter plot of Changes of CEPS-PERISCOPE Aggregate Index and Changes of Oxford Stringency Index, Mar-Nov 2020

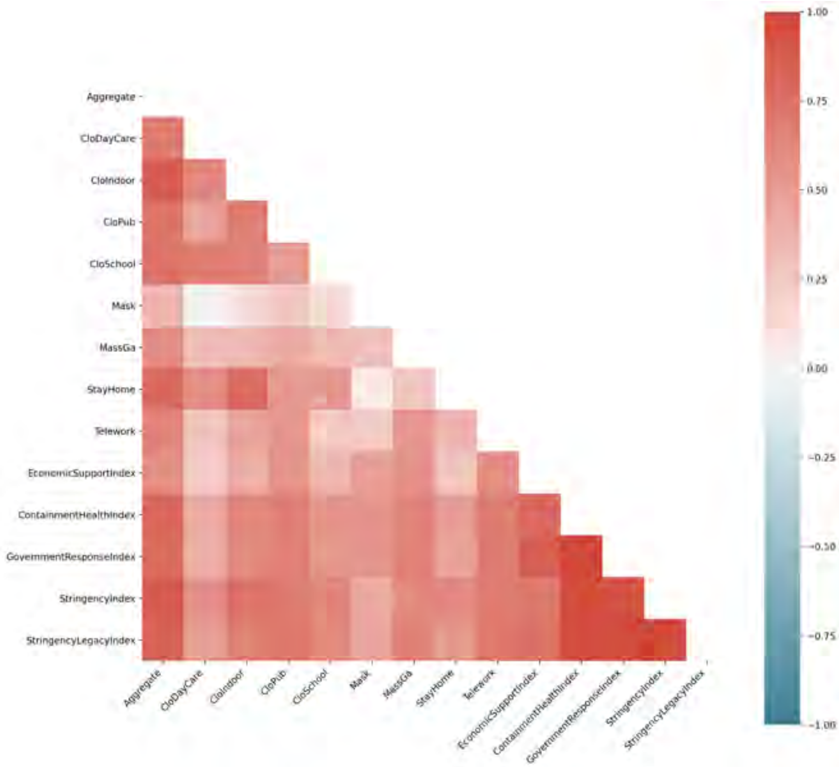


Figure 7: Correlation Heat-map of CEPS-PERISCOPE indices and Oxford Coronavirus Government Response Tracker, Jan-Nov 2020

Covid Economics 66, 28 January 2021: 115-137

6 Conclusions

Using information from the ECDC website we construct a numerical indicator of the restrictiveness of the NPIs imposed by European governments in eight different areas (closing schools, prohibiting gatherings, face coverings, etc.).

This new index can be used by researchers for many purposes. We illustrate two aspects. First, estimating the impact of different measures on the economy, we find that increases in the aggregate index tend to have a strong impact on growth, an impact that lasts for up to two months. Second, we find that tighter NPIs are followed by a slower growth of infection, indicating that the measures have been effective at lowering the reproduction rate of the virus. Our new index is somewhat different, but still highly correlated with the restrictiveness index provided by the University of Oxford.

Covid Economics 66, 28 January 2021: 115-137

References

Adeel, A. B., Catalano, M., Catalano, O., Gibson, G., Muftuoglu, E., Riggs, T., Sezgin, M. H., Shvetsova, O., Tahir, N., VanDusky-Allen, J. et al. (2020), 'COVID-19 policy response and the rise of the sub-national governments', *Canadian Public Policy* **46**(4), 565–584.

Amuedo-Dorantes, C., Kaushal, N. and Muchow, A. N. (2020), Is the cure worse than the disease? county-level evidence from the covid-19 pandemic in the United States, Technical report, National Bureau of Economic Research.

Cheng, C., Barceló, J., Hartnett, A. S., Kubinec, R. and Messerschmidt, L. (2020), 'COVID-19 government response event dataset (CoronaNet v. 1.0)', *Nature Human Behaviour* **4**(7), 756–768.

Chernozhukov, V., Kasahara, H. and Schrimpf, P. (2021), 'Causal impact of masks, policies, behavior on early Covid-19 pandemic in the U.S.', *Journal of Econometrics* **220**(1), 23 – 62.

Diamantopoulos, A. and Winklhofer, H. M. (2001), 'Index construction with formative indicators: An alternative to scale development', *Journal of Marketing Research* **38**(2), 269–277.

Edejer, T. T.-T., Hanssen, O., Mirelman, A., Verboom, P., Lolong, G., Watson, O. J., Boulanger, L. L. and Soucat, A. (2020), 'Projected health-care resource needs for an effective response to COVID-19 in 73 low-income and middle-income countries: a modelling study', *The Lancet Global Health* **8**(11), e1372–e1379.

Fana, M., Tolan, S., Torrejón, S., Urzi Brancati, C. and Fernández-Macías, E. (2020), The COVID confinement measures and EU labour markets, Technical report, Publications Office of the European Union.

Gayer, C., Marc, B. et al. (2018), A 'new modesty'? level shifts in survey data and the decreasing trend of 'normal'growth, Technical report, Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Covid Economics 66, 28 January 2021: 115-137

Hale, T., Angrist, N., Cameron-Blake, E., Hallas, L., Kira, B., Majumdar, S., Petherick, A., Phillips, T., Tatlow, H. and Webster, S. (2020), 'Variation in government responses to COVID-19', *BSG Working Paper Series* .

Hale, T., Atav, T., Hallas, L., Kira, B., Phillips, T., Petherick, A. and Pott, A. (2020), 'Variation in US states responses to COVID-19', *BSG Working Paper Series* .

Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press.

Karaivanov, A., Lu, S. E., Shigeoka, H., Chen, C. and Pamplona, S. (2020), Face masks, public policies and slowing the spread of COVID-19: evidence from Canada, Technical report, National Bureau of Economic Research.

Marozzi, M. (2016), 'Construction, robustness assessment and application of an index of perceived level of socio-economic threat from immigrants: A study of 47 European countries and regions', *Social Indicators Research* **128**(1), 413–437.

Mitze, T., Kosfeld, R., Rode, J. and Wälde, K. (2020), 'Face masks considerably reduce COVID-19 cases in Germany: A synthetic control method approach', *CESifo Working Paper No.8479* .

Yan, B., Zhang, X., Wu, L., Zhu, H. and Chen, B. (2020), 'Why do countries respond differently to COVID-19? a comparative study of Sweden, China, France, and Japan', *The American Review of Public Administration* **50**(6-7), 762–769.

Covid Economics 66, 28 January 2021: 115-137

A Appendix

A.1 SVAR model

Formally, we start from a standard Structural VAR model of order 1 (SVAR(1)) for the number of infected individuals in country i, $y_{i,t}$ in week t, and the policy response of the same country measured by the aggregate index, $x_{i,t}$:

$$\begin{pmatrix} 1 & -\beta_{y_i,0} \\ -\beta_{x_i,0} & 1 \end{pmatrix} \begin{pmatrix} y_{i,t} \\ x_{i,t} \end{pmatrix} = \begin{pmatrix} \mu_y \\ \mu_x \end{pmatrix} + \begin{pmatrix} \beta_{y_i,1} & \beta_{y_i,2} \\ \beta_{x_i,1} & \beta_{x_i,2} \end{pmatrix} \begin{pmatrix} y_{i,t-1} \\ x_{i,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \end{pmatrix} \quad (5)$$

Error terms are assumed to be i.i.d, $\varepsilon_{y_i,t} \sim \mathcal{N}(0, \sigma^2_{y_i})$ and $\varepsilon_{x_i,t} \sim \mathcal{N}(0, \sigma^2_{x_i})$.
In matrix form:

$$A_0 z_{i,t} = \mu + A_1 z_{i,t-1} + \varepsilon_t \quad (6)$$

with $\varepsilon_t \sim \mathcal{N}(0, \Sigma)$.
The term 'structural' is used here in a wider sense. The underlying 'model' that justifies our set-up is the fact that restrictive measures are usually introduced in response to an increase in infections (implying that the indicator should be function of lagged infections) and that these restrictions are meant to reduce infections (implying that infections should be a function of lagged restrictions). See Hamilton (1994) for examples of SVAR in economics.

To ensure identification of the structural VAR we impose a restriction on the contemporaneous coefficients in A_0. This seems reasonable given the weekly frequency of the data. Restrictions imposed in a given week (usually at the end of the week) cannot be expected to impact infections during that same week. Hence we assume that the number of infections does not react contemporaneously to the policy restrictions but only with a lag, that is $\beta_{y_i,0} = 0$.

The SVAR model can then be estimated equation by equation using Ordinary Least Squares. We check for stationarity of the results from the reduced form VAR:

$$z_{i,t} = \tilde{\mu} + \tilde{A}_1 z_{i,t-1} + \tilde{\varepsilon}_t \quad (7)$$

where $\tilde{\mu} = A_0^{-1}\mu$, $\tilde{A}_1 = A_0^{-1}A_1$ and $\tilde{\varepsilon}_t = A_0^{-1}\varepsilon_t$. Stationarity is achieved when the eigenvalues of $\tilde{A}_1$ lie within the unit circle (Hamilton, 1994).

Covid Economics 66, 28 January 2021: 115-137

A.2 Additional irf

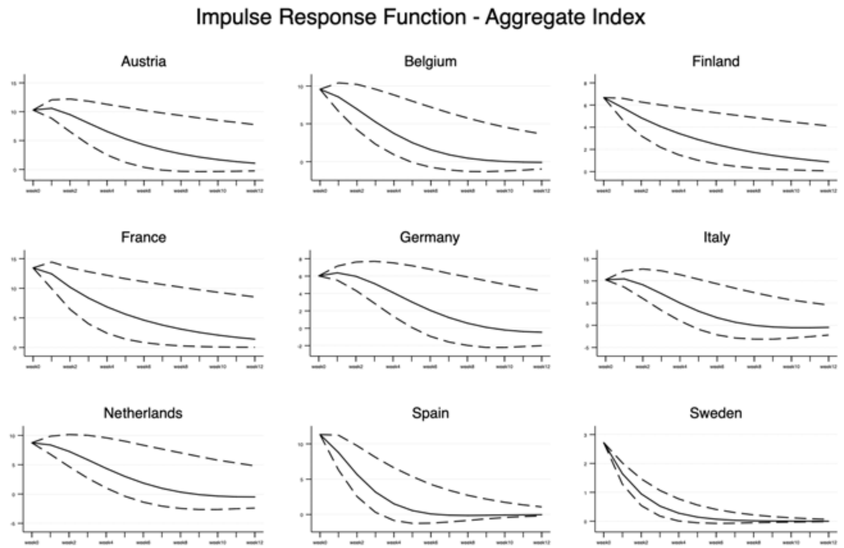


Figure 8

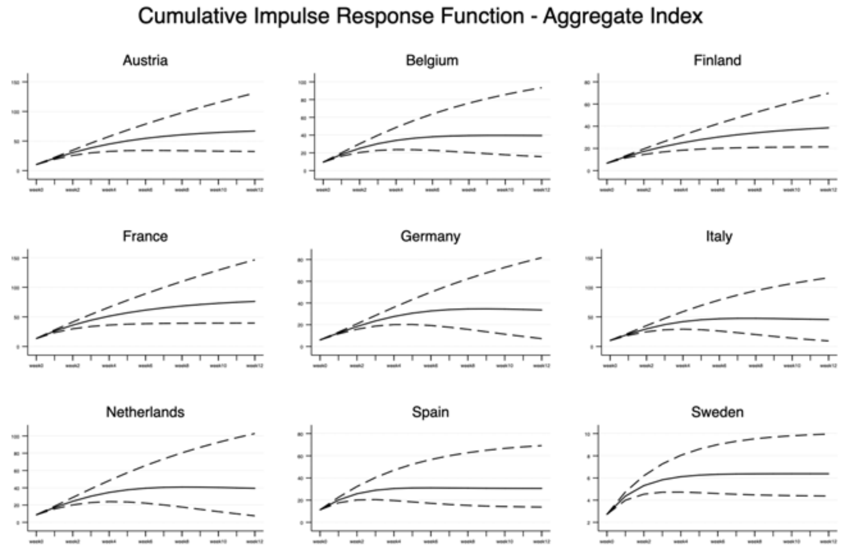


Figure 9