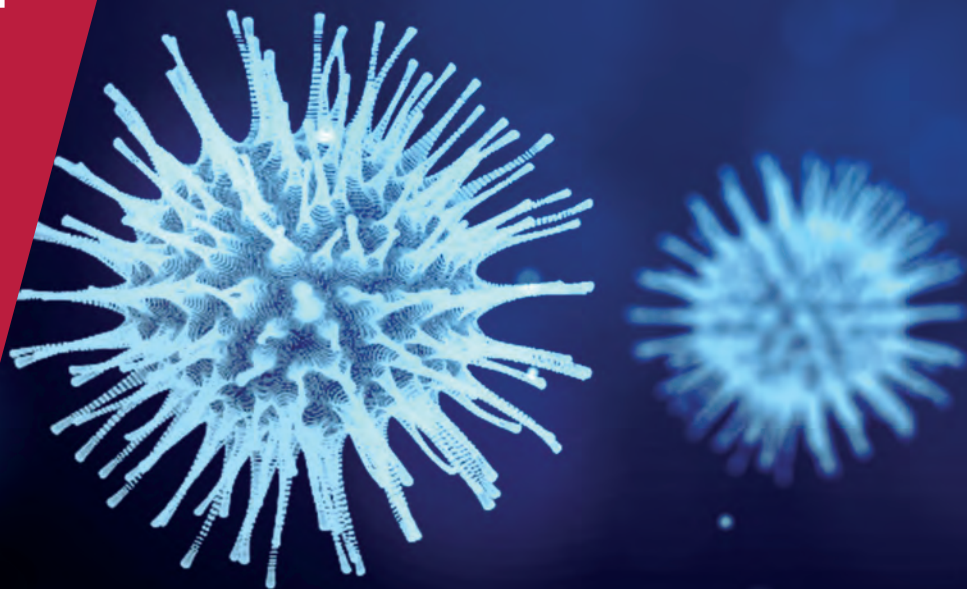


**CENTRE FOR
ECONOMIC
POLICY
RESEARCH**

CEPR PRESS



COVID ECONOMICS
VETTED AND REAL-TIME PAPERS

ISSUE 58
19 NOVEMBER 2020

**OVERDIAGNOSIS AND
UNDERTESTING**

Tinglong Dai and Shubhranshu Singh

AGE-VARYING EPIDEMICS

Marta Giagheddu and Andrea Papetti

CONSUMPTION IN SOUTH KOREA

Changyong Kim,
Estefania Santacreu-Vasut and
Edward Keunuk Shin

**ONLINE LEARNING WHEN
SCHOOLS ARE CLOSED**

Masato Ikeda and Shintaro Yamaguchi

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.

Copyright for the papers appearing in this issue of *Covid Economics: Vetted and Real-Time Papers* is held by the individual authors.

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.

Chair of the Board

Sir Charlie Bean

Founder and Honorary President

Richard Portes

President

Beatrice Weder di Mauro

Vice Presidents

Maristella Bott cini

Ugo Panizza

Philippe Martin

Hélène Rey

Chief Executive Officer

Tessa Ogden

Editorial Board

Beatrice Weder di Mauro, CEPR

Charles Wyplosz, Graduate Institute Geneva and CEPR

Viral V. Acharya, Stern School of Business, NYU and CEPR

Guido Alfani, Bocconi University and CEPR

Franklin Allen, Imperial College Business School and CEPR

Michele Belot, European University Institute and CEPR

David Bloom, Harvard T.H. Chan School of Public Health

Nick Bloom, Stanford University and CEPR

Tito Boeri, Bocconi University and CEPR

Alison Booth, University of Essex and CEPR

Markus K Brunnermeier, Princeton University and CEPR

Michael C Burda, Humboldt Universitaet zu Berlin and CEPR

Aline Bütikofer, Norwegian School of Economics

Luis Cabral, New York University and CEPR

Paola Conconi, ECARES, Universite Libre de Bruxelles and CEPR

Giancarlo Corsetti, University of Cambridge and CEPR

Fiorella De Fiore, Bank for International Settlements and CEPR

Mathias Dewatripont, ECARES, Universite Libre de Bruxelles and CEPR

Jonathan Dingel, University of Chicago Booth School and CEPR

Barry Eichengreen, University of California, Berkeley and CEPR

Simon J Evenett, University of St Gallen and CEPR

Maryam Farboodi, MIT and CEPR

Antonio Fatás, INSEAD Singapore and CEPR

Francesco Giavazzi, Bocconi University and CEPR

Christian Gollier, Toulouse School of Economics and CEPR

Timothy J. Hatton, University of Essex and CEPR

Ethan Ilzetzki, London School of Economics and CEPR

Beata Javorcik, EBRD and CEPR

Simon Johnson, MIT and CEPR

Sebnem Kalemli-Ozcan, University of Maryland and CEPR Rik Frehen

Tom Kompas, University of Melbourne and CEBRA

Miklós Koren, Central European University and CEPR

Anton Korinek, University of Virginia and CEPR

Michael Kuhn, Vienna Institute of Demography

Maarten Lindeboom, Vrije Universiteit Amsterdam

Philippe Martin, Sciences Po and CEPR

Warwick McKibbin, ANU College of Asia and the Pacific

Kevin Hjortshøj O'Rourke, NYU Abu Dhabi and CEPR

Evi Pappa, European University Institute and CEPR

Barbara Petrongolo, Queen Mary University, London, LSE and CEPR

Richard Portes, London Business School and CEPR

Carol Propper, Imperial College London and CEPR

Lucrezia Reichlin, London Business School and CEPR

Ricardo Reis, London School of Economics and CEPR

Hélène Rey, London Business School and CEPR

Dominic Rohner, University of Lausanne and CEPR

Paola Sapienza, Northwestern University and CEPR

Moritz Schularick, University of Bonn and CEPR

Flavio Toxvaerd, University of Cambridge
Christoph Trebesch, Christian-Albrechts-Universitaet zu Kiel and CEPR

Karen-Helene Ulltveit-Moe, University of Oslo and CEPR

Jan C. van Ours, Erasmus University Rotterdam and CEPR

Thierry Verdier, Paris School of Economics and CEPR

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.

<i>American Economic Review</i>	<i>Journal of Economic Growth</i>
<i>American Economic Review, Applied Economics</i>	<i>Journal of Economic Theory</i>
<i>American Economic Review, Insights</i>	<i>Journal of the European Economic Association*</i>
<i>American Economic Review, Economic Policy</i>	<i>Journal of Finance</i>
<i>American Economic Review, Macroeconomics</i>	<i>Journal of Financial Economics</i>
<i>American Economic Review, Microeconomics</i>	<i>Journal of International Economics</i>
<i>American Journal of Health Economics</i>	<i>Journal of Labor Economics*</i>
<i>Canadian Journal of Economics</i>	<i>Journal of Monetary Economics</i>
<i>Econometrica*</i>	<i>Journal of Public Economics</i>
<i>Economic Journal</i>	<i>Journal of Public Finance and Public Choice</i>
<i>Economics of Disasters and Climate Change</i>	<i>Journal of Political Economy</i>
<i>International Economic Review</i>	<i>Journal of Population Economics</i>
<i>Journal of Development Economics</i>	<i>Quarterly Journal of Economics</i>
<i>Journal of Econometrics*</i>	<i>Review of Corporate Finance Studies*</i>
	<i>Review of Economics and Statistics</i>
	<i>Review of Economic Studies*</i>
	<i>Review of Financial Studies</i>

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

Covid Economics

Vetted and Real-Time Papers

Issue 58, 19 November 2020

Contents

COVID-19 diagnosis and viral load reporting: A theory of overdiagnosis and undertesting <i>Tinglong Dai and Shubhranshu Singh</i>	1
The macroeconomics of age-varying epidemics <i>Marta Giagheddu and Andrea Papetti</i>	22
Trade-off between health and wealth? Insights from COVID-19 in South Korea <i>Changyong Kim, Estefania Santacreu-Vasut and Edward Keunuk Shin</i>	57
Online learning during school closure due to COVID-19 <i>Masato Ikeda and Shintaro Yamaguchi</i>	85

COVID-19 diagnosis and viral load reporting: A theory of overdiagnosis and undertesting

Tinglong Dai¹ and Shubhranshu Singh²

Date submitted: 5 November 2020; Date accepted: 13 November 2020

The ongoing discourse about COVID-19 testing revolves around undertesting (i.e., insufficient testing capacity relative to demand). An important yet little studied systematic issue is overdiagnosis (i.e., positive diagnoses for patients with negligible viral loads): recent evidence shows U.S. laboratories have adopted a hyper-sensitive diagnosis criterion for COVID-19 testing, such that up to an estimated 90% of positive diagnoses are for minuscule virus loads. Motivated by this situation, we develop a theory of testing for COVID-19 that explains both undertesting and overdiagnosis. We show that a laboratory has an incentive to inflate the diagnosis criterion, which generates a higher diagnosis-driven demand as a result of contact-tracing efforts, albeit while dampening demand from disease transmission. An inflated diagnosis criterion prompts the laboratory to build a higher testing capacity, which may not fully absorb the inflated demand, so undertesting arises. Finally, we examine a social planner's problem of whether to mandate the laboratory to report viral load along with its diagnosis, such that a physician or contact tracer can make informed triage decisions. The social planner may prefer not to mandate viral load reporting, because it induces a higher testing capacity and may help reduce disease transmission.

- ¹ Associate Professor of Operations Management and Business Analytics, Carey Business School, Johns Hopkins University; Core Faculty, Hopkins Business of Health Initiative (HBHI); Affiliated Faculty, Johns Hopkins School of Nursing; Affiliated Faculty, Johns Hopkins Institute for Data Intensive Engineering and Science (IDIES).
- ² Associate Professor of Marketing, Carey Business School, Johns Hopkins University; Affiliated Faculty, Department of Economics, Johns Hopkins University.

Copyright: Tinglong Dai and Shubhranshu Singh

1. Introduction

In a recent *JAMA* article, Cutler and Summers (2020) estimate the coronavirus disease 2019 (COVID-19) pandemic will cost the U.S. a total of \$16 trillion, over 90% of its annual economic output. A fateful contributor to this staggering loss is the “lost month,” that is, the time between late January and early March 2020, during which the U.S. failed to roll out abundant and usable COVID-19 testing, leading to widespread community infection without containment or detection (Shear et al. 2020). As of November 1, 2020, many parts of the U.S. continued to experience undertesting (i.e., a lack of laboratory testing capacity relative to the demand for testing).

Whereas the issue of undertesting has been widely recognized and investigated, another important issue — overdiagnosis — has not received due attention. By “overdiagnosis,” we do not mean patients are given wrong diagnoses. Rather, we mean patients with negligible viral loads are given positive diagnoses.¹ According to a *New York Times* article entitled “Your Coronavirus Test is Positive. Maybe It Shouldn’t Be” (Mandavilli 2020),

In three sets of testing data that include cycle thresholds, compiled by officials in Massachusetts, New York and Nevada, up to 90 percent of people testing positive carried *barely any* virus [emphasis added], a review by The Times found. On Thursday, the United States recorded 45,604 new coronavirus cases, according to a database maintained by The Times. If the rates of contagiousness in Massachusetts and New York were to apply nationwide, then perhaps only 4,500 of those people may actually need to isolate and submit to contact tracing.

Overdiagnosis of COVID-19 arises because when conducting the polymerase chain reaction (PCR) test, the gold standard of COVID-19 diagnosis,² U.S. laboratories rely on an unusually high magnification factor, as measured by the cycle threshold (CT), that is, the number of cycles necessary for spotting the coronavirus. (The PCR process amplifies DNA segments cycle by cycle. A total of 30 cycles results in 2^{30} , approximately 1 billion, copies of the original DNA segments, whereas a total of 40 cycles results in 2^{40} , approximately 1 *trillion*, copies.) According to Mandavilli (2020), “most tests set the limit at 40, a few at 37,” which means patients are given positive diagnoses if

¹ The concept of overdiagnosis has been widely used in medicine. In the context of cancer, Esserman et al. (2013) define overdiagnosis as a scenario “when tumors are detected that, if left unattended, would not become clinically apparent or cause death.” Welch et al. (2011) use the analogy of a turtle to illustrate that certain cancers progress so slowly that they are unlikely to have meaningful health impacts during one’s lifetime. More broadly, Brodersen et al. (2018) define overdiagnosis as “making people patients unnecessarily, by identifying problems that were never going to cause harm or by medicalising ordinary life experiences through.” Our use of the concept is aligned with the medical literature while incorporating the infection risk that a patient poses to others. Note that “overdiagnosis” differs from “overtesting;” the latter refers to over-utilization of diagnostic testing (Brodersen et al. 2018).

² We focus on the PCR test throughout the paper. In practice, a number of “rapid tests,” such as the ID NOW COVID-19 test developed by Abbott Laboratories, have been used in various situations. However, experts have expressed concerns about their relatively low sensitivity and high false-negative rates (Madrigal and Meyer 2020). Another reason we focus on the PCR test is that it is a standard diagnostic tool for many infectious diseases, including Ebola, HIV, SARS, and tuberculosis (Salis 2009), meaning our model has broader implications for future pandemics (or public health crises).

the PCR test takes up to 40 or 37 cycles. Yet, “tests with thresholds so high may detect not just live virus but also genetic fragments, leftovers from infection that pose no particular risk — akin to finding a hair in a room long after a person has left.” For this reason, experts agree that tests with CT above 35 are too sensitive, and that “a more reasonable cutoff would be 30 to 35.” Echoing the above *New York Times* article, in a July 17, 2020, interview with *This Week in Virology*, Dr. Anthony Fauci explained (TWiV 2020),

If you get a cycle threshold of 35 or more, the chance of it being replication competent are minuscule... It's very frustrating for the patients as well as for the physicians that someone comes in and they repeat their PCR and it's like 37 cycle threshold, but you almost never can culture virus from a 37, 38, even 36 cycle. It's just dead nucleotides. Period.

Even as they rely on a hyper-sensitive diagnostic standard in diagnosing whether a patient suffers from COVID-19, U.S. laboratories rarely, if ever, report the CT values along with a binary diagnosis (i.e., positive or negative), blinding physicians, contact tracers, and others who seek to assess the viral load behind the binary diagnosis. Physician leaders in the U.S. increasingly demand a departure from the lack of reporting of viral load, as evidenced in a recent *Science* article entitled “A Call for Diagnostic Tests to Report Viral Load” (Service 2020), which contends that including the CT value in the diagnosis can help (1) contact tracers “triage their efforts based on CT values” and (2) physicians “flag patients most at risk for severe disease and death.” The call has been echoed outside the U.S. In India, “many doctors are now telling patients that their COVID-19 test reports should mention the cycle threshold (CT) value, and not just the positive or negative outcome,” according to a *Times of India* article (Rao 2020).

In this paper, motivated by this situation, we develop a theory of laboratory testing for COVID-19 that explains both undertesting and overdiagnosis. A laboratory (hereafter, “lab”) sets the CT cutoff it uses to determine whether a patient has a positive or negative condition: if a patient has a detectable level of virus before reaching the CT cutoff, the patient is diagnosed as having a positive condition, and vice versa. In the absence of revenue incentives, the lab would choose a CT cutoff to minimize the patient’s total expected cost from providing a false-positive diagnosis as well as that from providing a false-negative diagnosis. In addition to choosing the CT cutoff, the lab chooses its future capacity level by weighing demand for testing resulting from both contact tracing and community spread.

When the lab is concerned about its future revenue, we show that, under reasonable parametric assumptions, it has an incentive to inflate the PCR diagnosis criterion, as specified by the CT cutoff. The inflated CT cutoff generates a higher demand for testing as a result of contact-tracing efforts. At the same time, the inflated CT cutoff reduces the chance of false-negative diagnoses, and thus helps reduce community spread; as a result, the lab faces dampened demand from disease

transmission. In view of this tradeoff, our further analysis reveals the lab is more likely to inflate the diagnosis criterion if the intensity of contact tracing is high relative to the expected contagiousness.

Our model allows the lab's diagnostic policy to interact with its capacity decision. We show an inflated diagnosis criterion incentivizes the lab to build a higher testing capacity. Yet, the increased capacity may not fully absorb the inflated demand. As a result, the phenomenon of undertesting (i.e., the probability that all demand can be satisfied is lower than when the lab does not inflate the diagnosis criterion) can arise despite the high testing capacity. In this sense, we have developed a unifying theory that explains both undertesting and overdiagnosis.

We also analyze a social planner's problem of whether to mandate the lab to report the viral-load information along with a binary diagnosis, such that a physician or contact tracer can better assess the patient's condition. We show not mandating the lab's reporting of viral-load information may be in the social planner's best interest for the following reason. When the lab is mandated to report the viral-load information along with each test result, other parties (including, e.g., physicians, contact tracers, and patients) can use the viral-load information to triage contact tracing efforts that contribute to additional demand for testing. Stated differently, by disclosing the viral-load information, the lab essentially gives up its demand-inducing power to others. Because other parties will rely on a lower CT cutoff in their decisions, the lab faces a lower total demand, all else being the same. Accordingly, the lab under-invests in building testing capacity. In addition, a lower CT cutoff increases the patient's likelihood of being tested false negative, which contributes to community spread. Therefore, the social planner may find it more desirable not to mandate reporting of the viral-load information.

Our paper contributes to the ongoing discourse of diagnostic testing by providing a unified theory of overdiagnosis and undertesting. It suggests that in the absence of direct public investment in testing capacity, both phenomena, as undesirable as they are, may necessarily arise and are indeed out of the social planner's rational choice. Our findings have important implications for the ongoing pandemic and future public health crises. Echoing the view expressed by Nobel laureate Gary Becker (2020, p. 55), our findings suggest governments may rethink their approach to building testing capacity predominantly through the private sector, given the oversized economic impact and immeasurable human cost of a global pandemic. Public-private partnerships and innovative contracting mechanisms aimed at disentangling labs' diagnostic testing decision from incentives to build testing capacity may be worth exploring.

2. Literature

Our paper is related to the literature on diagnostic services. This literature originated from Darby and Karni (1973) and initially focused on strategies to reduce fraud by credence-goods sellers (e.g.,

automobile mechanics, electronic specialists, and physicians). For example, Dranove (1988) studies the negative effect of unnecessary treatments on reputation, and Wolinsky (1993) shows reputation considerations can help discipline experts in these markets. Subsequently, Arora and Fosfuri (2005) study buyers' willingness to pay for diagnostic information that can be used for decision-making. Durbin and Iyer (2009) show the presence of side payments can facilitate truthful communication by a corruptible diagnostician. Jiang et al. (2014) study a diagnostic expert's pricing decision as a signal of her level of altruism. Dai and Singh (2020) study how a physician can use the diagnostic process to signal her diagnostic skills to peers and show undertesting can serve as a signal of diagnostic ability. Our paper contributes to this literature by studying the externality of individual diagnosis decisions, which has implications for both short- and long-term demand for diagnostic testing and poses a novel economic trade-off. In doing so, it enriches this literature by formalizing the interaction between diagnostic decision-making and testing capacity.³

Another related literature concerns the theme of information disclosure and emerges from the seminal work by Grossman (1981) and Milgrom (1981), which establishes full unraveling of private quality information. Subsequent work has revealed a myriad of drivers of *partial* information disclosure. For example, Jovanovic (1982) and Verrecchia (1983) argue the presence of disclosure costs can result in partial disclosure, whereas Dye (1985) shows that a manager who is believed to be imperfectly informed may be able to hide information. Bhardwaj et al. (2008) and Mayzlin and Shin (2011) reason that a limited disclosure bandwidth may force a player to disclose only a subset of information. Guo and Zhao (2009) show competition can reduce firms' incentive to disclose information. Similar to our paper, Zhang (2014) studies implications of a mandatory-disclosure requirement imposed by the policymaker. However, the focus of that paper is on the policy's effect on consumers' inference of quality.⁴ In our paper, the lab conceals patient information to create additional future demand for testing. Because this information-concealment behavior also helps reduce the spread of disease transmission, a mandatory-disclosure requirement may not be desired even from the social planner's perspective.

Finally, our paper contributes to the literature on the delegation of decision rights. Lal (1986) and Bhardwaj (2001) examine the implication of delegation of pricing decisions in the context of salesforce. Özer et al. (2018) study consequences of information sharing, advice provision, and delegation on cooperation between two parties. Dogan et al. (2018) examine the effect of automation on a principal's incentive to delegate its decision-making rights to a manager. In our paper, the

³ Our paper also contributes to a fast-growing literature that addresses the economics of the COVID-19 pandemic by being the first to examine diagnostic laboratories' decision-making. We refer interested readers to an online journal, *Covid Economics*, available at <http://bit.ly/Covid19Econ>, for an up-to-date sample of the literature.

⁴ Another related stream of literature (e.g., Gal-Or et al. 2007; Guo 2009, 2020; Sun and Tyagi 2020) studies information-disclosure incentives in the context of distribution channels.

lab decides whether to keep the right of deciding the patient's diagnosis or disclose the viral load on the test result and delegate the authority to physicians.

3. Model

There are two periods (1 and 2). Consider a mass 1 of individuals, a proportion ϕ of whom are infectious at the beginning of period 1. The remaining $(1 - \phi)$ are not infectious, either because they have never been infected or because they carry a negligible amount of viral residue. After a patient becomes infectious, symptoms (e.g., fever, cough, headache, fatigue, and new loss of taste or smell) may take several days to appear. In addition, some COVID-19 patients fully recover without ever showing any symptoms (Buitrago-Garcia et al. 2020). We assume the probability that an infectious patient shows COVID-19 symptoms is α . Equivalently, a proportion α of all infectious patients show symptoms. The remaining proportion $(1 - \alpha)$ of infectious patients are asymptomatic. COVID-19 symptoms are not unique. Patients may experience COVID-19 symptoms for a variety of minor and non-COVID-19 conditions. We assume a proportion β of individuals (where, $\beta < \alpha$) who are not infectious also show COVID-19-like symptoms. Patients with symptoms seek testing for the virus.

A lab performs a PCR test (hereafter, “test”) to determine if a patient is infectious. The test returns a CT value, denoted by x , a proxy for viral load. We assume x is continuous and normalize the range of x to $[0, 1]$.⁵ The value of x returned by the test is stochastic and depends on whether the patient is infectious. For an infectious patient, the value of x is a draw from the distribution with cumulative distribution function (cdf) $G(x) = 1 - (1 - x)^2$, whereas for a patient who is not infectious, the test returns an x that is a draw from the distribution with cdf $H(x) = x^2$. Both of these conditional distributions have supports of $[0, 1]$. The CT value (x) returned by the test is private information of the lab.

Policymakers may mandate the lab to disclose the CT value to physicians. The lab follows the policymakers' directions. If the policymaker mandates disclosure of the CT value, the physician ignores the lab's binary diagnosis, if any, and uses the CT value as the basis for diagnosing the patient. The physician sets a CT cutoff $\kappa_p \in [0, 1]$ — the subscript p stands for “physician” — and diagnoses a patient as positive if $x \leq \kappa_p$ and negative if $x > \kappa_p$.⁶ (We use “physician” as the proxy

⁵ Although a cycle number is, by definition, an integer, a patient's CT value can be fractional, because it measures the number of cycles at which the fluorescence passes the threshold, so it can be between two cycles. For example, according to a *New York Times* story (Grynbaum 2020) published on October 14, 2020, “Mr. Trump's P.C.R. test had a cycle threshold — a proxy for viral load — of 34.3, Dr. [Anthony] Fauci said.” The CT value led Dr. Fauci to conclude, “We can say with a high degree of confidence that he is not transmissible.”

⁶ We define a true-position condition as one in which the patient has a meaningful viral load such that further treatment and contact-tracing efforts are necessary. A true-negative condition does not necessarily mean the patient has zero virus; it simply means the patient does not have enough viruses to be infectious or develop severe conditions.

of individuals who act in the best interest of the patient. In real life, these individuals can include nurses, contact tracers, and caregivers.) In such a case, an infectious patient may be diagnosed negative (i.e., false negative) with probability $[1 - G(\kappa_p)]$. The infectious patient incurs a cost c_{FN} if she is diagnosed as negative; the cost includes, for example, the risk posed to the patient's family members and other close contacts. The patient's expected cost from a false-negative diagnosis is therefore $[1 - G(\kappa_p)] \cdot c_{FN}$. Similarly, a patient who is not infectious is diagnosed as positive with probability $H(\kappa_p)$. We represent the patient's cost in the event of a false-positive diagnosis by c_{FP} , which can include, for example, the patient's lost potential productivity and unnecessary isolation from family (Wu 2020). The patient's expected cost from a false-positive diagnosis is therefore $H(\kappa_p) \cdot c_{FP}$. The physician, who acts in the best interest of the patient, chooses a CT cutoff κ_p that minimizes the patient's expected cost from false-positive and false-negative diagnoses. In doing so, the physician equivalently maximizes patient utility, because the benefits of true-positive and true-negative diagnosis can be absorbed into c_{FN} and c_{FP} , respectively. It is straightforward to show $\kappa_p = c_{FN} / (c_{FN} + c_{FP})$.

In the absence of mandatory-disclosure requirements, the lab must decide how it will disclose the test results. The lab may set a CT cutoff $\kappa_l \in [0, 1]$ on its own and conclude the patient is positive if $x \leq \kappa_l$ and negative if $x > \kappa_l$. In this case, the lab does not disclose the CT value (x) or the CT cutoff (κ_l) it used to determine the test outcome. It only discloses whether the patient has a positive condition. Alternatively, the lab may voluntarily disclose the CT value and let the physician determine if the patient has a positive condition. After setting its viral-load disclosure policy, the lab starts testing patient samples.

The lab generates a revenue of r_t for each tested sample. Because r_t is usually determined by the insurance company and is not dependent on the lab's disclosure decision or the volume of tests processed, we assume it is exogenously given. In period 1, the lab has a limited and exogenously given testing capacity C_1 , which is insufficient to meet the overwhelming period-1 demand for testing. Therefore, a subset of patients is randomly selected from those seeking testing to match the testing capacity. In addition to setting the test-result disclosure policy in period 1, the lab determines its testing capacity C_2 for period 2 (which is also the terminal period). The cost of period-2 testing capacity is given by γC_2^2 , where $\gamma > 0$ is the capacity-cost parameter.

All patients with positive test results are quarantined; these patients include both true and false positives. Once quarantined, the true-positive patients do not infect others and false positive patients do not become infectious. However, many infectious patients are either not tested or tested false-negative in period 1. Because these patients are not quarantined, they come in contact with others and spread the virus. We assume each such patient infects R_0 other individuals. (In epidemiology, R_0 is known as the basic reproduction number and captures the infectiousness of the

disease.) We model R_0 as a random variable and its value is realized at the beginning of period 2. We denote by $F(\cdot)$ its cdf. For ease of analysis, we assume R_0 follows a uniform distribution with a support of $[0, \bar{R}_0]$.

In period 2, in addition to those with symptoms, individuals, who came in direct contact with those tested positive in period 1, seek to be tested. We refer to the demand from this group as contact-tracing demand (or diagnosis-induced demand). It is determined by the contact-tracing intensity ρ , which represents the number of individuals who seek testing because they came in direct contact with someone who was subsequently tested positive in period 1.⁷ Similar to period 1, if demand for testing is higher than the period-2 testing capacity, a subset of all the individuals seeking testing is randomly selected to match the testing capacity C_2 . We assume the lab has a discount factor of 1 for simplicity of analysis.

The timing of actions is the following. All the policy and capacity decisions are made in period 1, which has five stages. In the first stage, the policymaker decides whether the disclosure of the CT value is mandatory. Next, the lab makes CT-value-disclosure and period-2-capacity decisions. Third, period-1 patients are tested and the CT value or the diagnosis is shared with the physician. Fourth, if the lab discloses the CT value from the test, the physician determines (based on her own threshold) whether the patient is positive, whereas if the lab only shares the diagnosis and not the CT value, the physician follows the lab's diagnosis. Fifth, patients who test positive are quarantined. At the beginning of period 2, the demand for testing is realized based on contact-tracing intensity and the number of patients with symptoms. The next two stages of period 2 are similar to the third and fourth stages of period 1. Finally, payoffs are realized.

4. Analysis and Insights

In this section, we analyze our model and derive the lab's optimal diagnostic threshold and future testing capacity. We also analyze the social planner's problem of whether to require the lab to disclose the CT value (x).

To start with, we determine the realized demand for testing D_2 in period 2 for a given CT cutoff (κ_t) set by the lab in period 1. We denote the mass of patients with symptoms in period 1 by s_1 , which is given by $\phi\alpha + (1 - \phi)\beta$. All of these patients seek COVID-19 testing. However, because the period-1 testing capacity is limited (specifically, $C_1 < s_1$), only a subset of these patients are tested. A total of C_1 patients are randomly selected and tested. Therefore, among the patients who are tested in period 1, the probability that a patient is infectious is $\frac{\alpha\phi}{\alpha\phi + (1 - \phi)\beta}$.

⁷ Note that ρ measures demand for additional testing as an immediate result of a positive diagnosis. Such demand can come from official contact tracing efforts as well as individuals' voluntary testing decisions. In the rest of the paper, for ease of exposition, we use "contact tracing" to refer to all such efforts that contribute to diagnosis-induced demand.

For a given κ_l , the probability that the lab classifies a patient as positive (or infectious) is given by:

$$\begin{aligned} \Pr(x \leq \kappa_l) &= \frac{\alpha\phi}{\alpha\phi + (1-\phi)\beta} G(\kappa_l) + \left[1 - \frac{\alpha\phi}{\alpha\phi + (1-\phi)\beta}\right] H(\kappa_l) \\ &= \frac{\alpha\phi(2\kappa_l - \kappa_l^2) + (1-\phi)\beta\kappa_l^2}{\alpha\phi + (1-\phi)\beta}, \end{aligned}$$

and the probability that the patient tests negative is

$$\begin{aligned} \Pr(x > \kappa_l) &= \frac{\alpha\phi}{\alpha\phi + (1-\phi)\beta} [1 - G(\kappa_l)] + \left[1 - \frac{\alpha\phi}{\alpha\phi + (1-\phi)\beta}\right] [1 - H(\kappa_l)] \\ &= \frac{\alpha\phi(1 - \kappa_l)^2 + (1-\phi)\beta(1 - \kappa_l^2)}{\alpha\phi + (1-\phi)\beta}. \end{aligned}$$

Among the patients who are tested in period 1, the probability that a patient receives a false-positive diagnosis and a false-negative diagnosis, as functions of κ_l , are $\frac{(1-\phi)\beta}{\alpha\phi + (1-\phi)\beta} \kappa_l^2$ and $\frac{\alpha\phi}{\alpha\phi + (1-\phi)\beta} (1 - \kappa_l)^2$, respectively. As expected, a higher κ_l reduces the probability of a false-negative diagnosis but increases the probability of a false-positive diagnosis.

Among the C_1 patients who are tested in period 1, a fraction $\frac{\alpha\phi}{\alpha\phi + (1-\phi)\beta} (2\kappa_l - \kappa_l^2)$ both test positive and are actual carriers of the virus (true positives). Because only those patients who test positive are quarantined, a mass $\phi - \frac{\alpha\phi}{\alpha\phi + (1-\phi)\beta} (2\kappa_l - \kappa_l^2) C_1$ of patients will continue to spread the virus. Therefore, we can write the mass of newly infectious patients at the beginning period 2 as

$$\phi_2 = \left[\phi - \frac{\alpha\phi}{\alpha\phi + (1-\phi)\beta} (2\kappa_l - \kappa_l^2) C_1 \right] R_0.$$

The total mass of infectious patients at the beginning of period 2 is $\phi + \phi_2$, and a fraction α of these patients are symptomatic.⁸ In addition, a fraction β of $(1 - \phi - \phi_2)$ individuals (who are not infectious with the virus) also show symptoms. Therefore, the mass of patients with symptoms in period 2 is

$$s_2 = (\phi + \phi_2) \alpha + (1 - \phi - \phi_2) \beta.$$

All of these s_2 patients seek testing. Additional demand for testing is generated in period 2 as a result of contact-tracing efforts. Close contacts (e.g., family members, friends, and colleagues) of patients who tested positive in period 1 may also seek testing. Therefore, the total demand for testing in period 2

$$D_2 = s_2 + \rho C_1 \cdot \frac{\alpha\phi(2\kappa_l - \kappa_l^2) + (1-\phi)\beta\kappa_l^2}{\alpha\phi + (1-\phi)\beta}, \quad (1)$$

⁸ The underlying assumption is that no patients recover between period 1 and period 2. The rationale behind this simplifying assumption is the long time (up to six weeks or more) needed for patients to fully recover from infection. Patient recovery can be captured by assuming the mass of infectious patients at the beginning of period 2 is $(1 - \lambda)\phi + \phi_2$, where $\lambda \in [0, 1]$ is the recovery rate.

where the fraction is the probability that each of the C_1 patients undergoing testing receives a positive diagnosis in period 1, and ρ is the intensity of contact tracing. We can rearrange equation (1) to separate the effects of contact tracing and the contagious nature of the infection as

$$D_2 = z_1(\kappa_l) + z_2(\kappa_l)R_0, \tag{2}$$

where

$$z_1(\kappa_l) \equiv \phi\alpha + (1 - \phi)\beta + \rho C_1 \cdot \frac{\alpha\phi(2\kappa_l - \kappa_l^2) + (1 - \phi)\beta\kappa_l^2}{\alpha\phi + (1 - \phi)\beta},$$

$$z_2(\kappa_l) \equiv (\alpha - \beta) \left[\phi - \frac{\alpha\phi}{\alpha\phi + (1 - \phi)\beta} (2\kappa_l - \kappa_l^2) C_1 \right].$$

Note $z_1(\kappa_l)$ also includes the initial (and constant) testing demand of $\phi\alpha + (1 - \phi)\beta$ in addition to diagnosis-induced demand.

LEMMA 1. $z_1(\kappa_l)$ increases in κ_l , whereas $z_2(\kappa_l)$ decreases in κ_l .

The proof of Lemma 1, along with those of other technical results, is in the online appendix. Lemma 1 suggests a higher CT cutoff κ_l set by the lab increases the mass of patients who test positive. As a consequence, diagnosis-induced demand for testing increases. Additionally, a higher CT cutoff κ_l facilitates quarantine of a larger share of infectious patients (because true positives also increase). This dampens the spread of the virus and puts a downward force on the demand for testing in period 2.

In formulating the lab’s objective function, we omit its revenue from the first period, which is fixed. In period 2 (which is also the terminal period), the lab sets a CT cutoff of $c_{FN}/(c_{FN} + c_{FP})$ that minimizes the sum of the patient’s expected false-positive and false-negative costs (i.e., $c_{FP} \cdot H(\kappa_l) + c_{FN} \cdot [1 - G(\kappa_l)]$). Note both the CT-cutoff value and the patient’s expected cost of erroneous diagnosis from the second period are independent of the lab’s period-1 decisions. For ease of notation, we define

$$r \equiv r_t - \min_{\kappa_l} \{c_{FP} \cdot H(\kappa_l) + c_{FN} \cdot [1 - G(\kappa_l)]\}.$$

Thus, the lab sets κ_l and C_2 to maximize its objective function:

$$\pi(\kappa_l, C_2) = r \cdot \mathbb{E} \min\{C_2, z_1(\kappa_l) + z_2(\kappa_l) \cdot R_0\} - \gamma \cdot C_2^2 - C_1 \cdot \{c_{FP} \cdot H(\kappa_l) + c_{FN} \cdot [1 - G(\kappa_l)]\}, \tag{3}$$

where the first term represents the lab’s total expected revenue less patients’ expected cost of erroneous diagnosis during period 2, the second term is the cost of building capacity, and the third term captures patients’ expected cost of erroneous diagnosis during period 1.⁹

⁹ We focus on the case in which γ is not prohibitively high such that $C_2 \geq z_1(\kappa_l)$ in equilibrium. Otherwise, the term $\min\{C_2, z_1(\kappa_l) + z_2(\kappa_l) \cdot R_0\}$ in (3) will always be equal to C_2 and the lab’s problem becomes a trivial one.

At this point, understanding how the lab’s incentive to set the period-2 capacity C_2 depends on the period-1 CT cutoff κ_l is useful. The following lemma describes the lab’s optimal capacity choice for a given κ_l .

LEMMA 2. *For a given κ_l , the lab’s optimal capacity choice $C_2^*(\kappa_l)$ is*

$$C_2^*(\kappa_l) = \frac{r[z_1(\kappa_l) + z_2(\kappa_l)\bar{R}_0]}{r + 2z_2(\kappa_l)\bar{R}_0\gamma}, \tag{4}$$

which increases in κ_l if

$$\rho \geq \frac{\phi\alpha(\alpha - \beta)(1 - \kappa_l)\bar{R}_0}{\phi\alpha(1 - \kappa_l) + (1 - \phi)\beta\kappa_l}.$$

If the intensity of contact tracing is sufficiently large, the lab’s period-2 capacity increases in the CT cutoff. A large CT cutoff leads to many patients testing positive (including false positives). In the presence of a significant contact-tracing effort, a large number of positive diagnoses lead to a high diagnosis-driven demand for testing in period 2 and incentivize the lab to invest more in period-2 capacity. The effect of a larger expected R_0 is exactly the opposite. If the virus is highly contagious, a large demand for period-2 testing is organically generated. In this case, a higher CT cutoff, which increases the true-positive rate, imposes a large cost on the lab in the sense that a larger proportion of infectious patients are quarantined and stop contributing toward increasing period-2 demand for testing (i.e., they stop spreading the virus). However, if the virus is not highly contagious (expected R_0 is small), the lab becomes more motivated to set a higher CT cutoff, which creates larger contact-tracing demand but does not significantly hurt the organic demand for testing in period 2. Therefore, the lab becomes more motivated to invest in building period-2 capacity.

If the policymaker requires the lab to report the CT value on the test result, the lab does so and the physician sets the CT cutoff κ_p at $c_{FN}/(c_{FN} + c_{FP})$, which minimizes the sum of patient’s expected false-positive and false-negative costs, to determine if the patient is positive. Because the physician’s objective function and costs are common knowledge, both the lab and the policymaker precisely anticipate this κ_p . The lab sets its period-2 capacity at $C_2^*(\kappa_p)$, as given by (4). If the policymaker does not require the lab to report the CT value (x) from the test, the lab decides whether to disclose the CT value voluntarily or disclose only the diagnosis based on its own CT cutoff. If the lab chooses to voluntarily report the CT value on the test result, similar to the mandatory CT-value reporting case, the physician sets the CT cutoff κ_p at $c_{FN}/(c_{FN} + c_{FP})$ and the lab sets period-2 testing capacity at $C_2^*(\kappa_p)$. The following proposition characterizes the lab’s optimal decisions (κ_l^* and C_2^*) when the lab sets its own CT cutoff and reports only the test outcome (positive or negative).

PROPOSITION 1. *The lab's optimal capacity C_2^* satisfies*

$$C_2^* = \frac{[z_1(\kappa_l^*) + z_2(\kappa_l^*)\bar{R}_0]r}{r + 2z_2(\kappa_l^*)\bar{R}_0\gamma}, \tag{5}$$

where κ_l^* is the optimal CT cutoff and satisfies

$$\frac{r}{2\bar{R}_0 C_1} \int_0^{\frac{\bar{R}_0[r \cdot z_2(\kappa_l^*) - 2\gamma \cdot z_1(\kappa_l^*)]}{z_2(\kappa_l^*)(r + 2\gamma \bar{R}_0)}} [z_1'(\kappa_l^*) + z_2'(\kappa_l^*)R_0] dR_0 = c_{FP}\kappa_l^* - c_{FN}(1 - \kappa_l^*). \tag{6}$$

The above solution to the lab's problem (κ_l^* and C_2^*) helps us generate important insights about the lab's tradeoff in choosing its CT cutoff and testing capacity. The following proposition presents a comparison of the CT cutoff κ_l^* (which is optimal from the lab's perspective) and κ_p (which is ideal from the patient's perspective). It also compares the lab's period-2 testing-capacity decisions corresponding to the CT-cutoff values of κ_l^* and κ_p . Finally, it describes the lab's decision of whether to report the CT value in the absence of a mandatory reporting requirement.

PROPOSITION 2. *If $\rho \geq (\alpha - \beta)\bar{R}_0$, in period 1 the lab*

- (i) *sets a CT cutoff that is higher than what is optimal for the patient, that is, $\kappa_l^* > \kappa_p$;*
- (ii) *sets a period-2 capacity that is higher than the capacity under a CT cutoff of κ_p , that is, $C_2^* > C_2^*(\kappa_p)$; and*
- (iii) *prefers to set the CT cutoff itself and diagnose the patient rather than report the CT value on the test result.*

Before providing intuition behind Proposition 2, we remark about the condition $\rho \geq (\alpha - \beta)\bar{R}_0$. Recall from footnote 7 that ρ captures the demand for additional testing resulting from a positive diagnosis; the demand can come from official contact-tracing effort, but more importantly, voluntary testing based on an individual's assessment of his or her contact history even in the absence of formal contract tracing. Because the average U.S. household size is 2.5 (US Census 2019), and assuming an infectious individual comes in contact with at least one person outside his or her home, we can reasonably suggest ρ is at least $2.5 \times 2 - 1 = 4$. On the other hand, CDC's current best estimate of R_0 for COVID-19 is 2.5 and its estimated upper bound of R_0 is 4.0 (CDC 2020); Hilton and Keeling (2020) find most countries have an R_0 for COVID-19 falling between 2 and 4. According to the CDC's current best estimate, the percent of infections that are symptomatic (α) is 60% (CDC 2020). Although the value of β is hard to estimate, we conclude the condition $\rho > \alpha\bar{R}_0 > (\alpha - \beta)\bar{R}_0$ is satisfied under reasonable parametric assumptions.

In the terminal period 2, the CT cutoff is set at κ_p regardless of whether the lab or the physician sets it. The reason is that in period 2, the lab is not concerned about generating future demand for testing and, similar to the physician, acts in the best interest of the patient. However, in

period 1, the lab sets a higher CT cutoff than what physicians would consider appropriate (i.e., $\kappa_l^* > \kappa_p$). A high CT cutoff helps the lab boost diagnosis-induced demand but suppress demand from community spread. If the lab expects a sufficiently large number of close contacts of patients who tested positive to seek testing, the lab focuses on future demand creation by increasing CT cutoff, and thus engaging in overdiagnosis. This result and its intuition provides a rationale for the observed high CT cutoff that is currently set by labs and challenged by healthcare experts (Mandavilli 2020; Rao 2020; Service 2020). It also suggests that in the future, we can expect labs to either report the CT value on the test result or lower the CT-cutoff value when determining whether a patient tests positive.

The lab sets a higher period-2 testing capacity when setting its own period-1 CT cutoff at κ_l^* than the capacity chosen corresponding to the CT cutoff of κ_p . This result builds on part (i) of the proposition: $\kappa_l^* > \kappa_p$. The lab sets a higher CT cutoff to boost its future demand for testing. Yet, for demand to translate into revenue, it must invest in building testing capacity. Stated differently, an inflated CT cutoff is not entirely a negative force, because it incentivizes the lab to expand its testing capacity.

If the policymaker does not require mandatory CT-value reporting, the lab decides whether to voluntarily report it along with the binary diagnosis. We find that in this case, the lab is indifferent between reporting and not reporting the CT value as part of the test result in the terminal period 2. However, in period 1, motivated by its desire to enhance period-2 demand, the lab finds *not* reporting the CT value is optimal, because it allows the lab to inflate the CT cutoff (to a level above κ_p) when determining whether period-1 patients are COVID-19 positive. Given the lab's decision to inflate the period-1 CT cutoff, healthcare experts' outcry against the prevailing practice of not reporting the CT value on test results raises an important question: Why doesn't the policymaker intervene and require the lab to disclose CT values? The following proposition examines implications of a mandatory CT-value reporting requirement and offers a possible rationale behind the policymaker's lack of intervention despite concerns raised by healthcare experts.

PROPOSITION 3. *If $\rho \geq (\alpha - \beta)\bar{R}_0$, a mandatory CT-value reporting requirement imposed on the lab results in*

- (i) *a higher expected utility of the tested patient;*
- (ii) *a more widespread infection in the subsequent period; and*
- (iii) *a lower period-2 testing capacity yet a higher service level (which is the likelihood that no testing shortage would occur) than in the absence of a mandatory-reporting requirement.*

If the reporting of CT value — a proxy for viral load — is not mandatory, the lab does not report it in period 1 and inflates the CT cutoff above what is best from an individual patient's

perspective. A requirement to report the CT value as part of a test result in essence transfers the diagnostic-decision right from the lab to the physician. The physician, who acts in the best interest of the patient, sets the CT cutoff at κ_p . Therefore, as a result of imposing the mandatory CT-value reporting requirement, the expected utility of the tested patient increases.

One drawback of a mandatory viral-load reporting requirement is that it leads to more false-negative diagnoses, so it can potentially contribute to more widespread COVID-19 infection in the subsequent period. Although a higher CT cutoff imposes a cost on the tested patients, it facilitates quarantining of a larger proportion of infectious patients, and therefore reduces the spread of the virus. To the extent society is willing to bear the cost associated with quarantining a large number of individuals who may not be contagious, in order to reduce the spread of the virus, letting the lab set a higher CT cutoff may be better than requiring viral-load reporting.

Finally, if the lab is not mandated to report the viral load in period 1, it sets the CT cutoff at κ_l^* and the period-2 capacity at C_2^* , which is higher than $C_2^*(\kappa_p)$. (See the related Proposition 2 (ii) and its intuition.) Although $C_2^* > C_2^*(\kappa_p)$, it is insufficient to absorb the increase in demand resulting from the lab's inflated CT cutoff κ_l^* . As the lab increases its CT cutoff, it finds maintaining the same service level (by sufficiently increasing its capacity) increasingly costly to deliver. An insufficient increase in the period-2 capacity essentially increases the likelihood of continued testing shortage in period 2. Interestingly, overdiagnosis not only coexists with COVID-19 undertesting but also contributes to its increased likelihood.

5. Concluding Remarks

The main contribution of this paper is to shed light on the interaction between diagnostic decision making and capacity building when diagnoses have implications for future market demand for testing. When the lab provides diagnoses unaccompanied by fine-grained viral-load information, positive cases with nearly non-existent viral loads (i.e., “overdiagnosis”) may trigger contact-tracing efforts, and hence boost market demand for testing. At the same time, such positive diagnoses lead to quarantining of potentially infectious individuals on a wider scale, and thus helps decelerate disease transmission and dampen future organic demand for testing. By analyzing this tradeoff, we find scenarios exist in which the lab has an incentive to engage in overdiagnosis to generate higher diagnosis-driven future demand. Overdiagnosis leads to a higher testing capacity, which is nevertheless insufficient to absorb the inflated demand, so undertesting arises along with overdiagnosis.

Beyond this insight, our model has implications for a policymaker's decision regarding information-disclosure requirements. When the lab is mandated to disclose the viral-load information, it essentially forfeits its decision rights — in terms of the interpretation of the viral load, which may trigger subsequent interventions that shape future demand for testing — to physicians

and other stakeholders. We show that although the concept of information disclosure is broadly appealing, when the policymaker intends to induce the lab to build a high testing capacity and reduce viral transmission, *not* mandating such disclosure may be in the best interest of the policymaker. Although the lab and the policymaker have divergent objectives, a CT cutoff that is higher than what is ideal for an individual patient is desirable to both.

References

- Arora A, Fosfuri A (2005) Pricing diagnostic information. *Management Sci.* 51(7):1092–1100.
- Becker, G. S. (2020). Some economics of flu pandemics. *Chicago Booth Rev.* (Summer) pp. 54–55. <http://bit.ly/BeckerPandemic>. Originally published in 2009 on *The Becker-Posner Blog* under the title of “The economics of the flu epidemic.” <http://bit.ly/Becker09blog>. Both accessed October 28, 2020.
- Bhardwaj P (2001) Delegating pricing decisions. *Marketing Sci.* 20(2):143–169.
- Bhardwaj P, Chen Y, Godes D (2008) Buyer-initiated vs. seller-initiated information revelation. *Management Sci.* 54(6):1104–1114.
- Brodersen J, Schwartz LM, Heneghan C, O’Sullivan JW, Aronson JK, Woloshin S (2018) Overdiagnosis: What it is and what it isn’t. *BMJ EBM* 23(1):1–3.
- Buitrago-Garcia D, Egli-Gany D, Counotte MJ, Hossmann S, Imeri H, Ipekci AM, Salanti G, Low N (2020) Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis. *PLoS Med.* 17(9):e1003346.
- CDC (2020) COVID-19 pandemic planning scenarios. Centers for Disease Control and Prevention (September 10). <http://bit.ly/cdcestimates>. Accessed October 28, 2020.
- Cutler DM, Summers LH (2020) The COVID-19 pandemic and the \$16 trillion virus. *JAMA* 324(15):1495–1496.
- Dai T, Singh S (2020) Conspicuous by its absence: Diagnostic expert testing under uncertainty. *Marketing Sci.* 39(3):540–563.
- Darby MR, Karni E (1973) Free competition and the optimal amount of fraud. *J. Law Econom.* 16(1):67–88.
- Dogan M, Jacquillat A, Yildirim P (2019) Strategic automation and decision-making authority. *MIT Sloan School of Management Working Paper*.
- Dranove D (1988) Demand inducement and the physician/patient relationship. *Econom. Inquiry* 26(2):281–298.
- Durbin E, Iyer G (2009) Corruptible advice. *Amer. Econom. J. Microeconom.* 1(2):220–242.
- Dye RA (1985) Disclosure of nonproprietary information. *J. Accounting Res.* 23(1):123–145.
- Esserman LJ, Thompson IM, Reid B (2013) Overdiagnosis and overtreatment in cancer. *JAMA* 310(8):797–798.

- Gal-Or E, Gal-Or M, Dukes A (2007) Optimal information revelation in procurement schemes. *RAND J. Econom.* 38(2):400–418.
- Grossman SJ (1981) The informational role of warranties and private disclosure about product quality. *J. Law Econom.* 24(3):461–483.
- Grynbaum MM (2020) NBC faces blowback for holding Trump’s town hall opposite Biden’s. *New York Times* (October 14). <https://nyti.ms/31JhHKM>. Accessed October 21, 2020.
- Guo L (2020) Upstream exploitation and strategic disclosure. *Marketing Sci.* 39(5):923–938.
- Guo L, Zhao Y (2009) Voluntary quality disclosure and market interaction. *Marketing Sci.* 28(3):488–501.
- Hilton J, Keeling MJ (2020) Estimation of country-level basic reproductive ratios for novel Coronavirus (SARS-CoV-2/COVID-19) using synthetic contact matrices. *PLoS Comput. Biol.* 16(7):e1008031.
- Jiang B, Ni J, Srinivasan K (2014) Signaling through pricing by service providers with social preferences. *Marketing Sci.* 33(5):641–654.
- Johns Hopkins University (2020) COVID-19 dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). <https://coronavirus.jhu.edu/map.html>. Accessed October 28, 2020.
- Jovanovic B (1982) Truthful disclosure of information. *Bell J. Econom.* 13(1):36–44.
- Lal R (1986) Delegating pricing responsibility to the salesforce. *Marketing Sci.* 5(2):159–168.
- Madrigal AC, Meyer R (2020) Why Trump’s rapid-testing plan worries scientists: Experts were already divided on the right way to deploy new coronavirus tests. Then the White House barged ahead. *Atlantic* (October 9). <http://bit.ly/rapidcovidtesting>. Accessed October 28, 2020.
- Mandavilli A (2020) Your Coronavirus test is positive. Maybe it shouldn’t be. *New York Times* (August 29). <https://nyti.ms/31Vyv1B>. Accessed October 28, 2020.
- Mayzlin D, Shin J (2011) Uninformative advertising as an invitation to search. *Marketing Sci.* 30(4):666–685.
- Milgrom PR (1981) Good news and bad news: Representation theorems and applications. *Bell J. Econom.* 12(2):380–391.
- Özer Ö, Subramanian U, Wang Y (2018) Information sharing, advice provision, or delegation: What leads to higher trust and trustworthiness? *Management Sci.* 64(1):474–493.
- Rao S (2020) Covid-19 test reports must also state cycle threshold value: Doctors. *Times of India* (September 6). <http://bit.ly/toictvalue>. Accessed October 28, 2020.
- Salis AD (2009) Applications in clinical microbiology. Logan J, Logan JMJ, Edwards KJ, Saunders NA, eds. *Real-Time PCR: Current Technology and Applications* (Caister Academic Press, Norfolk, United Kingdom), 177–220.
- Service RF (2020) A call for diagnostic tests to report viral load. *Sci.* 370(6512) 22.

- Shear MD, Goodnough A, Kaplan S, Fink S, Thomas K, Weiland N (2020) The lost month: How a failure to test blinded the U.S. to Covid-19. *New York Times* (March 28). <https://nyti.ms/3joYBj5>. Accessed October 28, 2020.
- Sun M, Tyagi RK (2020) Product fit uncertainty and information provision in a distribution channel. *Production Oper. Management* 29(10):2381–2402.
- TWiV (2020) TWiV 641: COVID-19 with Dr. Anthony Fauci. *This Week in Virology* (July 17). <https://www.microbe.tv/twiv/twiv-641/>. Accessed October 28, 2020.
- US Census (2019) Historical households tables. United States Census Bureau (October 10). <http://bit.ly/censushhsize>. Accessed October 28, 2020.
- Verrecchia RE (1983) Discretionary disclosure. *J. Accounting Econom.* 5(April):179–194.
- Welch HG, Schwartz L, Woloshin S (2011) *Overdiagnosed: Making People Sick in the Pursuit of Health* (Beacon Press, Boston).
- Wolinsky A (1993) Competition in a market for informed experts' services. *RAND J. Econom.* 24(3):380–398.
- Wu KJ (2020) In coronavirus testing, false positives are more hazardous than they might seem. *New York Times* (October 25). <https://nyti.ms/34zqRf1>. Accessed October 28, 2020.
- Zhang J (2014) Policy and inference: The case of product labeling. *MIT Sloan School of Management Working Paper*.

Online Appendix to “COVID-19 Diagnosis and Viral Load Reporting: A Theory of Overdiagnosis and Undertesting”

Proof of Lemma 1. First, the first-order derivative of $z_1(\kappa_l)$ is

$$z'_1(\kappa_l) = 2\rho \cdot C_1 \cdot \frac{\alpha\phi(1 - \kappa_l) + (1 - \phi)\beta\kappa_l}{\alpha\phi + (1 - \phi)\beta} > 0.$$

Thus, $z_1(\kappa_l)$ increases in κ_l . Second, the first-order derivative of $z_2(\kappa_l)$ is

$$z'_2(\kappa_l) = -2(\alpha - \beta) \cdot C_1 \cdot (1 - \kappa_l) \cdot \frac{\alpha\phi}{\alpha\phi + (1 - \phi)\beta} < 0.$$

Thus, $z_2(\kappa_l)$ decreases in κ_l .

Q.E.D.

Proof of Lemma 2. Given κ_l , the first-order derivative of the lab’s objective function with respect to C_2 can be reorganized, by Leibniz’s rule, as

$$\begin{aligned} \frac{\partial \pi(\kappa_l, C_2)}{\partial C_2} &= r \cdot \frac{\partial \mathbb{E} \min\{C_2, z_1(\kappa_l) + z_2(\kappa_l) \cdot R_0\}}{\partial C_2} - 2\gamma C_2 \\ &= r \cdot \frac{\partial \left\{ \int_0^{\frac{C_2 - z_1(\kappa_l)}{z_2(\kappa_l)}} [z_1(\kappa_l) + z_2(\kappa_l) \cdot R_0] dF(R_0) + \int_{\frac{C_2 - z_1(\kappa_l)}{z_2(\kappa_l)}}^{\bar{R}_0} C_2 dF(R_0) \right\}}{\partial C_2} - 2\gamma C_2 \\ &= r \cdot \left[\frac{C_2}{z_2(\kappa_l)} - \frac{C_2}{z_2(\kappa_l)} + 1 - F\left(\frac{C_2 - z_1(\kappa_l)}{z_2(\kappa_l)}\right) \right] - 2\gamma C_2 \\ &= r \cdot \left[1 - F\left(\frac{C_2 - z_1(\kappa_l)}{z_2(\kappa_l)}\right) \right] - 2\gamma C_2. \end{aligned}$$

Using the first-order condition gives

$$F\left(\frac{C_2 - z_1(\kappa_l)}{z_2(\kappa_l)}\right) = 1 - \frac{2\gamma C_2}{r} \quad \text{at } C_2 = C_2^*(\kappa_l). \tag{OA1}$$

In addition, we can verify that the second-order condition is satisfied:

$$\frac{\partial^2 \pi(\kappa_l, C_2)}{\partial C_2^2} = -\frac{r}{z_2(\kappa_l)} f\left(\frac{C_2 - z_1(\kappa_l)}{z_2(\kappa_l)}\right) - 2\gamma < 0,$$

where $f(\cdot)$ is the probability density function (pdf) of R_0 . Thus, the lab’s objective function is concave in C_2 . Then, using the assumption that R_0 follows a uniform distribution, the first-order condition simplifies to (4).

We have from (4) that $C_2^*(\kappa_l)$ increases in κ_l if $z_1(\kappa_l) + \bar{R}_0 \cdot z_2(\kappa_l)$ increases in κ_l . The first-order derivative of $z_1(\kappa_l) + \bar{R}_0 \cdot z_2(\kappa_l)$ is

$$z'_1(\kappa_l) + \bar{R}_0 \cdot z'_2(\kappa_l) = \frac{2C_1}{\beta + (\alpha - \beta)\phi} \cdot \{ \phi\alpha[(\alpha - \beta)\bar{R}_0 - \rho] + (1 - \phi)\beta\rho \} \kappa_l - \phi\alpha[(\alpha - \beta)\bar{R}_0 - \rho].$$

The derivative $z'_1(\kappa_l) + \bar{R}_0 \cdot z'_2(\kappa_l) \geq 0$ if and only if

$$\{\phi\alpha[(\alpha - \beta)\bar{R}_0 - \rho] + (1 - \phi)\beta\rho\}\kappa_l - \phi\alpha[(\alpha - \beta)\bar{R}_0 - \rho] \geq 0,$$

which gives

$$\rho \geq \frac{\phi\alpha(\alpha - \beta)\bar{R}_0(1 - \kappa_l)}{\phi\alpha(1 - \kappa_l) + (1 - \phi)\beta\kappa_l}$$

and completes the proof. Q.E.D.

Proof of Proposition 1. Equation (5) follows from Lemma 2. To prove (6), note the first-order partial derivative of the lab's objective function (3) in terms of κ_l is

$$r \cdot \frac{\partial \mathbb{E} \min\{C_2, z_1(\kappa_l) + z_2(\kappa_l) \cdot R_0\}}{\partial \kappa_l} - C_1 \cdot \{c_{FP} \cdot (2\kappa_l) + c_{FN} \cdot [-2(1 - \kappa_l)]\}.$$

Using Leibniz integral rule, we have

$$\begin{aligned} \frac{\partial \mathbb{E} \min\{C_2, z_1(\kappa_l) + z_2(\kappa_l) \cdot R_0\}}{\partial \kappa_l} &= \frac{\partial \left\{ \int_0^{\frac{C_2 - z_1(\kappa_l)}{z_2(\kappa_l)}} [z_1(\kappa_l) + z_2(\kappa_l) \cdot R_0] dF(R_0) + \int_{\frac{C_2 - z_1(\kappa_l)}{z_2(\kappa_l)}}^{\bar{R}_0} C_2 dF(R_0) \right\}}{\partial \kappa_l} \\ &= C_2 \cdot \frac{\partial \frac{C_2 - z_1(\kappa_l)}{z_2(\kappa_l)}}{\partial \kappa_l} - C_2 \cdot \frac{\partial \frac{C_2 - z_1(\kappa_l)}{z_2(\kappa_l)}}{\partial \kappa_l} + \int_0^{\frac{C_2 - z_1(\kappa_l)}{z_2(\kappa_l)}} [z'_1(\kappa_l) + z'_2(\kappa_l)R_0] dF(R_0) \\ &= \int_0^{\frac{C_2 - z_1(\kappa_l)}{z_2(\kappa_l)}} [z'_1(\kappa_l) + z'_2(\kappa_l)R_0] dF(R_0). \end{aligned}$$

The first-order condition, along with (5), gives (6). Q.E.D.

Proof of Proposition 2. To prove Proposition 2, first, we establish the following claim:

CLAIM OA.1. *If $\rho \geq (\alpha - \beta)\bar{R}_0$, $z'_1(\kappa_l) + z'_2(\kappa_l)\bar{R}_0 \geq 0$ and $C_2^*(\kappa_l)$ increases in κ_l for $0 < \kappa_l < 1$.*

From the proof of Lemma 2, we have that $z'_1(\kappa_l) + z'_2(\kappa_l)\bar{R}_0 \geq 0$ if and only if

$$\rho \geq \frac{\phi\alpha(\alpha - \beta)\bar{R}_0(1 - \kappa_l)}{\phi\alpha(1 - \kappa_l) + (1 - \phi)\beta\kappa_l}.$$

Also,

$$\frac{\phi\alpha(\alpha - \beta)\bar{R}_0(1 - \kappa_l)}{\phi\alpha(1 - \kappa_l) + (1 - \phi)\beta\kappa_l} \leq (\alpha - \beta)\bar{R}_0 \quad \text{for all } 0 \leq \kappa_l \leq 1.$$

Therefore, $z'_1(\kappa_l) + z'_2(\kappa_l)\bar{R}_0 \geq 0$, if $\rho \geq (\alpha - \beta)\bar{R}_0$.

Next, note from Lemma 2 that

$$C_2^*(\kappa_l) = \frac{r[z_1(\kappa_l) + z_2(\kappa_l)\bar{R}_0]}{r + 2z_2(\kappa_l)\bar{R}_0\gamma}.$$

Under $\rho \geq (\alpha - \beta)\bar{R}_0$, $z'_1(\kappa_l) + z'_2(\kappa_l)\bar{R}_0 > 0$, so the numerator of the right-hand side of the above equation increases in κ_l . In addition, we know from Lemma 1 that its denominator always decreases in κ_l . Thus, if $\rho \geq (\alpha - \beta)\bar{R}_0$, $C_2^*(\kappa_l)$ increases in κ_l . Having established Claim OA.1, we now proceed to prove Proposition 2.

(i) From Lemma 1, $z'_2(\kappa_l) < 0$. Thus, for any realization of $R_0 \in [0, \bar{R}_0)$, we have from Claim OA.1 that,

$$z'_1(\kappa_l) + z'_2(\kappa_l)R_0 > z'_1(\kappa_l) + z'_2(\kappa_l)\bar{R}_0 > 0$$

if $\rho \geq (\alpha - \beta)\bar{R}_0$. Hence, the left-hand side of equation (6) must be strictly positive. Then, equation (6) gives

$$c_{FP} \cdot \kappa_l^* - c_{FN} \cdot (1 - \kappa_l^*) > 0. \tag{OA2}$$

In contrast, κ_p satisfies

$$c_{FP} \cdot \kappa_p - c_{FN} \cdot (1 - \kappa_p) = 0. \tag{OA3}$$

Because $c_{FP} \cdot H(\kappa_l) + c_{FN} \cdot [1 - G(\kappa_l)]$ increases in κ_l , comparing equations (OA2) and (OA3) gives $\kappa_l^* > \kappa_p$.

(ii) From Lemma 2 and Claim OA.1, $\frac{\partial C_2^*(\kappa_l)}{\partial \kappa_l} > 0$ for $\rho \geq (\alpha - \beta)\bar{R}_0$. In addition, from part (i) above, $\kappa_l^* > \kappa_p$. Therefore, $C_2^* = C_2^*(\kappa_l^*) > C_2^*(\kappa_p)$.

(iii) If the lab sets its own CT cutoff and testing capacity, the value of its objective function is $\pi(\kappa_l^*, C_2^*(\kappa_l^*))$. However, if the lab reports the viral load (CT value) on the test result, the physician sets the CT cutoff at κ_p . In this case, the lab sets capacity at $C_2^*(\kappa_p)$. Because the lab's objective function $\pi(\kappa_l, C_2)$ given in equation (3) is maximized at κ_l^* and $C_2^*(\kappa_l^*)$, it must be that $\pi(\kappa_l^*, C_2^*(\kappa_l^*)) > \pi(\kappa_p, C_2^*(\kappa_p))$. Therefore, the lab prefers to set κ_l itself and diagnose the patient than to report the CT value on the test result. *Q.E.D.*

Proof of Proposition 3. In the mandatory viral-load reporting scenario, the lab reports the CT value as part of the test result and the physician sets the CT cutoff at κ_p (in both periods 1 and 2). In this case, the lab sets period-2 capacity at $C_2^*(\kappa_p)$. However, in the absence of such mandatory-reporting requirements, the lab sets the period-1 CT cutoff at κ_l^* and period-2 capacity at $C_2^*(\kappa_l^*)$. Because period 2 is the terminal period, the period-2 CT cutoff is set at κ_p regardless of whether the lab or the physician sets it. Therefore, the comparison of mandatory and voluntary viral-load reporting is equivalent to the comparison of (1) period-1 CT cutoff κ_p and period-2 capacity $C_2^*(\kappa_p)$, and (2) period-1 CT cutoff κ_l^* and period-2 capacity $C_2^*(\kappa_l^*)$, respectively.

(i) Because the patient's expected cost from erroneous diagnosis $C_1 \cdot [c_{FP} \cdot H(\kappa) + c_{FN} \cdot [1 - G(\kappa)]]$ is minimized (equivalently, patient utility is maximized) at $\kappa = \kappa_p$, it is straightforward that a mandatory CT-value reporting requirement results in a higher expected utility of the tested patient than in the absence of a mandatory-reporting requirement.

(ii) The mass of newly infectious patients at the beginning of period 2 depends on the period-1 CT cutoff (κ) and is given by $\phi_2(\kappa) = \left[\phi - \frac{\alpha\phi}{\alpha\phi + (1-\phi)\beta} (2\kappa - \kappa^2) C_1 \right] R_0$. Note $\frac{\partial \phi_2}{\partial \kappa} < 0$ (i.e., a higher CT cutoff results in fewer infections). Because $\kappa_l^* > \kappa_p$, $\phi_2(\kappa_l^*) < \phi_2(\kappa_p)$. That is, a mandatory viral-load reporting requirement results in a more widespread infection in the subsequent period.

(iii) Finally, note that under a CT cutoff of κ_l , the lab's optimal service level is

$$\Pr(z_1(\kappa_l) + z_2(\kappa_l)R_0 \leq C_2) = F\left(\frac{C_2 - z_1(\kappa_l)}{z_2(\kappa_l)}\right),$$

which, according to the proof of Lemma 2, is equal to

$$1 - \frac{2\gamma C_2^*(\kappa_l)}{r}.$$

Because $\rho \geq \bar{R}_0(\alpha - \beta)$, we have from Claim OA.1 that $C_2^*(\kappa_l)$ increases in κ_l . In addition, $\kappa_l^* > \kappa_p$. Therefore, $C_2^*(\kappa_l^*) > C_2^*(\kappa_p)$, and the service level is lower corresponding to a CT cutoff of κ_l^* than κ_p . Q.E.D.

The macroeconomics of age-varying epidemics¹

Marta Giagheddu² and Andrea Papetti³

Date submitted: 12 November 2020; Date accepted: 13 November 2020

We incorporate age-specific socio-economic interactions in a SIR macroeconomic model to study the role of demographic factors for the COVID-19 epidemic evolution, its macroeconomic effects and possible containment measures. We capture the endogenous response of rational individuals who freely reduce consumption- and labor-related personal exposure to the virus, with interactions that can vary within and across ages, while fail to internalize the impact of their actions on others. The endogenous response amplifies the economic losses, but it implies that the individual behavioral response to the risk of infection is an important ally of the needed policy measures to contain the spread of the virus. Investigating the effect of different combinations of economic shutdown and age-targeted social distancing, we find that there are considerable economic benefits from measures targeting the elderly with higher mortality risk which are not part of the labor force. For any level of social distancing, the implied optimal economic shutdown generates small gains in terms of lives and large output losses over one-year time. These results are confirmed by calibrating the model to match real epidemic and economic data in the context of a scenarios exercise.

- 1 We are grateful to Gaetano Basso, Alessandro Borin and Karl Harmenberg for extensive comments and suggestions, to Max Elger for early advice. For helpful discussions, we thank the seminar participants at the Bank of Italy and the European Central Bank. The views expressed herein are those of the authors and do not necessarily reflect the views of the Bank of Italy or the Eurosystem. All remaining errors are our own.
- 2 Assistant Professor, Department of International Economics, The Johns Hopkins University School of Advanced International Studies.
- 3 Research Fellow, DG for Economics, Statistics and Research, Bank of Italy.

Copyright: Marta Giagheddu and Andrea Papetti

1 Introduction

Older people are more vulnerable to COVID-19 as shown in an increasing number of studies.¹ In addition, countries with many intergenerational contacts may see faster transmissions to elderly (Dowd et al., 2020) as indicated by the positive correlation across countries between the case fatality rate and the share of 30 to 49 year-old people living with their parents (Kuhn and Bayer, 2020). Containing the overall death toll of the disease could come from controlling the number of potential infectious contacts of those with a higher probability of dying.

Merging the epidemiological SIR models *à la* Kermack and McKendrick (1927) with macroeconomic models, Eichenbaum et al. (2020b) (ERT, henceforth) developed a “SIR macro” model. This model reveals that susceptible rational agents severely reduce their consumption and hours worked to lower their own probability of getting infected.² Meanwhile, they do not internalize their impact on the overall spread of the infection. To contain this externality, public policy measures are necessary.

As noted in Ellison (2020), allowing for variation in the contact rates by considering how sub-populations differ in their activity levels is a primary objective. We contribute via combining intergenerational and economic interactions developing a “SIR-age macro” model. Our reference model are ERT and Towers and Feng (2012) which we extend via allowing for age-specific economic interactions. We evaluate how this affects macroeconomic outcomes, and how these in turn depend on the design of age-specific containment policies.

We consider two age-groups: aged 70 or more (elderly) and the rest of the population (young). Susceptible agents can get infected. If infected, the individual either recovers (and cannot get infected again) or dies. The elderly face a higher mortality risk,³ they do not work and consume their fixed

¹See Dudel et al. (2020); Goldstein and Lee (2020) for evidence.

²See Farboodi et al. (2020) for the United States, Andersen et al. (2020) for Denmark and Sweden.

³Data on the South Korean fatality rates (supposedly the most reliable, given the world’s highest per capita test rates for COVID-19) and estimates of the infection fatality rates for Italy (Garibaldi et al., 2020), that we both use in our analysis, give that those aged 70 or more face a mortality probability upon infection of 22 to 28 times higher than the rest of the population. However, uncertainty remains on the actual level of the mortality risk (Goldstein and Lee, 2020).

pension. All individuals are hand-to-mouth. Interactions occur when consuming, working or for other residual activities. Each type of interaction has a different infection probability which, among other parameters, depends on the exogenous daily number of contacts (between and within age-groups).

The pandemic can be contained either, by means of a consumption tax controlled by the government (as in ERT), referred to as “economic shutdown”, or alternatively, by “social distancing”, reducing the number of interactions.

While the latter does not directly impair economic activity, the former makes consumption and hence production more expensive. In our model, social distancing can be attributed not only to mandated policies, but also to changes in behavior by individuals other than consumption and labor choices (e.g. the use of personal protective equipment).⁴

The no containment scenario is characterized by the highest death toll (0.82% of the population) and an output loss the first year of 3.86% as compared to the pre-epidemic steady state. Using a generalized social distancing lowers the corresponding output loss (1.47%) and death toll (0.21% of the population). Deviating from generalized social distancing via milder restriction on young interactions, results in higher deaths (0.47%) and higher output loss (2.85%). Via instead letting the elderly interact less with the young implies a 1.9% output loss (worse than generalized social distancing, but better than mild social distancing for young) and 0.11% deaths (better than generalized and mild on young distancing).

Without social distancing, an optimal economic shutdown is characterized by 0.43% deaths and 30.1% yearly output losses. Given different social distancing measures, the optimal economic shutdown generates 19.8%, 24.9% and 9.6% yearly output losses in the case of generalized social distancing, milder on young and stricter on old respectively. For any level of social distancing, the implied optimal economic shutdown generates small gains in terms of lives and large increases in terms of yearly output losses. To minimize output and death losses one should search for the policy mix that

⁴Cf. Baqaee et al. (2020) who also consider reduction in contacts both reflecting behavioral and mandated changes.

flattens the infection curve of the elderly the most, rather than the overall infection curve.

We evaluate the model's capacity of replicating actual data using Italy as one of the countries firstly affected by the epidemic at the global level. We calibrate our model, including the containment policies, to target the number of excessive deaths (as obtained in Galeotti et al. (2020)) and to capture the 5.3% quarterly output loss in the first quarter of 2020.⁵ The simulations suggest that about three weeks were necessary to flatten the curve of total deaths due to COVID-19 and that absent any government intervention more than 0.4% of the population would have died in two months. Given this epidemic evolution in line with the data, we evaluate the epidemic and its economic impact in a series of hypothetical post-lockdown scenarios corresponding to different types of social distancing for any given economic shutdown.

We recognize the uncertainty surrounding many of the model parameters. Our quantitative results are better interpreted in terms of a relative order of magnitude within the internal modeling consistency. Our model points to the preference of differentiating containment measures by age, acting more on social distancing of the elderly from the young than on homogeneous economic measures. Finally, while the recommendations stemming from the model points to the benefits of reducing the number of contacts between the elderly and the young, there are clearly human costs associated with long-term isolation that we are not considering.

This work contributes to the fast increasing literature on the economic consequences of the COVID-19 epidemic. The contributions in economic modeling has been ranging from *purely epidemiological models* (Acemoglu et al., 2020; Alvarez et al., 2020; Atkeson, 2020a,b; Berger et al., 2020; Chikina and Pegden, 2020; Favero et al., 2020; Fernandez-Villaverde and Jones, 2020; Rampini, 2020; Stock, 2020);⁶ through *epidemiological models with choices of rational economic agents* other than the social planner (Bodenstein et al., 2020; Brotherhood et al., 2020; Eichenbaum et al., 2020a,b;

⁵ISTAT (2020)

⁶Criticisms to the use of SIR models for policy evaluation are offered by Chang and Velasco (2020) and Von Thadden (2020).

Farboodi et al., 2020; Garibaldi et al., 2020; Glover et al., 2020; Jones et al., 2020; Kapička and Rupert, 2020; Kaplan et al., 2020; Krueger et al., 2020); to *purely economic models* (Faria-e-Castro, 2020; Gregory et al., 2020; Guerrieri et al., 2020; McKibbin and Fernando, 2020).

Other work considers age heterogeneity with respect to the COVID-19 epidemics. We differ from Glover et al. (2020) as, while not considering redistributive aspects, we allow for the likelihood of infection to increase with consumption focusing on the role of age-specific containment policies for aggregate health-output trade-offs. Compared to Brotherhood et al. (2020), we explicitly consider the empirical intergenerational contacts that prevail in “normal times”, tailoring our model to the Italian case for policy scenarios. Besides the revision of individual choices on consumption and labor to reduce the exposure for a given number of daily normal contacts, it might well be that individuals freely revise their number of contacts to a “new normal” for a while. This could be captured in our model by an exogenous reduction in the number of daily contacts and would therefore reduce the need of regulation.

We also extend on Acemoglu et al. (2020), Chikina and Pegden (2020), Favero et al. (2020), Gollier (2020a,b) and Rampini (2020) via specifically modeling age-specific interactions as economic ones capturing the endogenous consumption/labor response to the epidemic progression.

The rest of the paper is organized as follows. Section 2 describes the model. Section 3 details the main calibration of the model and the *laissez-faire* equilibrium in comparison also to ERT. Section 4 studies social distancing and the optimal economic shutdown, first separately and then as a policy mix. Section 5 applies the model to the Italian case to study different scenarios. Section 6 concludes.

2 The age-varying SIR macro model

To a purely epidemiological age-varying SIR model, building on [Eichenbaum et al. \(2020b\)](#),⁷ we add macroeconomic interactions affecting the number of infected people.

2.1 The age-varying SIR model with macroeconomic interactions

The population is grouped in two categories: young (y) and old (o). For age $a \in \{y, o\}$ and discrete time t the following set of equations holds:

$$S_{a,t+1} = S_{a,t} - T_{a,t} \tag{2.1}$$

$$I_{a,t+1} = I_{a,t} + T_{a,t} - (\pi_{a,r} + \pi_{a,d})I_{a,t} \tag{2.2}$$

$$R_{a,t+1} = R_{a,t} + \pi_{a,r}I_{a,t} \tag{2.3}$$

$$D_{a,t+1} = D_{a,t} + \pi_{a,d}I_{a,t} \tag{2.4}$$

$$N_{a,t} = S_{a,t} + I_{a,t} + R_{a,t} \tag{2.5}$$

$$N_t = \sum_a N_{a,t} \tag{2.6}$$

The number of newly infected people in each period in each category is given by:

$$T_{y,t} = \eta S_{y,t} \left[\pi_{y,1}^s z_{y,y} \frac{I_{y,t}}{f_y} c_{y,t}^i c_{y,t}^s + \pi_{y,2}^s z_{y,o} \frac{I_{o,t}}{f_o} c_{o,t}^i c_{y,t}^s + \pi_{y,3}^s z_{y,y} \frac{I_{y,t}}{f_y} n_{y,t}^i n_{y,t}^s + \pi_{y,4}^s \left(z_{y,y} \frac{I_{y,t}}{f_y} + z_{y,o} \frac{I_{o,t}}{f_o} \right) \right] \tag{2.7}$$

$$T_{o,t} = \eta S_{o,t} \left[\pi_{o,1}^s z_{o,o} \frac{I_{o,t}}{f_o} c_{o,t}^i c_{o,t}^s + \pi_{o,2}^s z_{o,y} \frac{I_{y,t}}{f_y} c_{y,t}^i c_{o,t}^s + \pi_{o,3}^s \left(z_{o,y} \frac{I_{y,t}}{f_y} + z_{o,o} \frac{I_{o,t}}{f_o} \right) \right] \tag{2.8}$$

where $z_{.,.}$ denotes the elements of the contact matrix:

$$Z = \begin{bmatrix} z_{y,y} & z_{y,o} \\ z_{o,y} & z_{o,o} \end{bmatrix} \tag{2.9}$$

⁷If we shut down the possibility of economic interactions the model described in the first subsection becomes a standard age-varying SIR model as employed in the epidemiological literature (see e.g. [Towers and Feng \(2012\)](#)).

describing the number of contacts between and within age groups.⁸ Denoting by f_o and $f_y = 1 - f_o$ the initial fraction of old and young respectively, this matrix must satisfy: $f_o z_{o,y} = (1 - f_o) z_{y,o}$. Each π_{\cdot}^s parameter captures the weight to each type of infectious interaction.⁹

The initial population is normalized to one, $N_0 = 1$. We assume that there is an initial shock ε to the total number of infected across the age groups according to: $I_{o,0} = \varepsilon f_o$ and $I_{y,0} = \varepsilon - I_{o,0}$.

2.1.1 Agents’ choices in the macroeconomy

The budget constraints for the young and the old individuals for $j \in \{s, i, r\}$ are:

$$\begin{aligned} (1 + \mu_{c,t})c_{y,t}^j &= w_t \phi^j n_{y,t}^j + \Gamma_t, & \phi^s = \phi^r = 1, & \phi^i < 1 \\ (1 + \mu_{c,t})c_{o,t}^j &= \bar{P} + \Gamma_t \end{aligned}$$

The elderly receive a constant pension transfer proportional to the steady state consumption of the young: $\bar{P} = \alpha c^s, 0 < \alpha < 1$.¹⁰

The government can set the consumption tax $\mu_{c,t}$ and distribute lump-sum transfers Γ_t according to the following budget constraint:

$$\mu_{c,t} C_t = \Gamma_t N_t \tag{2.10}$$

where

$$C_t = c_{y,t}^s S_{y,t} + c_{y,t}^i I_{y,t} + c_{y,t}^r R_{y,t} + c_{o,t}^s S_{o,t} + c_{o,t}^i I_{o,t} + c_{o,t}^r R_{o,t} \tag{2.11}$$

We assume the following utility function:

$$u(c, n) = \log c - \frac{\theta}{2} n^2$$

According to their type – young or old who are either susceptible, infected or recovered – individ-

⁸In this simple case, for example, the first row of Z denotes the number of contacts per period that a young makes with a young ($z_{y,y}$) and with an old ($z_{y,o}$).

⁹Setting $\pi_{y,1}^s = \pi_{y,2}^s = \pi_{y,3}^s = \pi_{o,1}^s = \pi_{o,2}^s = 0$ and $\pi_{y,4}^s = \pi_{o,3}^s = 1$ the model is a epidemiological age-varying SIR model without economic interactions. The ERT model without age variations is nested assuming that $z_{y,o} = z_{o,y} = z_{o,o} = 0$ and $f_y = 1$.

¹⁰Since our focus is on the short-run that pertains the outbreak of an epidemic we abstract from long-run issues such as transfers among individuals of different age classes as well as public debt sustainability.

uals satisfy the following dynamic programming.

Susceptibles.

$$U_{y,t}^s = u(c_{y,t}^s, n_{y,t}^s) + \beta [(1 - \tau_{y,t})U_{y,t+1}^s + \tau_{y,t}U_{y,t+1}^i] (1 - \delta_v) + \delta_v \beta U_{y,t+1}^r \quad (2.12)$$

$$U_{o,t}^s = u(c_{o,t}^s, 0) + \beta [(1 - \tau_{o,t})U_{y,t+1}^s + \tau_{o,t}U_{o,t+1}^i] (1 - \delta_v) + \delta_v \beta U_{o,t+1}^r \quad (2.13)$$

$$\tau_{y,t} = \frac{T_{y,t}}{S_{y,t}} \quad (2.14)$$

$$\tau_{o,t} = \frac{T_{o,t}}{S_{o,t}} \quad (2.15)$$

where δ_v is the per period probability of discovering a vaccine against the virus. As in ERT we assume that upon discovery the vaccine is administered to all the susceptibles in the country from the period of the discovery. Once a person is vaccinated this person becomes immune to the disease.

Infected.

$$U_{y,t}^i = u(c_{y,t}^i, n_{y,t}^i) + \beta [(1 - \pi_{y,d} - \pi_{y,r})U_{y,t+1}^i + \pi_{y,r}U_{y,t+1}^r] (1 - \delta_c) + \delta_c \beta U_{y,t+1}^r \quad (2.16)$$

$$U_{o,t}^i = u(c_{o,t}^i, 0) + \beta [(1 - \pi_{o,d} - \pi_{o,r})U_{o,t+1}^i + \pi_{o,r}U_{o,t+1}^r] (1 - \delta_c) + \delta_c \beta U_{o,t+1}^r \quad (2.17)$$

Recovered.

$$U_{y,t}^r = u(c_{y,t}^r, n_{y,t}^r) + \beta U_{y,t+1}^r \quad (2.18)$$

$$U_{o,t}^r = u(c_{o,t}^r, 0) + \beta U_{o,t+1}^r \quad (2.19)$$

Firms. There is a continuum of competitive representative firms of unit measure that produce consumption goods (Y_t) using hours worked (H_t) according to the technology $Y_t = AH_t$ to maximize profit:

$$\max_{H_t} \{AH_t - w_t H_t\}$$

which leads to the optimal condition: $w_t = A$.

Clearing.

$$C_t = AH_t + \bar{P}N_{o,t} \quad (2.20)$$

$$n_{y,t}^s S_{y,t} + \phi^i n_{y,t}^i I_{y,t} + n_{y,t}^r R_{y,t} = H_t \quad (2.21)$$

Welfare. When computing optimality of policy interventions we assume the following aggregate welfare in the first period of the epidemic:

$$U_0 = S_{o,0}U_{o,0}^s + I_{o,0}U_{o,0}^i + S_{y,0}U_{y,0}^s + I_{y,0}U_{y,0}^i \quad (2.22)$$

where the terms $U_{\cdot,0}^s$ and $U_{\cdot,0}^i$ represent the lifetime utilities of susceptibles and infected agents in each age groups.

3 *Laissez-faire* equilibrium

3.1 Contact matrix & parameter values

The initial share of young (those younger than 70) is set to $f_y = 0.825$.¹¹

The contact matrix Z in equation (2.9) is based on values from Mossong et al. (2017) contact survey data¹² and widely used in the epidemiological literature.¹³ The young have on average about 19.1 contacts per day with individuals of the same age and 1.3 contacts per day with the older group. The elderly have on average 6.3 contacts per day with the young and 1.4 contacts per day with other elderly.¹⁴ Notably, young individuals have more contacts and there is a tendency for within age group interactions.¹⁵

¹¹Compatible with the shares for Italy [United Nations World Population Prospects 2019](#)

¹²Dataset available via the package `socialmixr`. The main reference is Mossong et al. (2008). A contact is defined as “either skin-to-skin contact such as a kiss or handshake (a physical contact), or a two-way conversation with three or more words in the physical presence of another person but no skin-to-skin contact (a nonphysical contact)”.

¹³See e.g. Towers and Feng (2012) and in the COVID-19 related epidemiological literature Ferguson et al. (2020) and Figure 7 in Appendix C

¹⁴The contact matrix needs to respect a symmetric property such that $f_o z_{o,y} = f_y z_{y,o}$ which implies: $z_{o,y} = [f_y / (1 - f_y)] z_{y,o} = [0.825 / (1 - 0.825)] 1.337 = 6.303$.

¹⁵These patterns are confirmed in Appendix D where we show the contact matrix for the whole sample in the Mossong et al. (2017) survey data (namely, including also Germany, Luxembourg, Netherlands, Poland, United Kingdom, Finland, Belgium).

In the initial pre-infection steady state the population is composed only by susceptible individuals which yields:

$$\begin{aligned}n_y^s &= \theta^{-0.5} \\c_y^s &= wn_y^s = A\theta^{-0.5} \\c_o^s = \bar{P} &= \alpha c_y^s = \alpha A\theta^{-0.5}\end{aligned}$$

We assume $\alpha = 0.8$, i.e. the steady state consumption of an old individual is 80% of the steady state consumption of a young.¹⁶ Following Towers and Feng (2012), we assume that the “removal” rate γ (the rate at which an infected individual either recovers or dies) is equal across age classes. Similarly to ERT and Atkeson (2020b), we assume it takes 18 days to either recover or die for both ages, $\gamma = 7/18$.

Considering the case fatality rates by age reported by South Korean Ministry of Health and Welfare on April 5, 2020¹⁷ and interacting them with the demographic Italian shares in year 2019¹⁸ we set the ratio between the probability of dying and the removal rate to be 0.0045 and 0.1273 for the young and the old respectively.

We calibrate the parameter η in equations (2.7)–(2.8) relying on the SIR-age model, with no economic interactions. In this case η represents the transmission rate obtained from the expression for the *basic reproduction number* \mathcal{R}_0 ¹⁹ for a SIR-age model from Towers and Feng (2012):

$$\eta = \frac{\gamma \mathcal{R}_0}{\max\{\text{eig}(M)\}} \quad (3.1)$$

¹⁶This number reflects the life-cycle profile of consumption (see e.g. Fernandez-Villaverde and Krueger (2007)) with average consumption during retirement generally smaller than what one has during working-age periods.

¹⁷See https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030&act=view&list_no=366739&tag=&nPage=1

¹⁸See <https://population.un.org/wpp/Download/Standard/Population/>

¹⁹ \mathcal{R}_0 is defined as the average number of secondary infections produced by one infected individual during his/her entire period of infection in an entirely susceptible population.

where $\max\{\text{eig}(M)\}$ denote the largest eigenvalue of the M matrix:

$$M = \begin{bmatrix} z_{y,y} \frac{f_y}{f_y} & z_{y,o} \frac{f_y}{f_o} \\ z_{o,y} \frac{f_o}{f_y} & z_{y,o} \frac{f_o}{f_o} \end{bmatrix}$$

We set the \mathcal{R}_0 to 1.59 such that 60% of the initial population either recovers or die²⁰ implying $\eta = 0.0045$.

To calibrate all the $\pi_{\cdot,\cdot}^s$ parameters we revise the approach in ERT to account for the age-varying nature of our model. We assume that consumption and labor activities account for two-third of all the infection transmissions for each age class i.e. $\pi_{\cdot,\cdot}^s$ satisfy:²¹

$$\frac{\pi_{y,1}^s z_{y,y} (c_y)^2}{\Pi_y} = \frac{\pi_{y,2}^s z_{y,o} c_y c_o}{\Pi_y} = \frac{\pi_{y,3}^s z_{y,y} (n_y)^2}{\Pi_y} = 1/9,$$

$$\frac{\pi_{o,1}^s z_{o,o} (c_o)^2}{\Pi_o} = \frac{\pi_{o,2}^s z_{o,y} c_y c_o}{\Pi_o} = 1/6$$

where

$$\Pi_y = \pi_{y,1}^s z_{y,y} (c_y)^2 + \pi_{y,2}^s z_{y,o} c_y c_o + \pi_{y,3}^s z_{y,y} (n_y)^2 + \pi_{y,4}^s (z_{y,y} + z_{y,o})$$

$$\Pi_o = \pi_{o,1}^s z_{o,o} (c_o)^2 + \pi_{o,2}^s z_{o,y} c_y c_o + \pi_{o,3}^s (z_{o,y} + z_{o,o})$$

A second set of conditions returns limit-values for the number of people who either recover or die at the end of the epidemic:

$$\lim_{t \rightarrow \infty} \{R_{y,t} + D_{y,t}\} = 0.545 \qquad \lim_{t \rightarrow \infty} \{R_{o,t} + D_{o,t}\} = 0.055$$

which correspond to targeting $\lim_{t \rightarrow \infty} \{R_{y,t} + D_{y,t} + R_{o,t} + D_{o,t}\} = 0.6$.

The remaining parameters are set to the values assumed in ERT. In particular, $A = 39.835$,

²⁰As assumed by ERT and outlined by Angela Merkel in her March 11, 2020 speech <https://www.nytimes.com/2020/03/11/world/europe/coronavirus-merkel-germany.html>.

²¹Implying values for the shares of initial jump in the transmission probability due to the different causes of infection (consumption, work, residual). Consider for example equations (2.14) and (2.7). At the time of the initial infectious shock, given $I_{y,0} = \varepsilon f_y$, $I_{o,0} = \varepsilon f_o$, we have:

$$\tau_{y,0} = \varepsilon \eta [\pi_{y,1}^s z_{y,y} (c_y)^2 + \pi_{y,2}^s z_{y,o} c_y c_o + \pi_{y,3}^s z_{y,y} (n_y)^2 + \pi_{y,4}^s (z_{y,y} + z_{y,o})]$$

$\theta = 0.001275$ so that in the pre-epidemic steady state the average working week is 28 hours and the average weekly earnings is $\$58000/52$ US dollars. The discount factor $\beta = 0.96^{1/52}$ implies one life value equals 9.3 million 2019 dollars in the pre-epidemic steady state. The relative productivity of infected people is set to $\phi^i = 0.8$. The fraction of people initially infected ε is set to 0.1%. In the *laissez-faire* equilibrium there is no policy intervention i.e. $\mu_{c,t} = 0$ for all t and the frequency of contacts among individuals is set to the values in Figure 7.

3.2 Results: SIR-age macro model vs SIR-age model

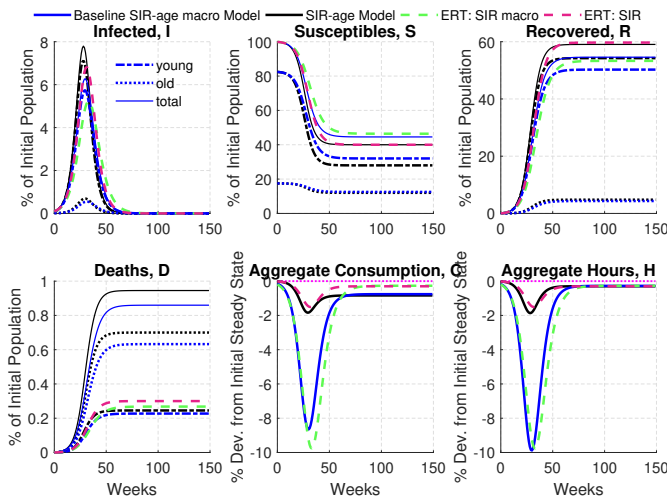


Figure 1: *Laissez-faire* SIR-age macro model vs SIR-age model vs Eichenbaum et al. (2020b) (ERT) models

Note. The SIR-age model is obtained by setting $\pi_{y,1}^s = \pi_{y,2}^s = \pi_{y,4}^s = \pi_{o,1}^s = \pi_{o,2}^s = 0, \pi_{y,3}^s = \pi_{o,3}^s = 1$.

Figure 1 shows three sets of results from different models: (i) the model from Eichenbaum et al. (2020b) (a SIR-macro model with no age differences); (ii) the “SIR-age macro model”, a SIR model with two age groups and economic interactions in the infection transmission (2.7)-(2.8), in a *laissez-faire* environment; (iii) the “SIR-age model”, a “mechanistic” epidemiological model without the endogenous behavioral response of individuals. In line with ERT, the baseline SIR-macro model predicts fewer deaths and a sharper recession than a model without the feedback

from the economy to disease diffusion due to that susceptibles reduce their consumption and hours worked to lower their probability of being infected. At the end of the transition, 0.86% of the initial population dies in the SIR-age macro model versus 0.95% in the SIR-age model (with the elderly representing 0.63% and 0.70% respectively).

These numbers are both considerably larger than the ERT baseline where 0.27% (0.3%) of the initial population dies in their SIR macro (SIR) model as they exclude those older than 70 while we focus explicitly on this age-class.²²

The recession in the SIR-age macro model is more than four times worse than in the SIR-age model. The average aggregate consumption in the first year of the epidemic falls by 3.86% (0.9%) in the SIR-age macro (SIR-age) model. From the peak-to-trough aggregate consumption decreases by 8.66% (1.84%) respectively. In the long-run aggregate consumption is permanently lowered by the death toll: 0.76% lower in the SIR-age macro model compared to 0.83% in the SIR-age model.²³ In spite of the permanent effect of the death toll, the recession is relatively reabsorbed, with aggregate consumption after one year standing at -1.7% (-0.88%) of its pre-infection level.²⁴

4 Containment policies

Based on the SIR-age-macro model we consider different possible ways of containing the epidemics spread: (1) a consumption tax μ_c inducing a reduction in the level of consumption and labor, referred to as “economic shutdown”; (2) the practice of “social distancing” which could be both mandated and behavioral implemented reducing the values of the entries $z_{..}$ in the contact matrix. Reductions in these contacts can be attributed not only to enforced measures, but also to

²²Recall that we assume that 60% of initial population either recovers or dies in the limit of the SIR-age model. This is also what ERT assume in their SIR model. Hence, with the same ending number of total susceptibles (and roughly with their same evolution, see row 1, column 2 of Figure 1) we obtain a much bigger death toll than ERT due to the higher case fatality rate of the elderly.

²³Steady state per capita consumption is given by $C/N = c_y^r[1 - (1 - \alpha)N_o/N]$. Young individuals consume the same in both the initial and final steady state (c_y^r) and the elderly are given the same fraction of consumption (αc_y^r). Since the infection reduces the proportion of elderly in the economy (N_o/N), per capita consumption in the final steady state will be (slightly) higher than in the initial steady state.

²⁴Compared to ERT, our SIR-age macro model predicts a milder recession reflecting the fact that only young individuals work in our economy while the elderly are always granted a certain level of consumption.

behavioral changes (other than consumption and labor) e.g. the use of personal protective equipment.

4.1 Social distancing

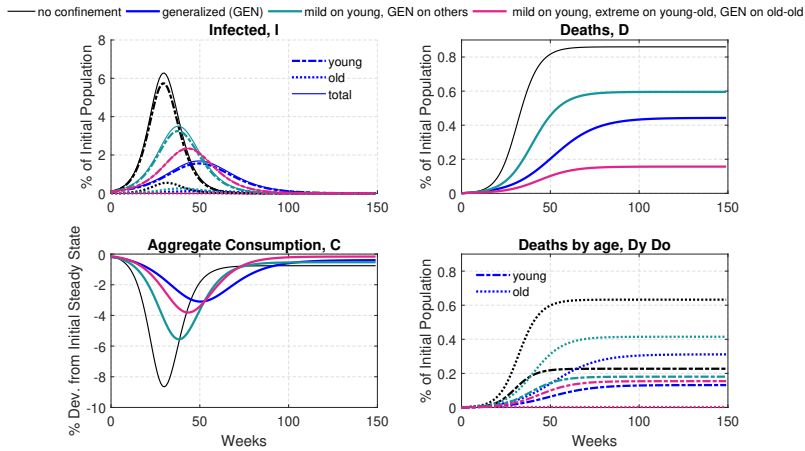


Figure 2: SIR-age macro model: different confinements

Note. The scenario “generalized (GEN)” assumes that each social contact is diminished by 20% with respect to the baseline (“no containment”). The scenario “mild on young, GEN on others” is the same as GEN but the social distancing between young individuals is 50% less severe than in GEN. The scenario “mild on young, extreme on young-old, GEN on old-old” is the same as the previous scenario but assumes that no contact is allowed between young and old individuals.

In Figure 2 we show the effect of different social distancing measures throughout the transition dynamics and we compare these scenarios with the baseline SIR-age macro model results (black lines) i.e. assuming freedom in social contacts.

A generalized policy applied to all individuals (blue lines) modelled as a generalized 20% cut in social contacts flattens the epidemic curve (halved death toll) and lowers aggregate consumption by 3% over the first year prolonging the recession as compared to the baseline. This reflects the smaller death toll (0.44% of the initial population dies versus 0.86% in the baseline) which is disproportionately borne by the old individuals.

The green line in Figure 2 shows the case in which the confinements measures restricting the

contacts among young are milder: i.e. social distancing is 50% less severe if it involves contacts among young individuals (while the generalized distancing applies to all other contacts). This implies that contacts among young individuals are 90% of the baseline and it generates a severer epidemic and a sharper recession whereas the elderly bear disproportionately more the brunt of the death toll.

Is it possible to impose a milder social distancing on contacts among young while containing the death toll and the recession imposing a total isolation of the elderly from the young as shown in the violet lines in Figure 2. In this case the death toll is significantly reduced (0.15% of the initial population dies by the end of the epidemic) with young individuals being relatively more represented in the death toll.

4.2 Shutdown of economic activity

A shutdown of economic activity is implemented (in absence of social distancing) increasing the consumption tax parameter $\mu_{c,t}$. The optimal economic shutdown μ_c is the one that maximizes the expected utility of all agents in the economy, i.e. our welfare function. The optimal level of shutdown is considered both in the case a vaccine is expected within one year from the onset of the epidemic $\delta_v = 1/52$, two years from the beginning of the epidemic $\delta_v = 0.5/52$ or when a vaccine is found with zero probability.

Two forces regulate the optimal shutdown level: (1) we want to minimize the number of deaths (this calls for a higher intensity of shutdown); (2) we want to maximize aggregate consumption (this calls for a lower intensity of the shutdown).

In Figure 3 we can see that when a vaccine is possible, it is optimal to immediately introduce severe containment measures. The closer in time the vaccine is expected to be, the more optimal it is to delay infections as a larger number of susceptible will then benefit from the vaccination (infection peaks in week 42, 45 and 48 when vaccine is available in two years, one year or never respectively). With a vaccine on the horizon the infection peaks at lower percentages of the initial population and the cumulative death toll is slightly higher (about 0.65% of the initial population

versus 0.6% in the case without vaccine) in line with less restrictive optimal shutdown throughout the epidemic.

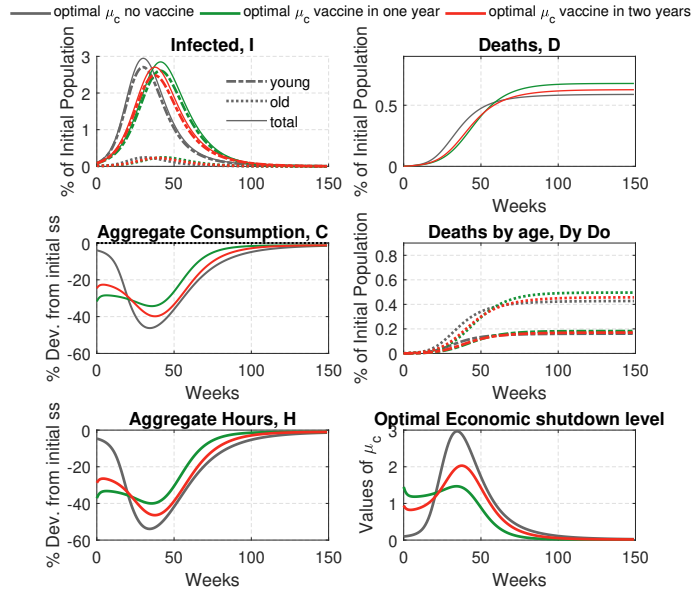


Figure 3: SIR-age macro model: Optimal shutdown by vaccine horizon

Intuitively, knowing that a vaccine will be available tomorrow, it is preferable to postpone consumption until when it is safe, hence the optimal shutdown curve is higher (green). This is associated with a larger initial drop in consumption the sooner the vaccine. It follows that the further away the vaccine is, the higher the immediate pay-off is realized from not constraining the economic activity.

Over time, as the epidemic spreads, the average level of contagion rises and the mortality-diminishing motive prevails over the consumption-rising motive in the utility maximization. This makes it optimal to rise the shutdown level, the more so, the lower the initial shutdown with consequent higher output losses up to 45% with respect to the pre-epidemic. Reducing the epidemic duration and intensity also has an amplifying positive effect for the macroeconomy in our framework as it endogenously encourages consumption.

The annual average output loss amounts to 29.6% and 31% in the case of optimal shutdown with vaccine in one year and in two years respectively, in line with ERT simulations.

4.3 Optimal shutdown with social distancing

The containment policy combination (restricting both the economic activity²⁵ and the contacts among people) can be proxied by a joint implementation of a uniform economic shutdown and social distancing policies.

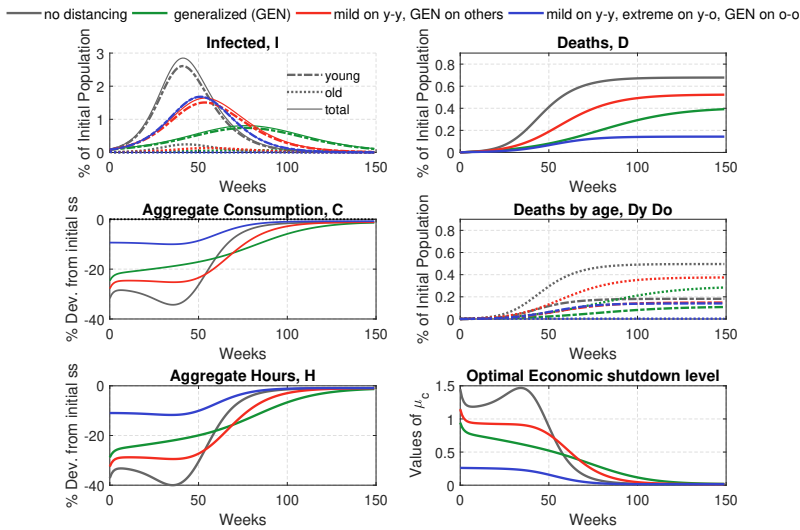


Figure 4: SIR-age macro model: Optimal shutdown $\mu_{c,t}$ with social distancing

Note. The optimal economic shutdown parameter μ_c is computed in the case in which a vaccine is discovered in one-year time from the outbreak of the epidemic for different social distancing scenarios. The scenario “generalized (GEN)” assumes that each social contact is diminished by 20% with respect to the baseline (“no distancing”). The scenario “mild on y-y, GEN on others” is the same as GEN but the social distancing among young individuals is 50% less severe than in GEN. The scenario “mild on y-y, extreme on y-o, GEN on o-o” is the same as the previous scenario but assumes that no contact is allowed between young and old individuals.

As both social distancing and economic shutdown contribute to flattening the infection curve

²⁵Across a broad set of industries with retail entertainment, restaurants and travel among the most affected.

with varying degrees of consumption losses, different social distancing measures of section 4.1 imply optimal different levels of economic shutdown displayed in Figure 4 where we have assumed that a vaccine becomes available on average in one year time after the beginning of the epidemic.²⁶

The policy minimizing consumption and death losses is not the one flattening the overall infection curve, but the policy mix flattening the elderly infections' curve the most.

The social distancing measure minimizing the optimal economic shutdown and the consumption reduction is the isolation of the old from the young population which also delivers the most favorable economic outcome over one year horizon (9.6% losses with respect to the pre-epidemic equilibrium). In this case the old isolate from the more socially active young dramatically cutting the contagion possibilities and therefore requiring weaker additional measures. In this case, consumption drops immediately by only 10% (as compared to e.g. the 25% with generalized social distancing) and the overall death toll is limited to 0.18% of the total population.

The scenario presenting the second lowest need of shutdown is a generalized ("GEN") reduction by 20% of social contacts with milder policies on the young-young contacts. As young agents have the largest number of social interactions (see Figure 7), the loosening of social distancing for these agents implies a higher level of optimal enforced shutdown.

When young agents are allowed to interact in a socio-economic context the rise in infection calls for stricter economic shutdown measures generating the highest average output loss (24.9% losses with respect to the pre-epidemic).

In addition when social distancing is not enforced it is necessary to rise the economic shutdown level over time, as contagion increases more. However, the larger rise in the number of infected occurring in this scenario imply a dampening endogenous response of young agents' consumption acting in the same direction of the required rise in the level of economic shutdown and harming economic activity up to 40% downwards.

²⁶In Figure 10 in Appendix D we observe the corresponding scenarios in absence of a vaccine possibility.

4.4 Containment policies comparison

To summarize what we learned from the analysis of the containment policies in the previous sections, we compare the different containment policies in terms of economic impact and death costs at the aggregate level. The comparison is done for the cases in which a vaccine is expected to be available in one year time from the epidemic outbreak.²⁷

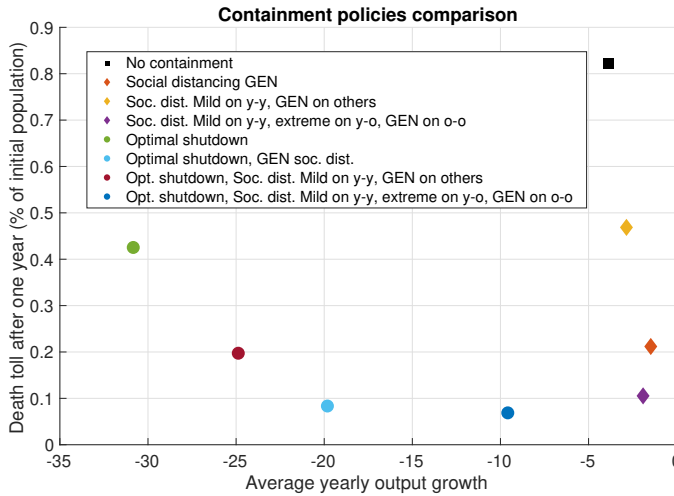


Figure 5: SIR-age macro: containment policies comparison

Note. The optimal economic shutdown parameter μ_c is computed in the case in which a vaccine is discovered in one year time from the outbreak of the epidemic for different social distancing scenarios. The scenario “generalized (GEN)” assumes that each social contact is diminished by 20% with respect to the baseline (“no distancing”). The scenario “mild on y-y, GEN on others” is the same as GEN but the social distancing among young individuals is 50% less severe than in GEN. The scenario “mild on y-y, extreme on y-o, GEN on o-o” is the same as the previous scenario but assumes that no contact is allowed between young and old individuals.

Figure 5 shows the deaths after one year (week 52) in percentage of the initial population (y-axis) and the average yearly output growth (x-axis) for different containment policies. The no containment scenario is characterized by the highest death toll of about 0.82% of the population and an economic loss of 3.86% of output as compared to the steady state.

²⁷Notice that we assume that the social distancing measures are in place forever. This assumption is irrelevant as long as the course of the epidemic spans over a year after which, as we assume, a vaccine can become available. In reality, of course, there is much uncertainty on both the length of the epidemic and the effective availability of a vaccine.

In comparison to a no containment case, a generalized social distancing would attain a lower average yearly output loss, 1.47% and a lower death toll after one year of 0.21% of the population. A deviation from this generalized social distancing featuring young people entertaining more social interactions would result in higher deaths (0.47%) and higher output loss (2.85%) as compared to generalized social distancing (however, it still represents an improvement in both dimensions as compared to no containment). A deviation from generalized social distancing via a reduction in older age groups contact with the younger is instead characterized by 1.9% output loss (worse than generalized social distancing, but better than mild social distancing for young) and 0.11% deaths as compared to initial population (better than both generalized and mild on young distancing).

An optimal shutdown measure is characterized by 0.43% deaths and 30.1% average yearly consumption losses, however in presence of social distancing the optimal shutdown will generate 19.8%, 24.9% and 9.6% average yearly output losses in the case of generalized social, milder on young and stricter on old social distancing respectively. The corresponding death tolls in the three cases are 0.08%, 0.2% and 0.07%. For any level of social distancing, the implied optimal economic shutdown generates small gains in terms of lives and large increases in terms of average output losses over one-year time.

5 Applying the model

Here we recalibrate the model to be consistent with the evolution of the epidemic as well as the economic shutdown and social distancing measures in Italy. We then investigate different possibilities for social distancing and economic shutdown for the progressive release of economic activity.

5.1 Re-calibrating the model

We consider the first week of 2020 as the *epidemic origin* (similarly to Favero et al. (2020)) and we recalibrate the model (as described in Appendix C) to match the number of effective deaths in week 10. According to our model it took two to three weeks to flatten the death curve after the

imposed *lockdown* began (at the beginning of week 11).²⁸ Absent the lockdown measures in Italy, March and April would have seen a higher death toll according to our model, killing more than 0.4% of the initial population in a matter of 2 months.²⁹

To be consistent with the death-toll observed in the data we assume that during the lockdown all contacts among individuals of any age group gets reduced to 46% of what would prevail in normal circumstances. To make our scenarios are also consistent with the official estimate of the Italian economic recession (real GDP) in the first quarter of 2020 (-5.3% in Q1 2020 with respect to the previous quarter, ISTAT (2020)) we set $\mu_c = 0.625$ in all periods of the lockdown phase.³⁰

5.2 Epidemic management scenarios

We assume that the government post-lockdown sets the intensity of the economic shutdown (μ_c) in proportion to the observed stock of contemporaneous infected people:

$$\mu_{c,t} = \beta_\mu (I_{y,t} + I_{o,t}) \quad (5.1)$$

aiming at capturing the ready-to-intervene attitude of the government.³¹

Assuming $\beta_\mu = 9$ in all periods t in the post-lockdown phase, we obtain that the annual average growth rate of output in 2020 is in the range of -11.4% to -8.6% for the three main social distancing scenarios analyzed in Figure 6. This result compares, for example, with the estimate of -9.1% provided by the International Monetary Fund (2020). For the second quarter of 2020 compared with the first quarter of the same year the model generates an economic recession in the range of -18.8% to -12.2%. For comparison, European Commission (2020) projects the equivalent figure to

²⁸Data on COVID-19 infections and deaths are observed since week 9 (starting February 23). Week 11 (starting March 8) marks the beginning of the general shutdown of “inessential” economic activities and of enforced generalized social distancing: *Lockdown*. In week 19 (starting May 3) the government started by decree a post-lockdown period with milder containment measures while keeping a close monitoring of the total stock of infected (adjusting the measures of economic shutdown accordingly). May 3 is also when we stopped retrieving data on observed total deaths.

²⁹The official and effective data point to a total death-toll of 0.047% and 0.070% of the initial population respectively by the end of week 18 (April 26 to May 2).

³⁰In particular, our model produces a recession in the first quarter of 2020 (compared to the previous quarter) of 5.3%, 5.2%, 4.9% corresponding to the blue, grey and green lines in Figure 8 in Appendix, respectively.

³¹Cf. “Phase 2, Press Conference of the Prime Minister, Giuseppe Conte”, April 26, 2020, available at <http://www.governo.it/node/14518>.

be -13.6%.

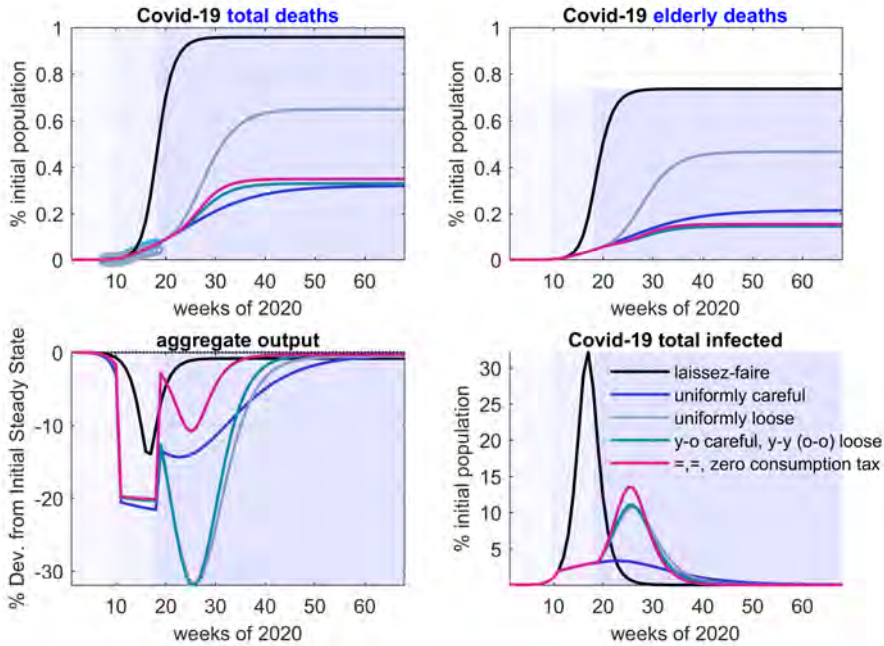


Figure 6: Different post-lockdown scenarios

We show the following possible scenarios for the post-lockdown phase:

- (0) The reference case is a *laissez-faire* equilibrium (black lines in Figure 6) where individuals maintain the same number of weekly contacts prevailing in the pre-epidemic times and the consumption tax is always set to zero. Our SIR-age macro model predicts that the economy would have faced a rapid recession with a trough at week 17 and an annual average growth rate of output in 2020 of -2.1% (vs -1.2% in the SIR-age model with no macroeconomic interactions). The cost of this milder recession is a high death-toll in the long-run of more than 0.9% of the initial population, i.e. more than 0.6 million lives.

The three post-lockdown social distancing policies are:

- (1) *Uniformly careful* (blue lines in Figure 6), as careful as in the lockdown phase. Compared to the *laissez-faire* equilibrium, the final death-toll in this case is about 3 times smaller while the annual average growth rate of output in 2020 is about 4 times more negative. About two-thirds of the final death toll occur among the elderly.
- (2) *Uniformly loose* (grey lines in Figure 6). The post lockdown social distancing is 50% looser than during the lockdown phase. Compared to scenario (1), this loosening doubles the final death toll, with more than two-thirds of deaths accounted by the elderly. The annual average growth rate of output in 2020 is -11.4%, worse than under scenario (1) (-8.6%) as the number of total infected increases slightly.
- (3) *Careful with young-old, loose with young-young and old-old contacts* (green lines in Figure 6). We assume that the contacts among young and old individuals are two-thirds less than in the lockdown phase; the contacts among individuals of the same age are correspondingly two-thirds more. Under this scenario, the final death-toll is roughly the same one prevailing in scenario (1): there are more contacts, but the composition is changed and the elderly now account for a smaller share of the final death toll (44%). A new wave which is similar to the one under scenario (2) is present due to more cases, but these new infectious cases are not so lethal as they mostly occur among the young individuals. This results in an annual growth rate of output at -10.6%. While part of this bigger loss of output is due to the fact that more infected individuals decrease aggregate productivity and induce individuals (who discount a higher probability of being infected) to cut back on consumption and hours worked, the main reason is that we assume that the government sets the consumption tax proportionally to the number of contemporaneous infected. This is confirmed by the violet line in Figure 6 showing what would happen if the government was to set the consumption tax to zero in all post-lockdown periods. In this case the output loss in the second quarter (quarter on quarter) would be about 2.4 times smaller (compare with the green line).³²

³²Throughout the simulations we keep the number of contacts among the elderly (“old-old contacts”) at the same level of looseness prevailing for contacts among the young. However, the old-old contact channel is not very important in

6 Concluding remarks

Combining an epidemiological framework (Towers and Feng, 2012) into a macroeconomic one (Eichenbaum et al., 2020b), our model suggests that differentiating the policies to contain the COVID-19 crisis by age would be optimal. Compared to uniform social distancing and economic shutdown measures, age-targeted measures reduce the death toll while containing the output losses.

We calibrated the model for two age-groups: those aged 70 or more and the rest of the population. An extension would be to consider more age-groups that would allow to extend the set of policies as well as analysis of distributional policies and compensation policies. While highlighting the differences in impact of an economic shutdown and social distancing scenarios, this model can be extended to incorporating additional economic trade-off channels of social distancing especially relevant when analyzing phases of economic reopening.

Other important aspects that are not yet tackled in our analysis are the limits imposed by the hospitalization capacity and the possibility of other targeted policies focusing on the revealed health status of individuals. Aspects as the human costs of prolonged social distancing between young individuals and the elderly or time-varying developments of the COVID-19 transmissibility are not considered in our framework and one should remember to put into context the interpretation of our results.

Given the uncertainty surrounding many of the model parameters, we recognize that our quantitative results are most valuable in terms of a relative order of magnitude. In addition, together with the economic literature complementing our analysis, we contribute to the rising view that differentiating containment measures by age, acting more on social distancing of the elderly from the young is preferable to uniform and untargeted economic measures.

the model at a macroeconomic level as the elderly are fewer in the macroeconomy (initially 16.7% of the population) and have comparably few contacts among themselves in normal times (see Figure 7).

References

- Acemoglu, D., Chernozhukov, V., Werning, I., and Whinston, M. D. (2020). A Multi-Risk SIR Model with Optimally Targeted Lockdown. Working Paper 27102, National Bureau of Economic Research.
- Alvarez, F. E., Argente, D., and Lippi, F. (2020). A Simple Planning Problem for COVID-19 Lockdown. *Covid Economics*, 14:1–32. 6 May. <https://cepr.org/sites/default/files/CovidEconomics14.pdf>.
- Andersen, A. L., Hansen, E. T., Johannesen, N., and Sheridan, A. (2020). Consumer responses to the covid-19 crisis: Evidence from bank account transaction data. *Covid Economics*, 7:92–118. 20 April. <https://cepr.org/sites/default/files/CovidEconomics7.pdf>.
- Atkeson, A. (2020a). How deadly is covid-19? understanding the difficulties with estimation of its fatality rate. Working Paper 26965, National Bureau of Economic Research.
- Atkeson, A. (2020b). What Will Be the Economic Impact of COVID-19 in the US? Rough Estimates of Disease Scenarios. NBER Working Papers 26867, National Bureau of Economic Research, Inc.
- Baqaei, D., Farhi, E., Mina, M. J., and Stock, J. H. (2020). Reopening scenarios. Working Paper 27244, National Bureau of Economic Research.
- Berger, D. W., Herkenhoff, K. F., and Mongey, S. (2020). An seir infectious disease model with testing and conditional quarantine. *Covid Economics*, 13:1–30. 4 May. <https://cepr.org/sites/default/files/CovidEconomics13.pdf>.
- Bodenstein, M., Corsetti, G., and Guerrieri, L. (2020). Social Distancing and Supply Disruptions in a Pandemic. *Finance and Economics Discussion Series 2020-031*.
- Brotherhood, L., Kircher, P., Santos, C., and Tertilt, M. (2020). An economic model of the covid-19 epidemic: The importance of testing and age-specific policies.
- Cereda, D., Tirani, M., Roviola, F., Demicheli, V., Ajelli, M., Poletti, P., Trentini, F., Guzzetta, G., Marziano, V., Barone, A., Magoni, M., Deandrea, S., Diurno, G., Lombardo, M., Faccini, M., Pan, A., Bruno, R., Pariani, E., Grasselli, G., Piatti, A., Gramegna, M., Baldanti, F., Melegaro, A., and Merler, S. (2020). The early phase of the COVID-19 outbreak in Lombardy, Italy. Available at <https://arxiv.org/abs/2003.09320>.
- Chang, R. and Velasco, A. (2020). Economic Policy Incentives to Preserve Lives and Livelihoods. *Covid Economics*, 14:33–56. 6 May. <https://cepr.org/sites/default/files/CovidEconomics14.pdf>.
- Chikina, M. and Pegden, W. (2020). Modeling strict age-targeted mitigation strategies for COVID-19. *arXiv preprint arXiv:2004.04144*.
- Dowd, J. B., Andriano, L., Brazel, D. M., Rotondi, V., Block, P., Ding, X., Liu, Y., and Mills, M. C. (2020). Demographic science aids in understanding the spread and fatality rates of covid-19. *Proceedings of the National Academy of Sciences*, 117(18):9696–9698.

- Dudel, C., Riffe, T., Acosta, E., van Raalte, A. A., Strozza, C., and Myrskyl, M. (2020). Monitoring trends and differences in COVID-19 case-fatality rates using decomposition methods: contributions of age structure and age-specific fatality. MPIDR Working Papers WP-2020-020, Max Planck Institute for Demographic Research, Rostock, Germany.
- Eichenbaum, M. S., Rebelo, S., and Trabandt, M. (2020a). The macroeconomics of testing and quarantining. Technical report, National Bureau of Economic Research.
- Eichenbaum, M. S., Rebelo, S., and Trabandt, M. (2020b). The Macroeconomics of Epidemics. NBER Working Papers 26882, National Bureau of Economic Research, Inc.
- Ellison, G. (2020). Implications of heterogeneous sir models for analyses of covid-19. *Covid Economics*, 53:1–32. 23 October. <https://cepr.org/sites/default/files/CovidEconomics53.pdf>.
- European Commission (2020). European economic forecast: – spring 2020. *European Commission INSTITUTIONAL PAPER 125*. Statistical Annex.
- Farboodi, M., Jarosch, G., and Shimer, R. (2020). Internal and external effects of social distancing in a pandemic. *Covid Economics*, 9:25–61. 24 April. <https://cepr.org/sites/default/files/CovidEconomics9.pdf>.
- Faria-e-Castro, M. (2020). Fiscal Policy during a Pandemic. *Covid Economics*, 2:67–101. 8 April. <https://cepr.org/sites/default/files/CovidEconomics2.pdf>.
- Favero, C. A., Ichino, A., and Rustichini, A. (2020). Restarting the economy while saving lives under Covid-19. *Centre for Economic Policy Research*. London. https://cepr.org/active/publications/discussion_papers/dp.php?dpno=14664.
- Ferguson, N., Laydon, D., Nedjati, G. G., Imai, N., Ainslie, K., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunuba, P. Z., Cuomo-Dannenburg, G., Dighe, A., Dorigatti, I., Fu, H., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Okell, L., van, E. S., Thompson, H., Verity, R., Volz, E., Wang, H., Wang, Y., Walker, P., Walters, C., Winskill, P., Whittaker, C., Donnelly, C., Riley, S., and Ghani, A. (2020). Report 9: Impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand. Technical report, Imperial College London.
- Fernandez-Villaverde, J. and Krueger, D. (2007). Consumption over the Life Cycle: Facts from Consumer Expenditure Survey Data. *The Review of Economics and Statistics*, 89(3):552–565.
- Fernndez-Villaverde, J. and Jones, C. I. (2020). Estimating and Simulating a SIRD Model of COVID-19 for Many Countries, States, and Cities. Working Paper 27128, National Bureau of Economic Research.
- Galeotti, A., Hacıoglu, S., and Surico, P. (2020). A rule of thumb to detect excess deaths? Lessons from Italy. Manuscript. April 28, available at https://www.dropbox.com/s/0uq5xc355mdlyi8/death_count_final_ghs.pdf?dl=0.

- Garibaldi, P., Moen, E. R., and Pissarides, C. A. (2020). Modelling contacts and transitions in the SIR epidemics model. *Covid Economics*, 5:1–20. 16 April. <https://cepr.org/sites/default/files/news/CovidEconomics5.pdf>.
- Glover, A., Heathcote, J., Krueger, D., and Ríos-Rull, J.-V. (2020). Health versus wealth: On the distributional effects of controlling a pandemic. *Covid Economics*, 6:22–64. 17 April. <https://cepr.org/sites/default/files/CovidEconomics6.pdf>.
- Goldstein, J. R. and Lee, R. D. (2020). Demographic Perspectives on Mortality of Covid-19 and Other Epidemics. *NBER Working Paper*, (27043). <https://www.nber.org/papers/w27043>.
- Gollier, C. (2020a). Cost-benefit analysis of age-specific deconfinement strategies. *Journal of Public Economic Theory*. Forthcoming. Prepublished in CEPR Covid Economics 24, 1-31 <https://cepr.org/sites/default/files/news/CovidEconomics24.pdf>.
- Gollier, C. (2020b). If the objective is herd immunity, on whom should it be built? *EconPol POLICY BRIEF*, 4(29).
- Gregory, V., Menzio, G., and Wiczer, D. G. (2020). Pandemic Recession: L or V-Shaped? *Covid Economics*, 15:88–109. 7 May. <https://cepr.org/sites/default/files/CovidEconomics15.pdf>.
- Guerrieri, V., Lorenzoni, G., Straub, L., and Werning, I. (2020). Macroeconomic implications of covid-19: Can negative supply shocks cause demand shortages? Technical report, National Bureau of Economic Research.
- International Monetary Fund (2020). World Economic Outlook, April 2020: The Great Lockdown. *International Monetary Fund*. Available at <https://www.imf.org/en/Publications/WEO/Issues/2020/04/14/weo-april-2020>.
- ISTAT (2020). Statistics Flash: I Quarter 2020 Quarterly National Accounts. May 29. Available at <https://www.istat.it/it/files//2020/05/Quarterly-national-accounts-Q1-2020.pdf>.
- Jones, C. J., Philippon, T., and Venkateswaran, V. (2020). Optimal Mitigation Policies in a Pandemic: Social Distancing and Working from Home. NBER Working Papers 26984, National Bureau of Economic Research, Inc.
- Kapička, M. and Rupert, P. (2020). Labor Markets during Pandemics. April 9, manuscript.
- Kaplan, G., Moll, B., and Violante, G. L. (2020). The great lockdown and the big stimulus: Tracing the pandemic possibility frontier for the u.s. Working Paper 27794, National Bureau of Economic Research.
- Kermack, W. and McKendrick, A. (1927). A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society of London A*, 115(772):700–721. <https://doi.org/10.1098/rspa.1927.0118>.

- Krueger, D., Uhlig, H., and Xie, T. (2020). Macroeconomic Dynamics and Reallocation in an Epidemic. *Covid Economics*, 5:21–55. 16 April. <https://cepr.org/sites/default/files/CovidEconomics5.pdf>.
- Kuhn, M. and Bayer, C. (2020). Intergenerational ties and case fatality rates: A cross-country analysis. Manuscript. Available at <https://www.wiwi.uni-bonn.de/kuhn/paper/COVID.pdf>.
- McKibbin, W. and Fernando, R. (2020). The global macroeconomic impacts of COVID-19: Seven scenarios. CAMA Working Papers 2020-19, Centre for Applied Macroeconomic Analysis, Crawford School of Public Policy, The Australian National University.
- Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., Massari, M., Salmaso, S., Tomba, G. S., Wallinga, J., Heijne, J., Sadkowska-Todys, M., Rosinska, M., and Edmunds, W. J. (2008). Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. *PLOS Medicine*, 5(3):1–1.
- Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., Massari, M., Salmaso, S., Tomba, G. S., Wallinga, J., Heijne, J., Sadkowska-Todys, M., Rosinska, M., and Edmunds, W. J. (2017). Polymod social contact data. Version 1.1.
- Rampini, A. A. (2020). Sequential Lifting of COVID-19 Interventions with Population Heterogeneity. Working Paper 27063, National Bureau of Economic Research.
- Rinaldi, G. and Paradisi, M. (2020). An empirical estimate of the infection fatality rate of covid-19 from the first italian outbreak. *medRxiv*.
- Stock, J. H. (2020). Data Gaps and the Policy Response to the Novel Coronavirus. *Covid Economics*, 3:1–11. 10 April. <https://cepr.org/sites/default/files/CovidEconomics3.pdf>.
- Towers, S. and Feng, Z. (2012). Social contact patterns and control strategies for influenza in the elderly. *Mathematical biosciences*, 240:241–9.
- Von Thadden, E.-L. (2020). A simple, non-recursive model of the spread of covid-19 with applications to policy. *Covid Economics*, 10:24–43. 27 April. <https://cepr.org/sites/default/files/news/CovidEconomics10.pdf>.

Appendix

A First order conditions

The first order conditions are:

Susceptibles

$$c_{y,t}^s : \frac{1}{c_{y,t}^s} - (1 + \mu_{c,t})\lambda_{y,t}^s + \eta \left(\pi_{y,1}^s z_{y,y} \frac{I_{y,t}}{f_y} c_{y,t}^i + \pi_{y,2}^s z_{y,o} \frac{I_{o,t}}{f_o} c_{o,t}^i \right) \lambda_{y,t}^\tau = 0 \quad (\text{A.1})$$

$$n_{y,t}^s : -\theta n_{y,t}^s + w_t \lambda_{y,t}^s + \eta \pi_{y,A}^s z_{y,y} \frac{I_{y,t}}{f_y} n_{y,t}^i \lambda_{y,t}^\tau = 0 \quad (\text{A.2})$$

$$\tau_{y,t} : \beta (U_{y,t+1}^i - U_{y,t+1}^s) (1 - \delta_v) = \lambda_{y,t}^\tau \quad (\text{A.3})$$

$$\lambda_{y,t}^s : (1 + \mu_{c,t})c_{y,t}^s = w_t n_{y,t}^s + \Gamma_t \quad (\text{A.4})$$

$$\lambda_{o,t}^s : c_{o,t}^s = \frac{\bar{P} + \Gamma_t}{1 + \mu_{c,t}} \quad (\text{A.5})$$

Infected

$$c_{y,t}^i : \frac{1}{c_{y,t}^i} - (1 + \mu_{c,t})\lambda_{y,t}^i = 0 \quad (\text{A.6})$$

$$n_{y,t}^i : -\theta n_{y,t}^i + \phi^i w_t \lambda_{y,t}^i = 0 \quad (\text{A.7})$$

$$\lambda_{y,t}^i : (1 + \mu_{c,t})c_{y,t}^i = w_t \phi^i n_{y,t}^i + \Gamma_t \quad (\text{A.8})$$

$$\lambda_{o,t}^i : c_{o,t}^i = \frac{\bar{P} + \Gamma_t}{1 + \mu_{c,t}} \quad (\text{A.9})$$

Recovered

$$c_{y,t}^r : \frac{1}{c_{y,t}^r} - (1 + \mu_{c,t})\lambda_{y,t}^r = 0 \quad (\text{A.10})$$

$$n_{y,t}^r : -\theta n_{y,t}^r + w_t \lambda_{y,t}^r = 0 \quad (\text{A.11})$$

$$\lambda_{y,t}^r : (1 + \mu_{c,t})c_{y,t}^r = w_t n_{y,t}^r + \Gamma_t \quad (\text{A.12})$$

$$\lambda_{o,t}^r : c_{o,t}^r = \frac{\bar{P} + \Gamma_t}{1 + \mu_{c,t}} \quad (\text{A.13})$$

B Computing the equilibrium

Following closely ERT, for a given sequence of containment rates $\{\mu_{c,t}\}_{t=0}^F$ for some final horizon F , guess sequences $\{n_{y,t}^s, n_{y,t}^i, n_{y,t}^r\}_{t=0}^F$ compute the sequence of the remaining unknown variables

in each of the following equilibrium equations:

$$\lambda_{y,t}^r = \frac{\theta n_{y,t}^r}{A} \tag{B.1}$$

$$c_{y,t}^r = [(1 + \mu_{c,t})\lambda_{y,t}^r]^{-1} \tag{B.2}$$

$$\Gamma_t = (1 + \mu_{c,t})c_{y,t}^r - An_{y,t}^r \tag{B.3}$$

$$c_{o,t}^r = \frac{\bar{P} + \Gamma_t}{1 + \mu_{c,t}} \tag{B.4}$$

$$u_{y,t}^r = \log c_{y,t}^r - \frac{\theta}{2}(n_{y,t}^r)^2 \tag{B.5}$$

$$u_{o,t}^r = \log c_{o,t}^r \tag{B.6}$$

Iterate backwards from the post-epidemic steady-state values of $U_{y,t}^r, U_{o,t}^r, U_{y,F}^r = u_{y,t}^r/(1 - \beta), U_{o,F}^r = u_{o,t}^r/(1 - \beta)$:

$$U_{y,t}^r = u_{y,t}^r + \beta U_{y,t+1}^r \tag{B.7}$$

$$U_{o,t}^r = u_{o,t}^r + \beta U_{o,t+1}^r \tag{B.8}$$

Calculate the sequence for remaining unknowns in the following equations:

$$\lambda_{y,t}^i = \frac{\theta n_{y,t}^i}{\phi^i A} \tag{B.9}$$

$$c_{y,t}^i = [(1 + \mu_{c,t})\lambda_{y,t}^i]^{-1} \tag{B.10}$$

$$c_{o,t}^i = \frac{\bar{P} + \Gamma_t}{1 + \mu_{c,t}} \tag{B.11}$$

$$u_{y,t}^i = \log c_{y,t}^i - \frac{\theta}{2}(n_{y,t}^i)^2 \tag{B.12}$$

$$u_{o,t}^i = \log c_{o,t}^i \tag{B.13}$$

$$c_{y,t}^s = \frac{An_{y,t}^s + \Gamma_t}{1 + \mu_{c,t}} \tag{B.14}$$

$$c_{o,t}^s = \frac{\bar{P} + \Gamma_t}{1 + \mu_{c,t}} \tag{B.15}$$

$$u_{y,t}^s = \log c_{y,t}^s - \frac{\theta}{2}(n_{y,t}^s)^2 \tag{B.16}$$

$$u_{o,t}^s = \log c_{o,t}^s \tag{B.17}$$

Given initial values $I_{o,0} = \varepsilon f_y, I_{y,0} = \varepsilon - I_{o,0}, S_{y,0} = (1 - \varepsilon)f_y, S_{o,0} = 1 - \varepsilon - S_{y,0}, N_{y,0} = f_y, N_{o,0} = f_o = 1 - f_y, N_0 = N_{y,0} + N_{o,0} = 1$, iterate forward the following equations for $t = 0, 1, \dots, F - 1$:

$$T_{y,t} = \eta S_{y,t} \left[\pi_{y,1}^s z_{y,y} \frac{I_{y,t}}{f_y} c_{y,t}^i c_{y,t}^s + \pi_{y,2}^s z_{y,o} \frac{I_{o,t}}{f_o} c_{o,t}^i c_{y,t}^s + \pi_{y,3}^s z_{y,y} \frac{I_{y,t}}{f_y} n_{y,t}^i n_{y,t}^s + \pi_{y,4}^s \left(z_{y,y} \frac{I_{y,t}}{f_y} + z_{y,o} \frac{I_{o,t}}{f_o} \right) \right] \quad (\text{B.18})$$

$$T_{o,t} = \eta S_{o,t} \left[\pi_{o,1}^s z_{o,o} \frac{I_{o,t}}{f_o} c_{o,t}^i c_{o,t}^s + \pi_{o,2}^s z_{o,y} \frac{I_{y,t}}{f_y} c_{y,t}^i c_{o,t}^s + \pi_{o,3}^s \left(z_{o,y} \frac{I_{y,t}}{f_y} + z_{o,o} \frac{I_{o,t}}{f_o} \right) \right] \quad (\text{B.19})$$

$$S_{y,t+1} = S_{y,t} - T_{y,t} \quad (\text{B.20})$$

$$S_{o,t+1} = S_{o,t} - T_{o,t} \quad (\text{B.21})$$

$$I_{y,t+1} = I_{y,t} + T_{y,t} - (\pi_{y,r} + \pi_{y,d}) I_{y,t} \quad (\text{B.22})$$

$$I_{o,t+1} = I_{o,t} + T_{o,t} - (\pi_{o,r} + \pi_{o,d}) I_{o,t} \quad (\text{B.23})$$

$$R_{y,t+1} = R_{y,t} + \pi_{y,r} T_{y,t} \quad (\text{B.24})$$

$$R_{o,t+1} = R_{o,t} + \pi_{o,r} T_{o,t} \quad (\text{B.25})$$

$$D_{y,t+1} = D_{y,t} + \pi_{y,d} T_{y,t} \quad (\text{B.26})$$

$$D_{o,t+1} = D_{o,t} + \pi_{o,d} T_{o,t} \quad (\text{B.27})$$

$$N_{y,t} = S_{y,t} + I_{y,t} + R_{y,t} \quad (\text{B.28})$$

$$N_{o,t} = S_{o,t} + I_{o,t} + R_{o,t} \quad (\text{B.29})$$

$$N_{t+1} = N_{y,t+1} + N_{o,t+1} \quad (\text{B.30})$$

Given the post-epidemic steady-state values of $U_{y,t}^i, U_{o,t}^i, U_{y,t}^s, U_{o,t}^s$:

$$U_y^i = \frac{u_y^i + \beta[\pi_{y,r}(1 - \delta_c) + \delta_c] U_y^r}{1 - \beta(1 - \pi_{y,d} - \pi_{y,r})(1 - \delta_c)}$$

$$U_o^i = \frac{u_o^i + \beta[\pi_{o,r}(1 - \delta_c) + \delta_c] U_o^r}{1 - \beta(1 - \pi_{o,d} - \pi_{o,r})(1 - \delta_c)}$$

$$U_{y,F}^i = (1 - \delta_c)^F U_y^i + [1 - (1 - \delta_c)^F] U_y^r$$

$$U_{o,F}^i = (1 - \delta_c)^F U_o^i + [1 - (1 - \delta_c)^F] U_o^r$$

$$U_y^s = \frac{u_y^s + \delta_v \beta U_y^r}{1 - \beta(1 - \delta_v)}$$

$$U_o^s = \frac{u_o^s + \delta_v \beta U_o^r}{1 - \beta(1 - \delta_v)}$$

$$U_{y,F}^s = (1 - \delta_v)^F U_y^s + [1 - (1 - \delta_v)^F] U_y^r$$

$$U_{o,F}^s = (1 - \delta_v)^F U_o^s + [1 - (1 - \delta_v)^F] U_o^r$$

iterate backwards:

$$U_{y,t}^i = u_{y,t}^i + \beta \left[(1 - \pi_{y,d} - \pi_{y,r}) U_{y,t+1}^i + \pi_{y,r} U_{y,t+1}^r \right] (1 - \delta_c) + \delta_c \beta U_{y,t+1}^r \quad (\text{B.31})$$

$$U_{o,t}^i = u_{o,t}^i + \beta \left[(1 - \pi_{o,d} - \pi_{o,r}) U_{o,t+1}^i + \pi_{o,r} U_{o,t+1}^r \right] (1 - \delta_c) + \delta_c \beta U_{o,t+1}^r \quad (\text{B.32})$$

$$\tau_{y,t} = T_{y,t}/S_{y,t} \tag{B.33}$$

$$\tau_{o,t} = T_{o,t}/S_{o,t} \tag{B.34}$$

$$U_{y,t}^s = u_{y,t}^s + \beta [(1 - \tau_{y,t})U_{y,t+1}^s + \tau_{y,t}U_{y,t+1}^i] (1 - \delta_v) + \delta_v \beta U_{y,t+1}^r \tag{B.35}$$

$$U_{o,t}^s = u_{o,t}^s + \beta [(1 - \tau_{o,t})U_{o,t+1}^s + \tau_{o,t}U_{o,t+1}^i] (1 - \delta_v) + \delta_v \beta U_{o,t+1}^r \tag{B.36}$$

Calculate the sequence of remaining unknowns by the following equations:

$$\lambda_{y,t}^\tau = \beta (U_{y,t+1}^i - U_{y,t+1}^s) (1 - \delta_v) \tag{B.37}$$

$$\lambda_{y,t}^s = \frac{\frac{1}{c_{y,t}^s} + \eta \left(\pi_{y,1}^s z_{y,y} \frac{I_{y,t}}{f_y} c_{y,t}^i + \pi_{y,2}^s z_{y,o} \frac{I_{o,t}}{f_o} c_{o,t}^i \right) \lambda_{y,t}^\tau}{(1 + \mu_{c,t})} \tag{B.38}$$

Finally, given aggregate consumption:

$$C_t = c_{y,t}^s S_{y,t} + c_{y,t}^i I_{y,t} + c_{y,t}^r R_{y,t} + c_{o,t}^s S_{o,t} + c_{o,t}^i I_{o,t} + c_{o,t}^r R_{o,t} \tag{B.39}$$

use a gradient-based method to adjust the guesses $\{n_{y,t}^s, n_{y,t}^i, n_{y,t}^r\}_{t=0}^{F-1}$ so that the following three equations hold with arbitrary precision:

$$(1 + \mu_{c,t})c_{y,t}^i - A\phi^i n_{y,t}^i - \Gamma_t = 0 \tag{B.40}$$

$$-\theta n_{y,t}^s + A\lambda_{y,t}^s + \eta \pi_{y,3}^s z_{y,y} \frac{I_{y,t}}{f_y} n_{y,t}^i \lambda_{y,t}^\tau = 0 \tag{B.41}$$

$$\mu_{c,t}C_t - \Gamma_t N_t = 0 \tag{B.42}$$

C Calibration details

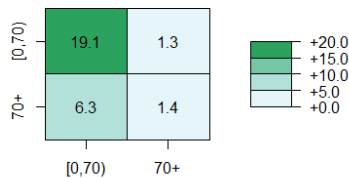


Figure 7: Daily contact matrix: 2 age-groups, Italy

Note. Elaboration on survey data from Mossong et al. (2017). The underlying matrix where each entry is multiplied by the relative demographic size of each age group is made symmetric employing the same demographic shares employed in the model, namely $f_y = 0.825$, $f_o = 1 - f_y$.

C.1 Calibration to the Italian case

The official data on COVID-19 deaths, published by the *Protezione Civile* of the Italian government, are available from February 24, 2020.³³ As noted by Galeotti et al. (2020), the official

³³We used data until May 3, 2020, retrieved from this data repository <https://github.com/pcm-dpc/COVID-19/tree/master/dati-andamento-nazionale> which is updated on a daily basis.

numbers might under report the actual death-toll of the virus in the Italian population. We build a counterfactual of the 2020 total deaths in absence of COVID-19 based on the trend in the preceding five years and obtain an estimate of the COVID-19 effective number of deaths on a weekly basis.³⁴ Throughout the calibration we aim at targeting these epidemiological data and we will consider as goodness of fit whether our model's simulation series will lay in-between those two data curves in the lockdown phase.

The first step of the calibration we did is to set the value of R_0 , the basic reproduction number, in the purely epidemiological ("SIR-age") model to a value consistent with the Italian case. While there is a notorious uncertainty surrounding this statistic, in an early study based on Lombardy (the Northern Italian region most hit by COVID-19) [Cereda et al. \(2020\)](#) report a value of 3.1 (95% CI, 2.9 to 3.2), which is the value we assume. Applying equation 3.1, this value of the reproduction number implies $\eta = 0.01129$. We depart from ERT's strategy of employing COVID-19 mortality rates on reported cases in South Korea, turning instead to the estimates of the infection fatality rate (IFR) by age for Italy provided by [Rinaldi and Paradisi \(2020\)](#). For the two age classes of our interest, their central estimates – once interacted with the Italian demographic shares – give the following probabilities of dying upon infection:

$$\frac{\pi_{y,d}}{\pi_{y,r} + \pi_{y,d}} = 0.00278 \qquad \frac{\pi_{o,d}}{\pi_{o,r} + \pi_{o,d}} = 0.06274$$

Given their reported estimates, we consider that the elderly are those aged 71 or more (71+) while the rest of the population figures as young. The initial demographic shares, using data from the [United Nations World Population Prospects 2019](#), are $f_y = 0.833$, $f_o = 1 - f_y = 0.167$.

Furthermore, we assume that it takes on average 14 days to either recover or die from the infection (cf. [Eichenbaum et al. \(2020a\)](#)). Hence, given that our model is calibrated at the weekly frequency, we set $\pi_{y,r} + \pi_{y,d} = \pi_{y,r} + \pi_{y,d} = \gamma = 7/14$.

Given the same main parameter values of section 3.1, we run the SIR-age model finding the following limiting values:

$$\lim_{t \rightarrow \infty} \{R_{y,t} + D_{y,t}\} = 0.8189 \qquad \lim_{t \rightarrow \infty} \{R_{o,t} + D_{o,t}\} = 0.1214$$

that we used as new values in the SIR-age macro model to calibrate the π_{\dots} . To do so we had to choose the size of the initial shock to the total amount of infected. We found that a value of $\varepsilon = 0.0000225$ gave a number of total deaths in the SIR-age macro model consistent with the value of total effective deaths in the first and second weeks observed in the data (see Figure 8). This amounts to assume that in the first week of 2020 about 1360 people were infected.

In the SIR-age macro model we always set the vaccination probability δ_v to $1/52$ which implies that it takes on average 52 weeks (i.e. 1 year) for the vaccine to become available.³⁵

³⁴The methodology exploits the data released by the [Italian Office for National Statistics \(ISTAT\)](#), available at <https://www.istat.it/it/archivio/240401> - the specific dataset used is "Dataset analitico con i decessi giornalieri", which comprises the daily total deaths (by any cause) in 2020 until April 4 in the 1689 Italian municipalities most affected by COVID-19. Contrary to [Galeotti et al. \(2020\)](#), we also impute a value to the effective number of deaths due to COVID-19 beyond April 4 (until May 3). We do so applying the last observed 'bias factor' (the ratio of the stock of COVID-19 effective deaths over the official ones) which is 1.48. We find that this statistic was very high, at 25, in the first week of official data release (February 23 to 29) and then decreased progressively: 8.6, 3.9, 2.45, 1.86 and 1.48 (in the last week from March 29 to April 4).

³⁵We do so because otherwise it would be hard to justify such a strict observed lockdown in Phase 1 (if not for other reasons such as the overrunning of the hospitalization capacity which is a factor that we do not consider in the

Figure 8 shows both the effective and the official data series for the total deaths caused by COVID-19 in Italy in the weeks corresponding to the period between February 24 and May 3, 2020, as percent of the initial population (according to [United Nations World Population Prospects 2019](#) the Italian population stood at 60.55 million people in 2019).

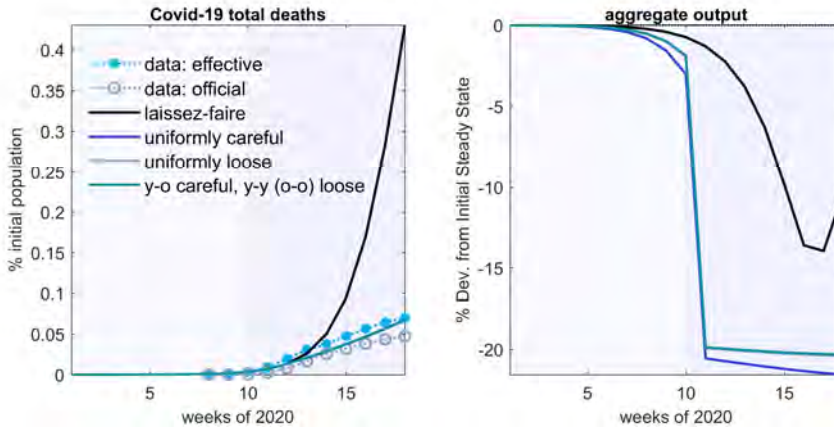


Figure 8: The lockdown phase calibration: data vs model

D Additional figures

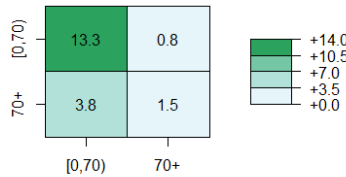


Figure 9: Daily contact Matrix: 2 age-groups, full sample

Note. Elaboration on survey data from [Mosson et al. \(2017\)](#). Full sample comprises: Italy, Germany, Luxembourg, Netherlands, Poland, United Kingdom, Finland, Belgium. The underlying matrix (where each entry is multiplied by the relative demographic size of each age group) is made symmetric employing the same demographic shares employed in the main text, namely $f_y = 0.825$, $f_o = 1 - f_y$.

current analysis). As also found by ERT, and as we have documented in the previous section, this assumption leaves essentially unaltered the resulting *laissez-faire* equilibrium outcomes. It matters, however, for the implied optimal economic shutdown.

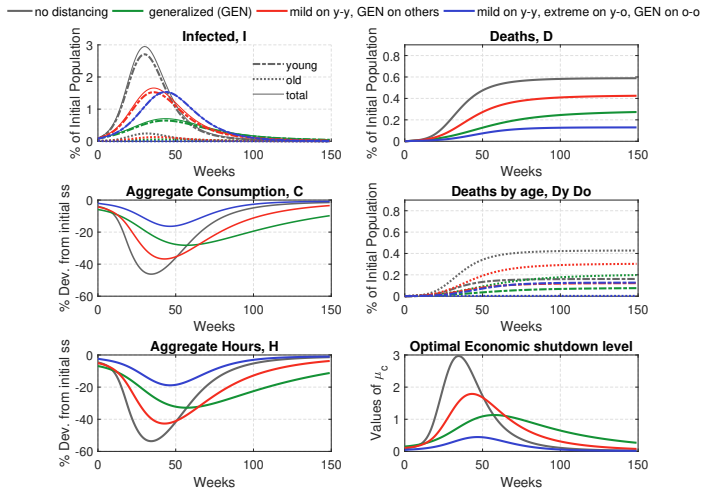


Figure 10: SIR-age macro model: Optimal shutdown with social distancing

Note. The optimal economic shutdown parameter μ_c is computed in the case in which a vaccine is never discovered for different social distancing scenarios. The scenario “generalized (GEN)” assumes that each social contact is diminished by 20% with respect to the baseline (“no distancing”). The scenario “mild on y-y, GEN on others” is the same as GEN but the social distancing among young individuals is 50% less severe than in GEN. The scenario “mild on y-y, extreme on y-o, GEN on o-o” is the same as the previous scenario but assumes that no contact is allowed between young and old individuals.

Trade-off between health and wealth? Insights from COVID-19 in South Korea¹

Changyong Kim,² Estefania Santacreu-Vasut³ and Edward Keunuk Shin⁴

Date submitted: 12 November 2020; Date accepted: 14 November 2020

Understanding the immediate consequences of the COVID-19 pandemic on consumer behaviour is essential for informing the policy makers on the economic cost of strict measures, such as population lockdowns and business shutdowns. Yet, estimating the effect of the health shock on consumption, net of policy restrictions, is challenging because such measures affect consumer choices. South Korea is an interesting case because its policy response in the early stages of the pandemic did not involve such restrictive measures. We exploit this fact to study the consequences of the health shock on consumption. Because the intensity of the pandemic varied greatly across administrative regions, we are able to quantify the direct effect of the health shock on consumption at the epicentre of the pandemic and to compare it with that in locations initially spared from the virus. Further, we quantify spillover effects from the epicentre to the periphery by studying changes in consumption outside of the epicentre. Our results show that consumers adjusted their response as a function of the local and national evolution of the pandemic, refraining from exposing themselves to the health risk in cities and sectors that are relatively more exposed to the virus. This implies that consumers' voluntary response to the pandemic can contribute to alleviate the trade-off between health and economic objectives, minimising the economic cost and mitigating the spread of the virus.

- 1 We are grateful to Daniel L. Hicks for helpful comments; to Sean Brame and Sam Nelberto Nicholas Quinto for their excellent research assistance. All errors remain our own.
- 2 Master's Student at Department of Economics, Seoul National University.
- 3 Associate Professor at Department of Economics, ESSEC Business School and THEMA.
- 4 Master's Student at Graduate School of International Studies, Seoul National University, and at Department of Economics, ESSEC Business School.

Copyright: Changyong Kim, Estefania Santacreu-Vasut and Edward Keunuk Shin

1. INTRODUCTION

On March 12th 2020, the World Health Organisation declared COVID-19 a pandemic. The management of the health crisis and of its resulting economic crisis have varied greatly across countries, partly due to different policy responses. Coercive policy choices involving forced lock-down and shutdown of non-essential businesses have been regarded as having detrimental consequences on consumption above and beyond the direct consequences of the health crisis. An alternative theory is that, regardless of the policy choice, consumption has fallen because consumers avoid in person interactions that are risky from a health perspective. Debate over whether a trade-off between economic and health policy objectives exists would be better informed by an understanding of these effects.

In this paper, we provide one piece of evidence from the first wave of the pandemic in South Korea that can shed light on the nature of the trade-off between health and wealth in the COVID-19 pandemic. South Korea is an interesting case because the initial policy measures adopted by the government did not legally restrict citizens mobility and business operations. Furthermore, cases were heavily concentrated at the outbreak of the pandemic in the metropolitan city of Daegu, leading to substantial variation in exposure to the health shock.

We use monthly credit card data at the level of metropolitan cities and provinces to study changes in consumption patterns following the outbreak of COVID-19 in South Korea. We exploit local variation in the intensity of the pandemic to measure the local and spillover effects of the health shock. In particular, we first focus on comparing the magnitude of the decline in consumption in the metropolitan city of Daegu compared with the decline in all other locations. We then study the extent to which the decline in locations other than Daegu is the result of local cases or of spillover cases.

We find significant differences in the contraction of consumption as a function of the relative exposure to the virus in different locations. In March, the fall in total consumption in Daegu, the epicentre of the pandemic, was one and a half larger than in Busan, an otherwise comparable city that was relatively spared by the virus. When studying the impact of the pandemic outside of the epicentre, our analysis suggests that contemporaneously, half of the decline in consumption is attributable to new local cases while half is attributable to new cases in other regions (spillover) including the epicentre of the pandemic. Interestingly, consumers downward adjustment is larger in consumption categories with higher exposure to the health risk. Finally, when using the number of new cases (local and spillover) in the previous month, we find that lagged local cases have the strongest impact on consumers' behaviour.

Our findings imply that in a context that did not implement restrictive measures such as the South Korean, consumers adjusted their consumption response to different local exposure of the health risk and selectively reallocated consumption away from categories with higher health risk. Our analysis of consumer behaviour outside the epicentre also shows that consumers reacted to the progression of the pandemic outside of their city and locally with anticipation. Overall, these findings suggest that South Korean consumers voluntary response to the pandemic allowed the country to minimise its economic cost and, at the same time, contributed to mitigate the propagation of the virus without the need of restrictive measures. The implication of this result for the debate regarding the pursue of health or economic objectives is that the voluntary response of well-informed citizens can contribute to minimise the size of the trade-off.

The paper is organised as follows. The next section presents stylised facts about the COVID-19 crisis in South Korea. We contextualise these facts through an international comparison and a discussion of related literature. Section 3 presents a chronology of policy measures adopted by the South Korean government. Section 4 presents the research design, section 5 data and section 6 the results. Section 7 presents extensions

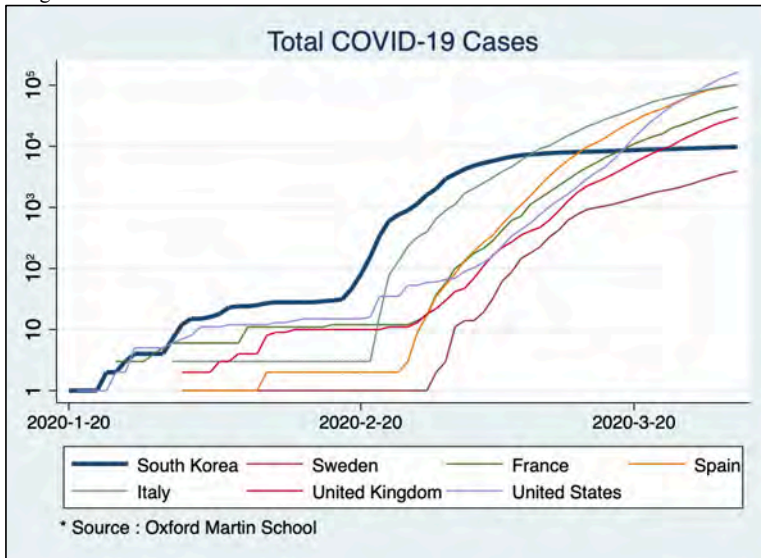
and robustness checks which may be skipped without losing the main message of the paper. We conclude by discussing policy implications.

2. STYLISED FACTS

2.1. International comparison and related literature

The first case of COVID-19 in South Korea was announced on January 20th 2020. Figure 1 plots total COVID-19 cases on a logarithmic scale for South Korea, Sweden, France, Spain, Italy, the UK and the US. While South Korea experienced an acceleration of COVID-19 on February 18th 2020, with a higher growth rate than the other countries, by mid-march, the country had managed to “flatten the curve”.

Figure 1. Total COVID-19 cases



The economic crisis triggered by the health crisis combines both supply and demand shocks, making it a unique crisis (Baldwin, 2020). At the macroeconomic level, the immediate consequences on global demand have been felt mostly in the decline of private consumption, which has been more impacted than private investment has, government spending, or trade. For a comprehensive literature review on COVID-19 overall economic consequences see, for example, Brodeur et al. (2020).

Figure 2 presents Bank of Korea (2020) data showing the percentage deviation of consumption in 2020Q1 and 2020Q2 compared to 2019Q4. In South Korea, consumption dropped by 6 % in 2020Q1 compared to 2019Q4. In 2020Q2 South Korea experienced a smaller drop in consumption vis-à-vis the previous quarter and vis-à-vis other countries.

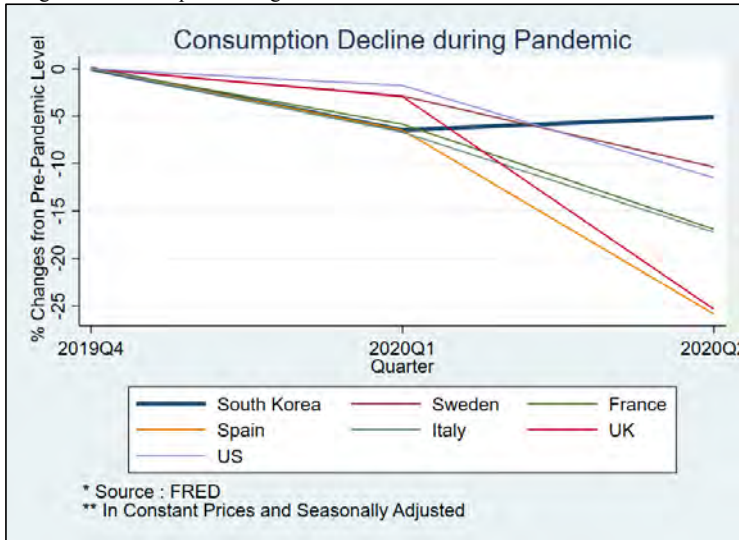
The policy responses to the COVID-19 outbreak adopted by governments varied greatly across countries. We next review these responses together with a discussion of related literature.

At one extreme of the policy spectrum, China implemented a strictly enforced lock-down. Chen, Qian and Wen (2020) study the impact of COVID-19 and policy response (with a focus on the before-after the lock-

Covid Economics 58, 19 November 2020: 57-84

down) taken by the Chinese government on offline consumption only. They use credit card records to study the drop in consumption and the changes in consumption patterns in China. They find that following that date, the sharpest decline concerned services, which dropped by 72 %. They argue that consumer behaviour changed as a result of the closing of non-essential service businesses and related restrictions.

Figure 2. Consumption change across countries.



Temporary lock-down measures were also adopted in continental Europe. For instance, France started official containment on March 17th 2020. Bounie, Camara and Galbraith (2020) use French transaction data to study consumers' response to COVID-19. They document a decline in offline consumption of 60 % and a decline in online consumption of 30 % after the official containment starting on March 17th 2020. They also show sectorial evidence that in some sectors (such as supermarket or clothing) online consumption increased, partly offsetting the decline in offline consumption.

In Spain, official national lock-down started on March 14th 2020. Carvalho et al. (2020a,b) study consumer before and after that date. Using individual transaction data, they document significant reallocation of consumption away from non-essential goods and towards essential goods and goods with low demand elasticity such as tobacco. Interestingly, although Spanish regions were heterogeneously impacted by the virus, they find no significant regional differences in consumption patterns. Indeed, changes in consumption seem related exclusively to the lock-down policy rather than to the exposure to the virus. At the city level, they do find evidence of incidence of the virus exposure on consumption patterns within Madrid.

In the UK, Chronopoulos, Lukas and Wilson (2020) show that declines in consumption took place in the first weeks of lockdown. They also show evidence of stockpiling behaviour as spending on groceries increased before the lockdown.

In Denmark, partial shutdown of the economy started on March 11th 2020. Andersen, Hansen, Johannesen and Sheridan (2020) use transaction-level individual data to estimate the changes in spending driven by the pandemic and by the shutdown policies. Interestingly, they document that in Denmark the fall in consumption is concentrated in sectors directly affected by the shutdown such as services, with little spillover effects to unrestricted sectors. They classify consumption categories as a function of how they were affected by

shutdown policies (closed, constrained or open). Interestingly, consumption did not change much before restrictions were put in place.

In the US, Baker, Farrokhnia, Meyer, Pagel and Yannelis (2020) also uses transaction-level household data to analyse US households and find evidence consistent with stockpiling behaviour and a decline in the consumption of services.

At the other extreme, countries like Sweden or South Korea have adopted less coercive measures, recommending social distance and not imposing any temporary business shut-downs but tracing contagion patterns closely and providing up-to-date information. In Sweden, the government did not impose restrictions on either businesses or consumers, avoiding stringent lockdown measures and restrictions on economic activity. Krueger, Uhlig and Xie (2020) study consumption shifts across sectors of the Swedish economy. Redirecting consumption from one sector to another could, according to the authors, be the key to mitigating epidemics. We next turn to the discussion of stylised facts of the COVID-19 crisis within South Korea.

2.2. Variation within South Korea

When the World Health Organization declared COVID-19 a pandemic, South Korea counted 7869 confirmed cases and 66 deaths. The spread of COVID-19 in South Korea accelerated in Daegu and relates to the 31st confirmed case. This individual, also identified as the super-spreader, tested positive on February 18th. It is interesting to understand what happened in the weeks before the test, thanks to the publicly available records. Back on February 6th this individual had a car accident in DongGu of Daegu, which led to a visit to Seronan Hospital. During that visit, the patient developed fever but left the hospital without being tested. The record shows that in the following days, this same individual visited Shincheonji Daegu Church at least twice (February 09th and February 16th). The sum of people visiting the Church during that period is about 1000. On February 10th the authorities advised the individual to be tested, but the individual refused. On February 15th the individual participated to a wedding ceremony in Queenbell Hotel in Dongu. Furthermore, authorities learned that the individual had concealed information regarding the whereabouts during the day of February 5th. When the individual was finally tested positive, the spread of the virus in Daegu had accelerated. Indeed, out of the subsequent 14 cases, at least 10 were from the same Shincheonji Daegu Church. This case sparked legislative action to sanction individuals who refused testing.

The metropolitan city of Daegu was, therefore, severely hit by COVID-19 both in terms of confirmed cases and deaths. Figure 3 maps the number of confirmed cases as of March 1st 2020. Figure 4 shows the evolution of new cases from January 2020 to June 2020 by region breakout. As both figures show, Daegu became the epicentre of the pandemic in February and March 2020.

One of the key aspects of the policy response to COVID-19 in South Korea has been the testing, tracking, and disclosing of information about individuals who tested positive, including the whereabouts of these individuals in the days before being tested positive. Doing so allows Korean residents to be tested if they have visited places concurrently visited by confirmed cases, and also to selectively avoid certain businesses during a particular period without imposing business shutdown. Thanks to this policy, individuals could trace the origin and nature of the massive outbreak that originated in the metropolitan city of Daegu. For example, the following text message was sent by the authorities to the population: “[YongIn city hall] 17th confirmed person occurrence (female/26 years old, Giheung-gu Shingal-dong resident), traffic line to be made publicly available after investigation on SNS and home page”. (*authors’ translation*)

Figure 3. Map of confirmed cases as of March 1st 2020 in South Korea.

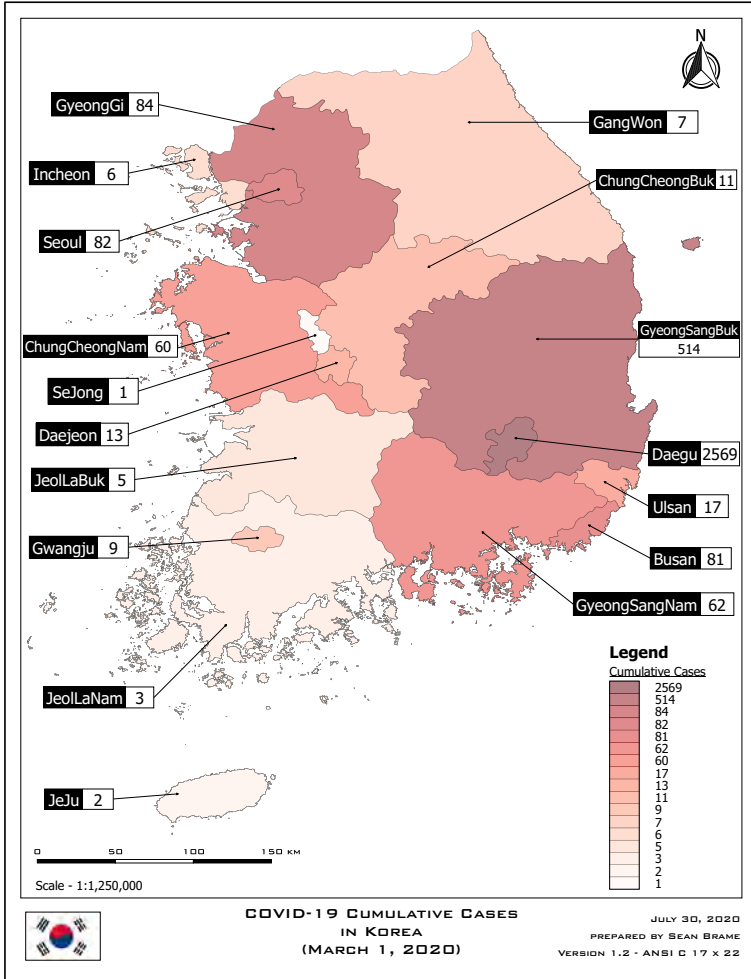
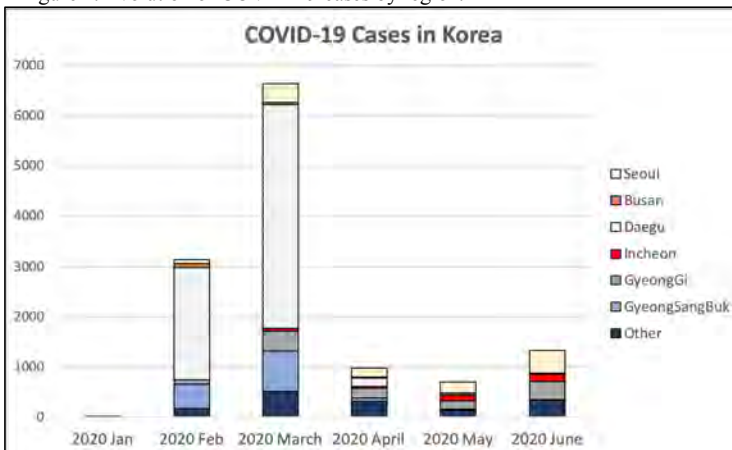


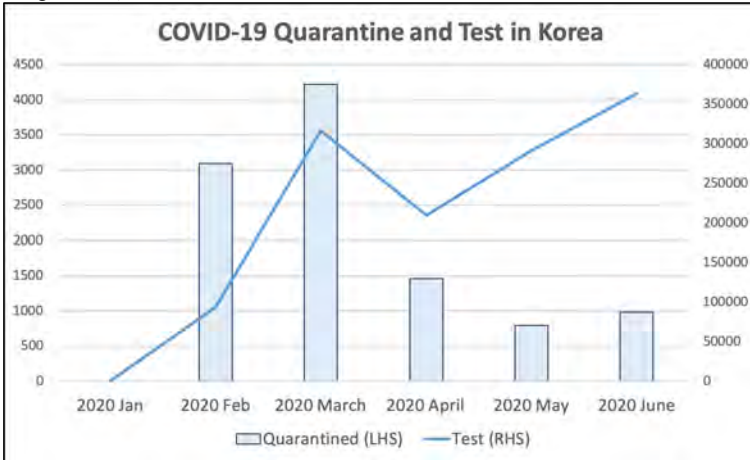
Figure 4. Evolution of COVID-19 cases by region.



Covid Economics 58, 19 November 2020: 57-84

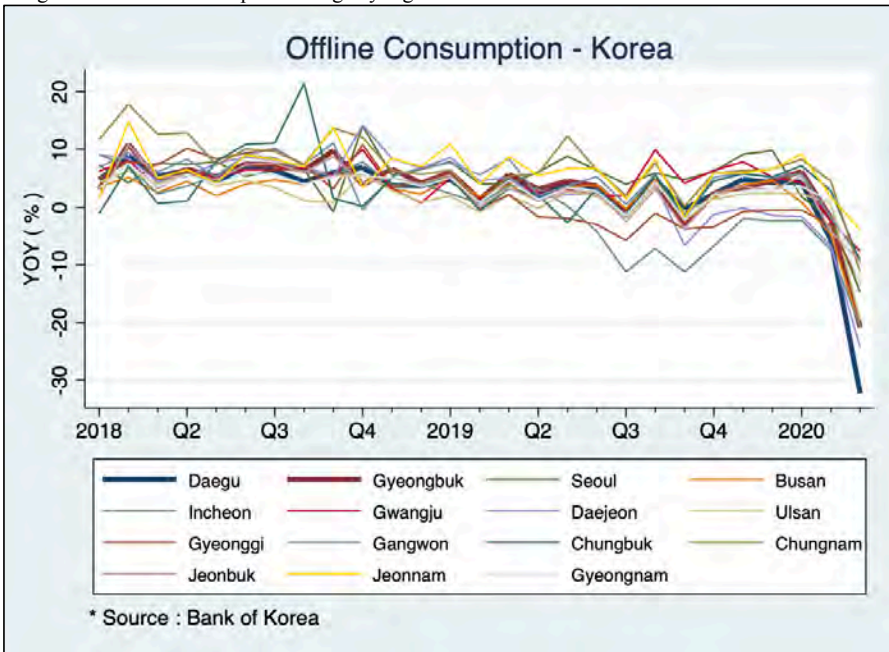
Figure 5 presents data on the number of monthly tests and quarantined individuals (among individuals tested positive). Tests per capita were among the world's highest during the initial stages of the pandemic.

Figure 5. Quarantine and tests in South Korea



Mirroring the heterogeneous health impact of the pandemic across cities in South Korea, the negative impact on consumption was also highly heterogeneous. Figure 6 plots offline consumption from January 2018 to March 2020 for the 15 regions (1 capital city, 6 metropolitan cities and 8 provinces). As the data shows, Daegu, the epicentre of the pandemic experienced the sharpest decline in offline consumption in February and March.

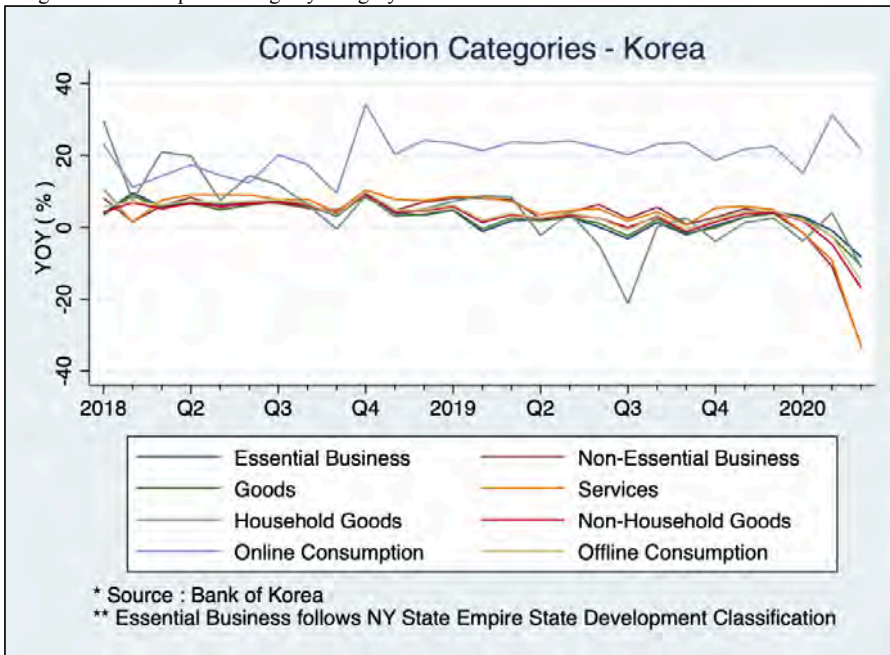
Figure 6. Offline consumption change by region



Covid Economics 58, 19 November 2020: 57-84

The heterogeneity of the consumers' response also concerned consumption categories. Figure 7 plots consumption from January 2018 to March 2020 for offline and online consumption. Regarding offline consumption, the distinction is also made between the following categories: Essential businesses, Non-Essential businesses, Services, Goods, Household goods, Non-Household goods. Similar to the experience documented in other countries, Non-Essential business consumption contracted the most, followed by Services and Non-Household goods. The weakest contraction concerned Essential businesses. Online consumption experienced an increase particularly in February.

Figure 7. Consumption change by category

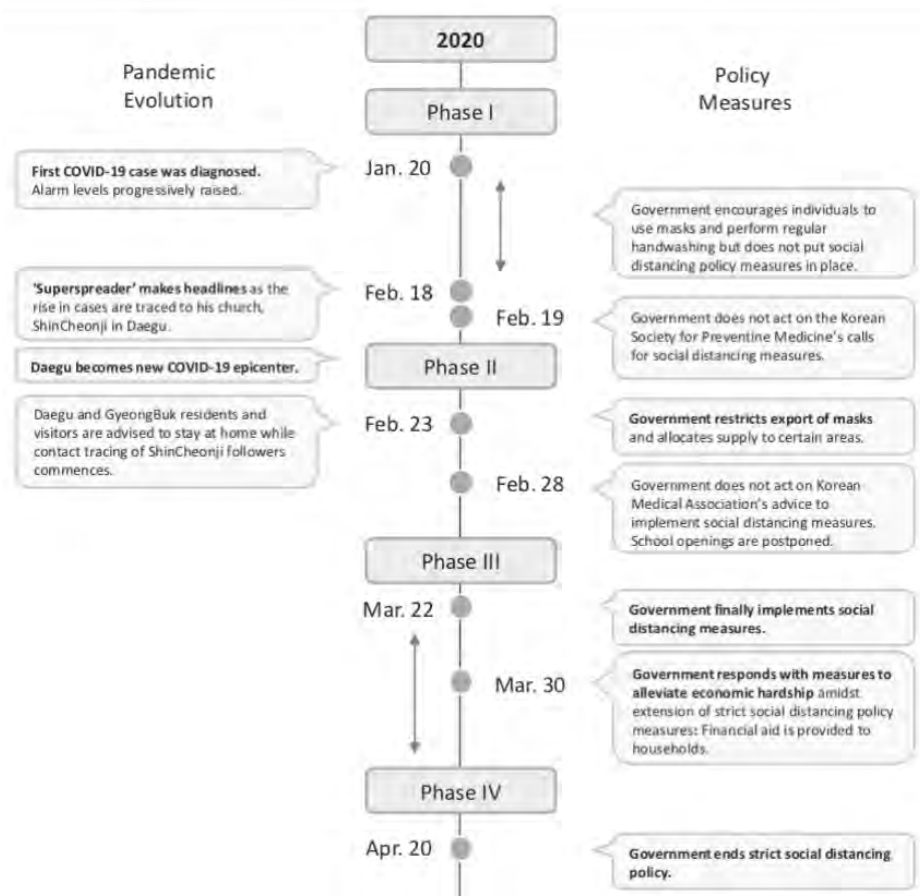


This paper exploits the heterogeneity in consumer response across space and across consumption categories in a policy context that did not impose restrictive measures. The next section describes the chronology of policy measures taken in South Korea.

3. CHRONOLOGY OF POLICY MEASURES IN SOUTH KOREA

The policy measures implemented by the health and public authorities in response to COVID-19 can be summarised by distinguishing four phases, as presented in Figure 8. The initial phase (phase I) starts on January 20th when the first COVID-19 case was diagnosed and lasts until February 23rd. This phase is characterised by a progressive raising of the alert levels and by a government stance that is based on advising individuals to use masks and regular hand-washing. On February 18th 2020, a total of 31 patients were diagnosed with COVID-19 in Korea. While the Korean Society for Preventive Medicine did mention "social distancing" on February 19th the government remained silent about it during this phase.

Figure 8. Chronology of COVID-19 in South Korea (January 2020 to April 2020)



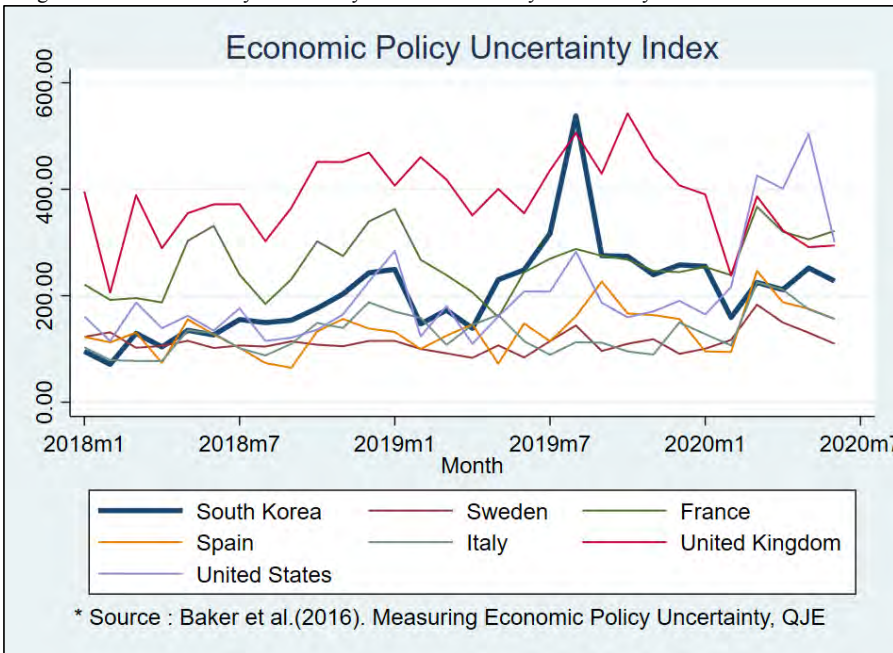
From February 23rd to March 22nd (phase II) the alert level was characterised as serious and the government took a more proactive attitude including advising Daegu and Gyeongbuk visitors to stay at home and get tested if they had any symptoms, the shutdown of schools and the tracing of ShinCheonJi Church followers. During the last week of February, the government also decided to restrict the export of masks and manage some of its allocation selectively. And on February 28th the Korean Medical Association advised the government to begin social distancing, but this was not adopted during this phase. The opening of schools was postponed.

March 22nd marks the start of a new phase (phase III) characterised by stronger measures. In particular, the government decided to implement a social distancing policy. This phase lasted until April 19th. Contact-intensive business such as karaoke bars, night-clubs, religious facilities and gyms were advised to be temporarily shut-down, but could continue operating conditional on following the government guidelines. These included to keep a minimum distance, to wear a mask, to check whether a visitor has a fever or some respiratory symptom at the entrance and to sterilise the facility at least once every 2 days. The government also adopted economic measures to alleviate the economic hardship through an assistance fund announced on March 30th. On April 4th the strong distancing policy was extended, and the government continued its recommendation to religious, athletic and entertainment activities to shut down. Furthermore, subway running hours were reduced. April 19th marks the end of this phase as it is the date when the government decided to end strong social distancing.

From April 20th to May 06th which we define as phase IV, the government changed the recommendation for the contact-intensive business from advising them to shut down, to voluntarily refrain their activities. Schools progressively reopened in May.

In this paper, we focus our analysis on the initial phases of the pandemic (Phase I and Phase II) when no social distancing measures were taken. Governmental recommendations towards social distancing started in late March. It is also important to note that the initial policy measures during our period of analysis were taken at the national level. Only starting from May, regional differences in the policy response have appeared in the way night bars, clubs, and entertaining rooms were either prohibited or recommended to shut down. Furthermore, economic policy uncertainty in South Korea was not particularly high during the period. Figure 9 plots the economic policy uncertainty (EPU) index from Baker et al. (2016) for South Korea and 6 other countries from January 2018 to July 2020. In South Korea, COVID-19 led to an increase in EPU from February onwards albeit smaller than in the United States or in the United Kingdom. Furthermore, the spike in uncertainty is smaller compared to that experienced in the country around July 2019 and related to the US-China trade war tensions.

Figure 9. Economic Policy Uncertainty Index from January 2018 to July 2020.



Last but not least, it is important to note that the communication strategy followed in South Korea was in line with best practices. Consider, for instance, existing guidelines, such as the Crisis and Emergency Risk Communication (CERC) written by the US Center for Disease Control. Two of the key pillars of a good communication strategy during a health crisis were followed in South Korea. First, a single person familiar with the health crisis acts as spokesperson and second, this person is a credible and trustworthy (preferably not a politician). In South Korea, the spokesperson was the Dr. Eun Kyeong JEONG, Commissioner of Korea Disease Control and Prevention Agency.

Covid Economics 58, 19 November 2020: 57-84

4. RESEARCH DESIGN

To shed light on the potential trade-off between economic and health objectives using the experience of South Korea, we follow two steps.

Step 1: Comparison of the centre of the pandemic with rest of the country

In this step, our goal is to estimate the effect of the COVID-19 shock on consumption by comparing outcomes in the city of Daegu (epicentre of the pandemic) from the outcomes in similar cities that were relatively spared from the pandemic, before and after the shock. Formally, we use a difference-in-differences (DiD) method which is based on a comparison of before-after and treatment-control. This method involves:

- to compute the difference in consumption before and after the start of the pandemic in the city of Daegu (epicentre of the pandemic), our treatment group.
- to compute the difference in consumption before and after the start of the pandemic in a control group (for instance in metropolitan cities that were relatively spared from the virus).
- to compute the difference between these two differences (double difference).

As discussed in the previous sections, we are interested in studying the immediate consumption changes post COVID-19. In particular, our focus is to study changes during phases I and II of the pandemic (see the chronology in Figure 8) which are phases when the government did not implement strict social distancing measures. Because these two phases span the months of February and March (from February 18th to March 22nd) and since we have monthly level data, we consider both February and March in our DiD. While this could be seen as a limitation of our analysis, an alternative interpretation is that February captures the initial shock of the pandemic while March captures its acceleration (see Figure 4).

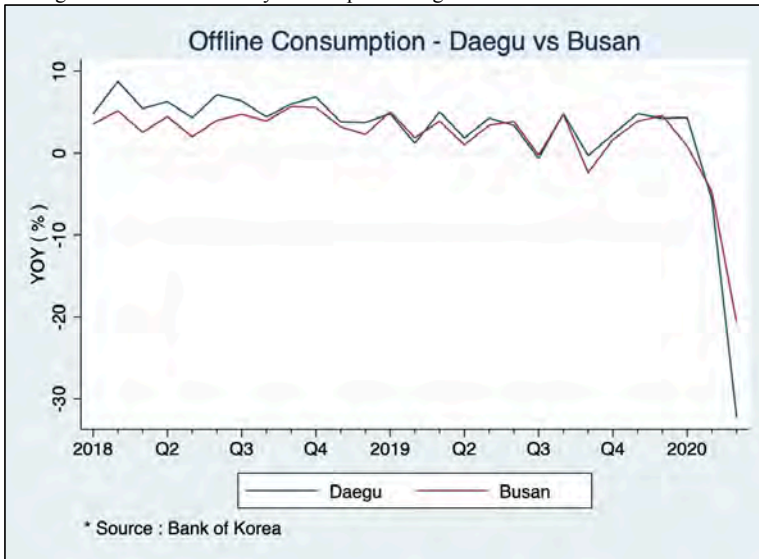
A key challenge in the design of a DiD is finding a proper comparison group (control group). Indeed, identification of a causal effect of the COVID-19 on consumption based on a DiD method relies on assuming that, absent COVID-19, Daegu would have followed the same time trend as its comparison group, an assumption formally known as the parallel trends assumption.

We argue that the metropolitan city of Busan is an ideal candidate for this comparison. Busan is the second largest city of Korea and Daegu is the fourth largest city. Per capita income (in real terms) as of 2018 is equal to 25,495 thousand KRW in Busan, slightly above that of Daegu which is equal to 21,993 thousand KRW. In addition, both cities have similar industrial structure, dominated by manufacturing which accounts for 23 % of GDP in Daegu and 18 % of GDP in Busan, followed by the Real Estate and Wholesale and Retail Sectors. Both cities have similar population densities. Importantly, both cities have similar medical capacities by several standards. For instance, both cities count 1,5 hospitals per thousand inhabitants. Daegu counts 0.40 physicians per thousand and 0.765 related doctors per thousand. Physicians were formally called by the minister of Health and Wealth (Neunghu-Park) to voluntarily work to battle of COVID-19. Busan counts 0.38 physicians per thousand and 0.725 related doctors per thousand.

Figure 10 presents the pairwise comparison of offline consumption year-on-year changes between Daegu and Busan from January 2018 to March 2020. Visual inspection of the figure suggests that consumption behaved in a very similar fashion in the two cities before COVID-19. Formally, this means that the assumption of parallel trends prior to COVID-19 is not rejected. In addition to the comparison of aggregate monthly consumption, we also study the reallocation of consumption. In particular, we also analyse the DiD for each of the following consumption categories between the two cities: Essential, Non-Essential, Services, Goods, Household and Non-Household goods, and Online consumption. For completeness, we also perform the DiD analysis between Daegu and each of the other metropolitan cities in the country.

As Figure 10 shows, the year-on-year change in offline consumption in February was twice as large in Daegu than in Busan and in March it was one and a half larger in Daegu than in Busan.

Figure 10. Offline monthly consumption Daegu versus Busan



Our sample covers the period from January 2018 until March 2020. This allows us to include a certain number of pre-COVID-19 periods, following best practice in DiD studies. Another assumption of DiD analysis is that there is no spillover effect between the treatment group and the control group, which is known as the Stable Unit Treatment Value. We devote the second part of our analysis to the study of spillover effects as presented next in step 2 of our research design.

Step 2: Quantifying the spillover effect from the epicentre to other locations.

A second step in our research design is to consider the possible spillover effects of the health shock from the epicentre to other locations. That is, even if relatively spared by the virus, consumers' behaviour in metropolitan cities near Daegu may have adjusted their behaviour. Indeed, as Figure 10 shows Busan experienced a year-on-year negative change of consumption equal to 20 % by March 2020. Because Busan is near Daegu (88 km), one possible explanation for this non-negligible decline in consumption in March 2020 is that the virus propagated to nearby cities (but with less severity) and therefore that it had a weaker but significant effect on consumers' behaviour. The DiD estimate would then capture the relative impact of the health shocks on the two cities. Figure 11 plots the correlation between average new cases in metropolitan cities and provinces and distance from the epicentre, Daegu. As the figure shows, distance from the epicentre does not correlate with the spread of the virus. Similar results are obtained when using average new deaths and cumulative cases.

An alternative explanation is that the decline of consumption in Busan reflects spillover effects from cases outside of Busan (and prominently from the epicentre). That is, consumer behavioural changes in Busan may not be directly related to the propagation of the virus in their own city but rather to precautionary or anticipatory behaviour in reaction to the propagation of the virus in other locations. Figure 12 plots the correlation between consumption growth in metropolitan cities and provinces and distance from Daegu for the

period of December 2019 to March 2020. As the figure shows, metropolitan cities closer to Daegu saw a sharper decline in consumption, suggesting the interest in further studying the spillover effects, which we do in our empirical analysis. Figure 12 also includes the correlation before the outbreak, showing that there is no relationship between distance from Daegu and the consumption pattern before the pandemic.

Figure 11. Average new cases and distance from Daegu.

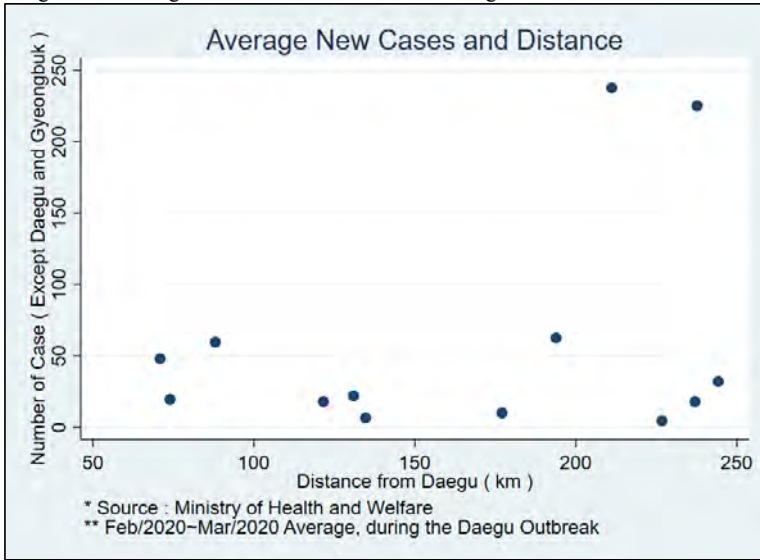
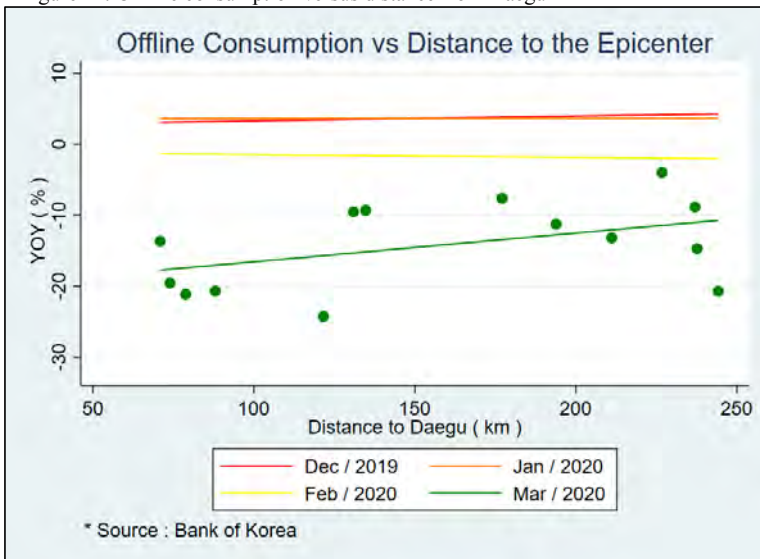
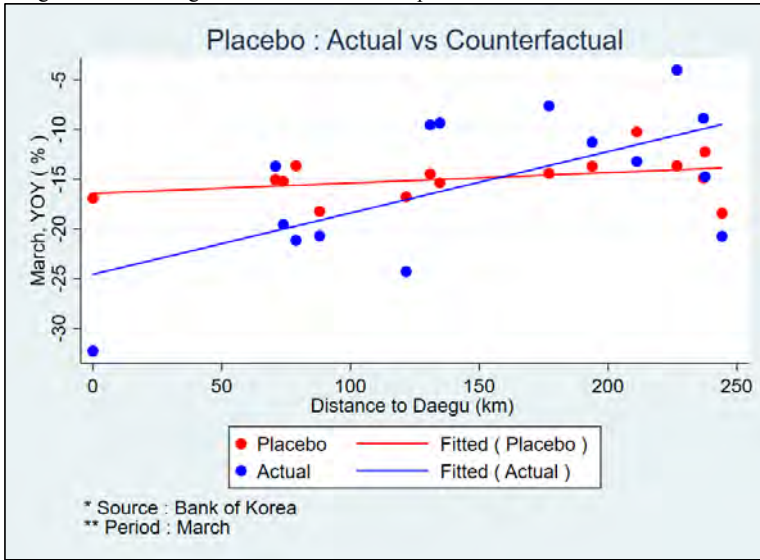


Figure 12. Offline consumption versus distance from Daegu



A potential concern arises if businesses vulnerable to COVID-19 are concentrated near the epicentre. To rule out this possibility, we create a counterfactual assuming that each sector grew at the nationwide average growth rate, and recalculate the regional growth rate of total monthly expenditure. As Figure 13 shows, differential exposure to sectors in regions closer to the epicentre does not explain the pattern of consumption and distance from Daegu.

Figure 13. Checking the role of sectorial composition



Overall, the preliminary evidence in figures (11)-(13) could suggest that the lack of relationship between the spread of the virus and distance from the epicentre, on the one hand, and the relationship between the decline in consumption and distance from the epicentre, on the other hand, are not attributed to luck or coincidence but rather to the voluntarily cautious behaviour of consumers.

To formally investigate the role of spillover effects we build a measure of spillovers called “Spillover” that is based on the weighted average number of cases outside of a given city or province. The weights are based on the relative commute intensity between a location and all the other locations as predicted by distance between locations and measured using total commute of individuals older than 12 years old between location pairs (regardless of the medium of transportation).

Using this measure, we then estimate the change in year-on-year consumption at the aggregate level (and by category) on the new cases at the local level and on the spillover cases. We also control for the level of uncertainty using the Economic Policy Uncertainty measure to capture a time-varying common uncertainty shock, as well as a location fixed effect. Because consumer behaviour can influence the spread of the pandemic, we also analyse the results using the lagged measures (by one month) of new local cases and spillover cases instead of the contemporaneous ones.

5. DATA

We include data on the capital city and 6 metropolitan cities and 8 provinces. These are the following: Seoul, Busan, Daegu, Incheon, Gwangju, Daejeon, Ulsan, Gyeongbuk, Gyeonggi, Gangwon, Chungbuk, Chungnam, Jeonbuk, Jeonnam, Gyeongnam. We study the period from January 2018 to March 2020 for the DiD analysis and the period from December 2019 to March 2020 for the spillover analysis.

5.1. COVID-19 data

We use monthly data on COVID-19 confirmed cases which we acquired from a site dedicated to the COVID-19 pandemic from the Korea Center for Disease Control and Prevention, Ministry of Health and Welfare, ranging from 2020 January to June.

5.2. Expenditure data

We use store-level data obtained from personal credit card payments, measured in million KRW and collected from the Economic Statistical System of the Bank of Korea. Using this data we calculate year-on-year growth to take account of the magnitude differences and seasonality effects.

The data readily allows us to distinguish online from offline expenditures as well as expenditures on goods versus expenditures on services. To further study changes in consumption patterns, including the reallocation of expenditure away from Non-Essential businesses, we follow the New York State guidelines to split our expenditure data between Essential and Non-Essential business. We also split our data between Household and Non-Household goods. The classification used is available in the Appendix.

5.3. Other data

To capture economic policy uncertainty, we use the Economic Policy Uncertainty measure from Baker et al. (2016).

The shortest distance on a sphere between each region was calculated based on the Google Maps coordinates of the city and province administrative buildings. This is consistent with previous literature that incorporates distance in pandemic analysis (Chin and Wilson, 2018; Bounie, Camara and Galbraith, 2020).

Statistics Korea's 2015 population census gives us commute data, based on the location of workplace and educational institute for those over 12 years old, irrespective of the means of transportation.

6. ESTIMATION RESULTS

6.1. DiD results

Table 1 presents the results concerning total, offline and online consumption DiD analysis.

Table 1: DiD estimations between Daegu and Busan

	Total	Offline	Online
Daegu*February	-2.143*** (0.654)	-2.159*** (0.609)	-30.11 (22.91)
Daegu*March	-12.35*** (0.654)	-12.72*** (0.609)	14.73 (22.91)
February	-7.610*** (0.466)	-7.862*** (0.403)	34.89 (22.67)
March	-23.74*** (0.466)	-23.87*** (0.403)	-12.11 (22.67)
Daegu	1.172* (0.654)	1.151* (0.609)	-14.06 (22.91)
Constant	3.205*** (0.466)	3.170*** (0.403)	30.78 (22.67)
Observations	54	54	54
R-squared	0.889	0.905	0.014

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is year-on-year change (%) of total, offline and online personal monthly credit card expenditure. The criteria for classification into disaggregated expenditure categories are available in the appendix. 'Daegu' is a dummy variable indicating Daegu (treatment group). 'Feb' and 'Mar' are dummy variables indicating the month of February and March 2020 respectively. To estimate the DiD we use OLS.

Sources: Bank of Korea.

Regarding total consumption (first column of Table 1), the coefficient for the month of February tells us that the expected mean change in year-on-year in Busan was equal to -7.610 percentage points and strongly significant. The coefficient for the month of March tells us that the expected mean change in year-on-year in Busan was equal to -23.74 percentage points and strongly significant. The difference-in-differences estimator, our main focus, tells us that the difference in the expected mean change in year-on-year consumption in Daegu versus Busan was equal to -2.143 percentage points in February and -12.35 percentage points in March. The results for total consumption are driven by the contraction in offline consumption (second column of Table 1) which is very similar in economic and statistical terms to that in total consumption. Indeed, results are not statistically significant for online consumption. These results are in line with the raw data presented in Figure 7.

Table 2 presents the results concerning offline consumption by category.

Table 2: DiD estimations between Daegu and Busan: Offline (by category)

	Essential	Non-Essential	Goods	Services	Household	Non-Household
Daegu*February	0.726 (0.894)	-5.277*** (0.749)	-2.531*** (0.751)	0.832 (0.840)	-25.50*** (3.052)	-0.880 (0.617)
Daegu*March	-9.914*** (0.894)	-16.49*** (0.749)	-13.43*** (0.751)	-10.29*** (0.840)	-22.45*** (3.052)	-12.48*** (0.617)
February	-1.138** (0.497)	-21.80*** (0.607)	-4.249*** (0.449)	-22.73*** (0.670)	21.09*** (1.695)	-9.934*** (0.412)
March	-11.57*** (0.497)	-46.44*** (0.607)	-16.99*** (0.449)	-47.89*** (0.670)	-0.808 (1.695)	-25.54*** (0.412)
Daegu	1.522* (0.894)	0.382 (0.749)	1.145 (0.751)	0.963 (0.840)	0.0121 (3.052)	1.181* (0.617)
Constant	2.255*** (0.497)	4.218*** (0.607)	2.215*** (0.449)	5.334*** (0.670)	-0.0555 (1.695)	3.080*** (0.412)
Observations	54	54	54	54	54	54
R-squared	0.571	0.956	0.795	0.941	0.163	0.912

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is year-on-year change (%) of offline personal monthly credit card expenditure by category. The criteria for classification into disaggregated expenditure categories are available in the appendix. 'Daegu' is a dummy variable indicating Daegu (treatment group). 'Feb' and 'Mar' are dummy variables indicating the month of February and March 2020 respectively. To estimate the DiD we use OLS.

Sources: Bank of Korea.

Regarding Essential business consumption (first column of Table 2), the coefficient for the month of February tells us that the expected mean change in year-on-year consumption in Busan was equal to -1.138 percentage points and strongly significant. The coefficient for the month of March tells us that the expected mean change in year-on-year consumption in Busan was equal to -11.57 percentage points and strongly significant. The contraction of Non-Essential business consumption (second column of Table 2) is much larger. The expected mean change in year-on-year consumption of Non-Essentials in Busan was equal to -21.80 percentage points and strongly significant. The coefficient for the month of March tells us that the expected mean negative change in year-on-year consumption in Busan was equal to -46.44 percentage points. The difference-in-differences estimator, our main focus, tells us that the difference in the expected mean change in year-on-year consumption in Daegu versus Busan concerning Essential business consumption was positive and equal to 0.726 percentage points in February, but not statistically significant. In March, the difference-in-differences estimator was statistically significant equal to -9.914 percentage points.

Regarding Non-Essential business consumption, the difference-in-differences estimator tells us that the difference in the expected mean change in year-on-year consumption in Daegu versus Busan was equal to -5.2 percentage points in February, and statistically significant. In March, the difference-in-differences estimator was statistically significant equal to -16.49 percentage points. Overall, the results concerning the Essential vs. Non-Essential distinction shows that consumers reallocated their expenditure away from Non-Essential goods. While this is consistent with what has been documented in other countries (see section 2), it is interesting to note that this was a voluntary reallocation as the government in South Korea did not impose the shutdown of Non-Essential business.

Regarding the distinction between Goods and Services, results in Table 2 tell us that the expected mean change in year-on-year in Goods in Busan was equal to -4.249 percentage points and strongly significant. The coefficient for the month of March tells us that the expected mean change in year-on-year consumption in Busan was equal to -16.99 percentage points and strongly significant. As observed in other countries, the contraction was much larger concerning Services. In particular, the expected mean change in year-on-year in Services in Busan in February was equal to -22.73 percentage points and strongly significant. The coefficient for the month of March tells us that the expected mean change in year-on-year in Busan was equal to -47.99

percentage points and strongly significant. The difference-in-differences estimator, our main focus, tells us that the difference in the expected mean change in year-on-year consumption in Daegu versus Busan concerning Goods consumption was equal to -2.531 percentage points in February, and statistically significant. In March, the difference-in-differences estimator was statistically significant equal to -13.43 percentage points.

Regarding the consumption of Services, the difference in the expected mean change in year-on-year consumption in Daegu versus Busan was equal to -10.29 in March. It was not significant and close to zero in February. Overall, the results show that Services were more severely affected in March, and that the contraction was one fifth larger in the epicentre of the pandemic.

Finally, despite the fact that the government did not impose population lockdown, it is interesting to investigate the consumption patterns concerning Household and Non-Household goods. An interesting results stand out: the expected mean change in year-on-year in Busan in the consumption of Household goods in February was positive and equal to 21.09 percentage points and strongly significant. Household goods include electronic devices and furniture. It is possible that consumers in Busan anticipated some of these expenditures as their expectation of time spent in the household may have increased with the pandemic. This pattern is not observed at the epicentre of the pandemic.

Table 3 replicates the results for the DiD analysis between Daegu (as treatment group) and each of the other metropolitan cities for total offline consumption. Results are similar across alternative control groups albeit with varying magnitude of the difference-in-differences estimator. The R-squared is highest concerning our favourite comparison between Daegu and Busan.

Table 3: DiD total consumption (offline) Daegu versus other cities

	Busan	Seoul	Incheon	Gwangju	Daejeon	Ulsan
Daegu*February	-2.159*** (0.609)	-1.781** (0.742)	-1.221 (1.322)	-2.482*** (0.686)	1.048 (1.074)	-8.161*** (0.634)
Daegu*March	-12.72*** (0.609)	-15.59*** (0.742)	-14.44*** (1.322)	-23.94*** (0.686)	-7.958*** (1.074)	-14.62*** (0.634)
February	-7.862*** (0.403)	-8.239*** (0.585)	-8.800*** (1.240)	-7.539*** (0.512)	-11.07*** (0.972)	-1.860*** (0.440)
March	-23.87*** (0.403)	-21.00*** (0.585)	-22.15*** (1.240)	-12.65*** (0.512)	-28.63*** (0.972)	-21.97*** (0.440)
Daegu	1.151* (0.609)	-1.954** (0.742)	2.902** (1.322)	-0.711 (0.686)	-0.0328 (1.074)	1.894*** (0.634)
Constant	3.170*** (0.403)	6.275*** (0.585)	1.419 (1.240)	5.032*** (0.512)	4.353*** (0.972)	2.427*** (0.440)
Observations	54	54	54	54	54	54
R-squared	0.905	0.863	0.669	0.860	0.778	0.892

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is year-on-year change (%) of total personal monthly credit card expenditure. Each column 'Daegu' is a dummy variable indicating Daegu (treatment group). 'Feb' and 'Mar' are dummy variables indicating the month of February and March 2020 respectively. To estimate the DiD we use OLS. The first column presents the DiD between Daegu and Busan as control group. Subsequent columns replicate the DiD estimation using each of the other metropolitan cities as control group. Data spans the period January 2018 to March 2020.

Sources: Bank of Korea.

Overall, the DiD analysis suggests that the contraction of offline consumption was more severe in March than in February, in line with the timeline of the evolution of the pandemic. Furthermore, the contraction was particularly strong in Daegu, the epicentre of the pandemic. The analysis of consumption categories also shows an endogenous shift in consumer behaviour away from Non-Essential goods, and from Services, in a policy context that did not impose either business shutdown nor population lockdown, suggesting that consumers reallocated consumption as a mechanism to minimise the exposure to the virus.

6.2. Spillover results

Table 4 presents our estimation of spillover effects for the sample of metropolitan cities (excluding Daegu). We present the results for total consumption, offline consumption and online consumption. Table 5 present the results for offline consumption by category (Essential, Non-Essential, Services, Goods, Household and Non-Household).

Table 4: Spillover (All metropolitan cities)

	Total	Offline	Online
Local	-0.0211* (0.00973)	-0.0393*** (0.00906)	-0.0524** (0.0135)
Spillover	-0.0401** (0.0102)	-0.0382** (0.00951)	0.0423* (0.0193)
EPU	-0.0115 (0.0247)	0.000448 (0.0265)	-0.647** (0.171)
Constant	5.120 (5.119)	0.238 (5.724)	169.0*** (39.23)
N	24	24	24
R-squared	0.585	0.632	0.706
FE	6	6	6

Notes: Robust standard errors clustered at the city level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is year-on-year change (%) of total, offline and online personal monthly credit card expenditure. Local is the monthly and city level number of new cases. ‘Spillover’ is the monthly and city level number of cases in other cities (weighted by the relative commuting intensity) as described in section 4. EPU is the nationwide monthly level Economic Policy Uncertainty index. We use an OLS linear specification with city fixed effects from December 2019 to March 2020.

Sources: Bank of Korea for the expenditure data, Korea Center for Disease Control and Prevention, Ministry of Health and Welfare for the COVID-19 monthly new cases, Baker et al. (2016) for the EPU data, and Google Maps and Statistics Korea for the construction of the ‘Spillover’ measure.

As Table 4 shows, the local number of new cases has a negative impact on all consumption categories, including online. Coefficients for spillover new cases are also negative, except for the positive impact on online consumption. For instance, an increase equivalent to 100 new cases outside the city leads to an overall decrease in total consumption of 4 percentage points. As the results show, the impact of spillover cases is larger in magnitude than that of local cases.

Table 5: Spillover (All metropolitan cities) by category (offline)

	Essential	Non-Essential	Goods	Services	Household	Non-Household
Local	-0.0312*** (0.00524)	-0.0686*** (0.0162)	-0.0237** (0.00804)	-0.0994*** (0.0134)	-0.0154 (0.0129)	-0.0444*** (0.00973)
Spillover	-0.0186*** (0.00438)	-0.0683** (0.0183)	-0.0287** (0.00800)	-0.0591** (0.0157)	-0.0222*** (0.00420)	-0.0377** (0.0100)
EPU	0.00568 (0.0296)	0.0240 (0.0340)	0.0215 (0.0284)	-0.00943 (0.0368)	-0.0769 (0.0513)	0.0165 (0.0303)
Constant	-0.302 (6.686)	-9.169 (6.969)	-4.349 (6.198)	-2.092 (8.092)	13.08 (11.33)	-4.328 (6.595)
N	24	24	24	24	24	24
R-squared	0.435	0.701	0.517	0.721	0.268	0.628
FE	6	6	6	6	6	6

Notes: Robust standard errors clustered at the city level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is year-on-year change (%) of offline personal monthly credit card expenditure by category. The criteria for classification into disaggregated expenditure categories are available in the appendix. Local is the monthly and city level number of new cases. ‘Spillover’ is the monthly and city level number of cases in other cities (weighted by the relative commuting intensity) as described in section 4. EPU is the nationwide monthly level Economic Policy Uncertainty index. We use an OLS linear specification with city fixed effects from December 2019 to March 2020.

Sources: Bank of Korea for the expenditure data, Korea Center for Disease Control and Prevention, Ministry of Health and Welfare for the COVID-19 monthly new cases, Baker et al. (2016) for the EPU data, and Google Maps and Statistics Korea for the construction of the ‘Spillover’ measure.

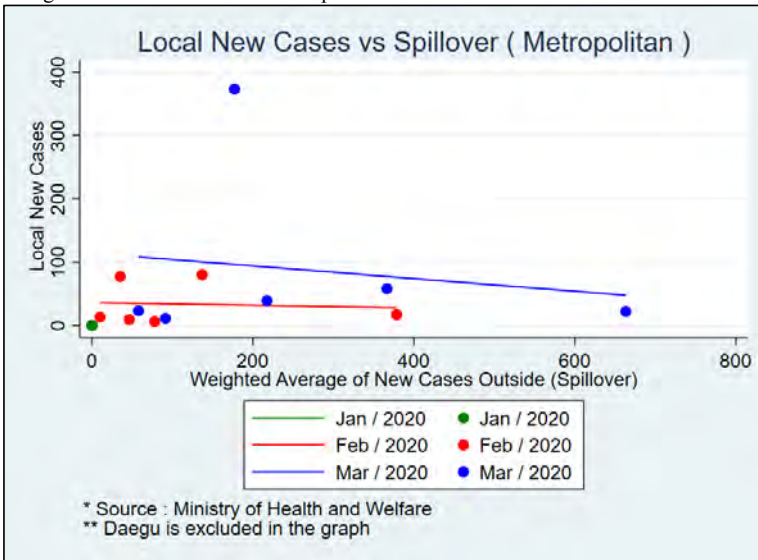
Table 5 shows results for offline consumption disaggregated by category. The economic interpretation of the coefficients is as follows. For an increase of 100 new cases in a given city there is a decrease in total offline consumption of about 3 percentage points. Interestingly, the magnitude doubles when considering Non-

Essential goods (consumption decreases by 6 percentage points for this category). The analysis of Essential versus Non-Essential business expenditures shows that spillover effects concern mainly Non-Essentials. This suggests that spillover effects are the result of consumers voluntarily changing their consumption behaviour, avoiding offline consumption of Non-Essentials.

Regarding the impact of nation-wide policy uncertainty, it has a negative and significant effect on household goods expenditure and online expenditure.

A caveat for the interpretation of our results is the potential collinearity between new local cases and new spillover cases. Figure 14 presents the correlation between local new cases and spillover cases, showing that, if anything, there is a negative correlation between the two in the month of March.

Figure 14. Local New Cases vs. Spillover cases



Another caveat for the interpretation of results so far which particularly concerns the impact of local new cases is that consumption patterns can themselves influence the spread of the virus locally. To deal with this possibility, we replicate our analysis using lagged values of local new cases.

Table 6 and Table 7 replicate the analysis in Table 4 and Table 5 respectively but using lagged measures of local new cases (instead of contemporaneous local cases) and lagged measures of spillover cases (instead of contemporaneous spillover cases).

As Table 6 and Table 7 show, the impact of lagged local new cases becomes more significant both statistically and in economic terms. In particular, an increase in local cases equal to 100 leads to an overall decline in consumption of 14.5 percentage points. Furthermore, the order of magnitude of the effect of an increase in 100 local cases on offline consumption is multiplied by a factor of 10. This strongly supports the hypothesis that consumers may have voluntarily engaged in adapting their behaviour to prevent the spread of the virus, taking into account data from the previous month regarding the local evolution of the pandemic as well as that in neighbouring cities. Interestingly, the magnitude of the effect of lagged spillover cases is of a similar order of magnitude to that using contemporaneous spillover cases.

Table 6: Spillover (Metropolitan cities) using lagged local and spillover cases.

	Total	Offline	Online
L.Local	-0.153*** (0.0380)	-0.186*** (0.0336)	0.111* (0.0533)
L.Spill	-0.0563*** (0.0111)	-0.0547*** (0.0101)	0.0496** (0.0137)
EPU	0.0204 (0.0159)	0.0366*** (0.00842)	-0.811*** (0.198)
Constant	-4.592 (3.192)	-10.58*** (1.925)	199.4*** (41.46)
N	18	18	18
R-squared	0.688	0.731	0.780
FE	6	6	6

Notes: Robust standard errors clustered at the city level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is year-on-year change (%) of total, offline and online personal monthly credit card expenditure. 'L.Local' is the previous month city level number of new cases. 'L.Spillover' is the previous month city level number of cases in other cities (weighted by the relative commuting intensity) as described in section 4. EPU is the nationwide monthly level Economic Policy Uncertainty index. We use an OLS linear specification with city fixed effects from December 2019 to March 2020.

Sources: Bank of Korea for the expenditure data, Korea Center for Disease Control and Prevention, Ministry of Health and Welfare for the COVID-19 monthly new cases, Baker et al. (2016) for the EPU data, and Google Maps and Statistics Korea for the construction of the 'Spillover' measure.

Table 7: Spillover (Metropolitan cities) using lagged cases (by category)

	Essential	Non-Essential	Goods	Services	Household	Non-Household
L.Local	-0.106** (0.0305)	-0.351*** (0.0490)	-0.125** (0.0354)	-0.391*** (0.0667)	-0.0857 (0.0485)	-0.195*** (0.0341)
L.Spill	-0.0365*** (0.00474)	-0.0839*** (0.0180)	-0.0471*** (0.00960)	-0.0714*** (0.0123)	-0.0231*** (0.00371)	-0.0561*** (0.0104)
EPU	0.0256* (0.0124)	0.0823** (0.0241)	0.0473*** (0.00915)	0.0481 (0.0321)	-0.0799 (0.0569)	0.0525*** (0.0105)
Constant	-6.136* (2.458)	-27.54*** (5.471)	-11.91*** (1.707)	-20.41** (7.196)	10.97 (11.55)	-15.16*** (2.340)
N	18	18	18	18	18	18
R-squared	0.603	0.767	0.638	0.814	0.414	0.735
FE	6	6	6	6	6	6

Notes: Robust standard errors clustered at the city level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is year-on-year change (%) of total personal monthly credit card expenditure for the first column and by category for subsequent columns. The criteria for classification into disaggregated expenditure categories are available in the appendix. 'L.Local' is the previous month city level number of new cases. 'L.Spillover' is the previous month city level number of cases in other cities (weighted by the relative commuting intensity) as described in section 4. EPU is the nationwide monthly level Economic Policy Uncertainty index. We use an OLS linear specification with city fixed effects from December 2019 to March 2020.

Sources: Bank of Korea for the expenditure data, Korea Center for Disease Control and Prevention, Ministry of Health and Welfare for the COVID-19 monthly new cases, Baker et al. (2016) for the EPU data, and Google Maps and Statistics Korea for the construction of the 'Spillover' measure.

7. EXTENSION AND ROBUSTNESS

7.1. EXTENSIONS

Our main analysis has focused on studying metropolitan cities. We did not include provinces because internal variation in exposure to the virus and in consumer behaviour within provinces may make our estimates less precise. This section replicates our main analysis at the province level. We obtain similar results. It is worth noting that the results are also unchanged if we estimate both metropolitan cities and provinces in the same analysis.

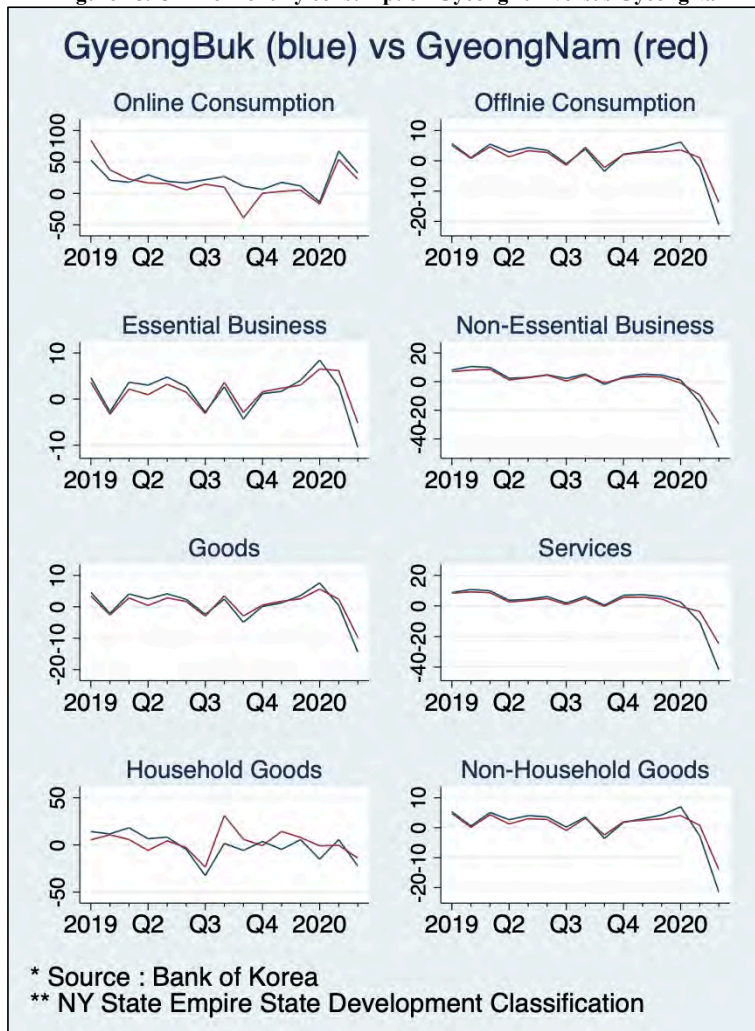
7.1.1. DiD for provinces

For our first step, we use the province of GyeongBuk as our treatment group and the province of GyeongNam as our control group. Indeed, as the map in Figure 3 shows, the province of GyeongBuk was hardest hit by the virus, while its incidence was smaller in the province of GyeongNam. To motivate this choice, we next

compare the two provinces before proceeding to a visual inspection of their consumption patterns before and during the COVID-19 outbreak. The province of GyeongBuk per capita GDP in 2018 equals 38,615 thousand KRW. In Gyeongnam per capita GDP equals, as of 2018, 31,881 thousand KRW. The industrial structure of these two provinces is also similar, and dominated by manufacturing, which represents 46 % of GDP in GyeongBuk and 39 % of GDP in GyeongNam. Public service comes next representing 8 % of GDP in both provinces. In terms of medical capacity, GyeongBuk counts 1 hospital per thousand inhabitants while GyeongNam counts 1,1 hospitals per thousand inhabitants. GyeongBuk counts 0.2 physicians per thousand and 0.450 related doctors per thousand. GyeongNam counts 0.27 physicians per thousand and 0.514 related doctors per thousand. Finally, in terms of demographics, both provinces exhibit similar age profiles with Gyeongbuk counting slightly higher percentage of the population older than 60 (17,7 %). Gyeongnam counts 14,1 % of its population older than 60 years.

Figure 15 present the pairwise comparison of consumption by category between Gyeongbuk and Gyeongnam from January 2018 to March 2020.

Figure 15. Offline monthly consumption GyeongBuk versus GyeongNam



Covid Economics 58, 19 November 2020: 57-84

Figure 15 shows similar patterns between GyeongNam (the most affected province) and GyeongBuk, our control group. In particular, the contraction of consumption in GyeongNam was more severe. Furthermore, in both provinces the contraction of consumption affected particularly Non-Essential and Services. Online consumption increased in both provinces.

7.1.2. Spillover results for provinces

This section replicates our analysis for South Korean provinces instead of metropolitan cities as in the main analysis. Table 8 includes results for total, offline and online consumption. Table 9 includes results for offline consumption by consumption category.

Table 8: Spillover (Provinces)

	Total	Offline	Online
Local	-0.00144 (0.00593)	-0.0101 (0.00554)	-0.00246 (0.0353)
Spillover	-0.0473*** (0.00988)	-0.0465*** (0.00971)	0.00367 (0.0442)
EPU	0.000234 (0.0163)	0.00265 (0.0167)	-0.517** (0.169)
Constant	2.656 (3.421)	3.132 (3.515)	115.8** (38.33)
N	28	28	28
R-squared	0.652	0.677	0.488
FE	7	7	7

Notes: Robust standard errors clustered at the province level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is year-on-year change (%) of total, offline and online personal monthly credit card expenditure. Local is the monthly and province level number of new cases. 'Spillover' is the monthly and province level number of cases in other provinces (weighted by the relative commuting intensity) as described in section 4. EPU is the nationwide monthly level Economic Policy Uncertainty index. We use an OLS linear specification with province fixed effects from December 2019 to March 2020. Sources: Bank of Korea for the expenditure data, Korea Center for Disease Control and Prevention, Ministry of Health and Welfare for the COVID-19 monthly new cases, Baker et al. (2016) for the EPU data, and Google Maps and Statistics Korea for the construction of the 'Spillover' measure.

Table 9: Spillover (Provinces) for offline consumption (by category)

	Essential	Non-Essential	Goods	Services	Household	Non-Household
Local	-0.00128 (0.00368)	-0.0320** (0.0108)	-0.00524 (0.00306)	-0.0317** (0.0124)	-0.00668 (0.00454)	-0.0117* (0.00581)
Spillover	-0.0290*** (0.00746)	-0.0805*** (0.0159)	-0.0349*** (0.00669)	-0.0776*** (0.0180)	-0.0558*** (0.00969)	-0.0471*** (0.0101)
EPU	0.00581 (0.0163)	0.0203 (0.0203)	0.0175 (0.0168)	-0.00977 (0.0210)	-0.0883** (0.0260)	0.0135 (0.0174)
Constant	3.955 (3.441)	-3.343 (4.261)	0.164 (3.640)	5.837 (4.246)	23.07*** (6.207)	0.962 (3.621)
N	28	28	28	28	28	28
R-squared	0.503	0.726	0.620	0.684	0.719	0.685
FE	7	7	7	7	7	7

Notes: Robust standard errors clustered at the province level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is year-on-year change (%) of offline personal monthly credit card expenditure for the first column and by category for subsequent columns. The criteria for classification into disaggregated expenditure categories are available in the appendix. Local is the monthly and province level number of new cases. 'Spillover' is the monthly and province level number of cases in other provinces (weighted by the relative commuting intensity) as described in section 4. EPU is the nationwide monthly level Economic Policy Uncertainty index. We use an OLS linear specification with province fixed effects from December 2019 to March 2020. Sources: Bank of Korea for the expenditure data, Korea Center for Disease Control and Prevention, Ministry of Health and Welfare for the COVID-19 monthly new cases, Baker et al. (2016) for the EPU data, and Google Maps and Statistics Korea for the construction of the 'Spillover' measure.

The local number of new cases has a negative impact on all consumption categories, but only statistically significant for Non-Essential expenditure, Non-Household and Services. The economic interpretation of the coefficient is as follows. For an increase of 100 new cases in a given province there is a decrease in Non-

Essential business expenditure slightly above 3 percentage points. Again, one caveat for the interpretation of results concerning the impact of local new cases is that consumption patterns can themselves influence the spread of the virus locally. To deal with this possibility, in the robustness check section we use lagged values of local new cases.

Coefficients for spillover new cases (weighted average of new cases outside in other provinces) are also negative, except for online consumption. Contrary to the previous case, they are all statistically significant across all consumption categories, except for online consumption. For instance, an increase equivalent to 100 new cases outside the city leads to an overall decrease in total consumption of 4.73 percentage points. This negative effect is particularly large for Non-Essential goods (8 percentage points), services (7.61 percentage points) and Non-Household goods (5.58 percentage points). As the results show, the impact of spillover cases is larger in magnitude than that of local cases across categories. Overall, these results suggest that spillover effects, which concern mostly Non-Essential business expenditures are the result of consumers voluntarily changing their consumption behaviour, avoiding offline consumption of Non-Essentials.

Regarding the impact of nation-wide policy uncertainty, it has a negative and significant effect only on online expenditure.

Finally, Table 10 and Table 11 replicate the previous analysis using lagged measures of local and spillover cases.

Table 10: Spillover (Provinces) using lagged cases

	Total	Offline	Online
L.Local	-0.125 (0.0788)	-0.156* (0.0668)	-0.0173 (0.183)
L.Spill	-0.0573* (0.0292)	-0.0524 (0.0299)	0.0754 (0.0789)
EPU	0.0390** (0.0137)	0.0449** (0.0122)	-0.642** (0.184)
Constant	-7.519** (2.681)	-7.820** (2.373)	136.8** (38.95)
N	21	21	21
R-squared	0.651	0.689	0.605
RegionFE	7	7	7

Notes: Robust standard errors clustered at the province level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is year-on-year change (%) of total, offline and online personal monthly credit card expenditure. 'L.Local' is the previous month province level number of new cases. 'L.Spillover' is the previous month province level number of cases in other provinces (weighted by the relative commuting intensity) as described in section 4. EPU is the nationwide monthly level Economic Policy Uncertainty index. We use an OLS linear specification with province fixed effects from December 2019 to March 2020.

Sources: Bank of Korea for the expenditure data, Korea Center for Disease Control and Prevention, Ministry of Health and Welfare for the COVID-19 monthly new cases, Baker et al. (2016) for the EPU data, and Google Maps and Statistics Korea for the construction of the 'Spillover' measure.

Table 11: Spillover (Provinces) using lagged cases for offline (by category)

	Essential	Non-Essential	Goods	Services	Household	Non-Household
L.Local	-0.0948 (0.0530)	-0.301** (0.102)	-0.115* (0.0491)	-0.302** (0.109)	-0.151* (0.0634)	-0.166* (0.0690)
L.Spill	-0.0432 (0.0228)	-0.0693 (0.0478)	-0.0492* (0.0235)	-0.0607 (0.0494)	-0.0437 (0.0373)	-0.0536 (0.0310)
EPU	0.0500** (0.0145)	0.0637*** (0.0141)	0.0660*** (0.0117)	0.0274 (0.0197)	-0.0488 (0.0328)	0.0589*** (0.0118)
Constant	-5.605* (2.730)	-17.85*** (3.294)	-10.71*** (2.199)	-7.406 (4.096)	11.45 (7.180)	-10.57*** (2.196)
N	21	21	21	21	21	21
R-squared	0.678	0.702	0.732	0.683	0.588	0.708
RegionFE	7	7	7	7	7	7

Notes: Robust standard errors clustered at the province level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is year-on-year change (%) of offline personal monthly credit card expenditure by category. The criteria for classification into disaggregated expenditure categories are available in the appendix. 'L.Local' is the previous month province level number of new cases. 'L.Spillover' is the previous month province level number of cases in other provinces (weighted by the relative commuting intensity) as described in section 4. EPU is the nationwide monthly level Economic Policy Uncertainty index. We use an OLS linear specification with province fixed effects from December 2019 to March 2020.

Sources: Bank of Korea for the expenditure data, Korea Center for Disease Control and Prevention, Ministry of Health and Welfare for the COVID-19 monthly new cases, Baker et al. (2016) for the EPU data, and Google Maps and Statistics Korea for the construction of the 'Spillover' measure.

Consistent with the findings concerning metropolitan cities, lagged local cases have the strongest effect on consumption compared to the impact of either contemporaneous local cases or spillover cases.

7.2. ROBUSTNESS CHECKS AND LIMITATIONS

7.2.1. Robustness checks

We have performed a series of robustness checks available upon request. First, we have checked that the DiD analysis is robust to the inclusion of pre-COVID-19 trend dummies. Second, we have checked robustness to our classification of consumption categories. In particular, our main analysis has excluded expenditures in utilities, as these are hard to change in a short period of time. We also excluded expenditures in financial services, as Seoul is disproportionately represented in this category. Yet, we have run the analysis including these two items and results are fundamentally unchanged. Third, we have checked the robustness of our results to the inclusion of unemployment rate as explanatory variable. Results are fundamentally unchanged, as unemployment rate did not spike during the months of February and March because individuals decreased their labour supply.

7.2.2. Limitations

One of the data limitations of our analysis is that we do not have evidence regarding cash transactions. In the last quarter of 2019, personal credit card payments represented 60 % of household consumption in the national accounts. Furthermore, the proportion of credit card payments out of consumption increased during the first quarter of 2020. This implies that our results may under-estimate the true magnitude of consumption changes. A second concern is that our expenditure data is based on the business location rather than on the consumer physical address. This is not a big concern for offline consumption but could be an issue for online consumption data. Yet, since August 2019 the share of online expenditures in Seoul has remained stable.

8. DISCUSSION AND CONCLUSION

In the management of the COVID-19 pandemic, societies and governments have been confronted to difficult choices, often at the intersection between economic and public health considerations. In doing so, the trade-off between health and economic objectives has prompted governments to adopt radically different policy responses, regardless of the democratic nature of their political institutions. Restrictive policy measures involving shutting down entire sectors considered as non-essential and restricting population mobility have contrasted with softer responses involving information release about the spread of the virus and relying on citizens self-management.

Critics of the restrictive approach argue that the economic cost is too large, while their supporters argue that plummeting consumption has resulted from consumers shying away from consuming in sectors with higher exposure to the health risk, rather than as a consequence of mobility restrictions. In other words, that absent those restrictions, the economic cost would have been of similar magnitude and nature.

This paper objective has been to advance evidence that can inform on the magnitude and nature of the trade-off by studying the case of South Korea in the months following the outbreak of the pandemic in the country. The experience of South Korea is valuable because it allows to measure consumers' immediate response to COVID-19 absent any restrictions and because the country was heterogeneously exposed to the health shock, which was concentrated in the metropolitan city of Daegu. We exploit this variation to estimate the consumer response at the epicentre and outside of it. We further investigate consumer response across categories. Since there were no business shutdowns doing so allows to measure the extent to which consumers voluntarily reallocated their expenditures in response to COVID-19.

Our analysis shows that consumers did adjust their consumption pattern both in terms of magnitude and reallocation, refraining from consuming non-essential goods, services, and shifting their consumption basket towards online consumption. We also find that consumers were sensible to the different risk levels in different locations, with smaller changes in consumption in places relatively spared by the virus. Overall, we interpret these results as implying that consumers' voluntary response to the pandemic contributed to minimise the economic cost and at the same time to mitigate the spread of the pandemic. The latter is supported by our evidence showing that consumers changed their consumption mostly in response to the evolution of local cases during the previous month and to a smaller extent to the evolution of cases outside of their own city.

Moving forward, policy-makers around the world choosing between health and economic objectives may incorporate two main lessons from our analysis. First, that consumers endogenous response to the health shock can help fight the pandemic at a lower cost. Second, that for this endogenous response to be effective, governments should invest in their communication strategy and immediate release of information regarding detailed geographic patterns of COVID-19 cases, allowing consumers to selectively avoid certain locations and businesses during short periods of time and to test themselves if cases have been diagnosed in concurrent locations. This requires the development of three important tangible and intangible resources: first, tests need to be available, second, citizens need to trust the information received, and third, governments need to manage privacy issues related to pandemic data management. Absent these, it is unlikely that voluntary consumer response can, by itself, reduce the magnitude of the trade-off effectively. The failure of the STOP-COVID app in France is certainly not encouraging news in this respect.

9. APPENDIX

9.1. Consumption categories

We define the following categories: (1) Online, (2) Offline, (3) Essential, (4) Non-Essential, (5) Goods, (6) Services, (7) Household, and (8) Non-Household.

(1) Online is simply Electronic Commerce because this consumption type contains all online consumption in this dataset. (2) Offline is Total, subtracted by (1) Online, Utility Charges / Personal and Professional Services, and Financial Services / Insurances. Utility Charges / Personal and Professional Services contains utility service bills, which does not reflect consumption. Financial Services' headquarters are concentrated in Seoul, which can overestimate consumption in Seoul and underestimate consumption elsewhere. (3) Essential follows 2020 June 29 Empire State Development's guidance for determine essential business, in response to COVID-19. Thus, (3) Essential Consumption Type includes Discount Stores / Other Superstores, Supermarkets, Convenience Stores, General Food Products, Health Additive Foods, Garments / Clothing Materials, Cosmetics, Fuels, Household Appliances / Communication Equipment, General Hospitals, Other Health Facilities, Domestic Motor Vehicles (New Products), Other Vehicles, Motor Vehicle Services, Airlines, Public Transportation, Books / Stationary, Accommodation. (4) Non-Essential is (2) Offline, subtracted by (3) Essential, and Other Goods and Services. (6) Services includes Motor Vehicles Services, Airlines, Restaurants / Other Food Services, and Education. (5) Goods is (2) Offline, subtracted by (5) Services, and Other Goods and Services. (7) Household includes Furniture, and Household Appliances / Communication Equipment. (8) Non-Household is (2) Offline, subtracted by (7) Household, and Other Goods and Services. Other Goods and Services are subtracted from all categories but (2) Offline because we do not know the exact composition, other than knowing that it is completely consumed offline.

REFERENCES

- Andersen, A. L., Hansen, E. T., Johannesen, N., & Sheridan, A. (2020). "Consumer responses to the COVID-19 crisis: Evidence from bank account transaction data". *Covid Economics Vetted and Real-time Papers*, (7), 92-118.
- Baldwin, R. (2020). "Keeping the lights on: Economic medicine for a medical shock". VoxEU.org, 13th March.
- Bank of Korea (2020). Press release from the Economic Statistics Department (82-2-759-5265, 5237.)
- Baker, S. R., Bloom, N., & Davis, S. J. (2016). "Measuring economic policy uncertainty". *Quarterly Journal of Economics*, 131(4), 1593-1636.
- Bounie, D., Camara, Y., & Galbraith, J. W. (2020). "Consumers' Mobility, Expenditure and Online-Offline Substitution Response to COVID-19: Evidence from French Transaction Data". Available at SSRN 3588373.
- Brodeur, A., Gray, D. M., Islam, A., & Bhuiyan, S. (2020). "A Literature Review of the Economics of Covid-19". IZA Discussion Paper No. 13411.
- Carvalho, V. M., Garcia J. R., Hansen S., Ortiz Á., Rodrigo T., Rodríguez Mora J. V., & Ruiz J. (2020a). "Tracking the COVID-19 crisis with high-resolution transaction data". Cambridge-INET Working Paper, 2009.
- Carvalho, V. M., Garcia J. R., Hansen S., Ortiz Á., Rodrigo T., Rodríguez Mora J. V., & Ruiz J. (2020b). "Tracking the COVID-19 crisis through the lens of 1.4 billion transactions". VoxEU.org, 27th April.
- Chen, H., Qian W., & Wen Q. (2020). "The Impact of the COVID-19 Pandemic on Consumption: Learning from High Frequency Transaction Data". Working paper.
- Chin, Y. M., & Wilson, N. (2018). "Disease risk and fertility: evidence from the HIV/AIDS pandemic". *Journal of Population Economics*, 31(2), 429-451.
- Chronopoulos, D. K., Lukas M., & Wilson J. O. S. (2020). "Consumer spending responses to the Covid-19 pandemic: An assessment of Great Britain". *Covid Economics Vetted and Real-time Papers*, (34), 145-186.
- Krueger, D., Uhlig, H., & Xie, T. (2020). "Macroeconomic Dynamics and Reallocation in an Epidemic". *Covid Economics Vetted and Real-time Papers*, (5), 21-55.

Online learning during school closure due to COVID-19¹

Masato Ikeda² and Shintaro Yamaguchi³

Date submitted: 10 November 2020; Date accepted: 11 November 2020

This paper estimates the effects of school closure on students' study time and the number of messages sent from teachers to students using an online learning service. We find that both study time and message numbers increased significantly from the beginning of the school closure but they returned to pre-COVID-19 levels when the state of emergency ended in late May 2020. In addition, we find that students with prior access to the online learning service at home and students at higher-quality schools increased their study time more than other students. However, we find no gender differences in these outcomes.

1 We gratefully appreciate SuRaLa Net Co.,Ltd. for providing data. We also thank contribution by Makiko Nakamuro to our research at an early stage, excellent research assistance by Sae Morita, and financial support from KAKENHI 20H01510.

2 Ph.D. candidate, University of Tokyo, Graduate School of Economics.

3 Professor, University of Tokyo, Graduate School of Economics.

Copyright: Masato Ikeda and Shintaro Yamaguchi

I. Introduction

At the peak of the educational disruption caused by COVID-19, over 1.5 billion students in roughly 190 countries missed out on learning at school¹. Given the scale and duration of school closures under COVID-19, the learning loss is likely to result in long-term negative consequences. During school closures, governments attempted to minimize the adverse effects. For instance, in at least 34 states in the US, local entities such as school districts and state departments of education partnered with public television stations to support online learning for students and teachers². The severity of the educational disruption, including whether government attempts to alleviate it were successful, has not yet been examined in detail. To contribute toward this gap in the literature, in this paper, we document how an online learning service enabled students to study during the COVID-19 pandemic in Japan.

The Japanese government began a nationwide school closure on March 2, 2020, which continued for three months until the end of May 2020. There is evidence that the school closure was unexpected and that many schools were unprepared for online learning. In this paper, we estimate the effect of school closure due to COVID-19 on the utilization of an online learning service by comparing user logs in 2020 with those from 2019. Our findings assist in understanding how an online learning service could mitigate learning loss during COVID-19.

We find that the school closure during COVID-19 increased the study time of students using the online learning service, with the effect being strongest at the beginning of the school closure period. It decreased gradually and disappeared in June 2020, when most schools reopened. This result suggests that the online learning service enabled students to study online and compensated for the missed classes during the closure period. In addition, we

¹ The specific number cited is from <https://en.unesco.org/covid19/educationresponse>

² More detailed information can be obtained from <https://apts.org/news/station-stories/public-media-education-partnerships-school-districts-governments-and-education-agencies>

find a positive effect on teachers' effort levels, measured by messages sent from teachers to students via the online learning service. Interestingly, the effect on teachers' effort is positively correlated with the effect on study time.

The contribution of this paper is to provide further evidence for the effects of online learning during the current COVID-19 pandemic. A few studies have investigated education under COVID-19, mostly with a focus on online learning (Bacher-Hicks et al., 2020; Chetty et al., 2020). Our study is most closely related to Chetty et al. (2020), in that both describe the utilization of an online education platform. Given that Figlio et al. (2013) find that online learning is a reasonable substitute for face-to-face learning, our results suggest that online learning services have mitigated the negative consequences of COVID-19 on education.

The remainder of the paper is organized as follows. In Section II, we discuss the related literature. Section III outlines the institutional background and the online learning service. In Section IV, we describe our data and Section V presents and discusses our results. Section VI offers concluding remarks.

II. Related Literature

Some studies based on observational data find negative effects of online lectures on learning (Bacolod et al., 2018; Bettinger et al., 2017; Xu & Jaggars, 2014). However, by comparing the estimates from observational and experimental data, Joyce et al. (2015) argue that the former are likely to overestimate the effect of face-to-face lectures. Figlio et al. (2013) show that an online lecture in a university increased students' grade points by up to 3 points out of 100 relative to the effect of face-to-face lectures, although it had adverse effects for males and academically weaker students. Overall, these findings suggest that an online lecture can be a reasonable substitute for a face-to-face lecture for many students. Note that the result from the literature may not be directly applicable to junior high school and high school students because these studies examine university students.

Several studies investigate the impacts of COVID-19 on education. One strand of literature deals with real-time data under COVID-19. Chetty et al. (2020) examine data from an online learning service largely used as part of a math curriculum in the US. They observe an acute drop of average study hours in the service during COVID-19, with a particularly strong negative effect for students in low-income areas. Aucejo et al. (2020) conduct a survey of 1,500 students at Arizona State University, asking how COVID-19 affected their studies. They find a decrease in study time and a negative effect on educational outcomes, including delays in graduating and greater dropout from courses. Both Chetty et al. (2020) and Aucejo et al. (2020) describe how students' study behavior is affected by COVID-19 and report that the total amount of study time decreased. Finally, Bacher-Hicks et al. (2020) use regional-level data on Google searches and find that the search intensity for online education material increased after the school closure and that the increase was more prominent in high-income regions.

We contribute to this strand of literature by documenting students' and teachers' activities in an online learning service during the COVID-19 pandemic. Although our study is most closely related to Chetty et al. (2020), in the present work, the online learning service is not incorporated into a math curriculum but is primarily offered as supplementary educational materials, which were available before COVID-19 (and are likely to be after). This difference accounts for why study time increased during school closure in our study, whereas Chetty et al. (2020) find a decrease in study time.

A second strand of literature examines the long-term effect of COVID-19 based on estimates from past incidents, such as the 2001 foot and mouth disease epidemic in England (Cook, 2020), and a combination of summer vacation, weather-related school closure, and absenteeism (Kuhfeld et al., 2020). Both Cook (2020) and Kuhfeld et al. (2020) predict a negative long-term effect of COVID-19. However, these studies may overestimate the effect

of COVID-19 because, as Kuhfeld et al. (2020) state, online education can mitigate the negative effects.

III. Background

In this section, we document the timeline of school closure due to COVID-19 in Japan to assist readers to understand the context. It should be noted that, in contrast to the situation in many other countries, the Japanese academic year begins in April and ends in March of the following year.

On February 27, 2020, Prime Minister Shinzo Abe made a public statement that the Japanese government requested all elementary schools, junior high schools, and high schools to close from March 2, 2020 until the beginning of spring break, which typically begins in the third or fourth week of March and lasts around two weeks, although the exact date and duration vary by region, school, and year. Thus, the statement implied roughly three weeks of school closure.

The announcement came as a complete surprise, as indicated by two pieces of evidence³ concerning the decision-making process of the government and the government school closure policy prior to the announcement. First, the Minister of Education, Culture, Sports, Science and Technology conceded in the Diet that Prime Minister Abe had only informed him of the school closure on the morning of February 27, the day of the public announcement. The Minister's comment reveals that only a few people close to Prime Minister Abe were involved in the decision and that the information was not shared with the education minister, let alone with other policymakers, or teachers, parents, and students.

Second, prior to the announcement, the government's school closure policy was, as requested by the Ministry of Education, Culture, Sports, Science and

³ The information in this section is obtained from the following three websites:

https://www.kantei.go.jp/jp/singi/novel_coronavirus/taisaku_honbu.html

<https://kokkai.ndl.go.jp/#/>

<https://www.pref.tottori.lg.jp/item/1204674.htm>

Technology, that each municipality's school board should keep schools open unless a positive case of COVID-19 was found in a school. Given that most schools had no positive cases at that time, the school boards were unlikely to expect imminent school closures. In addition, on February 25, the Ministry requested that school boards begin making contingency plans for possible school closures.

The nationwide school closure due to COVID-19 began on March 2 and lasted until the end of May in most provinces⁴. Following the government's announcement, nearly all schools managed to close on March 2, only four days after the announcement. At that time, the government stated that schools would be closed for about three weeks, until the beginning of spring break. On March 20, the government announced that the school closure would not be extended after spring break and released guidelines for reopening schools on March 24. Hence, schools were expected to reopen after spring break, which was the beginning of the new academic year.

However, schools did not reopen after spring break because on April 7 the government declared a state of emergency for seven provinces, which was later extended nationwide. Along with this declaration, the government asked municipalities to decide whether to reopen schools at their own discretion. This declaration resulted in a de facto continuation of the school closure. As shown in Table 1, most schools remained closed, or reclosed, after the declaration of the state of emergency.

⁴ The actual administrative term used by the government is prefecture, but we use province as it is more intuitive to most readers.

TABLE 1—RATES OF SCHOOL CLOSURE

	April 22	May 11	June 1
Public Schools			
Junior High Schools	95%	88%	1%
High Schools	99%	92%	8%
Private Schools			
Junior High Schools	97%	90%	8%
High Schools	98%	88%	6%

Note: The table shows the percentage of schools that were open each day on the dates specified.

Source: https://www.mext.go.jp/a_menu/coronavirus/mext_00007.html

Surala

Surala is an online learning service provider. Their service covers a variety of subjects, including math, Japanese, English, science, and social studies, and caters to all grades, from grade 1 to grade 12. Students can study materials above or below their own grade to review past materials or to prepare for university entrance exams. Subscription to the service can be on either an individual or a school basis. In this paper, we focus on users under school-wide contracts because this means that school and teacher characteristics are available. Students included in our data are enrolled in either junior high schools or high schools. Junior high school covers grades 7–9 and high school grades 10–12, but we omit grades 9 and 12 from our analysis, as discussed below.

The service provides three types of learning materials: lectures, drills, and tests. For the lecture material, students watch videos and answer quizzes during the lecture. As they can pause and rewind lectures, the time taken to complete a lecture varies between students. After each lecture, students are asked to solve drill questions. These differ from the quizzes during the lecture, which are primarily designed to draw students' attention and ensure they understand the explanations provided in the lecture. By contrast, drills are problem sets designed to enhance students' deeper understanding of the material.

A learning unit on Surala consists of a lecture and a drill. Several units relating to the same topic comprise one lesson. Finally, lessons are categorized into stages, depending on the level of advancement. Students may begin their study from any unit. Our data set includes user activity logs on units, but does not include any information on tests on Surala.

To manage students' learning, Surala issues an account for teachers, which enables them to observe how their students study online. For instance, a teacher examines students' understanding based on their progress and test scores with the learning service. In addition, teachers can send messages to students through the service, either to an individual student, a group of selected students, or to all students in the school.

IV. Data

User Activity Log

We now describe our data set and define the variables used in the analysis. Our main data set is drawn from user activity logs on Surala. It includes information on when and what each student studied, and selected demographic characteristics, including grade and gender, as well as a school identifier. The data set also includes the teachers' message log, which records the time and content of teachers' messages to students.

Sample Restrictions

We focus on students in junior high school (grades 7–9) and high school (grades 10–12), but we exclude the third-year students in both types of schools (i.e., those in grades 9 and 12) to avoid issues arising from attrition. Because the academic year ends in March and begins in April, most third-year students in both junior high school and high school move to a different school in April. Most junior high schools and high schools are separated, although some private schools provide a combined program.

We excluded observations for movers, students with missing information, and outliers. In the raw data, there are students who spent more than 10 hours on one unit, whereas the average time to complete a unit is about 10 minutes. As we suspect that these students paid little attention to the material, despite being logged on to Surala, we exclude their entire study records from the data. Specifically, students with weekly study time exceeding the 99.9th percentile are dropped from the sample. Note that excluding outliers does not substantially change the mean study time, but slightly decreases the standard errors of our estimates. In addition, we excluded individuals with missing school identifiers for analyses that use school characteristics.

With regard to the message data, our sample includes messages sent from teachers to individual students. Although teachers can also send messages to a group of selected students or to all students in the school, in our data, 77% of messages were sent to individual students. In addition, when we construct a school-level variable, such as the weekly average number of messages per school, we include schools that used the service in January in our sample.

Summary Statistics

Table 2 reports summary statistics for the key variables. The average time spent on completing a study unit is 8.98 minutes. During the period of analysis, 32% of students logged in to Surala in a given week. Among those who logged in, the average time for studying on Surala was 73.84 minutes. When we include students who did not log in (i.e., the majority), the average study time is 23.29 minutes. In our sample, 68% of students are in high school and 55% are male. Grades and gender were reported by the school as of April 2020.

We determine whether a student ever accessed the online learning service from home before school closure using the user log. Specifically, we consider that a student did not access the service from home prior to the school closure if he/she did not use the online learning service either after 8:00 p.m. or on the

weekend during the period from April to December 2019. According to this definition, 17% of students had no prior access to the online learning service from home.

Our data include 224 schools, of which 136 agreed to disclose their names. They consist of 41 junior high schools, 85 high schools, and 10 combined junior and senior high schools. For these schools, we used a measure of school quality published by a private firm. The measured quality was originally scaled to a mean of 50 and a standard deviation of 10 among the schools on their list⁶. For junior high schools, the list includes only selective private schools, whereas for high schools, both public and private schools are included. The average quality index for junior high schools in our sample is 40.30, whereas that for the high schools is 49.81.

There were 4,596 messages sent from teachers to students and 95% of them were sent in 2020. Among the schools from which at least one message was sent from a teacher in a given week, the average weekly number of messages sent from a school is 10.66 and the average weekly number of teachers who sent a message is 1.41. Among teachers who sent at least one message in a given week, the average weekly number of messages is 7.56.

Finally, we collect data on the date of school closure. As described in Section I, schools gradually reopened as the state of emergency was lifted in selected provinces on May 14 and in all provinces on May 25. Figure 1 shows the percentage of schools in our data that were closed. Most of them closed on March 2 and reopened from June 1, and this pattern is common across junior high schools and high schools.

⁶ The data were downloaded from minkou.jp in June 2020. A marketing research company claims that this website is the most viewed of those providing information on schools in Japan and we confirmed during informal interviews that teachers, students, and their parents often referred to the website. Although other firms publish high school quality data, the data are very similar between firms.

TABLE 2—DESCRIPTIVE STATISTICS

	Observations	Mean	SD	Median	Min	Max
Time spent for one unit (min)	1,722,701	8.98	11.37	5.35	0.02	141.85
Indicator for log-in in a given week	624,175	0.32	0.47	0.00	0.00	1.00
Weekly study time conditional on log-in (min)	192,469	73.84	93.15	40.95	0.05	814.75
Weekly unconditional study time (min)	610,225	23.29	62.56	0.00	0.00	814.75
Indicator for high school	21,454	0.68	0.47	1.00	0.00	1.00
Indicator for male	21,454	0.55	0.50	1.00	0.00	1.00
No access from home	16,424	0.17	0.38	0.00	0.00	1.00
Number of students per school	224	95.78	160.45	22.50	1.00	1032.00
Average study time Apr–Dec 2019 (min)	164	596.67	681.36	373.46	11.57	4121.09
School quality index for junior high school	51	40.30	6.85	39.00	31.00	59.60
School quality index for high school	95	49.81	7.02	50.50	35.50	67.00
Indicator for messages sent in 2020	4,596	0.95	0.22	1.00	0.00	1.00
Weekly number of messages at the school level	431	10.66	27.47	3.00	1.00	264.00
Weekly number of teachers online at the school level	431	1.41	1.00	1.00	1.00	11.00
Weekly number of messages at the teacher level	608	7.56	13.98	2.00	1.00	104.00

Notes: The time spent for one unit is defined by the average study time per unit. Indicator for log-in in a given week is defined by the dummy variable indicating whether a student studies at least once a week. We use the school-level indicator obtained from external sources as a school quality index. The average study time April–December 2019 is defined by the average study time per student in each school in April–December 2019. The weekly number of messages at the school level is defined by the number of messages that teachers in a given school send in a given week. The weekly number of teachers at the school level is defined by the number of teachers who use messages in a given week. The weekly number of messages at teacher level is defined by the total number of messages sent by one teacher in a given week.

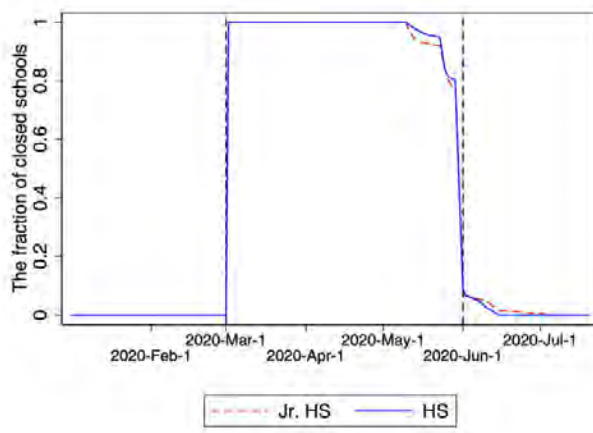


Figure 1. Fraction of Closed Schools in the Sample

Note: Jr. HS and HS denote the fractions of closed junior high schools and high schools, respectively.

V. Results

Students' Study Time

In this section, we show how students' study time and teachers' messaging changed in response to the COVID-19 school closure. Figure 2 presents the average weekly study time in 2019 and 2020. We use study time in 2019 as a comparison to indicate what would have happened in 2020 if the COVID-19 pandemic had not occurred. The vertical dashed line (in red) indicates the beginning of the nationwide school closure. For the first week of January 2020 to the last week of February, there is little difference in study time between 2019 and 2020. However, the online study time surges in the first week of March 2020 relative to the same week in 2019. The average study time in 2020 continues to be longer than that of 2019 until the first week of June, when most schools reopened.

Figure 3 shows the growth of the average weekly study time from 2019 to 2020 with a 95% confidence interval. From January 1 until the end of February, there is no statistically significant difference. However, from the

start of the school closure period, the study time in 2020 is significantly longer than the corresponding time in 2019. In fact, based on Table 3, the study time in 2020 is 22 minutes longer per week, which roughly amounts up to two 45-minutes-classes in one month. We observe a statistically significant growth in study time until the end of April. Although there is a growth in study time until the first week of June, the growth is statistically insignificant in May.

According to the weekly log-in rate in Table 2, only 32% of students study online each week. We expect that a large portion of students remain unaffected by school closure because they do not have access to the online learning service at home. Next, to determine the main driver of the overall effect, we examine the extensive and intensive margins of changes in study time. Figures 4 and 5 show changes in the extensive margin, defined by a weekly log-in indicator. They show that there is a marginally significant effect of school closure from March to the beginning of April. That is, the effect on the extensive margin disappears slightly earlier than the overall effect.

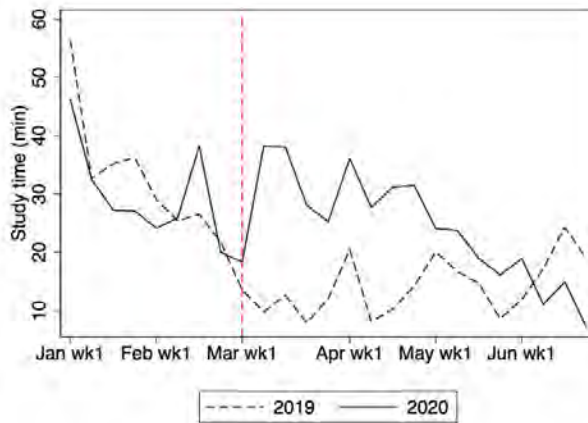


Figure 2. Average Weekly Study Time

Note: The figure shows the average study time for students in grades 7, 8, 10, and 11.

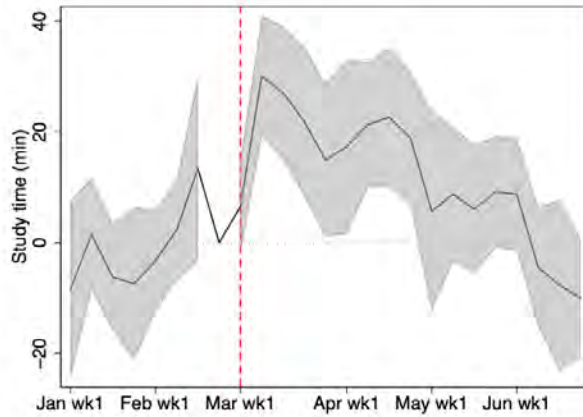


Figure 3. Change in Weekly Study Time from 2019 to 2020

Note: The shaded region is the 95% confidence interval. Study time is shown for students in grades 7, 8, 10, and 11.

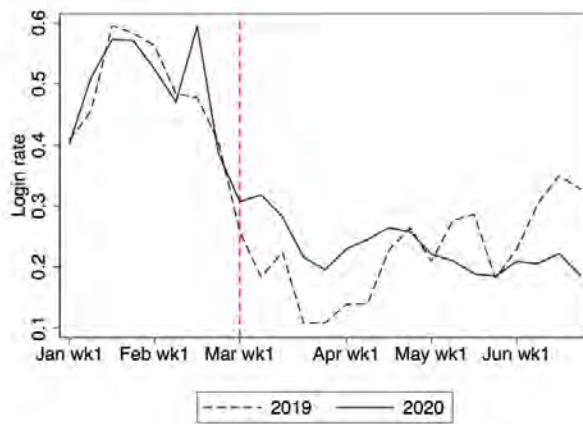


Figure 4. Weekly Log-in Rate

Note: The figure shows the log-in rate for students in grades 7, 8, 10, and 11. Log-in is defined as studying at least once in a given week.

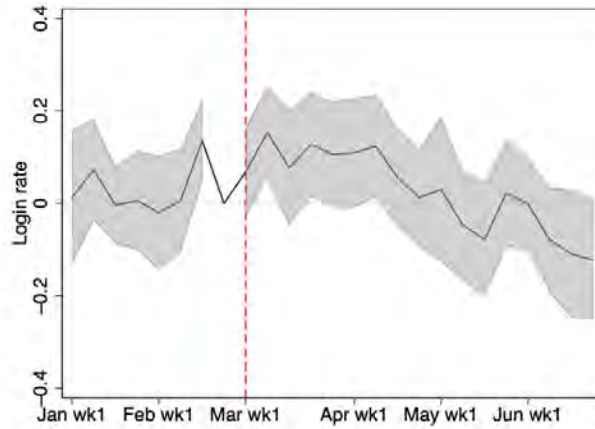


Figure 5. Change of Weekly Log-In Rate from 2019 to 2020

Notes: The shaded region is the 95% confidence interval. The figure shows the log-in rate for students in grades 7, 8, 10, and 11. Log-in is defined as studying at least once in a given week.

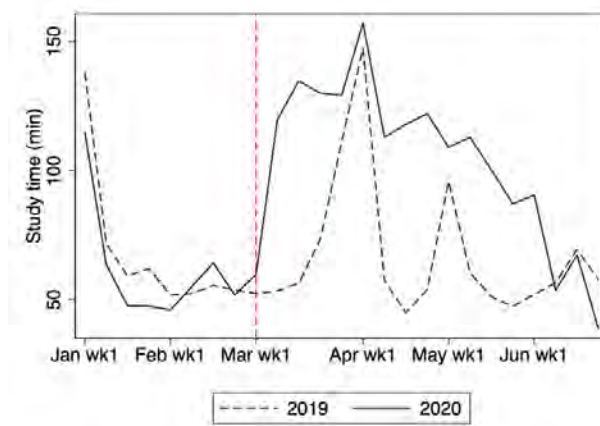


Figure 6. Average Weekly Study Time Conditional on Log-in

Note: Average study time conditional on log-in for students in grades 7, 8, 10, and 11. Only students who study at least once in a given week are included.

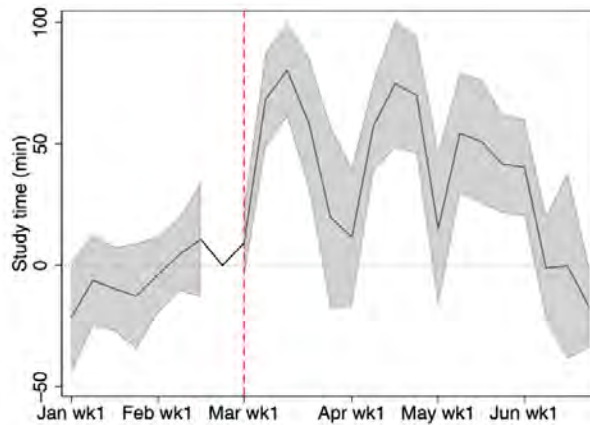


Figure 7. Change of Weekly Study Time Conditional on Log-in from 2019 to 2020

Note: The shaded region is the 95% confidence interval. The figure shows the average study time conditional on log-in for students in grades 7, 8, 10, and 11. Only students who study at least once in a given week are included.

Figures 6 and 7 show the changes in the intensive margin, defined by the average weekly study time conditional on log-in. In 2019, there are spikes at the end of both March and April. These periods correspond to the end of spring break and the one-week-long national holiday (Golden Week), respectively. A Surala manager explained the spikes by the fact that students study intensively to finish homework due after the spring break and the holidays. Regardless of the spikes, a significant positive effect for the intensive margin persists until the end of May. That is, the effect for the intensive margin persists longer than the overall effect. In fact, monthly analyses in Table 3 also shows that the magnitude of the effect on intensive margin is almost as large as the previous month while the effect on overall effect decreases by half in May.

Finally, Table 3 summarizes results from analyses of monthly log-in rate and study time, which shows patterns consistent with the weekly analyses. Note that, in these analyses, study time is defined by weekly average study time in a given month. In Table 3, study time, as well as the intensive and

extensive margin, responses to school closure in March, and the difference continues to be present by the time of schools reopening, June. For instance, the effect on unconditional study time in March is more than three times longer comparing to the effect in May. One difference is that we detect the statistically significant effect on unconditional study time in May. Overall, the analyses in Table 3 provide similar results to weekly analyses, with some improvement in precision.

TABLE 3—MONTHLY-LEVEL ANALYSES OF STUDY RECORDS

	January	February	March	April	May	June
Panel A. Unconditional						
2019	37.93 [3.462]	21.792 [2.054]	10.595 [1.744]	14.627 [2.052]	12.963 [1.681]	20.142 [3.350]
2020	31.431 [2.674]	25.618 [2.890]	32.386 [3.999]	30.102 [3.743]	19.435 [2.782]	11.057 [1.937]
Difference	-6.499*** [2.967]	3.827 [3.237]	21.791** [3.827]	15.476** [3.317]	6.472*** [2.949]	-9.085*** [3.345]
Panel B. Log-in Rate						
2019	-	0.756 [0.032]	0.384 [0.048]	0.476 [0.047]	0.433 [0.044]	0.457 [0.054]
2020	-	0.802 [0.028]	0.51 [0.046]	0.514 [0.050]	0.362 [0.042]	0.327 [0.046]
Difference	-	0.046 [0.033]	0.126*** [0.052]	0.038 [0.048]	-0.071 [0.048]	-0.131*** [0.050]
Panel C. Conditional						
2019	37.936 [3.463]	28.836 [2.340]	27.598 [3.635]	30.752 [2.754]	29.971 [2.855]	44.039 [3.989]
2020	31.438 [2.675]	31.954 [3.334]	63.461 [5.289]	58.632 [5.432]	53.771 [4.503]	33.846 [4.142]
Difference	-6.498*** [2.968]	3.118 [3.844]	35.863** [5.064]	27.881** [5.344]	23.801** [4.981]	-10.193* [6.077]

Note: Unconditional represents the study time for students in grades 7, 8, 10, and 11. Conditional represents the study time conditional on log-in for students in grades 7, 8, 10, and 11. Monthly log-in is defined as studying at least once in a given month. Study time is defined as the weekly average study time within each month. The log-in rate for January is blank because the estimation sample is conditioned on the log-in in January. The symbols, *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Source: Author calculations.

Teachers' Messaging

Not only students but also teachers responded to the school closure. Figure 8 shows the total number of messages sent from school each week. Before the school closure, teachers sent virtually no messages, presumably because they communicated with their students in person. The number of messages after

school closure increased in the months of March, April, and May. The largest increase occurred in the middle of April, whereas study time was longest in March. In fact, as in Table 3 and 4, the effect on overall study time in March is roughly three times larger than that in May, whereas the effect on message is five times less. As March is the end of the school year, students may not have required much support from teachers, given that they mostly review materials taught in class at this time. By contrast, in April, at the beginning of the new school year, students studying new materials may require more assistance from teachers, leading them to send more messages in April.

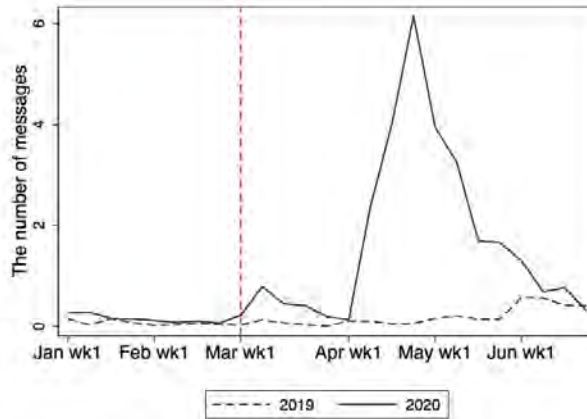


Figure 8. Total Number of Messages from Teachers Per Week

Note: The figure shows the weekly average number of messages from teachers to students per school.

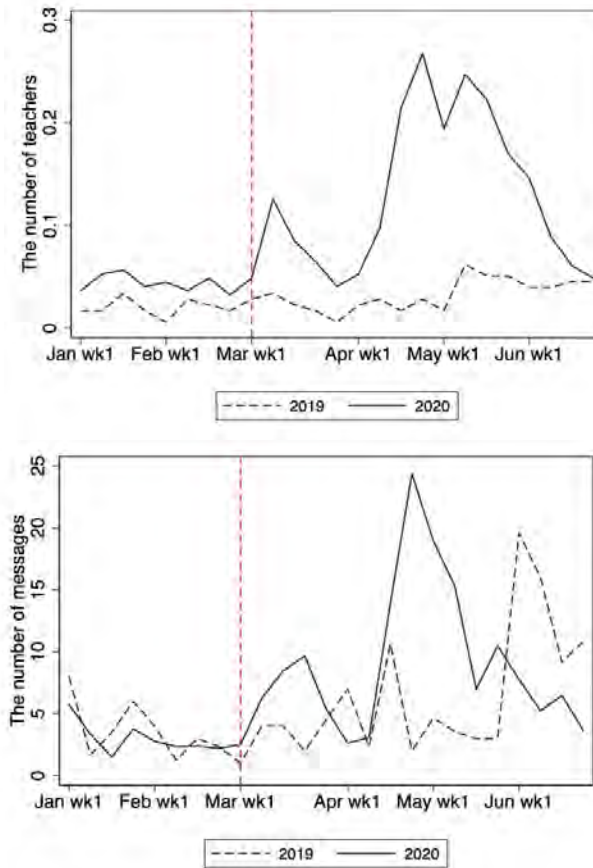


Figure 9. Numbers of Teachers Online (top) and Messages Sent Per Teacher

Note: Teachers are defined as being online if they send at least one message in a given week. The number of messages per teacher is defined as the average number of messages per teacher conditional on being online.

Figure 9 presents the number of teachers online and the number of messages sent per teacher. The upper panel of Figure 9 shows the number of teachers who sent at least one message each week. While there were more teachers online in 2020 in any week, the movements in the number of teachers online are parallel between 2019 and 2020 for January and February. There was a rise in the number of teachers online in the second week of March 2020, with a

further rise in the second week of April when the new academic year began. However, the number of teachers online fell to the pre-COVID-19 level in the second week of June 2020. The lower panel of Figure 9 shows the average number of messages per active teacher online. This number moves in a similar fashion to the number of active teachers online, although the pattern is clearer in the upper panel. The magnitude of the change is largest in April, which is consistent with the effect on the total number of messages shown in Figure 8. Overall, therefore, the changes in the aggregate numbers of messages were driven by both extensive and intensive margins.

TABLE 4—MONTHLY-LEVEL ANALYSES OF MESSAGE DATA

	January	February	March	April	May	June
Panel A.						
Number of Messages						
2019	0.041 [0.018]	0.035 [0.013]	0.037 [0.019]	0.04 [0.017]	0.11 [0.036]	0.453 [0.200]
2020	0.092 [0.024]	0.059 [0.018]	0.288 [0.075]	1.83 [0.868]	1.456 [0.485]	0.278 [0.112]
Difference	0.052* [0.030]	0.024 [0.022]	0.252*** [0.077]	1.79** [0.868]	1.346** [0.487]	-0.175 [0.229]
Panel B.						
Number of Teacher Online						
2019	0.01 [0.004]	0.02 [0.005]	0.01 [0.004]	0.013 [0.005]	0.034 [0.008]	0.034 [0.008]
2020	0.029 [0.006]	0.026 [0.007]	0.045 [0.008]	0.096 [0.026]	0.133 [0.030]	0.043 [0.015]
Difference	0.019** [0.007]	0.006 [0.008]	0.035*** [0.009]	0.084** [0.027]	0.099** [0.031]	0.009 [0.017]
Panel C.						
Messages Per Teacher						
2019	3.417 [0.935]	2.125 [0.713]	3.375 [1.256]	4.5 [1.645]	3.644 [0.788]	7.423 [4.604]
2020	2.414 [0.394]	2.04 [0.236]	5.476 [1.192]	13.899 [7.219]	9.538 [3.507]	4.115 [1.659]
Difference	-1.002 [0.930]	-0.085 [0.665]	2.101 [1.624]	9.399 [7.378]	5.893 [3.553]	-3.308 [5.405]

Note: Number of messages indicates the weekly average number of messages from teachers to students per school. Per school number of teachers online represents the number of teachers who send at least one message in a given period. Messages per teacher represent the average number of messages per teacher conditional on being online. All of these three measures are weekly averages within each month. The symbols, *, **, and *** denote significant at the 10%, 5%, and 1% levels, respectively.

Source: Author calculations.

Correlation between Study Time and Messages

Next, we examine the relationship between students' study time and messages from teachers. One might expect that teachers paying attention to

students by sending messages to them would encourage students to study more. Although we cannot test this hypothesis, we can examine the association between students' study time and teachers' messaging. Note that we exclude outliers, which we define as changes in study time exceeding 146 minutes (95th percentile) or changes in number of messages of more than 1,337 (99th percentile). The results including the outliers are reported in Appendix V.

As Figure 10 shows, we find a positive correlation between changes in the school-level average study time and changes in the number of messages at the school level. The former is calculated by taking the difference in school-level average study time from March to May in 2019 and that in 2020, noting that school level means the average study time per student in each school. Similarly, a change in the number of messages at the school level is measured as the difference between the total number of messages from March to May in 2019 and that in 2020 in each school.

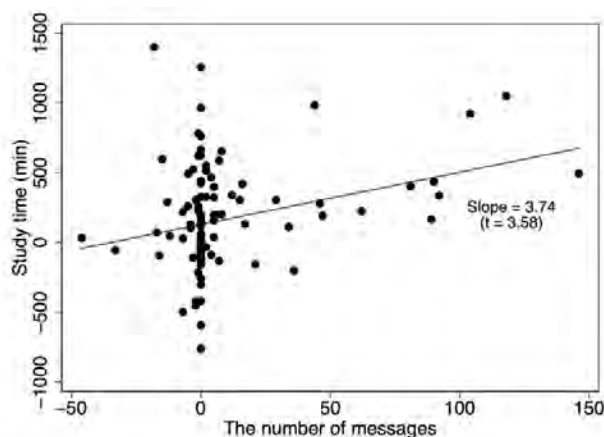


Figure 10. Correlation between Study Time and Messages

Note: Observations are excluded if the change in study time is greater than 146 minutes (the 95th percentile) or if the change in the number of messages exceeds 1,337 (the 99th percentile). The graph with the full sample is provided in Appendix V. The number of observations is 123 and the unit of observation is the school. Study time represents the average study time within each school in March, April, and May 2020. The number of messages represents the total number of messages sent from teachers to students in the same period.

In Figure 11, we observe a positive correlation between changes in study time and changes in the number of teachers online at the school level. We define a teacher being online as a teacher who sent at least one message during the period from March to May in 2019 or during 2020. School level means that we consider the number of teachers online in each school.

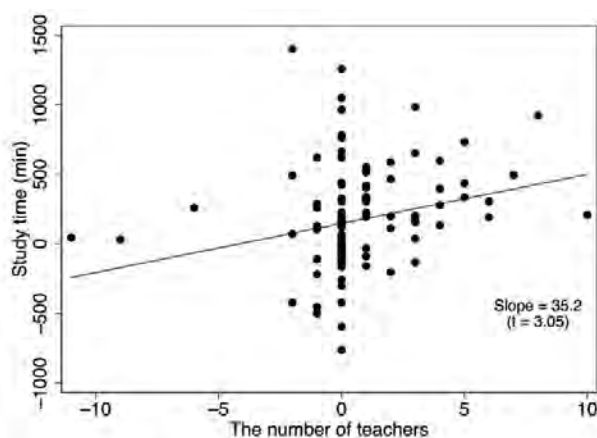


Figure 11. Correlation between Study Time and the Number of Teachers Online

Note: Observations are excluded if they involve a change in study time that exceeds 1,891.319 (99th percentile) or a change in the number of messages that is greater than 11 (the 99th percentile). A graph based on the full sample is provided in Appendix V. The number of observations is 123 and the unit of observation is the school. Study time represents the average study time within each school in March, April, and May. The number of teachers online represents the number of teachers who sent a message at least once in March, April, and May.

Heterogeneity

In this section, we examine the heterogeneous effects of COVID-19 on study time by prior access to the online services from home, school quality, and student gender. In Appendix VI, we report the heterogeneous effects by grade, region, past utilization, and in-class utilization. The analysis compares study time in 2020 between two groups in each case. For instance, we describe

heterogeneity by school quality by comparing the weekly average study time of a high-quality school in 2020 with that of a low-quality school in 2020.

By Prior Access from Home and School Quality

Chetty et al. (2020) find that the study time of students from different income areas varies during school closure. We suspect that this difference arises from variations in internet access at home and study habits. In the following, we examine heterogeneity in prior access to the online study services from home and school quality.

In Figure 12, we compare the study time of students with and without prior access to online services from home. We consider that a student has no prior access to online services from home if he/she never logged in after 8:00 p.m. on weekdays or at any time on the weekend from April 2019 to the end of December 2019. Although the study time of students with prior access to online services from home increased after the school closure, the study time of those with no prior access decreased. As shown in Figure 13, the difference decreased over time and continued to be statistically significant until the beginning of April.

We do not consider the large difference between the two groups at the beginning of January to be problematic for our identification because it is likely to have arisen from the company, SuRaLa Net Co.,Ltd., promoting its services and the New Year holiday. The company's promotion involved the students participating in a tournament, in which they were ranked by study times. Because students who could use the service from home would have found it easier to increase their study time in response to this promotion, they were likely to use the service more than students without access. In addition, as the New Year holiday occurs during the first week of January, students could not access the online service through school facilities at this time. This may explain why students with no prior access from home tend to study less at

the beginning of January. However, except for this period, the trend in study time before the school closure period is similar between the two groups.

In summary, we observe that students with prior access to online services from home utilized the service more under COVID-19 than did students with no prior access from home.

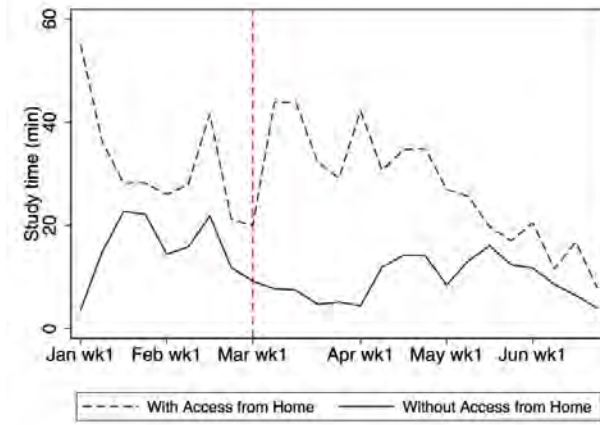


Figure 12. Average Weekly Study Time by Prior Online Learning Access from Home in 2020

Note: The figure shows the study time by prior access to online learning from home for students in grades 7, 8, 10, and 11. Studying at home is defined by accessing the online services after 8:00 p.m. or on the weekend. A student who never studied at home from April to December in 2019 is defined as one with no prior access from home. Note that students who never used the service from April to December in 2019 are excluded.

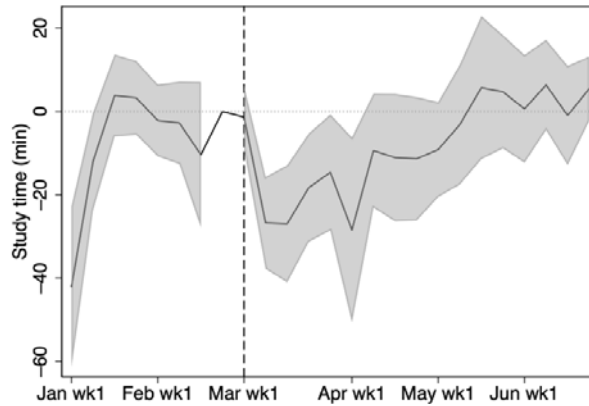


Figure 13. Average Weekly Study Time by Prior Access from Home in 2020

Note: The shaded region is the 95% confidence interval. The figure shows the study time for students in grades 7, 8, 10, and 11 with no prior access to online learning from home minus the study time for students in grades 7, 8, 10, and 11 with access from home. Studying at home is defined by accessing the online services after 8:00 p.m. or on the weekend. A student who never studied at home from April to December in 2019 is defined as one with no prior access from home. Note that students who never used the service from April to December in 2019 are excluded.

Second, we examine students by the quality of their schools. We consider a school to be high quality if its quality index is above the median. Figures 14 and 15 describe heterogeneity with respect to school quality. We find that students from higher-quality schools consistently studied more during the school closure period, although the difference is not statistically significant.

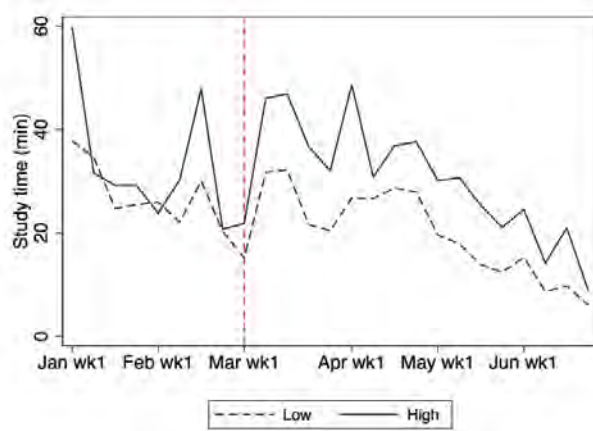


Figure 14. Average Weekly Study Time by School Quality in 2020

Note: The figure shows the study time by school quality for students in grades 7, 8, 10, and 11. School quality is defined by the level of the school, which is obtained from external sources. Low and high describe schools for which the quality level is below and above the median, respectively.

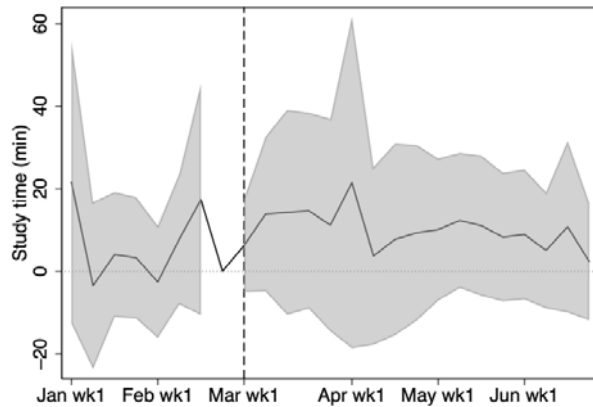


Figure 15. Heterogeneity by School Quality

Note: The shaded region is the 95% confidence interval. The figure shows the study time for students in grades 7, 8, 10, and 11 in high-quality schools minus the study time for students in grades 7, 8, 10, and 11 in low-quality schools. School quality is defined by the level of school, which is obtained from external sources. Low and high describe schools for which the quality level is below and above the median, respectively.

By Gender

Figlio et al. (2013) show that male students tend to struggle with online learning; therefore, we examine any differences in study time between male and female students. Figures 16 and 17 show that there is no statistically significant difference in the average weekly study time between male and female students in 2020.

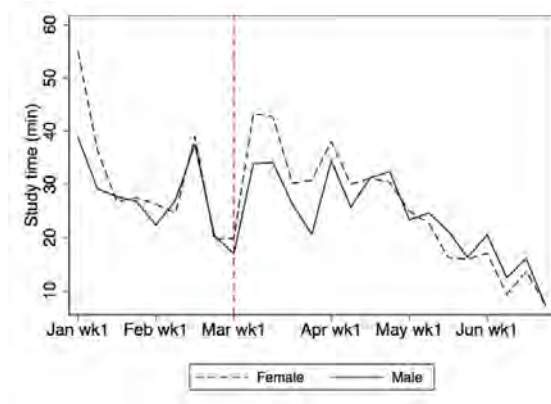


Figure 16. Average Weekly Study Time by Gender in 2020

Note: The figure shows the study time by gender for students in grades 7, 8, 10, and 11.

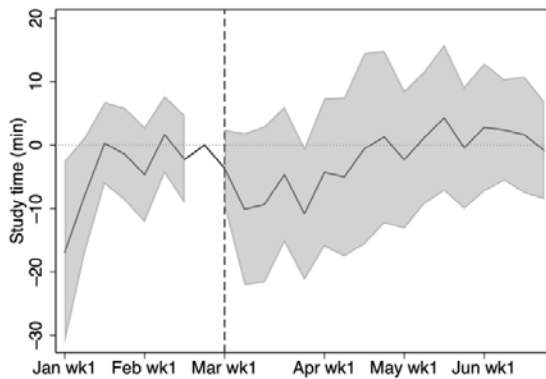


Figure 17. Heterogeneity by Gender

Note: The shaded region is the 95% confidence interval. The figure shows the study time for male students in grades 7, 8, 10, and 11 minus the study time for female students in grades 7, 8, 10, and 11.

We also examine heterogeneity across several other variables, namely grade, region, previous usage, and in-class utilization. For most of these, we find no sizable difference across the two groups; however, the one exception is heterogeneity by grade. In fact, we find a statistically significant difference between high school and junior high school students. In addition, the result concerning heterogeneity by past utilization of the online learning service suggests that the schools with more experience of the service are more likely to use it both before and after the school closure period in 2020. More details of these results can be found in Appendix VI.

VI. Summary and Concluding Remarks

This paper documents the effect of the school closure under COVID-19 on students' study time and teachers' inputs in an online learning service. We find that online study time significantly increased during the school closure and that it returned to the pre-COVID-19 level when the school closure lifted at the end of May 2020. In addition, we find that teachers sent more messages to students via the online service during the school closure than before or after. We note that the effects of school closure are heterogeneous. Specifically, students with access to the online learning service from home and students at higher-quality schools increased their study time more than other students.

Our finding suggests that an online learning service may help students to study during school closure. Thus, the government may want to consider introducing online learning tools in preparation for future possible school closures. Further, we note that policy makers should be aware that a lack of internet and/or personal computer access can raise inequality in learning during school closure.

A potential limitation of our study is that we do not have data on students' learning activities outside the service. For instance, students who studied intensively offline would not have suffered learning losses under COVID-19

school closures, despite not studying via the online learning service. Future research on the impact of COVID-19 on education should complement our results by examining a more comprehensive measure of the impact, such as long-term educational attainments.

References

- Aucejo, E. M., French, J. F., Araya, U. M. P., & Zafar, B. (2020). The impact of covid-19 on student experiences and expectations: evidence from a survey. *Journal of Public Economics*, 191, 104271.
- Bacher-Hicks, A., Goodman, J., & Mulhern, C. (2020). *Inequality in Household Adaptation to Schooling Shocks: Covid-Induced Online Learning Engagement in Real Time* (No. W27555). National Bureau of Economic Research. <https://doi.org/10.7249/WRA811-1>
- Bacolod, M., Mehay, S., & Pema, E. (2018). Who succeeds in distance learning? Evidence from quantile panel data estimation. *Southern Economic Journal*, 84(4), 1129–1145. <https://doi.org/10.1002/soej.12264>
- Bettinger, E. P., Fox, L., Loeb, S., & Taylor, E. S. (2017). Virtual Classrooms: How Online College Courses Affect Student Success. *American Economic Review*, 107(9), 2855–2875. <https://doi.org/10.1257/aer.20151193>
- Chetty, R., Friedman, J. N., Hendren, N., Stepner, M., & The Opportunity Insights Team. (2020). *How Did COVID-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data* (No. W27431). National Bureau of Economic Research. <https://www.nber.org/papers/w27431>
- Cook, W. (2020). School disruption and pupil academic outcomes – evidence from the 2001 foot and mouth disease epidemic in England. *Covid Economics*, 40, 214–228.

- Figlio, D., Rush, M., & Yin, L. (2013). Is it live or is it internet? experimental estimates of the effects of online instruction on student learning. *Journal of Labor Economics*, 31(4), 763–784. <https://doi.org/10.1086/669930>
- Joyce, T., Crockett, S., Jaeger, D. A., Altindag, O., & O’Connell, S. D. (2015). Does classroom time matter? *Economics of Education Review*, 46, 64–77. <https://doi.org/10.1016/j.econedurev.2015.02.007>
- Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the potential impacts of COVID-19 school closures on academic achievement. *EdWorkingPaper*, 20–226. <https://doi.org/10.26300/cdrv-yw05>
- Xu, D., & Jaggars, S. S. (2014). Performance Gaps between Online and Face-to-Face Courses: Differences across Types of Students and Academic Subject Areas. *The Journal of Higher Education*, 85(5), 633–659. <https://doi.org/10.1080/00221546.2014.11777343>

Appendix I

In this Appendix, we describe additional institutional information: more detailed description of the online learning service on which we focus in this paper, and additional comment on Japanese academic calendar.

First, Surala is run by a private company, SuRaLa Net Co.,Ltd., which is independent of any public entities. The company start off, primarily focusing on the business-to-business market, in 2007. In 2012, the company extend its business to the business-to-consumer market.

The service is adopted by various entities in different grades such as high school, elementary school, individuals, and private tutors. In our analysis, we mainly focus on school-level utilization of the service in junior high school and high school. In the case where a school makes a contract with the company, the school ask the company to issue accounts for student. Then, school distributes those accounts to the students and instructs them the way to study in this service. Students can freely login the service any time after distribution of their account. All the materials in subjects under the contract is available to the students.

The service offers study materials in five subjects: Math, English, Japanese, science, and social study. Some subjects are not available to certain grades. For instance, there is no high school-level science and social study materials available.

In the service, three types of functions are available: lectures, drills, and tests. In the lecture part, students watch recorded video and answer brief quizzes prepared by a teacher during the lecture. The drills are offered not only in the form of multiple-choice form as in brief quizzes in the lecture but also in various different ways such as open-ended questions and dictation. Finally, difficulty of the drills is adjusted depending on the level of individual ability.

The platform offers three types of tests. The first test is a general assessment of their academic ability. The second test is a mock exam for their mid-term and final exams. The third test is a short quiz taking around ten minutes. In the second and third test, the scope of topics in the test can be specified so that students can selectively study certain topics.

As of March in 2018, 151 schools, most of which is a private school, are under contract. The price of the service is around 7500 ~ 10,000 yen (roughly 71 ~ 95 USD) per month in individual case depending on the number of subjects and the type of contracts. Although this price should be higher than the price for school-level contract, the budget for technological advancement in public school, for instance less than 10,000 yen (roughly 95 USD) per year in the case of governmental policy in 2014, cannot afford adoption of the service. Furthermore, before COVID-19, Japanese educational system did not allow elementary school and junior high school to incorporate a class solely based on online materials into their official curriculum. In high school, only the limited number of classes based on online materials can be incorporated into their official curriculum.

School uses the service in and out of class. Some school, for example, can use this platform to teach English in class while other school can use it as homework for Math class. We have data on the ways of their usage and the purpose before and after school closure. However, even within schools categorized as the same way of usage in this data, there are quite a little variation in their actual study time. Thus, we do not clearly observe the purpose of usage in each school.

Second, even though nationwide school closure due to COVID-19 began from March 2 and ended in the end of May, the effective number of lost school days is less than three months. This is because there are three events in which schools would not have classes even in the absence of school closure: spring break, transition of academic year, and national holidays.

First, spring break begins typically from the third or fourth week of March and lasts in the first week of April. Second, most schools do not conduct classes for ceremonial and administrative procedures at the end and beginning of an academic year. Finally, in the beginning of May, there are three consecutive national holidays, from Monday to Wednesday. This means that the number of classes hold in this week is less than other weeks. Combining all of these events, roughly 8~9 weeks of classes were not taken place in school due to COVID-19.

Appendix II

The service provider SuRaLa Net Co.,Ltd. conducted a survey for schools using Surala to see how schools changed the ways they use the service during school closure. The survey was responded by a teacher representing a school.

Figure III-1 presents the purposes of using the service. Schools chose one of the following three alternatives: preparation, review, and compensatory education. In pre-COVID-19 period, most (77%) schools used Surala for review, meaning that students study materials that have recently been taught in class. The share of review drops to 47% during school closure, but the share of schools that use Surala for preparation for new materials rose from 4% to 40%. This shift from review to preparation suggests that schools tried to make up for lost class using Surala during school closure. Lastly, the share of compensatory education, meaning that students study materials taught in a lower grade, slightly increased from 13% to 20%.

Figure III-2 shows when students use the service in the day. Before school closure, students are almost equally likely to use the service in either class, after class, or at home. During school closure, students use Surala either in class, in time assigned by teachers, or any time in the day. Vast of majority (89%) of schools did not specify the time when students study using Surala. Their students are free to choose when to use Surala in the day during school closure.

Finally, Figure III-3 presents how students use the service. In most schools, teachers assign units to students both before and during school closure. The share of schools that use Surala for quizzes and review, meaning that schools instruct students first to take a test and to review the mistakes, slightly increased from 4% to 12%. Overall, we do not find significant change in these dimensions of usage between before and during school closure.

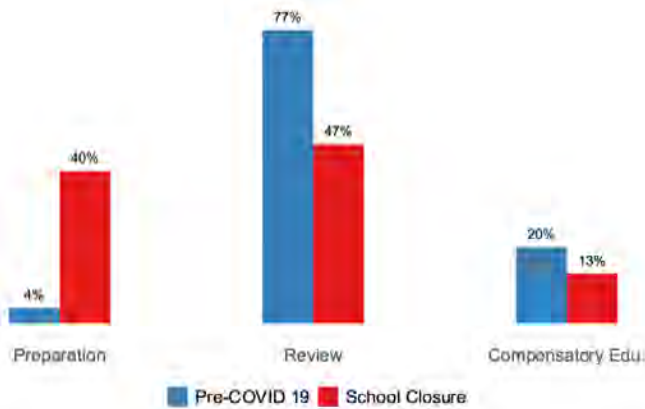


Figure II-1. Purpose of Using Surala

Notes: This graph is created based on the information obtained from the interview to each school. The number of observations for Pre-COVID 19 and School Closure are 188 and 156, respectively. Preparation means the platform is used to let students study materials covered in class before they take classes. Compensatory Edu. Means the platform is used to study topics covered in class long time ago.

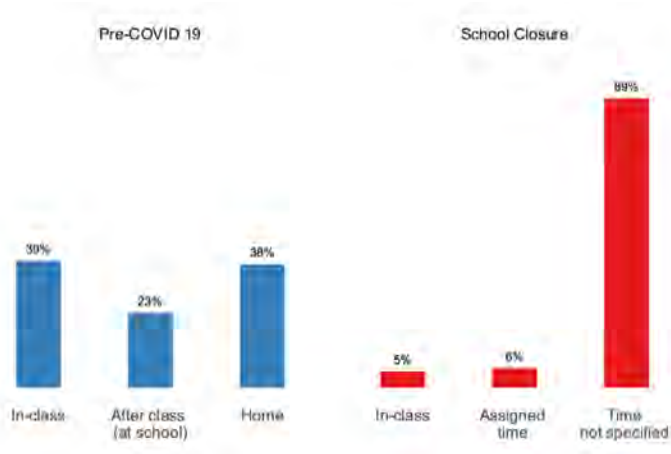


Figure II-2. Occasion Where Schools Use Surala

Notes: This graph is created based on the information obtained from the interview to each school. The number of observations for Pre-COVID 19 and School Closure are 188 and 163, respectively. Home means the platform is used homework. Assigned time means teachers tell students when to study in the platform. Time not specified means students are free to choose when to study in the platform.

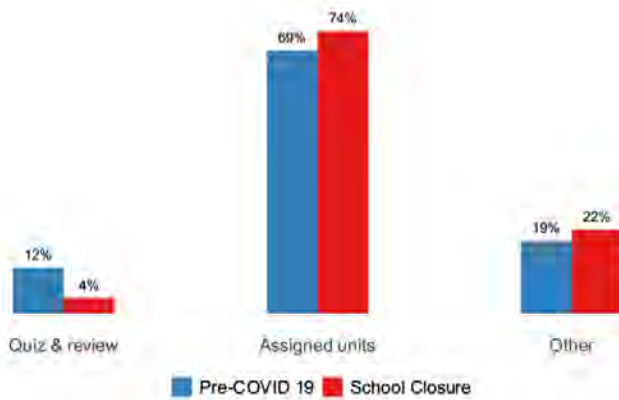


Figure II-3. How to Use Surala

Notes: This graph is created based on the information obtained from the interview to each school. The number of observations for Pre-COVID 19 and School Closure are 188 and 163, respectively. Quiz & review means students first take quiz and study only the part they are not familiar with. Assigned units means teachers assign certain units to students.

Appendix III

In this appendix, we discuss another variable which may capture teachers' input. In the online learning service, teacher can set target for students. The company provide the information of when the target is set, when the target is supposed to be finished, and when the target is actually finished.

Overall, we find target is mostly used in the beginning of winter break, December, and the beginning of academic calendar, April. Even though we observe slight increase after school closure, the amount of increase is quite small comparing to the beginning of winter break and the beginning of academic calendar. Thus, we find it difficult to detect the effect of school closure on this variable except we observe almost no target usage in the beginning of academic calendar in 2020.

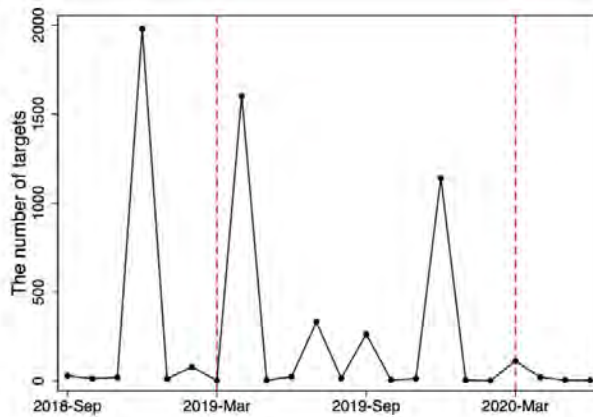


Figure III-1. Weekly average number of targets based on set date

Notes: The unit of observation is school. The date of target is defined by the day when the target is set by a teacher.

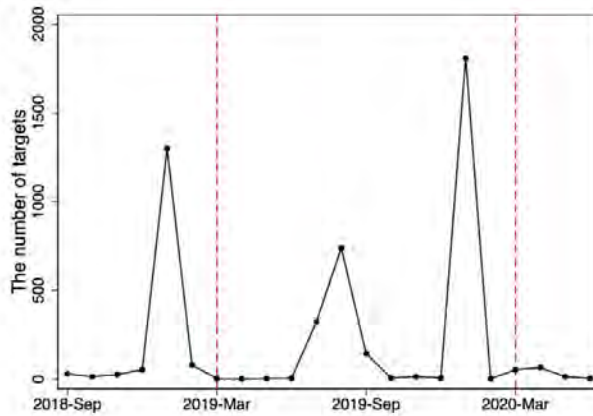


Figure III-2. Weekly average number of targets based on end date

Notes: The unit of observation is school. The date of target is defined by the day when the target is supposed to be completed.

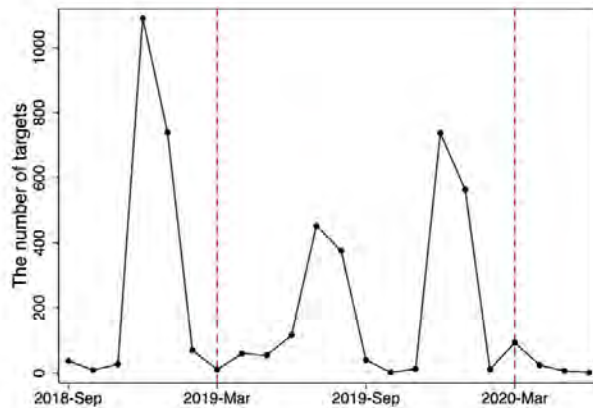


Figure III-3. Weekly average number of targets based on complete date

Notes: The unit of observation is school. The date of target is defined by the day when a student finishes the target set by a teacher. We exclude students who did not complete the assigned target.

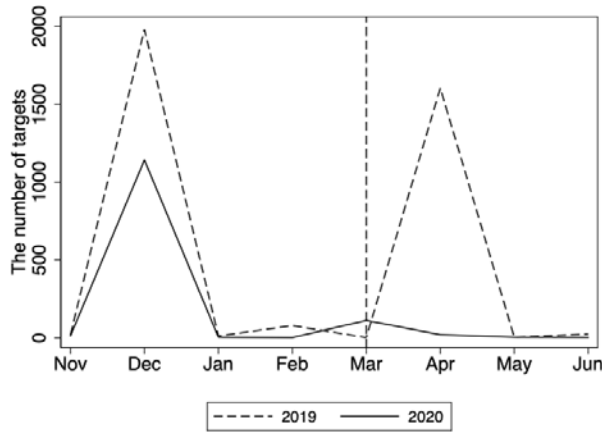


Figure III-4. Difference in weekly average number of targets based on set date

Notes: The unit of observation is school. The date of target is defined by the day when the target is set by a teacher.

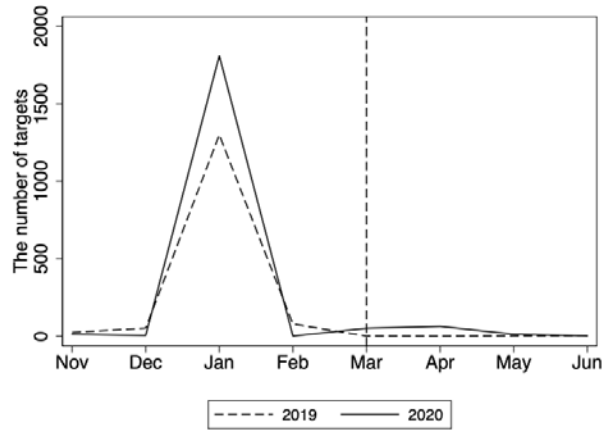


Figure III-5. Difference in weekly average number of targets based on end date

Notes: The unit of observation is school. The date of target is defined by the day when the target is supposed to be completed.

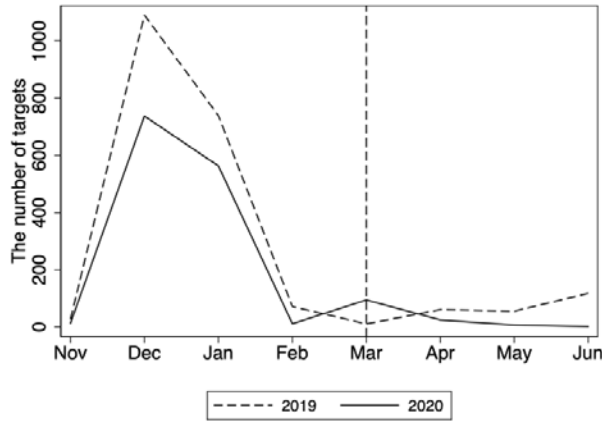


Figure III-6. Difference in weekly average number of targets based on complete date

Notes: The unit of observation is school. The date of target is defined by the day when a student finishes the target set by a teacher. We exclude students who did not complete the assigned target.

Appendix IV

In Section II, we discuss correlation between study time and messaging with the graphs without outliers. In this Appendix, we present the graphs with full sample.

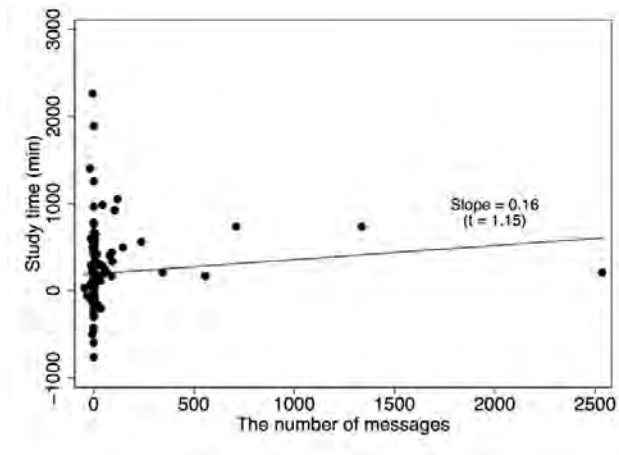


Figure IV-1. Correlation between Study Time and Messages

Notes: The number of observations is 127. The unit of observation is school. Study time represents average study time within each school in March, April, and May. The number of messages represents the total number of messages sent from teachers to students in March, April, and May.

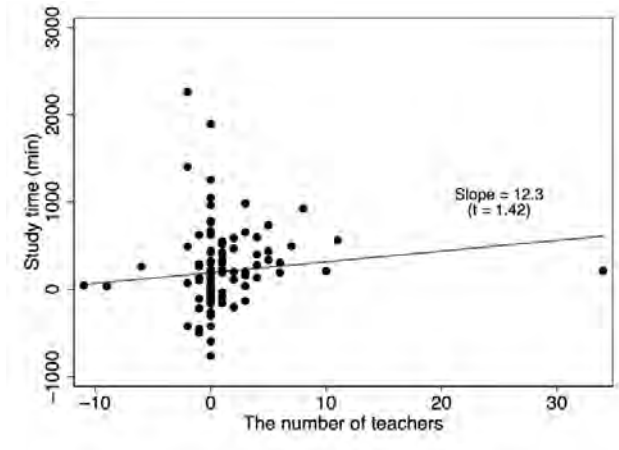


Figure IV-2. Correlation between Study Time and the number of teachers online

Notes: The number of observations is 127. The unit of observation is school. Study time represents average study time within each school in March, April, and May. The number of teachers online represents the number of teachers who use at least once in March, April, and May.

Appendix V

In this Appendix, we report heterogeneity analyses concerning variables not presented in Section V. To be more specific, we present heterogeneous effects by grade, by school quality, by region, by previous usage, and by in-class utilization.

By Grade

We examine heterogeneity of the responses to school closure by comparing the growth of study time from 2019 to 2020 between groups. There are a few reasons why responses may be different between Jr. high school and high school. First, high school and junior high school are different in contents and intensity of study. Second, we expect different selection into our sample. For instance, students enter public junior high school unless they choose to take entrance exam for private junior high school while both private and public high school require entrance exams. Finally, a manager of the company mentioned that students in the junior high schools in our sample tended to have better background than students in the high school. This difference may be particularly salient during the school closure because students cannot use the service without their own computer at home when their schools are closed.

Figure VI-1 shows the change of study time for Jr. high school students, while Figure VI-2 shows that for high school students. For both Jr. high school and high school students, the growth of study time is significant in March and April, although the magnitude is greater for Jr. high school students. We also directly compare the study time in 2020 between groups in Figure VI-3 and VI-4. In April and May, Jr. high school students increase their study time significantly more than high school students do.

A possible explanation behind this heterogeneity is differences in students' background. Japanese compulsory education covers Jr. high school, but not high school. Majority attends public Jr. high school, but students from relatively wealthy families tend to go to private Jr. high school. By contrast,

private institutions are common among high schools and not necessarily for wealthier students. Although we do not observe the wealth of students' families, we expect that Jr. high school students in our sample have better family socioeconomic status than high school students in our sample.

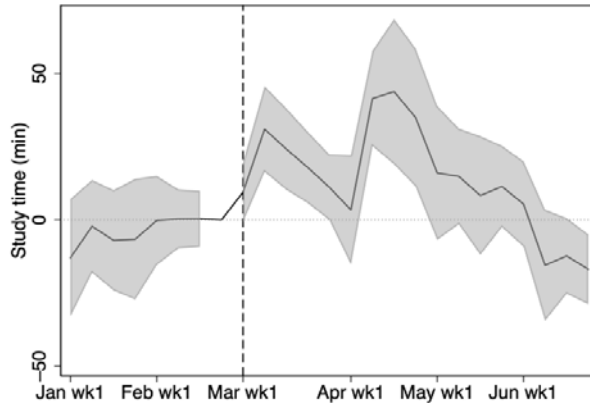


Figure V-1. Change of Weekly Study Time from 2019 to 2020 for Jr. HS

Notes: The shaded region is 95% confidence envelope. Study time for students in grades 7 and 8.

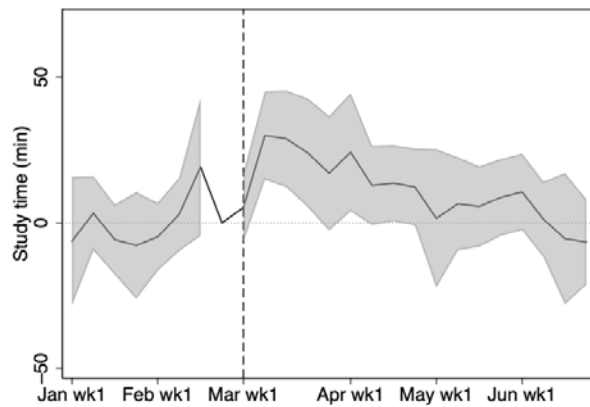


Figure V-2. Change of Weekly Study Time from 2019 to 2020 for HS

Notes: The shaded region is 95% confidence envelope. Study time for students in grades 10 and 11.

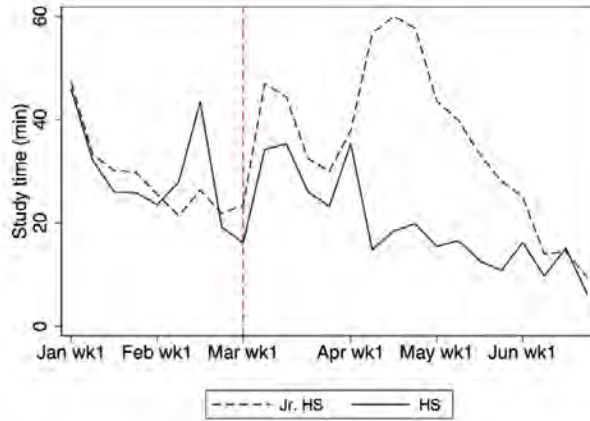


Figure V-3. Average Weekly Study Time by Grade in 2020

Note: Study time for students in grades 7, 8, 10 and 11 in 2020.

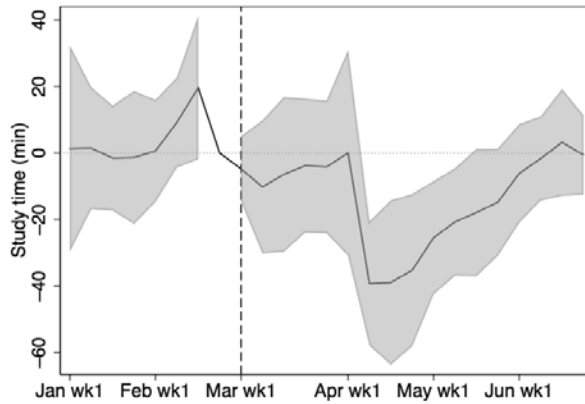


Figure V-4. Difference in the Growth of Study Time Between Jr. HS and HS Students

Notes: The shaded region is 95% confidence envelope. Study time for students in grades 10 and 11 in 2020 minus study time for students in grades 7 and 8 in 2020.

By School Quality

In Figure VI-5 and VI-6, we find different pattern for high school and junior high school even though the pattern is not statistically significant. Only

Covid Economics 58, 19 November 2020: 85-134

heterogeneity across school quality in high school in March is marginally statistically significant.

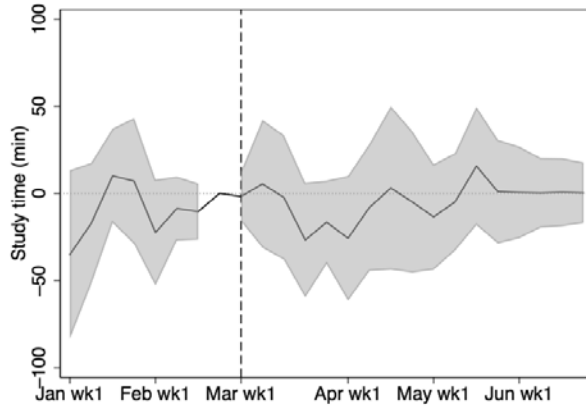


Figure V-5. Heterogeneity by School Quality for Jr. HS

Notes: The shaded region is 95% confidence envelope. Study time for students in grades 7 and 8 in high quality school minus study time for students in grades 7 and 8 in low quality school. School quality is defined by the level of school obtained from external sources.

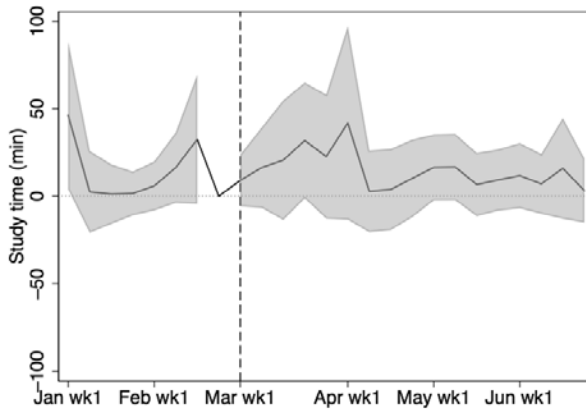


Figure V-6. Heterogeneity by School Quality for HS

Notes: The shaded region is 95% confidence envelope. Study time for students in grades 10 and 11 in high quality school minus study time for students in grades 10 and 11 in low quality school. School quality is defined by the level of school obtained from external sources.

By region

Figure VI-7 and VI-8 shows regional difference in average weekly study time. In this analysis, we define four cities (Tokyo, Aichi, Osaka, Kanagawa) as major cities. The fraction of students in major cities is around 46%. In major cities, we observe a marginally statistically significant difference between two categories. From March to April, students in major city tend to study longer than the other group.

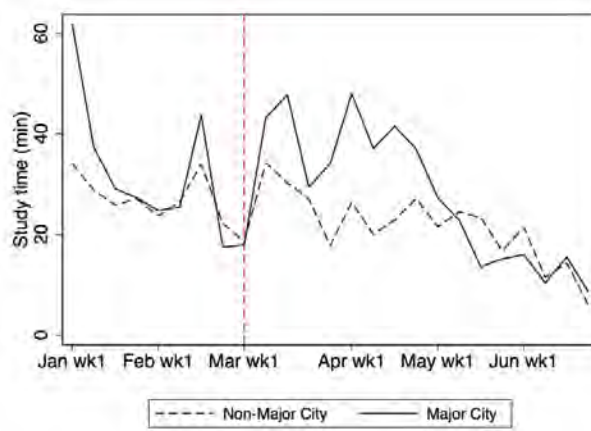


Figure V-7. Average Weekly Study Time by Region in 2020

Notes: Study time by region for students in grades 7, 8, 10, and 11 in 2020. Four cities in Japan, Tokyo, Osaka, Aichi, and Kanagawa, are defined to be major cities.

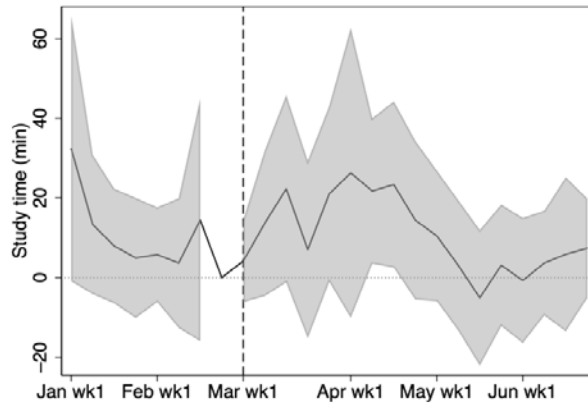


Figure V-8. Heterogeneity by Region

Notes: The shaded region is 95% confidence envelope. Study time for students in grades 7, 8, 10, and 11 in a school located in a major city minus study time for students in grades 7, 8, 10, and 11 in a school located in a non-major city. Four cities in Japan, Tokyo, Osaka, Aichi, and Kanagawa, are defined to be major cities.

By the Previous Usage

We examine heterogeneity with respect to the past utilization, specifically past study time and in-class utilization before school closure. We expect that different experience before COVID-19 may affect utilization during school closure. For instance, schools with large amount of past utilization may be affected by school closure more intensely because they are more familiar with the service. On the other hand, they may have little room for further increase in the amount of utilization.

Figures VI-9, VI-10, VI-11 and VI-12 describe heterogeneity with respect to the past study time. We define utilization category based on the school-level average study time from April 1st to December 31st in 2019. The fraction of students who belong to school with high utilization is around 83%. Overall, schools with higher past utilization consistently study longer in the service after COVID-19, as well as before COVID-19. As in Figure 18, however, there is no statistically significant difference between two group. Figure 15 and 16 show heterogeneity in high school and junior high school separately. In

high school, there is a marginally statistically significant effect in March while, in junior high school, there is a statistically significant effect in April. Both of the pattern observed here is consistent with the pattern observed in the analysis of the heterogeneity across grade. For instance, the effect in high school is strongest in March.

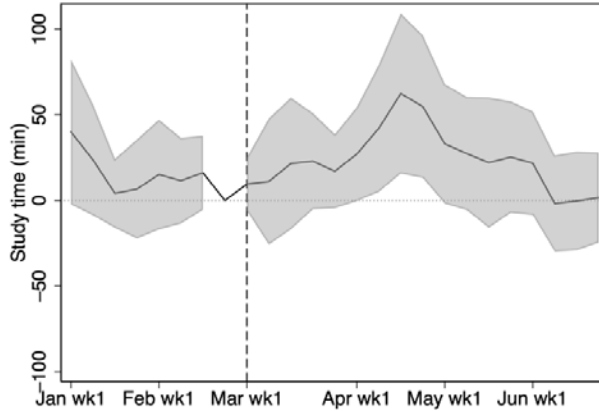


Figure V-9. Heterogeneity by the Previous Usage for Jr. HS

Notes: The shaded region is 95% confidence envelope. Study time for students in grades 7 and 8 in a school with high usage minus study time for students in grades 7 and 8 in a school with low usage. The previous usage is defined by the average study time for students in a given school from April 2019 to December 2019. Low and High are defined by a student belonging to a school with the previous usage below median and above median, respectively.

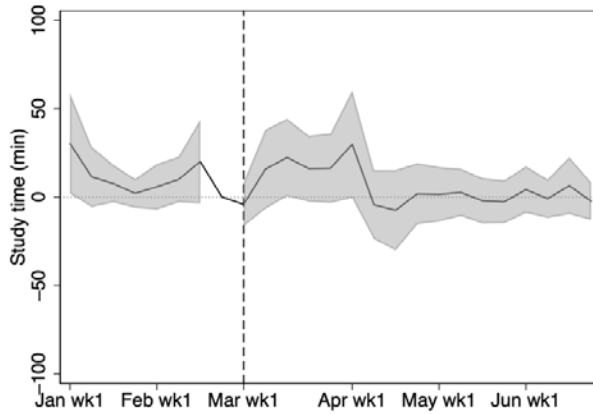


Figure V-10. Heterogeneity by the Previous Usage for HS

Notes: The shaded region is 95% confidence envelope. Study time for students in grades 10 and 11 in a school with high usage minus study time for students in grades 10 and 11 in a school with low usage. The previous usage is defined by the average study time for students in a given school from April 2019 to December 2019. Low and High are defined by a student belonging to a school with the previous usage below median and above median, respectively.

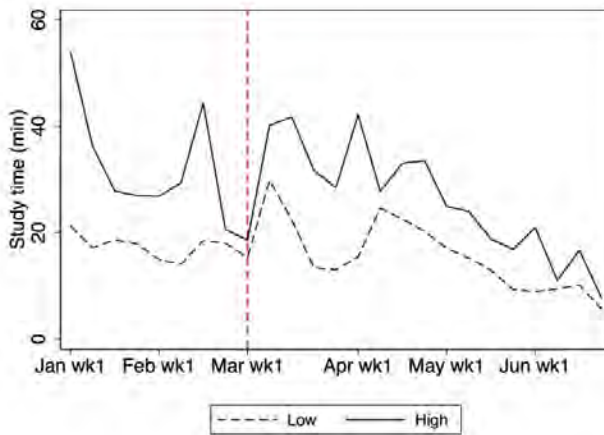


Figure V-11. Average Weekly Study Time by the Previous Usage in 2020

Notes: Study time by the previous usage for students in grades 7, 8, 10, and 11 in 2020. The previous usage is defined by the average study time for students in a given school from April 2019 to December 2019. Low and High are defined by a student belonging to a school with the previous usage below median and above median, respectively.

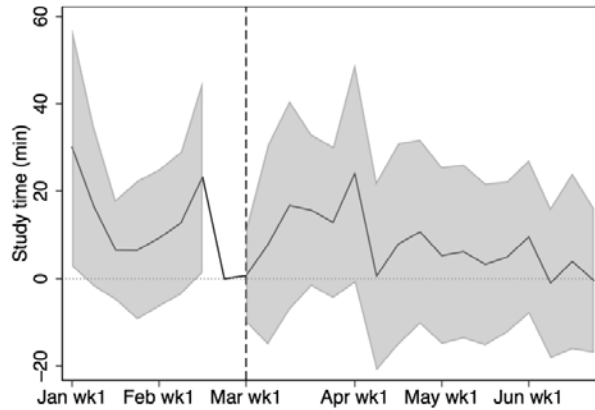


Figure V-12. Heterogeneity by the Previous Usage

Notes: The shaded region is 95% confidence envelope. Study time for students in grades 7, 8, 10, and 11 in a school with high usage minus study time for students in grades 7, 8, 10, and 11 in a school with low usage. The previous usage is defined by the average study time for students in a given school from April 2019 to December 2019. Low and High are defined by a student belonging to a school with the previous usage below median and above median, respectively.

By in-class utilization

Figure VI-13 and VI-14 shows the difference in average weekly study time across the way school utilize the platform before school closure. In-class utilization of the platform before school closure is defined based on the information obtained through a survey. The fraction of students in a school with in-class utilization is around 43%. There is no statistically significant difference between two groups.

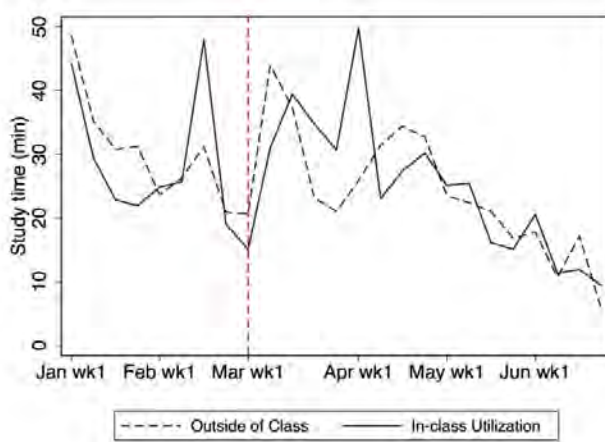


Figure V-13. Average Weekly Study Time by in-class utilization in 2020

Notes: Study time by whether the platform is used in class for students in grades 7, 8, 10, and 11 in 2020.

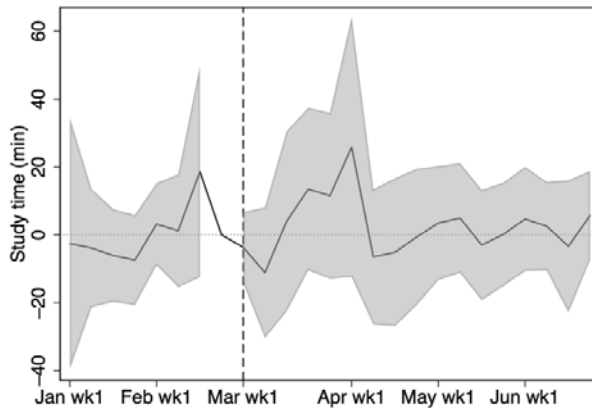


Figure V-14. Heterogeneity by in-class utilization

Notes: The shaded region is 95% confidence envelope. Study time for students in grades 7, 8, 10, and 11 in a school using the platform in class minus study time for students in grades 7, 8, 10, and 11 in a school using the platform outside class.