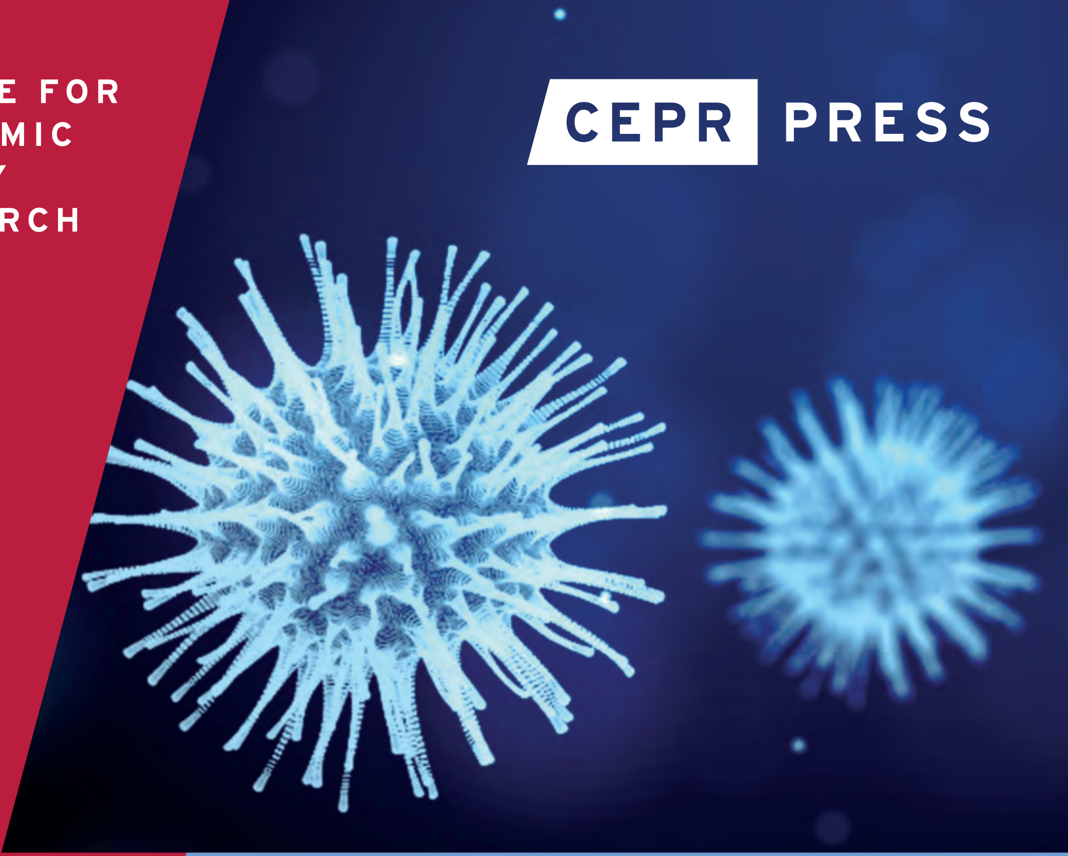
CENTRE FOR
ECONOMIC
POLICY
RESEARCH

CEPR PRESS



COVID ECONOMICS

VETTED AND REAL-TIME PAPERS

ISSUE 33
30 JUNE 2020

Covid Economics
Vetted and Real-Time Papers

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

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

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

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 Botticini
	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
Abi Adams-Prassl, University of Oxford and CEPR
Jérôme Adda, Bocconi University 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
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
Rachel Griffith, IFS, University of Manchester and CEPR
Timothy J. Hatton, University of Essex and CEPR
Ethan Ilzetzki, London School of Economics and CEPR
Beata Javorcik, EBRD and CEPR
Sebnem Kalemli-Ozcan, University of Maryland and CEPR Rik Frehen
Erik Lindqvist, Swedish Institute for Social Research (SOFI)
Tom Kompas, University of Melbourne and CEBRA
Miklós Koren, Central European University and CEPR
Anton Korinek, University of Virginia and CEPR
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
Paul Seabright, Toulouse School of Economics 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.

American Economic Review

American Economic Review, Applied Economics

American Economic Review, Insights

American Economic Review, Economic Policy

American Economic Review, Macroeconomics

American Economic Review, Microeconomics

American Journal of Health Economics

Canadian Journal of Economics

Economic Journal

Economics of Disasters and Climate Change

International Economic Review

Journal of Development Economics

*Journal of Econometrics**

Journal of Economic Growth

Journal of Economic Theory

*Journal of the European Economic Association**

Journal of Finance

Journal of Financial Economics

Journal of International Economics

*Journal of Labor Economics**

Journal of Monetary Economics

Journal of Public Economics

Journal of Political Economy

Journal of Population Economics

*Quarterly Journal of Economics**

Review of Economics and Statistics

*Review of Economic Studies**

Review of Financial Studies

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

Covid Economics
Vetted and Real-Time Papers

Issue 33, 30 June 2020

Contents

CEPR PRESS

Covid Economics **Issue 33, 30 June 2020**

What should we spend to save lives in a pandemic? A critique of the value of statistical life[1]

Matthew D. Adler[2]

Date submitted: 27 June 2020; Date accepted: 27 June 2020

The value of statistical life (VSL) is a risk-to-money conversion factor that can be used to accurately approximate an individual's willingness-to-pay for a small change in fatality risk. If an individual's VSL is (say) $7 million, then she will be willing to pay approximately $7 for a 1-in-1-million risk reduction, $70 for a 1-in-100,000 risk reduction, and so forth. VSL has played a central role in the rapidly emerging economics literature about COVID-19. Many papers use VSL to assign a monetary value to the lifesaving benefits of social-distancing policies, so as to balance those benefits against lost income and other policy costs. This is not surprising, since VSL (known in the U.K. as "VPF": value of a prevented fatality) has been a key tool in governmental cost-benefit analysis for decades and is well established among economists. Despite its familiarity, VSL is a flawed tool for analyzing social-distancing policy—and risk regulation more generally. The standard justification for cost-benefit analysis appeals to Kaldor-Hicks efficiency (potential Pareto superiority). But VSL is only an approximation to individual willingness to pay, which may become quite inaccurate for policies that mitigate large risks (such as the risks posed by COVID-19)—and thus can recommend policies that fail the Kaldor-Hicks test. This paper uses a simulation model of social-distancing policy to illustrate the deficiencies of VSL. I criticize VSL-based cost-benefit analysis from a number of angles. Its recommendations with respect to social distancing deviate dramatically from the recommendations of a utilitarian or prioritarian social welfare function. In the model here, it does indeed diverge from Kaldor-Hicks efficiency. And its relative valuation of risks and financial costs among groups differentiated by age and income lacks intuitive support. Economists writing about COVID-19 need to reconsider using VSL.

1 Many thanks to Maddalena Ferranna, Marc Fleurbaey, Michael Livermore, Lisa Robinson, Cass Sunstein, and Nicolas Treich for comments. The usual disclaimer applies.

2 Richard A. Horvitz Professor of Law and Professor of Economics, Philosophy and Public Policy, Duke University.

Copyright: Matthew D. Adler

Covid Economics 33, 30 June 2020: 1-45

Covid Economics 33, 30 June 2020: 1-45

I. Introduction

Fatality risk regulation is the bread and butter of the U.S. regulatory state. Administrative agencies in the U.S. government are required to prepare written documents accompanying proposed major rules that evaluate the proposed rules using cost-benefit analysis (CBA)[1]; and the monetized benefit of fatality risk reduction, as per these documents, is the largest category of monetized benefit.[2]

In particular, fatality risk reduction is central to the mission of the Environmental Protection Agency (EPA). Anti-pollution rules enacted by this agency, over the last five decades, have yielded large decreases in individuals' annual fatality risks; and much of the monetized fatality-risk-reduction benefit from federal regulation is attributable to the EPA.[3] But other federal agencies also enact regulations that reduce fatality risks and evaluate these regulations by means of CBA, including a monetization of the risk-reduction benefit. These include the Department of Transportation, the Occupational Safety and Health Administration, the Food and Drug Administration, the Department of Health and Human Services, and others.[4]

In general, CBA evaluates governmental policy by predicting, to the extent feasible, a policy's effects on the components of individual well-being (income, health, fatality risk, environmental quality, leisure, etc.) and then measuring these effects on a monetary scale.[5] The value of statistical life (VSL) is, in turn, the linchpin concept for monetizing risk reduction.[6] VSL can be thought of as a conversion factor for translating an individual's risk change into a monetary equivalent. Assume Suzy's VSL is \$7 million. Then Suzy's willingness to pay for a small risk reduction is approximately the reduction multiplied by \$7 million (so that she is willing to pay \$7 for a 1-in-1 million reduction, \$70 for a 1-in-100,000 reduction, and \$700 for a 1-in-10,000 reduction).

Although VSL is, in principle, heterogeneous—Suzy's VSL might well be different from Jana's or Wu's—the practice in the U.S. government is to use a population-average VSL, now on the order of \$10 million.[7] Note that the assumption of a single VSL leads to a simple formula for valuing lifesaving. If a policy reduces expected deaths in the U.S. population next year by ΔD, then the sum total of individual risk reduction is ΔD; and if a single VSL is being used, the

[1] Executive Order No. 12,866, 3 C.F.R. 638 (1994).
[2] Office of Management and Budget (2017, p. 11).
[3] *Id*. at 10-16; Robinson (2007).
[4] See Viscusi (2018, ch. 2); Robinson (2007).
[5] See, e.g., Adler (2019); Adler and Posner (2006).
[6] See generally Viscusi (2018).
[7] On the U.S. government's use of population-average VSL, see Robinson (2007); U.S. Department of Health and Human Services (2016); U.S. Department of Transportation (2016); U.S. Environmental Protection Agency (2010). For the \$10 million figure, see Viscusi (2018, p. 28).

aggregate willingness to pay for the policy's risk-reduction benefit is just $\Delta D \times \text{VSL}$. This calculation now occurs in numerous agency cost-benefit documents.

Other governments also employ VSL in governmental cost-benefit analysis, including the U.K. The concept is the same, but the terminology may differ. In the U.K., the term "value of a prevented fatality" is used instead of "value of statistical life," and the corresponding abbreviation is "VPF" rather than "VSL."[8]

Although U.S. governmental monetization of risk policy via VSL was, initially, controversial, that controversy has largely faded, and VSL is quite firmly entrenched in U.S. regulatory practice. It is also very well established in the academic literature. Many articles in applied economics employ behavioral evidence, in particular wage differentials for riskier occupations, or stated-preference evidence, to estimate VSL.[9] There is also a large theory literature about VSL.[10]

The COVID-19 pandemic has brought VSL to the fore. Among the flood of academic working papers posted in immediate response to the pandemic, quite a number employ VSL to estimate the benefits of social-distancing requirements.[11] Many opinion pieces and news articles have also discussed VSL.[12] This is not unexpected. Since VSL is now a cornerstone of U.S. governmental risk analysis, and the focus of decades of academic work, it is hardly surprising that academic and public discussion about COVID-19 policy would resort to VSL.

But VSL—despite its familiarity—is flawed. VSL is flawed as a component of administrative routines for evaluating proposed regulations. It gives us a misleading picture, for example, of the benefits of EPA's anti-pollution regulations or the Department of Transportation's safety requirements. VSL is also a flawed basis for thinking about COVID-19 policy. Or so I shall argue here.

The deficits are the same in the two cases. VSL's shortcomings as applied to COVID-19 policy are no different from its shortcomings in evaluating anti-pollution rules, vehicle safety requirements, workplace safety regulations, and so forth. In the years before the pandemic,

Covid Economics 33, 30 June 2020: 1-45

[8] The U.K. also uses a considerably lower value for VSL than the U.S. See Viscusi (2018, p. 38) on international differences in VSL.
[9] For overviews of the empirical literature on VSL, see Aldy & Viscusi (2007); Cropper, Hammitt and Robinson (2011); Kniesner and Viscusi (2019); Krupnick (2007); OECD (2012); Viscusi and Aldy (2003); Viscusi (2018).
[10] See, e.g., Eeckhoudt and Hammitt (2001); Evans and Smith (2010); Hammitt (2000, 2007); Hammitt et al. (2020); Johansson (2002); Jones-Lee et al. (2015).
[11] See, e.g., Alvarez, Argente and Lippi (2020); Bairoliya and Imrohoroğlu (2020); Barnett-Howell and Mobarak (2020); Béland, Brodeur and Wright (2020); Bethune and Korinek (2020); Greenstone and Nigam (2020); Gros et al. (2020); Pretnar (2020); Robinson, Sullivan and Shogren (2020); Scherbina (2020); Thurnstrom et al. (2020); Ugarov (2020); Wilson (2020). But see Pindyck (2020), criticizing the use of VSL; Hall, Jones and Klenow (2020), using VSL to calibrate a utilitarian social welfare function.
[12] See, e.g., Corcoran (2020); Henderson and Lipow (2020); Hilsenrath and Armour (2020) Ingraham (2020); Jenkins (2020); Masur and Posner (2020); Sunstein (2020).

various scholars, including myself, published careful academic critiques of VSL.[13] These critiques didn't prompt intensive academic or public conversation, because fatality risk regulation—in the years before the pandemic—wasn't generally the topic of intensive academic and public conversation. Certainly it was not for the first two decades of the 21st century, up until the terrible coronavirus outbreak of 2020. The topics on the front burner were instead, e.g., terrorism, globalization, inequality, climate change, and democratic breakdown.

But the pandemic's death toll, and the massive economic losses of shutdowns and social distancing, force us to think about balancing lives against livelihood. Risk policy is now at the center of our conversations. And so, too, we should use this moment to think about how we think about risk policy. Is VSL really the best we can do? No, it is not—or so I'll claim.[14]

I'll consider three different versions of VSL: (1) textbook VSL, (2) population-average VSL, and (3) the "value of statistical life year" (VSLY). Textbook VSL is as presented in the academic literature on CBA and VSL. VSL, here, is an individual's marginal rate of substitution between survival probability and money, and varies among individuals. As mentioned above, U.S. governmental agencies depart from VSL and instead monetize VSL with a constant VSL, specifically a population average. VSLY is a construct proposed by some scholars as a way to circumvent the flaws of textbook and population-average VSL.[15]

Various kinds of policy choices have bedeviled governments during the COVID-19 pandemic, but perhaps the most important and salient such choice concerns social distancing policy. Measures to reduce close physical proximity between individuals in the population can be expected to lower the fatality (and also morbidity) impacts of the pandemic, but at the cost of lost income as well as losses with respect to other dimensions of well-being. This Article's strategy will be to illustrate the flaws of the three versions of VSL—textbook, population-average, and VSLY—with reference to social-distancing policy. (As already mentioned, the academic papers that have appeared in response to the pandemic and that employ VSL mainly do so with respect to social distancing policy.)[16]

My main benchmark for evaluating VSL will be *utilitarianism*. Utilitarianism has been a dominant school of ethical thought for hundreds of years, since the work of Jeremy Bentham.

---

[13] See Adler (2016a); Adler (2017); Adler (2019, ch. 5); Adler, Hammitt and Treich (2014); Adler, Ferranna, Hammitt and Treich (2019); Broome (1978); Bronsteen, Buccafusco and Masur (2013); Broughel (2020); Dolan et al. (2008); Grüne-Yanoff (2009).

[14] See also Hammitt (2020), critically assessing the use of VSL to evaluate social-distancing policy.

[15] On VSLY, see Aldy and Viscusi (2007); Hammitt (2007); Hammitt et al. (2020); Jones-Lee et al. (2015); Kniesner and Viscusi (2019); Viscusi (2018, ch. 5).

[16] A second type of policy choice that was much discussed early in the pandemic was triage—in particular, how to allocate lifesaving equipment to seriously ill COVID-19 patients (namely, ventilators) under scarcity. The academics engaged in that debate were principally health ethicists, who tend not to favor CBA; and thus CBA and VSL played little (if any) role in their analyses. Still, VSL-based CBA is a global methodology for assessing *any* kind of policy involving fatality risk and certainly could be applied to triage. Thus applied, it would be no less problematic than VSL-based CBA applied to social distancing. So as not to try the reader's patience in belaboring the difficulties of the three types of VSL with respect to both social distancing and triage, I focus here on the former.

Covid Economics 33, 30 June 2020: 1-45

Covid Economics 33, 30 June 2020: 1-45

Utilitarianism has endured through generations of academic and public debates about ethics, and continues to play a vibrant role at both levels. To be sure, utilitarianism has hardly gone unchallenged. It has endured over centuries of ethical *debate*. In recent years, the chief opponents of utilitarianism within academic ethics have been non-consequentialists, in particular "contractualists" and deontologists. However, I'm going to leave aside non-consequentialism and focus on a consequentialist critique of VSL—specifically, a critique from the vantage point of utilitarianism.[17]

Utilitarianism and CBA are *not* the same. There are key, structural, dissimilarities between the two that can yield divergent recommendations in many different policy domains, including fatality risk policy.

A secondary benchmark for my assessment of VSL will be *prioritarianism*.[18] Prioritarianism is a variation on utilitarianism that has emerged in ethics over the last several decades, and (as suggested by the name) gives extra weight to the well-being of the worse off.[19] It is appealing to those who are impressed by utilitarianism's consequentialist structure and attention to individual well-being, but believe that utilitarianism's exclusive focus on the sum total of well-being is too narrow. Instead, prioritarianism considers *both* the sum total *and* the distribution of well-being.

In developing my critique of VSL-based CBA from the standpoint of utilitarianism and prioritarianism, I will rely upon a simulation model that is built upon the U.S. population survival curve and income distribution, and that will be used as a testbed for balancing the benefits and costs of social distancing policy.

The standard defense of CBA appeals to Kaldor-Hicks efficiency (potential Pareto superiority).[20] Assume that CBA picks policy option P^*, as opposed to a second option P—which could be the status quo, or alternatively the option chosen by a competing policy framework such as utilitarianism or prioritarianism. Then there is in principle a change ΔT to the tax-and-transfer system such that P^* together with ΔT is Pareto superior to P. ΔT would take the form of increased taxes upon (or lowered transfer payments to) those better off with P^* than P, and increased transfer payments to (or lowered taxes upon) those better off with P than P^*. In recent years, work by Louis Kaplow has extended the Kaldor-Hicks defense of CBA, by showing

[17] Non-consequentialists have struggled to formulate a systematic account of the ethics of risk. For a recent attempt, see Frick (2015). Utilitarianism experiences no such difficulty. Further, CBA itself is consequentialist. It will be especially illuminating, I hope, to set forth a critique of VSL that is not merely a rehash of the criticisms that non-consequentialists level against all forms of consequentialism—but instead is a critique from within the consequentialist camp, namely from the standpoint of utilitarianism and, secondarily, prioritarianism.

[18] Utilitarianism and prioritarianism are operationalized for policy assessment as social welfare functions (SWFs). On utilitarian and prioritarian SWFs, see generally Adler (2019). On these SWFs applied specifically to risk regulation, see Adler (2016a); Adler (2017); Adler (2019, ch. 5); Adler, Hammitt and Treich (2014); Adler, Ferranna, Hammitt and Treich (2019).

[19] Parfit (2000) is the seminal philosophical text on prioritarianism. For defenses, and overviews of the subsequent philosophical literature, see Adler (2012, ch. 5); Holtug (2010, 2017).

[20] See Adler and Posner (2006), discussing and criticizing this defense.

that ΔT exists not merely in a lump-sum tax system but even in a system (such as ours) where taxes are levied on income and thus increased taxes have a distortionary effect.[21]

The Kaldor-Hicks defense of CBA has been vigorously challenged. These challenges are set forth in a substantial academic literature debating the Kaldor-Hicks criterion.[22] Perhaps the strongest challenge is this: the change ΔT to the tax system is a purely *potential* change. Either ΔT is *actually* implemented together with P^*, in which case *any* Pareto-respecting assessment methodology (e.g., utilitarianism) will prefer P^* plus ΔT to P; or ΔT is *not* implemented together with P^*, in which case some are worse off with the policy and the Pareto criterion is not applicable. Indeed, in actual U.S. practice, administrative rules and new statutory provisions are rarely coupled with tax-and-transfer changes designed to compensate those who are made worse off; and regulatory agencies are instructed to engage in CBA regardless of whether they anticipate such a compensatory change.

For purposes of this Article, however, I will place to one side these well-known critiques of Kaldor-Hicks efficiency. Assume that the reader *accepts* the Kaldor-Hicks criterion. Should she, therefore, accept VSL-based CBA as a tool for evaluating fatality risk regulation and, more specifically, COVID-19 policy?

No. VSL-based CBA can *deviate* from the Kaldor-Hicks criterion. An individual's VSL, multiplied by the change in fatality risk that would result from a given policy, is only an *approximation* to her willingness to pay for that change. This approximation becomes poorer as the change becomes larger. Thus CBA with textbook VSL, used to assess policies that involve significant changes in individuals' risks, can readily favor policies that are *not* Kaldor-Hicks efficient relative to alternatives that CBA disprefers. And CBA with population-average VSL or VSLY can deviate from Kaldor-Hicks efficiency even for small changes in individuals' risks. We'll see, specifically, that all three approaches assign a positive score to a range of social distancing policies, preferring these policies to the status quo—even though the policies are not Kaldor-Hicks efficient relative to the status quo.

The structure of the Article is as follows. Part II provides the conceptual framework. Part III uses the simulation model based upon the U.S. survival curve and income distribution to illustrate the implications of textbook VSL (as defined in Part II) with respect to social distancing policy, as compared to utilitarianism and prioritarianism, and to assess whether VSL conforms to the Kaldor-Hicks criterion. Part IV does the same for CBA using population average VSL, and Part V for VSLY.

The aim of the Article, it should be stressed, is *not* to provide guidance with respect to social distancing policy. Rather, its aim is *methodological*: to argue that VSL-based CBA provides flawed guidance. The simulation model is used as a diagnostic tool, to bring to light the

[21] See Kaplow (1996; 2004; 2008).
[22] See Adler (2012, pp. 98-104), citing sources. For a response to Kaplow, see Adler (2019, pp. 225-33); Adler (2017).

Covid Economics 33, 30 June 2020: 1-45

deficiencies of VSL-based CBA—to demonstrate how the guidance it provides is different from that of utilitarianism, prioritarianism, and Kaldor-Hicks efficiency.

II. The Concept of VSL

Roughly speaking, VSL is an individual's marginal rate of substitution between money (income, consumption, wealth) and survival probability. It captures how money and survival trade off in terms of the individual's utility.

Providing a more precise definition of VSL depends on how utility is conceptualized. The conceptual framework I'll be using in this paper, a discrete-time and multi-period model, is as follows. An individual's life is divided into periods numbered 1, 2, … T, with T, the maximum number of periods that any individual lives. An individual either dies immediately after the beginning of the period or does not die and survives, at least, until the end of the period.

The number of the current period, for individual i, is A_i. (For example, if periods are years, and individual i is currently age 20, then $A_i = 21$.)

In the status quo, individual i has a lifetime profile of survival probabilities $(p_i^{A_i},..., p_i^T)$. These are conditional probabilities: namely, p_i^t is the probability that individual i survives to the end of period t, given that she is alive at the beginning.

Individual i also has a status quo lifetime profile of income and non-income attributes. $((y_i^1, b_i^1),...,(y_i^{A_i-1}, b_i^{A_i-1}),(y_i^{A_i}, b_i^{A_i}),...,(y_i^T, b_i^T))$. If t is a past period ($t < A_i$), y_i^t is the income that individual i earned in period t and b_i^t denotes the non-income attributes that she had in that period. If t is the current period or a future period ($t \geq A_i$), then y_i^t is the income which i will earn in period t if she survives to its end, and b_i^t is the bundle of non-income attributes which she'll have in that period if she survives to its end.

For short, let ($\mathbf{p}_i$, $\mathbf{y}_i$, $\mathbf{b}_i$) denote individual's status quo lifetime profile of survival probabilities, incomes, and non-income attributes.

Individual i has a lifetime utility function. Let $U_i(\cdot)$ denote expected utility. Individual i's status quo expected utility, at present, is
$U_i((p_i^{A_i}, p_i^{A_i+1},..., p_i^T),((y_i^1, b_i^1),...,(y_i^{A_i}, b_i^{A_i}),...,(y_i^T, b_i^T)))$ or, for short, $U_i(\mathbf{p}_i, \mathbf{y}_i, \mathbf{b}_i)$.

Note that the framework is quite general. Periods can be arbitrarily long or short. Non-income attributes of any sort (health, leisure, happiness, public goods, etc.) can be included in the period bundles. An individual's consumption in each period might be set equal to her income; or, instead, the individual might be supposed to have access to intertemporal financial markets, allowing her to save and to borrow against anticipated income, and to engage in saving and borrowing in either a perfectly or imperfectly rational manner. Finally, the lifetime utility function might be temporally additive or non-additive.

Covid Economics 33, 30 June 2020: 1-45

Assume that individual i's survival probability in period s is increased, relative to the status quo, by Δp_i^s. Period s is either the current period or a future period ($s \geq A_i$). Let $m(\Delta p_i^s)$ denote the reduction to individual i's current status quo income that just suffices to make her indifferent to this improvement in survival probability.[23] Then VSL_i^s , individual's i's VSL with respect to changes in survival probability in period s, is the limit of the ratio of $m(\Delta p_i^s)$ to Δp_i^s: $VSL_i^s \equiv \lim_{\Delta p_i^s \to 0} \dfrac{m(\Delta p_i^s)}{\Delta p_i^s}$. Equivalently, VSL_i^s is the marginal rate of substitution between survival probability in period s and current income.

(1) $$VSL_i^s = \frac{\partial U_i / \partial p_i^s}{\partial U_i / \partial y_i^{A_i}} \Big|_{(\mathbf{p}_i, \mathbf{y}_i, \mathbf{b}_i)}$$

The numerator in this fraction is the status quo marginal utility of period s survival probability ; the denominator is the status quo marginal utility of current income. VSL_i^s is the ratio of these marginal utilities.[24]

CBA is standardly defined as the sum of individual "compensating variations" or "equivalent variations."[25] A given governmental policy is some departure from the status quo. An individual's "compensating variation" for the policy is the change to her current income with the policy—the current income that she would have, were the policy to be implemented—that would just suffice to make her indifferent between the policy and the status quo. An individual's "equivalent variation" for the policy is the change to her current status quo income that would just suffice to make her indifferent between the policy and the status quo.[26]

VSL is a useful construct, for purposes of CBA, because it can be used to define an individual monetary valuation for a policy that is a good approximation to her compensating

[23] In other words, if s is the current period, $m(\Delta p_i^s)$ is such that
$U_i((p_i^{A_i} + \Delta p_i^S, p_i^{A_i+1}, ..., p_i^T), ((y_i^1, b_i^1), ..., (y_i^{A_i} - m(\Delta p_i^s), b_i^{A_i}), ..., (y_i^T, b_i^T))) = U_i((p_i^{A_i}, ..., p_i^T), ((y_i^1, b_i^1), ..., (y_i^{A_i}, b_i^{A_i}), ..., (y_i^T, b_i^T)))$
. If s is a future period, $m(\Delta p_i^s)$ is such that:
$U_i((p_i^{A_i}, ..., p_i^s + \Delta p_i^s, ..., p_i^T), ((y_i^1, b_i^1), ..., (y_i^{A_i} - m(\Delta p_i^S), b_i^{A_i}), ..., (y_i^T, b_i^T))) = U_i((p_i^{A_i}, ..., p_i^T), ((y_i^1, b_i^1), ..., (y_i^{A_i}, b_i^{A_i}), ..., (y_i^T, b_i^T)))$

[24] In one-period models, VSL is often defined as the marginal rate of substitution between survival probability and *wealth* rather than income. See, e.g., Eeckhoudt and Hammitt (2001). However, in a multiperiod model in which individuals have exogenous incomes in each period, it is more straightforward to define VSL in terms of this exogenous attribute rather than to construct an endogenous wealth attribute and then define VSL in terms of that. See Adler et al. (2019). If individuals are modelled as behaving myopically (consuming income and other resources available in each period, rather than saving and borrowing), then income equals wealth except perhaps in the first period (an inheritance). If individuals save and borrow, then VSL defined in terms of income will reflect opportunities to save and borrow.

Further, the sum of compensating variations whether defined in terms of income or wealth signals a potential Pareto improvement (Kaldor-Hicks efficiency), and thus VSL whether defined in terms of income or wealth coheres with the traditional justification for CBA (namely that VSL for small changes is a good approximation to the compensating variation).

[25] See, e.g., Adler and Posner (2006); Boadway (2016); Freeman (2003, ch. 3).

[26] See Appendix D for formulas for the equivalent and compensating variations.

variation or equivalent variation—*if* the policy is a small change from the status quo. In general, I'll use the symbol "MV" to denote this VSL-based monetary valuation of a policy: one that well approximates the individual's compensating and equivalent variations if the policy is a small change from the status quo. And I'll use the term "VSL-CBA" to mean the version of CBA that assesses policies by summing these MV values.

In particular, imagine that a policy increases individual i's period s survival probability by Δp_i^s. Then $\text{MV}_i = (\Delta p_i^s)VSL_i^s$. Alternatively, imagine that a policy increases individual i's period s survival probability by Δp_i^s and increases her current income by $\Delta y_i^{A_i}$. Then $\text{MV}_i = (\Delta p_i^s)VSL_i^s + \Delta y_i^{A_i}$.[27]

These formulas show why VSL can be thought of as a conversion factor that translates risk changes into monetary equivalents. If a policy changes i's period s survival probability by Δp_i^s, the monetary equivalent for that risk change is $(\Delta p_i^s)VSL_i^s$. This monetary equivalent is added to the monetary equivalents for the other policy impacts on individual i to determine her overall monetary valuation (MV_i) of the policy. In the simple case where the policy only changes i's period s survival probability, MV_i is equal to the monetary equivalent of the change: $(\Delta p_i^s)VSL_i^s$. Consider next a policy that changes i's current income by $\Delta y_i^{A_i}$. MV_i in this case is nothing other than $\Delta y_i^{A_i}$.[28] Finally, in the case where a policy changes both an individual's survival probability and her current income, we have the formula $\text{MV}_i = (\Delta p_i^s)VSL_i^s + \Delta y_i^{A_i}$. The first term is the monetary equivalent for the risk change Δp_i^s, i.e., the risk change multiplied by VSL_i^s; the second term is the monetary equivalent for the income change $\Delta y_i^{A_i}$, which is, trivially, just $\Delta y_i^{A_i}$. MV_i is the sum of these two.

Several other points about VSL are worth noting at this juncture. First, VSL is heterogeneous. An individual's VSL with respect to a current or future period depends upon her age, current and future (and perhaps past) income, and current and future (and perhaps past) non-income attributes. These can all, of course, vary among individuals, and so VSL can vary as well.

Second, an individual's MV is only an approximation to her compensating variation or equivalent variation, and this approximation may well be quite poor for policies that produce a large change in the individual's survival probabilities or non-income attributes. In particular, the component of MV for valuing a change to an individual's period s survival probability—$(\Delta p_i^s)VSL_i^s$—reflects the marginal impact of income on the individual's expected utility. The

[27] Similar formulas can be defined for policies that change non-income attributes, but because the model of social distancing below will focus on changes to risks and incomes, I won't spell out such formulas here.
[28] Note that MV_i in this simple case is not merely an approximation for the compensating and equivalent variation, but indeed exactly equal to both.

individual's willingness to pay for a non-marginal reduction in risk, or her willingness to accept in return for a non-marginal increase in risk, will not generally be equal to or even well approximated by $(\Delta p_i^s)VSL_i^s$ if her marginal utility of income is not constant and Δp_i^s is sufficiently large.

For an illustration of this point, consider a simple version of the conceptual framework here in which each individual is endowed with a lifetime vector of survival probabilities and incomes. The period length is one year; non-income attributes are ignored; and the individual consumes her income each period. Using this model, and the most recent available U.S. survival curve, I calculated equivalent variations, compensating variations, and MV values for a 50 year old individual earning \$60,000/year with an additive lifetime utility function equaling the sum of log income.[29] I assumed that the individual's current-year survival probability was as per the survival curve (.99597), minus 0.01. In short, in the status quo the individual's ordinary current-year survival probability has been lowered by 1% by virtue of some significant, short-term fatality risk (such as a pandemic). Under all the assumptions just stated, the individual's VSL for a current-year change in survival probability is \$9,081,485.

Table 1 displays the individual's equivalent variations, compensating variations, and MVs for a policy that improves her survival probability by amounts ranging from 1 in 1 million to 1%. (MV is simply \$9,081,485 times the risk change.) As Table 1 shows, MV becomes an increasingly poor approximation to the equivalent or compensating variation as the risk change from the policy increases. At the 1% level (a policy that wholly eliminates the acute current-year risk), the MV is \$90,814.85, while her equivalent variation is more than twice that amount, and her compensating variation roughly half.

Table 1: Equivalent Variations, Compensating Variations, and VSL-based Monetary Valuations (MVs)

Probability Change	*Equivalent Variation*	*Compensating Variation*	*MV*	*Diff. between Equiv. Variation and MV as % of MV*	*Diff. between Comp. Variation and MV as % of MV*
1 in 1 million	\$9.08	\$9.08	\$9.08	.01%	.01%
1 in 100,000	\$90.88	\$90.75	\$90.81	.08%	.08%
1 in 10,000	\$915.06	\$901.22	\$908.15	.76%	.76%
1 in 1000	\$9804.79	\$8419.70	\$9081.48	7.96%	7.29%
5 in 1000	\$67,885.45	\$31,742.11	\$45,407.42	49.50%	30.09%
1 in 100	\$212,578.16	\$46,590.54	\$90,814.85	134.08%	48.70%

The third and final observation about VSL worth noting at this juncture is closely related to the second. Because an individual's MV is only an *approximation* to her equivalent and

29 Temporal additivity and having utility be logarithmic in income are both quite standard assumptions in economics. Appendix D gives explicit formulas for equivalent and compensating variations applicable to the example here.

Covid Economics 33, 30 June 2020: 1-45

compensating variations, VSL-CBA—the sum of MVs—need not conform to Kaldor-Hicks efficiency. A policy may be ranked above the status quo by VSL-CBA and yet not be Kaldor-Hicks efficient relative to the status quo.

To continue with the above example: Imagine that there are a certain number of at-risk individuals in the population similarly situated to the individual above; and that, for each such individual, there are 1000 potential cost-bearers who aren't at risk but would bear the cost of any policy to reduce or mitigate the risk. Each at-risk individual has an MV of $90,814.85 for a policy that eliminates the 1% risk, but a compensating variation of only $46,590.54. She is willing to pay no more than $46,590.54 to eliminate the risk. Imagine, now, that the per-capita cost of a policy to eliminate the risk is between $46.59 and $90.81. Such a policy is approved by VSL-CBA (it has a positive sum of MVs), but is *not* Kaldor-Hicks efficient. There is no scheme of compensating payments from the at-risk individuals to the cost-bearers which, if adopted together with the policy, makes everyone better off than the status quo.

The conceptual framework set out here can also be used to highlight the difference between VSL-CBA and utilitarianism. While VSL-CBA ranks policies according to the sum of individuals' MV_i values, utilitarianism assigns each policy a score equaling the sum of individuals' expected utilities and ranks policies in the order of these scores.[30] Let ΔU_i denote the difference in i's expected utility between a given policy and the status quo. Utilitarianism says that the policy is better than the status quo if the sum of these ΔU_i values is positive; by contrast, VSL-CBA says that the policy is better than the status quo if the sum of the MV_i values is positive.

To illustrate the difference, consider a policy that changes i's period s survival probability by small amount Δp_i^s and her current income by a small amount $\Delta y_i^{A_i}$. In this case:

(2) $$\Delta U_i \approx (\Delta p_i^s)\frac{\partial U_i}{\partial p_i^s}|_{(\mathbf{p}_i,\mathbf{y}_i,\mathbf{b}_i)} + (\Delta y_i^{A_i})\frac{\partial U_i}{\partial y_i^{A_i}}|_{(\mathbf{p}_i,\mathbf{y}_i,\mathbf{b}_i)}$$

The first term in this equation is the approximate change in expected utility that results from the risk change; it is Δp_i^s multiplied by individual i's status quo marginal utility of period s survival probability. The second term is the approximate change in expected utility that results from the income change; it is $\Delta y_i^{A_i}$ multiplied by the status quo marginal utility of current income.

---

[30] I focus in this Article on utilitarianism understood as the sum of individuals' von-Neumann/Morgenstern (vNM) utilities. $U_i(\cdot)$ is a vNM utility function. If individuals have homogeneous preferences with respect to lifetime bundles of income and non-income attributes, and lotteries over these bundles, then utilitarianism is implemented by choosing any one of the vNM utility functions representing the common preferences and setting $U_i(\cdot)$ to be this function. If individuals have heterogeneous preferences, then we take a vNM utility function representing each preference in the population and rescale it using scaling factors. In this case, $U_i(\cdot)$ is the rescaled vNM utility function representing the preferences of individual i. See Adler (2019, ch. 2 and Appendix D); Adler (2016b).

Covid Economics 33, 30 June 2020: 1-45

Consider now the formula for MV$_i$. Recalling that VSL is the ratio of the marginal utilities of survival probability and income, we have:

(3) $$MV_i = \Delta p_i^s \frac{\partial U_i / \partial p_i^s}{\partial U_i / \partial y_i^{A_i}} |_{(\mathbf{p}_i, \mathbf{y}_i, \mathbf{b}_i)} + \Delta y_i^{A_i}$$

Comparing the formulas for ΔU_i and MV$_i$, two differences emerge. (1) While the ΔU_i formula multiplies Δp_i^s by the marginal utility of period *s* survival probability, the MV$_i$ formula does the same but then divides by the marginal utility of current income. Consider, for example, two individuals facing the very same small change in current survival probability. The relative valuation of these two changes, according to utilitarianism, depends upon the individuals' relative marginal utilities of current survival probability. By contrast, according to VSL-CBA, the relative valuation of the two changes depends upon these marginal utilities of survival probability *and* the two individuals' relative marginal utilities of current income. (2) While the ΔU_i formula multiples $\Delta y_i^{A_i}$ by marginal utility of current income, the MV$_i$ formula simply adds $\Delta y_i^{A_i}$ without adjustment. This difference illustrate why utilitarianism is sensitive to income distribution,[31] while VSL-CBA is not. Imagine that a policy redistributes a small Δy in current income from individual *i* to individual *j*. VSL-CBA sees this policy as a wash: MV$_i$ decreases by Δy, while MV$_j$ increases by Δy, and thus the sum of these monetary valuations does not change. By contrast, the relative magnitudes of ΔU_i and ΔU_j depend upon the individuals' relative marginal utilities of current income. If individual *i* has lower marginal utility of income than *j*, ΔU_i will be smaller in magnitude than ΔU_j, and so utilitarianism will see the income redistribution as an improvement.

Finally, the difference between *prioritarianism* and VSL-CBA can also be illustrated in the conceptual model being deployed here. Prioritarianism sums expected *transformed* utilities, using a *concave* transformation function (so as to give greater weight to the worse off).[32] Let ΔG_i denote the difference in *i*'s expected transformed utility between a given policy and the status quo; the prioritarian score for the policy is the sum of these ΔG_i values. Consider once more a policy that changes *i*'s period *s* survival probability by small amount Δp_i^s and her current income by a small amount $\Delta y_i^{A_i}$. In this case:

(4) $$\Delta G_i \approx (\Delta p_i^s) \frac{\partial G_i}{\partial p_i^s} |_{(\mathbf{p}_i, \mathbf{y}_i, \mathbf{b}_i)} + (\Delta y_i^{A_i}) \frac{\partial G_i}{\partial y_i^{A_i}} |_{(\mathbf{p}_i, \mathbf{y}_i, \mathbf{b}_i)}$$

[31] See Robinson, Hammitt and Zeckhauser (2016), documenting inattention to distribution in agency cost-benefit analyses of environmental, health and safety regulations.
[32] This is so-called "ex post" prioritarianism, as opposed to "ex ante" prioritarianism, a different version of prioritarianism under uncertainty. This Article focuses on the "ex post" approach. Ex ante prioritarianism violates a stochastic dominance axiom, and for that reason is quite problematic. See generally Adler (2019, chs. 3-4).

While the formula above for MV$_i$ values a risk change (Δp_i^s) by multiplying by the term $\frac{\partial U_i}{\partial p_i^s}$ and then dividing by a marginal income utility term $\frac{\partial U_i}{\partial y_i^{A_i}}$, the prioritarian formula instead multiplies by the term $\frac{\partial G_i}{\partial p_i^s}$—this captures the effect of the risk change on expected transformed utility – and does not divide by a marginal income utility term. Further, like utilitarianism, and unlike VSL-CBA, prioritarianism is sensitive to the distribution of income. The income change for an individual, $\Delta y_i^{A_i}$, is multiplied by a term $\frac{\partial G_i}{\partial y_i^{A_i}}$, capturing the effect of income on expected transformed utility.

III. VSL and Pandemic Policy: A Simulation Model

A. *A Simulation Model*

In order to compare VSL with utilitarianism and prioritarianism in assessing pandemic policy, I construct a simulation model based upon the U.S. population survival curve and income distribution. The population is divided into five income quintiles (denoted as "low," "moderate," "middle," "high," and "top"). I assume that an individual remains in the same quintile her entire life. An individual, at birth, is endowed with a survival curve and lifetime income profile. The survival curve for the middle quintile is the U.S. population survival curve, while the curves for the other quintiles are adjusted up or down to match observed differences in life expectancy by income.

The lifetime income profile, for a given quintile, is constructed as follows. Based upon U.S. governmental data, I estimate average after-tax-and-transfer individual income, by quintile, to be as follows: \$21,961; \$30,118; \$41,349; \$57,538; and \$134,840. An individual's income at a given age is then scaled up or down from this baseline amount, in accordance with data concerning the time path of earnings.

I assume that the current population consists of seven age groups: individuals aged 20, 30, 40, 50, 60, 70, and 80. Each age group, in turn, has five income quintiles. Thus the population consists of 35 age-income cohorts. The proportions of the population in the seven age bands are chosen to match the actual age distribution of the current U.S. population.

Although in reality the COVID-19 pandemic increases individuals' fatality risks over multiple years, I simplify the analysis by assuming that this risk manifests itself as an increased fatality risk only in the current year. The baseline current-year survival probability for each age-income cohort is taken from the survival curve for her income quintile, *except* that her survival chances have been made worse by the COVID-19 pandemic. The baseline (status quo) captures the case in which the population is faced with the pandemic and government does nothing to

Covid Economics 33, 30 June 2020: 1-45

CEPR PRESS

address it; social distancing policies will, then, reduce the COVID-19 risk, at some cost to individual incomes.

The COVID-19 risk for each age group is taken from the now-famous Imperial College report of March 2020, whose alarming estimates of the potential deaths from the virus triggered the adoption of shutdown policies in the U.K. and U.S. This report estimated infection fatality ratios (IFRs) by age group and also estimated that 81% of the population in each country would be infected absent governmental intervention.[33] Using the Imperial College IFRs by age and the 81% infected estimate, I assigned the current year COVID-19 risks given in Table 2 to the seven age groups.

Table 2: Risk of Dying from COVID-19 Absent Policy to Suppress or Mitigate Pandemic

Age group	*Baseline (No Intervention) Risk of Dying from COVID-19 in Current Year*
20	.024%
30	.065%
40	.122%
50	.486%
60	1.782%
70	4.131%
80	7.533%

I ignore non-income attributes and assume that an individual's consumption in each year equals her income. I assume a common, temporally additive, logarithmic, lifetime utility function: utility is the sum of the logarithm of consumption each year, minus a term to reflect the subsistence level of consumption.

In this setup, the utilitarian value of a policy is the sum of individuals' expected utilities (according to the common utility function just described). For prioritarianism, I use an Atkinson social welfare function with a moderate level of priority to the worse off, γ (the priority parameter) = 1.5.

More details about the simulation model, the lifetime utility function, and the utilitarian and prioritarian social welfare functions are provided in the Appendix.

VSLs for the 35 age-income cohorts, for a current-year change in survival probability, were calculated using the above assumptions about survival curves and lifetime income profiles for each of the cohorts and the common lifetime utility function, and are displayed in Table 3:

[33] Ferguson et al. (2020).

Covid Economics 33, 30 June 2020: 1-45

Table 3: Cohort VSLs

	Income: Low	*Moderate*	*Middle*	*High*	*Top*
Age: 20	$1,379,256	$2,181,761	$3,391,051	$5,240,842	$15,369,839
30	$3,163,107	$5,005,939	$7,784,849	$12,026,744	$35,235,573
40	$3,254,671	$5,191,073	$8,122,805	$12,593,803	$37,122,979
50	$2,454,637	$3,967,977	$6,274,797	$9,791,316	$29,183,439
60	$1,494,663	$2,458,852	$3,943,065	$6,207,337	$18,786,610
70	$755,476	$1,272,810	$2,080,198	$3,313,544	$10,231,995
80	$330,329	$570,613	$951,330	$1,533,270	$4,826,901

The average VSL across the 35 cohorts (weighted for the different proportions of the age groups) is $8,635,355. This is reasonably consistent with estimates of the U.S. population average VSL, typically in the range of $10 million. Observe also that the pattern of VSL in Table 3 has two features observed in much empirical work: VSL increases with income;[34] and its time profile has a "hump" shape—first increasing and then decreasing with age.[35]

The VSL information in Table 3 can be expressed in a different way. Consider an increment Δp in a given cohort's current-year fatality risk (increment meaning either a reduction or increase). The social value of that increment, according to VSL-CBA, is $\text{MV}_c = (\Delta p)\text{VSL}_c$. Let c^+ be some reference cohort. Then the relative value of an increment in fatality risk for cohort c, as compared to the value for the reference cohort, is just $\text{VSL}_c/\text{VSL}_{c+}$. Table 4a expresses the VSL information in this fashion, with an 80-year old, low income group as the reference cohort.

Table 4a: Social value of risk increment according to VSL-CBA
(relative to social value of risk increment for 80-year-old, low-income cohort)

	Income: Low	*Moderate*	*Middle*	*High*	*Top*
Age: 20	4.2	6.6	10.3	15.9	46.5
30	9.6	15.2	23.6	36.4	106.7
40	9.9	15.7	24.6	38.1	112.4
50	7.4	12.0	19.0	29.6	88.3
60	4.5	7.4	11.9	18.8	56.9
70	2.3	3.9	6.3	10.0	31.0
80	1.0	1.7	2.9	4.6	14.6

For example, Table 4a has the number 15.7 in the cell for the 40-year-old, moderate income cohort. This indicates that the MV for a given risk increment Δp accruing to a member of the 40-

[34] See, e.g., Kniesner and Viscusi (2019); Viscusi (2018, ch. 6).
[35] See, e.g., Aldy and Viscusi (2007); Viscusi (2018, ch. 5).

year-old, moderate-income cohort is 15.7 times the MV for the very same risk increment Δp accruing to a member of the 80-year-old, low-income cohort. The entry in the cell for the 20-year-old, top-income cohort is 46.5. So the MV for a risk increment of Δp to this cohort is 46.5 times that for the same increment to an 80-year-old, low-income individual; and 46.5/15.7 = 2.96 times that for the same increment to a member of the 40-year-old, moderate income cohort.

Tables 4b and 4c display the analogous information for utilitarianism and prioritarianism.[36] These show the relative social value of an increment Δp in each cohort's current-year fatality risk, again with 1 indicating its value to an 80-year-old low-income individual, according to utilitarianism (Table 4b) and prioritarianism (Table 4c). The social value of a risk increment to a cohort member as per utilitarianism is not the individual's MV, but rather the change in the individual's expected utility; the social value as per prioritarianism is the change in the individual's expected transformed utility.

Table 4b: Social value of risk increment according to utilitarianism
(relative to social value of risk increment for 80-year-old, low-income cohort)

	Income: Low	*Moderate*	*Middle*	*High*	*Top*
Age: 20	8.6	9.9	11.2	12.5	15.6
30	7.4	8.6	9.7	10.8	13.5
40	5.9	6.9	7.9	8.8	11.0
50	4.4	5.2	6.0	6.7	8.6
60	3.0	3.6	4.3	4.8	6.2
70	1.9	2.3	2.7	3.1	4.2
80	1.0	1.3	1.6	1.8	2.5

Table 4c: Social value of risk increment according to prioritarianism
(relative to social value of risk increment for 80-year-old, low-income cohort)

	Income: Low	*Moderate*	*Middle*	*High*	*Top*
Age: 20	31.5	29.6	27.9	26.3	23.2
30	19.7	18.6	17.7	16.7	14.8
40	11.9	11.4	11.0	10.5	9.5
50	7.0	6.9	6.8	6.5	6.0
60	4.0	4.0	4.0	4.0	3.7
70	2.1	2.2	2.3	2.2	2.2
80	1.0	1.1	1.1	1.1	1.1

[36] The numbers in tables 4a, 4b and 4c as well as 5a, 5b, and 5c are rounded to 1 decimal place (except for prioritarian income numbers, in Table 5c, for top-income cohorts)

Covid Economics 33, 30 June 2020: 1-45

As emerges from these tables, the relative values of risk increment assigned by VSL are quite different from those assigned by either utilitarianism and prioritarianism.[37] Within each income quintile, the utilitarian value of a risk increment decreases monotonically as individuals get older; saving a younger individual produces a larger gain in life expectancy and thus (within each income quintile) a larger gain in expected utility. Prioritarian values also decrease monotonically—yet more sharply than utilitarian values, reflecting priority to younger cohorts (who are worse off with respect to expected lifetime utility.) By contrast, the VSL-based values do not decrease monotonically with age. Rather, they have the classic "hump" shape. VSL-based values reflect not merely the gain to expected utility, but also the expected marginal utility of income—which changes over time because income changes over time within each quintile, first increasing with age and then decreasing. (See Table 6.)

Within each age band, the utilitarian value of a risk increment increases moderately with income (social values for top-income individuals are roughly twice those of low-income individuals of the same age). Because the utility of each year alive increases with income, the utilitarian value of extending a higher-income individual's life by a year is greater than that of extending a lower-income individual's life by a year. VSL-based values also increase with income in each age band, but much more dramatically than utilitarian values. VSL-based values are more highly skewed to the rich than utilitarian values, because the denominator of VSL is expected marginal utility of income—which decreases within each age band as income increases. Prioritarian values generally decrease or stay constant with income, reflecting priority for the worse off: lower-income individuals within an age band are worse off with respect to lifetime utility.

Tables 5a, 5b, and 5c display the same information on the income side. These show the relative social value of a \$1 increment in a cohort's current income, again with an 80-year-old, low-income cohort as the reference cohort. The number in each cell is the social value of a \$1 income increment, as a multiple of the value of a \$1 income increment for an 80-year-old, low-income individual.[38]

[37] See Adler (2019, ch.5), for similar tables showing the relative risk valuations of utilitarianism, prioritarianism, and CBA, and for a longer discussion of why the three methodologies have different such valuations.

[38] Because the effect of an income increment on an individual's expected utility and expected transformed utility is non-linear, the relative value of income increments for the cohorts displayed in Tables 5b and 5c depends upon the magnitude of the increment (here, \$1). MV values are linear in income increments, and so the constant pattern in Table 5a will hold true for any increment, not merely \$1.

By contrast, it can be shown that the effect of a current-year change in survival probability on individuals' expected utility and expected transformed utility *is* linear. Thus the relative values stated in Tables 4b and 4c are independent of the size of the increment. MV too, by construction, is linear in risk reduction, and thus the relative values in Table 4a (cohort VSLs, normalized by dividing by the VSL of the 80-year-old, low-income cohort) are also independent of the size of the increment.

Covid Economics 33, 30 June 2020: 1-45

Covid Economics 33, 30 June 2020: 1-45

Table 5a: Social value of \$1 income increment according to VSL-CBA
(relative to social value of \$1 increment for 80-year-old, low-income cohort)

	Income: Low	*Moderate*	*Middle*	*High*	*Top*
Age: 20	1	1	1	1	1
30	1	1	1	1	1
40	1	1	1	1	1
50	1	1	1	1	1
60	1	1	1	1	1
70	1	1	1	1	1
80	1	1	1	1	1

Table 5b: Social value of \$1 income increment according to utilitarianism
(relative to social value of \$1 increment for 80-year-old, low-income cohort)

	Income: Low	*Moderate*	*Middle*	*High*	*Top*
Age: 20	2.1	1.5	1.1	0.8	0.3
30	0.8	0.6	0.4	0.3	0.1
40	0.6	0.4	0.3	0.2	0.1
50	0.6	0.4	0.3	0.2	0.1
60	0.7	0.5	0.4	0.3	0.1
70	0.8	0.6	0.4	0.3	0.1
80	1.0	0.7	0.5	0.4	0.2

Table 5c: Social value of \$1 income increment according to prioritarianism
(relative to social value of \$1 increment for 80-year-old, low-income cohort)

	Income: Low	*Moderate*	*Middle*	*High*	*Top*
Age: 20	3.2	1.8	1.1	0.7	0.20
30	1.1	0.6	0.4	0.2	0.07
40	0.8	0.5	0.3	0.2	0.05
50	0.7	0.4	0.3	0.2	0.05
60	0.8	0.5	0.3	0.2	0.05
70	0.9	0.5	0.3	0.2	0.06
80	1.0	0.6	0.4	0.2	0.08

The VSL-based social value of an income increment is constant (the number 1 is in every cell.) Recall Equation 3 above. An increment Δy to an individual's income produces the very same change to her MV (namely, by Δy), regardless of who receives it. By contrast, utilitarian values decrease with income within each age band; this reflects the diminishing marginal utility of income. Prioritarian values decreases yet more dramatically, reflecting both the diminishing marginal utility of income and priority for the worse off.

Given the fairly stark differences between the relative valuations of risk increment and income increment, as per VSL-CBA, and those relative valuations as per utilitarianism and prioritarianism, it is not surprising that the recommendations of VSL-CBA with respect to social distancing policy are different from those of the latter two methodologies—as we'll now see.

B. *Social Distancing Policy*

The current-year incomes of the various cohorts are displayed in Table 6. To model social distancing policy, I assume that an 80% reduction in "GDP" (the sum of total current income across the cohorts) will completely eliminate the current-year COVID-19 risk (as given in Table 2).[39]

Table 6: Cohort Incomes

	Income: Low	*Moderate*	*Middle*	*High*	*Top*
Age: 20	$8,331	$11,425	$15,686	$21,827	$51,152
30	$22,098	$30,306	$41,607	$57,896	$135,680
40	$28,426	$38,984	$53,522	$74,476	$174,536
50	$28,681	$39,334	$54,003	$75,145	$176,103
60	$24,930	$34,189	$46,939	$65,316	$153,069
70	$19,719	$27,043	$37,128	$51,664	$121,075
80	$14,757	$20,238	$27,784	$38,662	$90,605

I posit a concave function from GDP reduction to the reduction in COVID-19 risk.[40] Specifically, I set the COVID-19 risk reduction to be the square root of the GDP reduction, scaled so that a 80% reduction wholly eliminates the risk. (See Appendix.). A concave rather than linear or convex function is employed on the assumption that GDP reductions have a diminishing marginal impact on COVID-19 risk. The first 1% GDP loss produces a larger drop in COVID-19 risk than the second 1%, and so on.

The assumptions just described are the base case for analysis. To allow for the possibility that the base case may be too pessimistic or optimistic about the costs of eliminating COVID-19 risk, I consider two alternative cases: a 40% reduction in GDP wholly eliminates the COVID-19

[39] This assumption is loosely based on Acemoglu et al. (2020, Figure 7). Note: after the analyses for this Article were run, Acemoglu et al. released a new version of their paper, which (as of the date of this Article, June 25, 2020) is the posted version. (Acemoglu et al. 2020b). The new Acemoglu et al. estimates of the GDP reduction needed to wholly eliminate the COVID-19 risk are more optimistic—closer to 40%. See Acemoglu et al. (2020b, Figure 5.1). Because the intention of this Article is not to provide guidance with respect to COVID-19 policy, but rather to compare VSL-based CBA with utilitarianism and prioritarianism, I have retained the 80% figure roughly based on Acemoglu (2020 Figure 7) as the base case for analysis. In any event, the 40% figure is covered here as the "optimistic case" and still shows significant differences between VSL-based CBA, on the one hand, and utilitarianism and prioritarianism, on the other.

[40] See Acemoglu et al. (2020 Figure 7).

Covid Economics 33, 30 June 2020: 1-45

risk (optimistic), and an 80% reduction only reduces the COVID-19 risk by half. I also consider a third alternative ("convex case"): a 80% reduction in GDP wholly eliminates the COVID-19 risk, and the function from GDP reduction to COVID-19 reduction is convex (increasing marginal impact of GDP reduction on the risk reduction) rather than concave.

On the cost side, I consider two alternative cases: "regressive incidence" and "proportional incidence." "Regressive incidence" means a π% reduction in GDP is borne more heavily (in proportional terms) by the three lowest income groups within each age, than by the two higher quintiles.[41] Proportional incidence means that individuals in all five quintiles incur a π% reduction in income when GDP is reduced by π%.

The causal linkages between social distancing policy and GDP reduction, on the one hand, and between social distancing policy and reduction in COVID-19 risk, on the other, are quite uncertain—and certainly were highly uncertain at the beginning of the pandemic, which is the timing of the policy choice being modeled here. (The full spectrum of COVID-19 risk-reduction options, including the option of wholly eliminating the risk via sufficiently stringent social-distancing measures, were only available to policymakers at the outset of the pandemic.) The assumptions I have adopted are not meant as best estimates, but rather as plausible, stylized facts that will illustrate the properties of the various assessment tools with respect to social distancing policy. As already explained, the aim of this Article is methodological: not to provide guidance with respect to social-distancing policy, but to use the simulation model to illustrate the features of VSL.

As illustrated in Table 7, the various cohorts have starkly opposing interests with respect to social distancing policy. Table 7 shows the cohort breakeven GDP reduction (the largest reduction such that its expected utility is not lower than baseline[42]), depending on whether the cost incidence of the reduction is regressive or proportional. (In each cell, the number above the slash is the cohort breakeven GDP reduction assuming regressive incidence; the number below the slash is for proportional incidence.)

[41] Specifically, my assumptions regarding regressive incidence are as follows. A π% reduction in GDP lowers income for individuals in the first three income quintiles (Low, Moderate and Middle) by π/80 of the difference between baseline income and subsistence income ($1000). Thus, an 80% reduction in GDP has the benefit of wholly eliminating the COVID-19 risk, but at the cost of reducing individuals in the first three quintiles to subsistence income. As for the two higher-income groups: a π% reduction in GDP reduces their income by $(\pi/80)p$, where p is chosen so that—given the above assumption about losses by the first three income quintiles—the overall reduction in GDP is *π%*. With the income data being used, $p = 0.7187$.

On these assumptions, each of the first three income quintiles experiences a larger percentage reduction in income than the top two quintiles, for $0 < \pi \leq 80$. "Regressive" is a slight misnomer, since the lowest quintile experiences a slightly lower percentage reduction than the second, in turn slightly lower than the third.

[42] Results for the simulation model were calculated for integer values of GDP reduction, and so the cohorts may be better off than baseline at these integer "breakeven" values. It would also interesting to know the cohort-*optimal* values of GDP reduction, but determining these values for all thirty-five cohorts would be quite laborious.

Covid Economics 33, 30 June 2020: 1-45

Table 7: Cohort Breakeven GDP Reductions

	Low	*Moderate*	*Middle*	*High*	*Top*
20	0/0	0/0	0/0	0/0	0/0
30	0/1	0/1	1/1	2/2	4/3
40	1/2	2/3	2/4	6/4	9/7
50	12/17	16/22	19/28	39/33	53/45
60	45/58	52/68	58/75	80/80	80/80
70	61/76	67/80	71/80	80/80	80/80
80	63/75	69/80	73/80	80/80	80/80

Older cohorts prefer a larger GDP reduction than younger ones. Within each age band, richer individuals prefer a larger GDP reduction than younger ones. This latter effect is dampened by shifting from regressive to proportional incidence, but not eliminated.

The choice of social distancing policy is a matter of balancing the (starkly) opposing interests of the various cohorts. VSL-CBA reaches a very different point of equipoise in this balancing, as compared with utilitarianism and prioritarianism.

Table 8 gives the optimal degree of reduction in GDP, according to the three methodologies—in the base case; under alternative assumptions (more optimistic or less optimistic) regarding the efficacy of social distancing policy in eliminating COVID-19 risk; and in the convex case. [43]

Table 8: Optimal reduction in GDP

	VSL-CBA	Utilitarianism	Prioritarianism
Base Case- Regressive Incidence	27%	11 %	12 %
Base Case – Proportional Incidence	27%	13%	14%
Optimistic Case—Regressive Incidence	40%	19%	19%
Optimistic Case – Proportional Incidence	40%	21%	23%
Pessimistic Case—Regressive Incidence	7%	4%	4%
Pessimistic Case – Proportional Incidence	7%	4%	4%
Convex Case—Regressive Incidence	80%	0%	0%
Convex Case—Proportional Incidence	80%	0%	0%

[43] For utilitarianism and prioritarianism, the optimal reduction was calculated exactly rather than using the marginal approximations stated in equations 2 and 4 above.

Let's first compare VSL-CBA to utilitarianism. In the base case, utilitarianism recommends a 11% reduction with regressive incidence and a 13% reduction with proportional incidence. The utilitarian social cost to a given reduction in income is lower if that loss is borne by higher-income groups (see Table 5b) and so shifting from regressive to proportional incidence increases the utilitarian optimum. VSL-CBA is insensitive to cost distribution (see Equation 3 and Table 5a), and so the optimum is a 27% reduction regardless of whether incidence is regressive or proportional. In either event, VSL-CBA's optimal reduction of GDP is *much* larger (27% versus 11%/13%) than the utilitarian optimum. Shifting to the optimistic and pessimistic cases does not change the pattern. VSL-CBA continues to prefer a much larger reduction than utilitarianism (40% versus 19%/21%; 7% versus 4%/4%).

Finally, in the convex case, the disparity between VSL-CBA and utilitarianism is quite extreme. The two methodologies choose corner solutions, but different ones: utilitarianism (with either regressive or proportional cost incidence) recommends no reduction in GDP at all, while VSL-CBA recommends wholly eliminating the risk at the cost of an 80% GDP reduction.

Why is VSL-CBA more stringent with respect to social distancing than utilitarianism?[44] There are two reasons, the first somewhat subtle. Let π denote the degree of GDP reduction, and let $\Delta p_c(\pi)$ denote the reduction in cohort c's fatality risk as a function of π; $\Delta y_c(\pi)$ the reduction in cohort c's income as a function of π; and $\Delta U_c(\pi)$ the increase in cohort c's expected utility. As per Equation 2, $\Delta U_c(\pi) \approx \frac{\partial U_c}{\partial p_c}\Delta p_c(\pi) - \frac{\partial U_c}{\partial y_c}\Delta y_c(\pi)$. As per Equations 1 and 3, the VSL-based monetary value of a π reduction in GDP is as follows:
$MV_c(\pi) = \Delta p_c(\pi) VSL_c - \Delta y_c(\pi) = \Delta p_c(\pi)\left(\frac{\partial U_c/\partial p_c}{\partial U_c/\partial y_c}\right) - \Delta y_c(\pi)$. Multiplying both sides of this equation by $(\partial U_c/\partial y_c)$, we have that $MV_c(\pi)\frac{\partial U_c}{\partial y_c} = \Delta p_c(\pi)\frac{\partial U_c}{\partial p_c} - \Delta y_c(\pi)\frac{\partial U_c}{\partial y_c}$, or that $MV_c(\pi) \approx \Delta U_c(\pi) k_c$, with $k_c = 1/\left(\frac{\partial U_c}{\partial y_c}\right)$. In short, a cohort's VSL-based monetary valuation of a policy is approximately its change in expected utility, times a weighting factor equaling the inverse of the marginal utility of income.

The optimal value of π, according to utilitarianism, is the value that maximizes the sum of ΔU_c across cohorts; while its optimal value, according to VSL-CBA, is the value that maximizes the sum of MV_c across cohorts. The k_c weighting factor increases with income. Thus VSL-CBA in this optimization gives greater weight to utility impacts on higher-income cohorts, and less weight to utility impacts on lower-income cohorts, as compared to utilitarianism.

[44] For a related discussion of why CBA and utilitarianism can diverge with respect to risk regulation, see Armantier and Treich (2004).

Note now that, with regressive and even with proportional incidence, richer cohorts tend to prefer a more restrictive social distancing policy than poorer cohorts. (See Table 7.)[45] Thus the effect of this skew towards the interests of higher-income cohorts is that VSL-CBA ends up with a more restrictive policy than utilitarianism.

The second reason that VSL-CBA is more stringent with respect to social distancing than utilitarianism is that VSL-CBA overstates the benefits of "large" policies, which reduce risks at the cost of significant changes in individual incomes.[46]

Turning now to prioritarianism: although in many contexts utilitarianism and prioritarianism can differ, here they converge. In all cases (base, optimistic, pessimistic, and convex, with proportional or regressive incidence), the optimum reduction in GDP as per

[45] Why do richer cohorts prefer a more restrictive social distancing policy than poorer cohorts? Shutdown policy is (approximately) a uniform risk reduction across income groups. Thus, on the risk side, each increment in π produces a larger increase in the expected utility of a richer group as compared to a poorer group. (See Table 4b.) On the cost side, whether each increment in π yields a smaller or larger loss in the expected utility of richer groups depends upon cost incidence. With a logarithmic utility function and proportional incidence, a given increment in π yields the same utility cost for different income groups within each age band. With regressive incidence, a given increment yields a larger utility cost for lower income groups. In either event, the *net* utility impact of an increment in π (utility benefit from risk reduction minus utility cost from income loss) is always larger for richer than smaller groups. Thus richer groups tend to prefer a larger π than poorer groups, as illustrated in Table 7.

[46] Consider the additive utility function used in the simulation; see Appendix B. Let A_c denote the number of the current period for cohort c. As in the text, let p_c denote the cohort's current survival probability (shorthand for $p_c^{A_c}$) and let y_c denote the cohort's current income ($y_c^{A_c}$). So as to avoid using the symbol π to mean two different things, let $\delta_c(t; A_c) = \prod_{s=A_c}^{t} p_c^s$ here denote the probability of a cohort member surviving to the end of period t.

$\Delta U_c(\pi) = \left[\Delta p_c(\pi) u(y_c - \Delta y_c(\pi)) + \Delta p_c(\pi) \sum_{t=A_c+1}^{T} \frac{\delta_c(t; A_c)}{p_c} u(y_c^t) \right] - p_c \left(u(y_c) - u(y_c - \Delta y_c(\pi)) \right)$. Note that, with this additive model, $\frac{\partial U_c}{\partial p_c} = u(y_c) + \sum_{t=A_c+1}^{T} \frac{\delta_c(t; A_c)}{p_c} u(y_c^t)$ and $\frac{\partial U_c}{\partial y_c} = p_c u'(y_c)$. Assume that $u(\cdot)$ is strictly increasing and also strictly concave (diminishing marginal income utility). Observe that the term in brackets in the formula for $\Delta U_c(\pi)$ is less than $\frac{\partial U_c}{\partial p_c} \Delta p_c(\pi)$ and that $p_c(\pi)\left(u(y_c) - u(y_c - \Delta y_c(\pi)) \right)$ is greater than $\frac{\partial U_c}{\partial y_c} \Delta y_c(\pi)$ by the strict concavity of $u(\cdot)$.

Thus , $\frac{\partial U_c}{\partial p_c} \Delta p_c(\pi) - \frac{\partial U_c}{\partial y_c} \Delta y_c(\pi) = \Delta U_c(\pi) + E$, with E a positive error term the magnitude of which increases with $\Delta y_c(\pi)$. In turn, $MV_c(\pi) = (\Delta U_c(\pi) + E)k_c$. In short, MV_c is skewed upwards relative to ΔU_c, and increasingly so as Δy_c becomes large.

prioritarianism is the same as, or quite close to, the utilitarian optimum—and thus much lower than VSL-CBA's recommendation.[47]

In short, relative to the utilitarian benchmark and to a prioritarian benchmark, VSL-CBA is significantly more restrictive.[48]

Can CBA's recommendations with respect to shutdown policy be defended on alternative grounds—by appeal to Kaldor-Hicks efficiency? I calculated cohorts' actual willingness to pay/accept amounts at the CBA optimum in the base case (27% reduction). At the optimum, all the cohorts aged 40 and younger are worse off than the status quo, as are the three age-50 cohorts at low, moderate and middle income. All the cohorts aged 60 and above are better off than the status quo, as are the two higher-income age-50 cohorts.[49] Total willingness-to-pay of the better-off cohorts does in fact exceed the total willingness-to-accept of the worse-off cohorts. So the optimum is Kaldor-Hicks efficient relative to the no-intervention baseline.

However, many of the non-optimal policies that CBA scores as an improvement relative to the status quo are *not* Kaldor-Hicks efficient. In the model here, *every* reduction in GDP (from 1% to 80%) is assigned a positive score (positive sum of MVs), relative to the baseline of 0% reduction. In other words, every reduction is seen by CBA as an improvement over the status quo. However, reductions at or above 37% are not Kaldor-Hicks efficient. In particular, CBA prefers to wholly eliminate the COVID-19 risk, at the cost of 80% of GDP, as compared to the no-intervention status quo; but at 80% reduction, only 6 of the 35 cohorts are better off than

---

[47]Prioritarianism is more concerned than utilitarianism with reducing risks among the young. (Compare Tables 4b and 4c.) Baseline COVID-19 risk increases with age, and thus the effect of any COVID-19 policy (in the model here) is to produce a smaller risk reduction for younger than older cohorts. These effects roughly cancel, and so prioritarianism ends up with roughly the same social distancing recommendation as utilitarianism.

As mentioned, the prioritarian results in the social-distancing analysis are for "ex post" prioritarianism. As a sensitivity analysis, I computed optimal GDP reduction for "ex ante" prioritarianism in the base case. Ex ante prioritarianism prefers a 9% reduction given regressive incidence and an 11% reduction given proportional incidence.

[48] A third possible source of the divergence between VSL-CBA and utilitarianism with respect to social distancing policy is the fact that the two methodologies, in the model here, have different patterns of valuation of risk reduction with age. Utilitarian values decrease monotonically, while VSL-CBA values do not. (Compares Tables 4a and 4b.) To test whether this divergence helps explain the divergent social-distancing recommendations, I undertook an alternative analysis in which the time path of income within each quintile is constant. At every age, individuals receive the very same income, depending on quintile: $21,961 (low); $30,118 (moderate); $41,349 (middle); $57,538 (high); and $134,840 (top). In this case, I find that utilitarianism with regressive incidence recommends a 12% GDP reduction, and with proportional incidence a 14% GDP reduction, while VSL-CBA recommends a 29% reduction. Thus the differing patterns of risk valuation by age do not appear to contribute to the VSL-CBA/utilitarianism divergence with respect to social distancing.

[49] These willingness-to-pay/accept amounts are the individuals' compensating variations (equilibrating changes to the incomes that individuals would have with the policy). It is the sum of compensating variations which tracks whether a policy is Kaldor-Hicks efficient relative to the status quo: if the sum is positive, there is a scheme of transfers which, if costlessly implemented together with the policy, would make everyone better off than in status quo.

Compensating variations were calculated assuming regressive incidence. With progressive incidence, a different array of groups might be better off than the status quo (compare the numbers above and below the slash in each cell in Table 7), but the sum total of compensating variations would be the same.

Covid Economics 33, 30 June 2020: 1-45

Covid Economics 33, 30 June 2020: 1-45

the status quo (high and top-income individuals aged 60, 70 and 80) and their total willingness to pay for 80% shutdown is insufficient to compensate the 29 cohorts who are worse off.

IV. Population-Average VSL

The previous two parts focused on textbook VSL-based CBA ("VSL-CBA"). VSL varies among individuals (Equation 1) and these individual-specific VSLs are used to compute MVs.

As already mentioned, governmental agencies in the U.S. employ a single population-average VSL. I'll refer to CBA with a population-average VSL as "VSL^{avg}-CBA." It uses this single conversion factor to calculate MVs.

While MV_i calculated with individual i's VSL is a good approximation to her compensating variation and equivalent variation for a small change in survival probability and other attributes, relative to the status quo, this is not true of MV_i calculated with a population-average VSL. The point is well illustrated with the simulation model presented in Part III, which I'll continue to use in this Part. The cohort-specific VSLs are given in Table 3. The population-average VSL (weighting by the proportion of the various age groups in the population) is \$8,635,355. The cohort-specific VSLs are generally significantly different (larger or smaller) than this average. For example, a 60-year old middle-income individual has a VSL of \$3,943,065. That individual's equivalent variation and compensating variation for a 1-in-1 million reduction in current fatality risk is approximately \$3.94, and for a 1-in-100,000 reduction is approximately \$39.43. But his MV according to VSL^{avg}-CBA is \$8.64 in the first case and \$86.35 in the second.

We saw earlier that VSL-CBA need not conform to Kaldor-Hicks efficiency. The divergence between VSL-CBA and Kaldor-Hicks efficiency occurs as MV_i diverges from i's equivalent and compensating variations. VSL^{avg}-CBA also need not conform to Kaldor-Hicks efficiency, and can diverge even for small changes from the status quo (since MV_i calculated with population-average VSL is not a good approximation to the equivalent or compensating variation even for a small change). For example, imagine that there are 1 million individuals in the 40-year old, low-income cohort (with a cohort VSL of \$3,254,671), and a policy will reduce each individual's risk by 1 in 100,000, at a total cost of \$50 million. VSL^{avg}-CBA approves the policy, since the population-average VSL of \$8,635,355 multiplied by 1-in-100,000 summed over the million individuals equals \$86.35 million, which exceeds the cost of \$50 million. But individuals in the cohort are actually willing to pay, in total, only \$32.55 million for the policy. Whoever the cost bearers might be, \$32.55 million would be insufficient to fully compensate them.

Table 4d is the analogue to tables 4a, 4b, and 4c. It gives the social value of an increment in current-year fatality risk according to VSL^{avg}-CBA, with 1 indicating the value of a risk increment to the member of the 80-year-old, low-income cohort.

Table 4d: Social value of risk increment according to VSL^{avg}-CBA
(relative to social value of risk increment for 80-year-old, low-income cohort)

	Income: Low	*Moderate*	*Middle*	*High*	*Top*
Age: 20	1	1	1	1	1
30	1	1	1	1	1
40	1	1	1	1	1
50	1	1	1	1	1
60	1	1	1	1	1
70	1	1	1	1	1
80	1	1	1	1	1

The entry in every cell in table 4d is the same (1), because a risk increment Δp to any cohort is assigned the very same value, namely Δp multiplied by $8,635,355.

On the income side, VSL^{avg}-CBA values an income increment the same way as VSL-CBA—a Δy increment to anyone's income changes her MV by Δy—and so the relative value of income increments is given by the same table for both methodologies, namely table 5a.

The risk-increment and income-increment tables show that VSL^{avg}-CBA diverges radically from utilitarianism and prioritarianism. Nor, as just explained, can it be supported with reference to Kaldor-Hicks efficiency.

Considerations of political feasibility may well push governments towards VSL^{avg}-CBA. In the U.S., a tentative move by the EPA in 2003 in the direction of adjusting VSL by age prompted a harsh political reaction by interest groups representing older individuals.[50] But such political constraints of course don't constitute a normative justification for VSL^{avg}-CBA.

In some respects, the methodology's schedule of relative risk valuations (table 4d—every group has the same value of risk reduction) is intuitively attractive. Surveys eliciting citizens' views regarding the allocation of health care show no support for preferring the rich.[51] In this respect, Table 4d is more intuitive than Table 4a (VSL-CBA) or Table 4b (utilitarianism). Moreover, Table 4d conforms to the intuition that every life has equal value—that there should be no differentiation whatsoever in deciding whose life to save or risk to reduce. Yet surveys of citizens also find support for a conflicting intuition, namely that the young should receive priority over the old.[52] Is it really the case that society should be indifferent between reducing an 80-year-old's risk of dying by Δp and reducing a 20-year-old's risk of dying by the same amount? So intuitive support for Table 4d can hardly be said to be rock-solid.

[50] This episode is discussed in Viscusi (2018, p. 93).
[51] See Dolan et al. (2005).
[52] See *id*.

On the cost side, VSL^{avg}-CBA, like VSL-CBA, is neutral to the distribution of the costs of risk reduction (table 5a). This is not especially intuitive.

Turning to the model social-distancing policy: VSL^{avg}-CBA is dramatically *more* restrictive than VSL-CBA, let alone utilitarianism and prioritarianism. In the base case, it recommends an 80% reduction, while VSL-CBA recommends 27%. In the pessimistic case, it recommends a 25% reduction, while VSL-CBA recommends 7%. (See Table 10.) VSL^{avg}-CBA advises a substantially greater reduction in risk and loss of income than VSL-CBA because the benefits of social distancing are skewed towards older cohorts, whose cohort-specific VSLs are generally less than the population average. (VSL^{avg}-CBA reaches the same result as VSL-CBA in the optimistic and convex cases.)

Recall that VSL-CBA's recommendation (a 27% reduction) is Kaldor-Hicks efficient relative to the status quo. VSL^{avg}-CBA's recommendation (an 80% reduction) is not.

V. VSLY

The value of statistical life year (VSLY) is VSL divided by the life expectancy gained from a risk reduction. Life expectancy gained might be discounted at the market interest rate or the individual rate of time preference, or undiscounted. The conceptual points here apply to both discounted and undiscounted VSLY.[53] The specific illustrations, drawn, from the simulation model, involved undiscounted VSLY.

At the individual level, VSL with respect to a particular time period can be expressed as such, or as VSLY. Consider a policy that increases individual i's period s survival probability by Δp_i^s. Recall that individual i's VSL for that period, VSL_i^s, is equal to her marginal rate of substitution between survival probability in period s and current income (Equation 1). MV_i in this case equals $(\Delta p_i^s)VSL_i^s$, which is a good approximation to her equivalent variation and compensating variation for the policy if Δp_i^s is small.

Let LE_i^s be the difference between (a) i's current life expectancy if her probability of surviving period s conditional on being alive at the beginning of s is 1, and (b) her current life expectancy if her probability of surviving period s conditional on being alive at the beginning of s is 0. Note that $(\Delta p_i^s)LE_i^s$ is the gain to current life expectancy from a Δp_i^s increase in period s survival probability. Let's now define $VSLY_i^s$ as VSL_i^s / LE_i^s. Then, by construction, $MV_i = (\Delta p_i^s)VSL_i^s = (\Delta p_i^s)LE_i^s \times VSLY_i^s$. Individual i's MV for a period s risk reduction can be expressed *either* as her VSL for that period multiplied by the risk reduction, *or* as the gain to life expectancy from the risk reduction, multiplied by her VSLY for that period. This period-and-individual specific VSLY is well-defined, but has no advantages (or disadvantages!) over the

[53] The conceptual points here about the relation between VSL and VSLY are not novel; see Hammitt (2007) and Jones-Lee et al. (2015).

period-and-individual specific VSL, since monetary valuations calculated either way are identical.

It might be thought that we could calculate a period-*invariant* individual VSLY (let's call it $VSLY_i$), such that an individual's VSL with respect to the current period or any future period is just equal to her life expectancy gain from surviving that period, multiplied by $VSLY_i$. That is: for any period *s*, $VSL_i^s = VSLY_i \times LE_i^s$. In other words, an individual's willingness to pay for a small risk reduction in the current period or any future period is the gain to life expectancy, multiplied by a constant individual VSLY. However, there is nothing in the theory of VSL to ensure that this is true. For example, imagine that an individual consumes income when she receives it, and that the time path of future income is variable. Between period s^* and s^{**}, her income will be higher than after s^{**}. If so, the individual will be willing to pay *more* in current income for a given increase in life expectancy secured through an increase in period s^* survival probability, than for the same increase in life expectancy secured through an increase in period s^{**} survival probability. There is *no* period invariant $VSLY_i$ such that $VSL_i^{s^*} = VSLY_i \times LE_i^{s^*}$ and $VSL_i^{s^{**}} = VSLY_i \times LE_i^{s^{**}}$.[54]

In short, at the individual level VSLY is either equivalent to VSL, or undefined.

At the population level, however, we can define a population-average VSLY that is well defined and that gives rise to a form of CBA—for short, "$VSLY^{avg}$-CBA"—that differs from both CBA with individual-specific VSLs (VSL-CBA) and CBA with a population-average VSL (VSL^{avg}-CBA). $VSLY^{avg}$ for a given period is just the average of individual VSLYs for that period. $VSLY^{avg}$-CBA assigns a monetary valuation to an individual's risk reduction in period *s* equaling the increase in life expectancy multiplied by $VSLY^{avg}$.

We can use the simulation model to illustrate. Table 9 gives the life-expectancy gains from preventing the current death of an individual in each of the 35 cohorts (that is, LE_i^s with *s* the current period).

[54] In the case at hand, $VSL_i^{s^*} / LE_i^{s^*} > VSL_i^{s^{**}} / LE_i^{s^{**}}$. On a related point, Hammitt (2007, p. 236) observes: "If [as is often observed] an individual's VSL first rises then falls with age, then her VSLY cannot be constant over her lifespan. Life expectancy typically decreases with age."

Table 9: Life expectancy gains (years) from saving a cohort member in the current year

	Income: Low	*Moderate*	*Middle*	*High*	*Top*
Age: 20	54.3	56.8	58.9	60.1	62.1
30	45.1	47.5	49.5	50.6	52.5
40	36.1	38.4	40.2	41.3	43.1
50	27.4	29.5	31.2	32.2	33.9
60	19.6	21.4	23.0	23.8	25.3
70	12.8	14.2	15.4	16.2	17.4
80	7.1	8.2	9.1	9.6	10.5

For each cohort, we divide the VSL (Table 3) by the life expectancy gain in Table 9 to arrive at a cohort VSLY, and then average these (weighting for the proportion of the different age groups in the population) to arrive at the population-average VSLY, which is $240,676.

An individual's monetary valuation calculated using the population average VSLY need not approximate her equivalent or compensating variation, even for small changes. This is because the individual's VSLY for the specific period (here, the current period) need not be the same as the population-average VSLY for that period. For example, VSL for a member of the 20 year old, high-income cohort is $5,240,842. Her VSLY is $5,240,842/60.1 = $87,202. Her equivalent variation and compensating variation for a 1-in-100,000 increase in survival probability is approximately $52, which is her VSL times 1-in-100,000 or, equivalently, her VSLY times 1-in-100,000 times the life expectancy gain from preventing her death in the current period (60.1 years). However, the population-average VSLY ($240,676) multiplied by 1-in-100,000 multiplied by 60.1 years gives a monetary value of $145.

Thus VSLY^{avg}-CBA need not conform to Kaldor-Hicks efficiency, even for small risk changes (like VSL^{avg}-CBA, but by contrast with VSL-CBA).

The relative social values of risk reduction implied by VSLY^{avg}-CBA are presented in Table 4e. As in the previous such tables, the values are normalized so that 1 indicates the value of risk reduction to a member of the 80-year-old, low-income cohort. These relative valuations are just the relative life expectancy gains in Table 9.

Covid Economics 33, 30 June 2020: 1-45

Table 4e: Social value of risk increment according to $VSLY^{avg}$-CBA
(relative to social value of risk increment for 80-year-old, low-income cohort)

	Income: Low	*Moderate*	*Middle*	*High*	*Top*
Age: 20	7.6	8.0	8.3	8.4	8.7
30	6.3	6.7	6.9	7.1	7.4
40	5.1	5.4	5.6	5.8	6.0
50	3.8	4.1	4.4	4.5	4.8
60	2.8	3.0	3.2	3.3	3.6
70	1.8	2.0	2.2	2.3	2.4
80	1.0	1.1	1.3	1.3	1.5

These relative valuations are quite different from those in Tables 4b (utilitarianism) and 4c (prioritarianism); and on the income side $VSLY^{avg}$-CBA has a constant valuation (Table 5a), so is quite different from utilitarian and prioritarian valuations of income increments (Tables 5b and 5c). In short, $VSLY^{avg}$-CBA is grounded neither in utilitarianism, nor prioritarianism, nor (as we've just seen) Kaldor-Hicks efficiency. Why adopt it?

$VSLY^{avg}$ –CBA conforms to the intuition that the social value of risk reduction should be lower for older individuals, since they have fewer life years remaining. (The valuations in Table 4e decrease according to life expectancy remaining.) If one has this intuition, $VSLY^{avg}$-CBA will be seen to be an improvement over VSL^{avg}-CBA. Further, it is dramatically less biased to the rich than VSL-CBA (Table 4a), and significantly less so than utilitarianism (Table 4b). Values increase with income in each age band in Table 4e only because higher income is associated with a more favorable survival curve and thus higher life expectancy—and not for the utilitarian reason that a year of life at higher income is better for well-being.

Note, however, that prioritarianism can avoid *any* bias towards the rich (see Table 4c) and in this sense has intuitive advantages over $VSLY^{avg}$-CBA.

With respect to social distancing policy, $VSLY^{avg}$-CBA is somewhat less restrictive than VSL-CBA in the base case. It recommends a 23% reduction rather than 27%. However, this recommendation is significantly more restrictive than utilitarianism and prioritarianism, which (recall) recommend at most 14% depending on incidence. In the pessimistic case, $VSLY^{avg}$-CBA is again somewhat less restrictive than VSL-CBA (6% versus 7%). In the optimistic and convex cases the two concur.

As regards Kaldor-Hicks efficiency, the recommendation of $VSLY^{avg}$-CBA (23% reduction) is Kaldor-Hicks efficient relative to the status quo. However, $VSLY^{avg}$-CBA scores every reduction as an improvement over the status quo. Recall that reductions at or above 37% are not Kaldor-Hick efficient.

Covid Economics 33, 30 June 2020: 1-45

Table 10 summarizes the recommendations of all the methodologies in all the cases, and Table 11 summarizes the results with respect to Kaldor-Hicks efficiency.

Table 10. Social Distancing Policy: Optimal Degree of GDP Reduction

	VSL-CBA	VSL^{avg}-CBA	$VSLY^{avg}$-CBA	Utilitarianism	Prioritarianism
Base Case- Regressive Incidence	27%	80%	23%	11 %	12 %
Base Case – Proportional Incidence	27%	80%	23%	13%	14%
Optimistic Case— Regressive Incidence	40%	40%	40%	19%	19%
Optimistic Case – Proportional Incidence	40%	40%	40%	21%	23%
Pessimistic Case—Regressive Incidence	7%	25%	6%	4%	4%
Pessimistic Case – Proportional Incidence	7%	25%	6%	4%	4%
Convex Case— Regressive Incidence	80%	80%	80%	0%	0%
Convex Case— Proportional Incidence	80%	80%	80%	0%	0%

Table 11. Kaldor-Hicks Efficiency (Base Case)

	VSL-CBA	VSL^{avg}-CBA	$VSLY^{avg}$-CBA
Optimal degree of GDP reduction	27%	80%	23%
Is the optimum Kaldor-Hicks efficient relative to the status quo?	Yes	No	Yes
Does the methodology assign a positive score (indicating that the GDP reduction is an improvement over the status quo) to GDP reductions that are not Kaldor-Hicks efficient relative to the status quo?	Yes (all reductions at or above 37%)	Yes (all reductions at or above 37%)	Yes (all reductions at or above 37%)

Covid Economics 33, 30 June 2020: 1-45

VI. Conclusion

This Article has critically examined the use of VSL-based CBA as a methodology for setting fatality risk reduction policy. I have done so via close consideration of a policy problem that, sadly, is all too timely: trading off the costs and benefits of social distancing to mitigate COVID-19. I consider three versions of VSL: textbook, population-average, and VSLY.

Strikingly, I find that all three recommend policies that *deviate* from Kaldor-Hicks efficiency (potential Pareto superiority), the traditional touchstone of CBA. All three versions of VSL conclude that a range of social distancing policies would be better than the status quo (in my simulation, policies at or above a 37% reduction of GDP), even though these policies are not in fact Kaldor-Hicks efficient relative to the status quo. CBA with population-average VSL *optimizes* at a policy that is not Kaldor-Hicks efficient relative to the status quo.

Thus these VSL-based methodologies *lack* grounding in the standard normative justification offered for CBA. Nor can they find normative support elsewhere. Utilitarianism, the oldest example of a systematic welfarist approach to normative reasoning, recommends quite different social-distancing policies—as does prioritarianism, a newer variant of welfarism. As for intuition: textbook VSL is quite counterintuitive, because it places a dramatically higher value on risk reductions for richer individuals; population-average VSL fails to differentiate with respect to age; and all three versions on the cost side are completely insensitive to the incidence of the costs of social distancing policy.

My own view is that social distancing policy, and risk regulation more generally, should be set with reference to a utilitarian or prioritarian social welfare function.[55] To be sure, this position implicates the long-running debate about the role of distributional considerations in non-tax policies.[56] The purist view, here, is that *all* non-tax policies should be designed to maximize the size of the "pie"—to be precise, the sum total of compensating variations relative to the status quo—and that the tax-and-transfer system should be used to share the "pie," to everyone's benefit.

There are various difficulties with the purist view,[57] above all this: the changes to existing tax-and-transfer laws required to render a non-tax policy Pareto-superior to the status quo—*universally* beneficial—are often merely hypothetical, given the actual political economy of the tax system. In the case of social distancing policy, a pie-maximizing-and-sharing combination of policies would mean quite stringent and prolonged social distancing requirements combined with significant taxes on older individuals and substantial payments to younger ones. This doesn't seem to be in the offing. Pie-maximization without pie-sharing may well leave some groups much worse off than the status quo.

[55] See Adler (2019, ch. 5); Adler (2017); Adler, Hammitt and Treich (2014): Adler, Ferranna, Hammitt and Treich (2019).
[56] See Kaplow (1996, 2004, 2008).
[57] See Adler (2019, pp. 225-33); Adler (2017).

In any event, a key takeaway from this Article should be that the debate about the role of distributional considerations in non-tax policy is *orthogonal* to VSL. The proponent of pie-maximization-and-sharing needs a methodology that approves policies if and only if they are Kaldor-Hicks efficient relative to the status quo, and that optimizes at the largest sum of compensating variations relative to the status quo. Such a methodology is not VSL-CBA, nor VSL^{avg}-CBA, nor $VSLY^{avg}$-CBA.

Covid Economics 33, 30 June 2020: 1-45

References

Acemoglu, Daron, et al. 2020. "A Multi-Risk SIR Model with Optimally Targeted Lockdown." NBER Working Paper No. 27102 (dated May 2020). Note: This paper was originally available at http://www.nber.org/papers /w27102, but has now been replaced at that link by Acemoglu et al. (2020b) immediately below. I have retained an electronic copy of Acemoglu et al. (2020) and can make it available on request.

Acemoglu, Daron, et al. 2020b. "Optimal Targeted Lockdowns in a Multi-Group SIR Model." NBER Working Paper No. 27102 (described on NBER site as a June 2020 revision, although dated May 2020 on the title page). Available at http://www.nber.org/papers /w27102.

Adler, Matthew D. 2012. *Well-Being and Fair Distribution*: *Beyond Cost-Benefit Analysis*. New York: Oxford University Press.

Adler, Matthew D. 2016a. "Benefit-Cost Analysis and Distributional Weights: An Overview." *Review of Environmental Economics and Policy* 10: 264-85.

Adler, Matthew D. 2016b. "Extended Preferences." In Matthew D. Adler and Marc Fleurbaey, eds., *The Oxford Handbook of Well-Being and Public Policy*, pp. 476-517. New York: Oxford University Press.

Adler, Matthew D. 2017. "A Better Calculus for Regulators: From Cost-Benefit Analysis to the Social Welfare Function." Duke Law School Public Law and Legal Theory Working Paper No. 2017-19. Available at https://papers.ssrn.com/sol3/papers.cm?abstract_id=2923829.

Adler, Matthew D. 2019. *Measuring Social Welfare*: *An Introduction*. New York: Oxford University Press.

Adler, Matthew D., Maddalena Ferranna, James K. Hammitt, and Nicolas Treich. 2019. "Fair Innings? The Utilitarian and Prioritarian Value of Risk Reduction over a Whole Lifetime." Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3493976.

Adler, Matthew D., James K. Hammitt, and Nicolas Treich. 2014. "The Social Value of Mortality Risk Reduction: VSL versus the Social Welfare Function Approach." *Journal of Health Economics* 35: 82-93.

Adler, Matthew D., and Eric A. Posner. 2006. *New Foundations of Cost-Benefit Analysis*. Cambridge, MA: Harvard University Press.

Aldy, Joseph E., and W. Kip Viscusi. 2007. "Age Differences in the Value of Statistical Life: Revealed Preference Evidence." *Review of Environmental Economics and Policy* 1: 241-60.

Alvarez, Fernando E., David Argente, and Francesco Lippi. 2020. "A Simple Planning Problem for Covid-19 Lockdown." *Covid Economics* 14: 1-32. Available at https://cepr.org/content/covid-economics-vetted-and-real-time-papers-0.

Armantier, Olivier, and Nicolas Treich. 2004. "Social Willingness to Pay, Mortality Risks, and Contingent Valuation." *Journal of Risk and Uncertainty* 29: 7-19.

Bairoliya, Neha, and Ayşe Imrohoroğlu. 2020. "Macroeconomic Consequences of Stay-at-Home Policies During the COVID-19 Pandemic." *Covid Economics* 13: 71-90. Available at https://cepr.org/content/covid-economics-vetted-and-real-time-papers-0.

Barnett-Howell, Zachary, and Ahmed Mushfiq Mobarak. 2020. "The Benefits and Costs of Social Distancing in Rich and Poor Countries." Working paper. Available at https://arxiv.org/abs/2004.04867.

Béland, Louis-Phillippe, Abel Brodeur, and Taylor Wright. 2020. "Covid-19, Stay-at-Home Orders and Employment: Evidence from CPS Data." IZA Discussion Paper No. 13282. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3608531.

Bethune, Zachary A., and Anton Korinek. 2020. "Covid-19 Infection Externalities: Trading Off Lives vs. Livelihoods." NBER Working Paper No. 27009. Available at http://www.nber.org/papers/w27009.

Boadway, Robin. 2016. "Cost-Benefit Analysis." In Matthew D. Adler and Marc Fleurbaey, eds., *The Oxford Handbook of Well-Being and Public Policy*, pp. 47-81. New York: Oxford University Press.

Broome, John. 1978. "Trying to Value a Life." *Journal of Public Economics* 9: 91-100.

Bronsteen, John, Christopher Buccafusco, and Jonathan Masur. 2013. "Well-Being Analysis vs. Cost-Benefit Analysis." *Duke Law Journal* 62: 1603-89.

Broughel, James. 2020. "Rethinking the Value of Life: A Critical Appraisal of the Value of a Statistical Life." Center for Growth and Opportunity, Policy Paper. Available at https://www.thecgo.org/research/rethinking-the-value-of-life/.

Chetty, Raj, et al. 2016. "The Association between Income and Life Expectancy in the United States, 2001-2014." *JAMA* 315: 1750-66.

Corcoran, Terence. May 15, 2020. "The Price of Life: Lockdown Costs are Real, but are the Benefits?" National Post (Canada).

Cropper, Maureen, James K. Hammitt, and Lisa A. Robinson. 2011. "Valuing Mortality Risk Reductions: Progress and Challenges." *Annual Review of Resource Economics* 3: 313-36.

Dolan, Paul, et al. 2005. "QALY Maximisation and People's Preferences: A Methodological Review of the Literature." *Health Economics* 14: 197-208.

Dolan, Paul, et al. 2008. "Valuing Lives and Life Years: Anomalies, Implications, and an Alternative." *Health Economics, Policy, and Law* 3: 277-300.

Eeckhoudt, Louis R., and James K. Hammitt. 2001. "Background Risks and the Value of Statistical Life." *Journal of Risk and Uncertainty* 23: 261-79.

Evans, Mary F., and V. Kerry Smith. 2010. "Measuring How Risk Tradeoffs Adjust with Income." *Journal of Risk and Uncertainty* 40: 33-55.

Ferguson, Neil M., et al. 2020. "Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce Covid-19 Mortality and Healthcare Demand." Imperial College, London, working paper. Available at https://doi.org/10.25561/77482.

Freeman, A. Myrick. 2003. *The Measurement of Environmental and Resource Values*, 2d ed. Washington, D.C: Resources for the Future.

Frick, Johann. 2015. "Contractualism and Social Risk." *Philosophy and Public Affairs* 43: 175-223.

Greenstone, Michael, and Visham Nigam. 2020. "Does Social Distancing Matter?" *Covid Economics* 7: 1-22. Available at https://cepr.org/content/covid-economics-vetted-and-real-time-papers-0.

Gros, Claudius, et al. 2020. "Containment Efficiency and Control Strategies for the Corona Pandemic Costs." Working paper. Available at https://arxiv.org/pdf/2004.00493.pdf.

Grüne-Yanoff, Till. 2009. "Mismeasuring the Value of Statistical Life." *Journal of Economic Methodology* 16: 109-23.

Hall, Robert E., Charles I. Jones, and Peter J. Klenow. 2020. "Trading Off Consumption and COVID-19 Deaths." Working paper. Available at https://web.stanford.edu/~chadj/Consumption_v_Covid.pdf.

Hammitt, James K. 2000. "Valuing Mortality Risk: Theory and Practice." *Environmental Science and Technology* 34: 1396-1400.

Hammitt, James K. 2007. "Valuing Changes in Mortality Risk: Lives Saved versus Life Years Saved." *Review of Environmental Economics and Policy* 1: 228-40.

Hammitt, James K. 2020. "Valuing Mortality Risk in the Time of Covid-19." Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3615314.

Hammitt, James K., et al. 2020. "Premature Deaths, Statistical Lives, and Years of Life Lost: Identification, Quantification, and Valuation of Mortality Risks." *Risk Analysis* 40: 674-95.

Henderson, David R., and Jonathan Lipow. June 16, 2020. "The Data are In: It's Time for Major Reopening." *Wall Street Journal*.

Hilsenrath, Jon, and Stephanie Armour. March 23, 2020. "As Economic Toll Mounts, Nation Ponders Trade-Offs." *Wall Street Journal*.

Holtug, Nils. 2010. *Persons, Interests, and Justice*. Oxford: Oxford University Press.

Holtug, Nils. 2017. "Prioritarianism." *Oxford Research Encyclopedia of Politics*, http://dx.doi.org/10.1093/acrefore/9780190228637.013.232.

Ingraham, Christopher. March 30, 2020. "Moderate Social Distancing Yields $8 Trillion in Economic Benefits, Study Finds." *Washington Post*.

Jenkins, Holman W., Jr. March 27, 2020. "Yes, We Need Cost-Benefit Analysis." *Wall Street Journal*.

Johansson, Per-Olov. 2002. "On the Definition and Age-Dependency of the Value of Statistical Life." *Journal of Risk and Uncertainty* 25: 251-63.

Jones-Lee, Michael, et al. 2015. "Valuing Gains in Life Expectancy: Clarifying Some Ambiguities." *Journal of Risk and Uncertainty* 51: 1-21.

Kaplow, Louis. 1996. "The Optimal Supply of Public Goods and the Distortionary Cost of Taxation." *National Tax Journal* 49: 513-33.

Kaplow, Louis. 2004. "On the (Ir)Relevance of Distribution and Labor Supply Distortion to Government Policy." *Journal of Economic Perspectives* 18: 159-75.

Kaplow, Louis. 2008. *The Theory of Taxation and Public Economics*. Princeton: Princeton University Press.

Kniesner, Thomas J., and W. Kip Viscusi. 2019. "The Value of a Statistical Life." Working paper. Available at http://ssrn.com/abstract_id=3379967. Forthcoming in *Oxford Research Encyclopedia of Economics and Finance.*

Krupnick, Alan. 2007. "Mortality-risk Valuation and Age: Stated Preference Evidence." *Review of Environmental Economics and Policy* 1: 261-82.

Masur, Jonathan, and Eric Posner. April 20, 2020. "Cost-Benefit Analysis Supports Continuing the National Shutdown." *The Regulatory Review.*

OECD. 2012. *Mortality Risk Valuation in Environment, Health and Transport Policies*. Paris: OECD Publishing.

Office of Management and Budget. 2017. "2017 Report to Congress on the Benefits and Costs of Federal Regulations and Agency Compliance with the Unfunded Mandates Reform Act." Available at https://www.whitehouse.gov/omb/information-regulatory-affairs/reports/#ORC.

Parfit, Derek. 2000. "Equality or Priority?" In Matthew Clayton and Andrew Williams, eds., *The Ideal of Equality*, pp. 81-125. Houndmills: Palgrave.

Pindyck, Robert S. 2020. "Covid-19 and the Welfare Effects of Reducing Contagion." NBER Working Paper 27121. Available at https://www.nber.org/papers/w27121.pdf.

Pretnar, Nick. 2020. "The Intergenerational Welfare Implications of Disease Contagion." Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3624857.

Robinson, Lisa A. 2007. "How U.S. Government Agencies Value Mortality Risk Reduction." *Review of Environmental Economics and Policy* 1: 283-99.

Robinson, Lisa A., James K. Hammitt, and Richard J. Zeckhauser. 2016. "Attention to Distribution in U.S. Regulatory Analyses." *Review of Environmental Economics and Policy* 10: 308-28.

Covid Economics 33, 30 June 2020: 1-45

Robinson, Lisa A., Ryan Sullivan, and Jason F. Shogren. 2020. "Do the Benefits of COVID-19 Policies Exceed the Costs? Exploring Uncertainties in the Age-VSL Relationship." Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3626710.

Scherbina, Anna. 2020. "Determining the Optimal Duration of the Covid-19 Suppression Policy: A Cost-Benefit Analysis." Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3562053.

Sunstein, Cass R. March 26, 2020. "This Time the Numbers Show We Can't Be Too Careful." *Bloomberg Opinion*.

Thurnstrom, Linda, et al. 2020. "The Benefits and Costs of Using Social Distancing to Flatten the Curve for Covid-19." Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3561934.

Ugarov, Alexander. 2020. "Inclusive Costs of NPI Measures for Covid-19 Pandemic." Working paper. Available at https://www.medrxiv.org/content/10.1101/2020.03.26.20044552v1.

U.S. Department of Health and Human Services. 2016. *Guidelines for Regulatory Impact Analysis*. Available at https://aspe.hhs.gov/pdf-report/guidelines-regulatory-impact-analysis.

U.S. Department of Transportation. 2016. "Revised Departmental Guidance on Valuation of a Statistical Life in Economic Analysis." Available at https://www.transportation.gov/regulations/economic-values-used-in-analysis.

U.S. Environmental Protection Agency. 2010. (Updated 2014). *Guidelines for Preparing Economic Analyses*. Available at https://www.epa.gov/environmental-economics/guidelines-preparing-economic-analyses.

Viscusi, W. Kip. 2018. *Pricing Lives*: *Guideposts for a Safer Society*. Princeton: Princeton University Press.

Viscusi, W. Kip, and Joseph E. Aldy. 2003. "The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World." *Journal of Risk and Uncertainty* 27: 5-76.

Wilson, Linus. 2020. "Estimating the Life Expectancy and Value of Statistical Life (VSL) Losses from Covid-19 Infections in the United States." Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3580414.

CEPR PRESS

Appendix

A. *Simulated Population*

The population is divided into 35 cohorts: seven age groups (individuals age 20, 30, 40, 50, 60, 70, and 80), each divided into five income quintiles ("Low," "Moderate," "Middle," "High" and "Top"). For the Middle income quintile, the survival probability in each year of life (the probability of surviving to the end of that year, conditional on being alive at the beginning) is set equal to the survival probability at each age in the most recent available U.S. population survival curve.[58] Survival probabilities for the other groups are adjusted so as to roughly match the estimates of life expectancy for different income groups in Chetty et al. (2016).[59] These survival probabilities for the five quintiles are then reduced to account for COVID-19 risk; see below, Appendix C.

Incomes by quintiles are based on the most recent available official data on average after-tax-and-transfer household annual income. By quintile, these incomes are: $35,000; $48,000; $65,900; $91,700; and $214,900. The incomes were divided by the square root of average household size (2.54) to arrive at estimates of after-tax-and-transfer individual income: $21,961; $30,118; $41,349; $57,358; $134,840.[60]

Data about the age distribution of income was used to estimate a time profile of income.[61] I estimated time factors for each year of life, and multiplied the quintile incomes above by the time factors to arrive at the income in that quintile in that year of life. The time factors are as follows (rounded to no decimal places): Ages 0 to 24: 38%. Ages 25 to 29: 84%. Ages 30 to

---

[58] U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, *National Vital Statistics Reports* (vol. 68, No. 7, June 24, 2019), Table 1 (U.S. life table for 2017).

[59] The mortality risk at each age is 1 minus the survival probability. For income quintiles other than Middle income, mortality risks were taken from the U.S. survival curve, and then adjusted by a multiplicative factor in each year. The adjusted survival probabilities are, then, 1 minus the adjusted mortality risks. The multiplicative factors for the Low, Moderate, High and Top quintiles were, respectively 1.5, 1.2, 0.9, and 0.75. These multiplicative factors were chosen so that the ratio between life expectancy at age 40 for individuals in that quintile, and life expectancy at age 40 for 50th percentile individuals, was approximately equal to the ratio as estimated by Chetty et al. (2016).

[60] The most recent data about after-tax-and-transfer household income (for 2016) was taken from: Congressional Budget Office, *Projected Changes in the Distribution of Household Income, 2016 to 2021* (December 2019; available at www.cbo.gov/publication/55941). See *id*. Appendix B, Table B-1. The recommendation to estimate individual income from household income by dividing by the square root of household size is taken from *id*., Appendix A. The estimate of average household size at 2.54 was taken from the 2017 data (to match the survival curve) provided in U.S. Census Bureau, Current Population Survey, "Historical Households Tables," Table HH6 ("Average Population Per Household and Family"), available at https://www.census.gov/data/tables/time-series/demo/families/households.html.

[61] Mean income for various age bands (for 2016) was taken from U.S. Census Bureau, Current Population Survey, "Tables for Personal Income," Table PINC-01 ("Selected Characteristics of People 15 Years and Over, by Total Money Income, Work Experience, Race, Hispanic Origin, and Sex"), available at https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-pinc/pinc-01.2016.html.
The age bands were as follows: 15 to 24; 25 to 29; 30 to 34; 35 to 39; 40 to 44; 45 to 49; 50 to 54; 55 to 59; 60 to 64; 65 to 69; 70 to 74; 75 and over. The income in each band was divided by mean income for all individuals 15 and over ($46,550 as per this table) to arrive at the time factor for each year of life within the age band. The time factor for individuals under 15 was set equal to that for individuals in the age band 15 to 24.

Covid Economics 33, 30 June 2020: 1-45

34: 101%. Ages 35 to 39: 121%. Ages 40 to 44: 129%. Ages 45 to 49: 130%. Ages 50 to 54: 131%. Ages 55 to 59: 124%. Ages 60 to 64: 114%. Ages 65 to 69: 97%. Ages 70 to 74: 90%. Ages 75 and over: 67%.

For example, the lifetime income profile for the Low quintile is such that an individual at age 62 will receive income equaling 114% of the quintile income of $21,961. The lifetime income profile for the Top income quintile is such that an individual at age 38 will receive an income equaling 121% of the quintile income of $134,840. And so forth.

The percentages of the current population in the seven age groups were estimated to be as follows (rounding to no decimal places): age 20: 18%; age 30: 18%; age 40: 16%; age 50: 18%; age 60: 16%; age 70: 10%; age 80: 5%. These percentages were taken from estimates of the current age distribution of the U.S. population.[62]

B. *Utility Function; Utilitarian and Prioritarian SWFs; VSL*

The formulas here for individual expected utility as a function of age, lifetime income profile, and lifetime profile of survival probabilities, and the formulas for the utilitarian and prioritarian SWFs and VSL, are those given in Adler, Ferranna, Hammitt and Treich (2019), except for not including a utility discount factor and changes in notation. This is in turn one version of the general model of utility and VSL presented in Part II. In this version, utility is temporally additive; non-income attributes are ignored; income is consumed when received (myopic consumption); and individuals have homogeneous preferences, captured in a common utility function.

There is a fixed population of N individuals. Each individual's life is divided into periods numbered 1, 2, …, T, with T the maximum number of periods that any individual lives. An individual either dies immediately after the beginning of the period or does not die and survives, at least, until the end of the period.

The number of the current period, for individual i, is A_i. In the status quo, denoted with the subscript B ("baseline"), an individual i has a lifetime profile of survival probabilities: $(p_{i,B}^{A_i},\ldots,p_{i,B}^{T})$. An individual also has a status quo lifetime income profile: $(y_{i,B}^{1},\ldots,y_{i,B}^{A_i-1},y_{i,B}^{A_i},\ldots,y_{i,B}^{T})$. A given policy P, denoted with the subscript P, endows each individual with a new lifetime profile of survival probabilities $(p_{i,P}^{A_i},\ldots,p_{i,P}^{T})$ and income profile $(y_{i,P}^{1},\ldots,y_{i,P}^{A_i-1},y_{i,P}^{A_i},\ldots,y_{i,P}^{T})$.[63]

[62] U.S. Census Bureau, Current Population Survey, "Age and Sex Tables; Age and Sex Composition in the United States," Table 1 (2017) ("Population by Age and Sex"), available at https://www.census.gov/topics/population/age-and-sex/data/tables.2017.html. Each age group in the simulation was assigned the population percentage of the age bracket including that age from the table, and these percentages were then scaled up proportionally to sum to 100%.

[63] Policies don't change individuals' past incomes, so $y_{i,B}^{t}=y_{i,P}^{t}$ for $t<A_i$.

In what follows, the B and P subscripts will be removed if the same generic formula applies both to individuals' baseline profiles of survival probabilities and income amounts, and to individuals' policy profiles of survival probabilities and income amounts. $(\mathbf{p}_i, \mathbf{y}_i)$ is shorthand for an individual profile: $(\mathbf{p}_i, \mathbf{y}_i) = ((p_i^{A_i},\ldots,p_i^T),(y_i^1,\ldots,y_i^T))$. $(\mathbf{p}_{i,B},\mathbf{y}_{i,B})$ and $(\mathbf{p}_{i,P},\mathbf{y}_{i,P})$ are, respectively, individual i's baseline profile and her profile with policy P.

The survival probabilities, denoted with p, are conditional probabilities: p_i^t is the probability that individual i survives to the end of period t, given that she is alive at the beginning. Related probabilities can be derived from these. Let $\pi_i(t; A_i)$, $t \geq A_i$, denote i's probability of surviving to the end of period t, given that she is currently alive at the beginning of A_i. Then $\pi_i(t;A_i) = \prod_{s=A_i}^{t} p_i^s$. Let $\mu_i(t; A_i)$, $t \geq A_i - 1$, denote i's current probability of surviving to the end of period t and then dying during the next period. (In other words, this is the current probability of living exactly t periods.) If $t = A_i - 1$, $\mu_i(t; A_i) = 1 - p_i^{A_i}$, If $t \geq A_i$, $\mu_i(t; A_i) = (1-p_i^{t+1})\pi_i(t;A_i)$.

For a past period ($t < A_i$), y_i^t is the income that i earned in period t. For the current period or a future period ($t \geq A_i$), y_i^t is the income which i will earn in t if she survives to its end.

The common period utility function is denoted as $u(\cdot)$. Let $V_i^s(\mathbf{y}_i)$ denote the individual's realized lifetime utility if she lives exactly s periods. $V_i^s(\mathbf{y}_i) = \sum_{t=1}^{s} u(y_i^s)$. Individual i's current expected utility, denoted as $U_i(\cdot)$,[64] is as follows: $U_i(\mathbf{p}_i,\mathbf{y}_i) = (1-p_i^{A_i})V_i^{A_i-1}(\mathbf{y}_i) + \sum_{t=A_i}^{T} \mu_i(t;A_i)V_i^t(\mathbf{y}_i)$. It can also be shown that $U_i(\mathbf{p}_i,\mathbf{y}_i) = \sum_{t=1}^{A_i-1} u(y_i^t) + \sum_{t=A_i}^{T} \pi_i(t;A_i)u(y_i^t)$.

The utilitarian social welfare function, denoted W^U, assigns a score to a policy or to the baseline as a function of individuals' profiles of survival probabilities and incomes, and ranks policies relative to each and the baseline as a function of these scores. The same is true for the ex post prioritarian social welfare function, denoted W^{EPP}.
$W^U((\mathbf{p}_1,\mathbf{y}_1),\ldots,(\mathbf{p}_N,\mathbf{y}_N)) = \sum_{i=1}^{N} U_i(\mathbf{p}_i,\mathbf{y}_i)$. W^{EPP} is, more precisely, a family of social welfare functions—each defined by some strictly increasing and strictly concave transformation function $g(\cdot)$. $W^{EPP}((\mathbf{p}_1,\mathbf{y}_1),\ldots,(\mathbf{p}_N,\mathbf{y}_N)) = \sum_{i=1}^{N} \left[(1-p_i^{A_i})g(V_i^{A_i-1}(\mathbf{y}_i)) + \sum_{t=A_i}^{T} \mu_i(t;A_i)g(V_i^t(\mathbf{y}_i)) \right]$

[64] The "i" subscripts on $U(\cdot)$ and $V(\cdot)$ could be dropped, since these are common utility functions.

In the simulation model, policies are combinations of changes to individuals' current survival probabilities and current incomes, relative to baseline. The formulas here for VSL and VSL-CBA are for this specific case. VSL$_i$, in this case, is the marginal rate of substitution between individual i's current survival probability and current income. $VSL_i = \frac{\partial U_i / \partial p_i^{A_i}}{\partial U_i / \partial y_i^{A_i}}|_{(\mathbf{p}_{i,B};\mathbf{y}_{i,B})}$. Using the formula for expected utility above, we have that $VSL_i = \frac{\sum_{t=A_i}^{T} \frac{\pi_{i,B}(t;A_i)}{p_{i,B}^{A_i}} u(y_{i,B}^t)}{p_{i,B}^{A_i} u'(y_{i,B}^{A_i})}$, with $u'(\cdot)$ the first derivative.

Let ($\mathbf{p}_i$,$\mathbf{y}_i$) differ from i*'s* baseline profile only with respect to her current survival probability and/or income amount, if at all. $\Delta p_i^{A_i} = p_i^{A_i} - p_{i,B}^{A_i}$. $\Delta y_i^{A_i} = y_i^{A_i} - y_{i,B}^{A_i}$. Then the individual's VSL-based monetary valuation of this profile, MV_i, is as follows: $MV_i(\mathbf{p}_i, \mathbf{y}_i) = \Delta p_i^{A_i} VSL_i + \Delta y_i^{A_i}$. VSL-CBA—like utilitarianism and prioritarianism—assigns a score to a policy or to the baseline as a function of individuals' profiles of survival probabilities and incomes, and ranks policies relative to each other and the baseline as a function of these scores. This score is the sum of MV amounts. $W^{VSL-CBA}((\mathbf{p}_1,\mathbf{y}_1),...,(\mathbf{p}_N,\mathbf{y}_N)) = \sum_{i=1}^{N} MV_i(\mathbf{p}_i,\mathbf{y}_i)$. Note that MV$_i$($\mathbf{p}_{i,B}$,$\mathbf{y}_{i,B}$) = 0 and so the score assigned by VSL-CBA to the baseline is 0. A policy is better than/worse than/equally good as baseline iff its score is positive/negative/zero.

In the simulation model, the periodization is annual and the period utility function is logarithmic. $u(y) = \log(y) - \log(y^{zero})$, with y^{zero} = \$1000. For an explanation of why a choice of y^{zero} is needed to specify the utility function, see Adler (2019, pp. 292-94); Adler (2017, pp. 71-72). y^{zero} can be thought of as subsistence income. Extending a life by one period with income y^{zero} leaves lifetime utility unchanged. The level of \$1000 was chosen based roughly on the World Bank extreme poverty level of \$1.90/day. See Adler (2017, pp. 71-72).

The Atkinson prioritarian SWF, which has attractive axiomatic properties, uses a power function. $g(V) = \frac{1}{1-\gamma} V^{1-\gamma}$, $\gamma > 0$. See Adler (2019, pp. 154-58, 274-75). As γ increases, priority for the worse off increase. I choose γ = 1.5, which has the effect of generally nullifying a utilitarian preference to reduce the risks of those at higher incomes. See Table 4c.

Baseline incomes for individuals in the 35 cohorts are as stated in Appendix A above; baseline survival probabilities are adjusted for the COVID-19 risk, as explained in Appendix C.

VSL^{avg}-CBA, as explained in the main text, uses the same formula as VSL-CBA except for using the population-average VSL to compute individuals' MV amounts.

Let L_i denote individual i's life expectancy with a given vector of survival probabilities. $L_i(\mathbf{p}_i) = (1 - p_i^{A_i})(A_i - 1) + \sum_{t=A_i}^{T} \mu_i(t; A_i)t = (A_i - 1) + \sum_{t=A_i}^{T} \pi_i(t; A_i)$. Let $LE_i^{A_i}$ or, for short, LE_i denote the difference between i's baseline life expectancy if she survives the current period and her baseline life expectancy if she dies during the period. $LE_i = \sum_{t=A_i}^{T} \frac{\mu_{i,B}(t; A_i)}{p_{i,B}^{A_i}} t - (A_i - 1)$. As above, let $(\mathbf{p}_i, \mathbf{y}_i)$ differ from i's baseline profile only with respect to her current survival probability and/or income amount, if at all. $\Delta p_i^{A_i} = p_i^{A_i} - p_{i,B}^{A_i}$. $\Delta y_i^{A_i} = y_i^{A_i} - y_{i,B}^{A_i}$. Then the difference in i's life expectancy between policy and baseline, $L_i(\mathbf{p}_i) - L_i(\mathbf{p}_{i,B}) = \Delta p_i^{A_i} LE_i$. $VSLY_i^{A_i}$ or, for short, $VSLY_i$ equals VSL_i/LE_i.

Let VSLY^{avg} be the population average VSLY. Then VSLY^{avg}-CBA sums monetary valuations defined in terms of this population average.

$$MV_i^{VSLY^{avg}}(\mathbf{p}_i, \mathbf{y}_i) = \Delta p_i^{A_i}(LE_i)(VSLY^{avg}) + \Delta y_i^{A_i}$$

C. *Social Distancing Policy*

Let $o_{i,B}$ denote individual i's baseline (no-governmental intervention) risk of dying during the current year from COVID-19. $o_{i,B}$ for the various cohorts is given in Table 2 (based upon the IFRs from Ferguson et al. (2020) and the estimate that, without intervention, 81% of the population would be infected). $o_{i,P}$ is the individual's risk of dying during the current year from COVID-19 after policy intervention P.

Let p_i* denote individual i's current-year survival probability, but for the pandemic. (These survival probabilities, for the various cohorts, are based upon the U.S. survival curve as described in Appendix A, first paragraph.). Individual i's baseline current-year survival probability with the COVID-19 risk, $p_{i,B}^{A_i}$, is set as follows: $p_{i,B}^{A_i} = p_i^*(1 - o_{i,B})$. Intuitively, i survives the current year only if she is not killed by non-COVID-19 causes and is not killed by COVID-19. Assuming these probabilities are independent—to be sure, a simplification—we have the formula here. Similarly, $p_{i,P}^{A_i} = p_i^*(1 - o_{i,P})$.

I model COVID-19 reduction as a concave function of GDP reduction. If GDP is reduced by π%, $o_{i,P} = o_{i,B}(1 - e\sqrt{\pi / 80})$. In the base case ($e$ =1), 80% reduction wholly eliminates the risk. In the optimistic case (e = √2), a 40% reduction wholly eliminates the risk. In the pessimistic case (e = 0.5), an 80% reduction eliminates half the risk. In the convex case,

$$o_{i,P} = o_{i,B}(1 - (\frac{\pi}{80})^2).$$

On the cost side, regressive incidence is as follows. After a π% reduction in GDP: (1) if individual i is in the Low, Moderate, or Middle income quintiles, $y_{i,P}^{A_i} = y_{i,B}^{A_i} - (\frac{\pi}{80})(y_{i,B}^{A_i} - 1000)$; and (2) if individual i is in the High or Top income quintiles, $y_{i,P}^{A_i} = y_{i,B}^{A_i}(1-(\frac{\pi}{80})q)$, with q chosen based upon cohort baseline incomes so that the reduction in total income is π%. In the model here, $q = 0.7187$.

Proportional incidence is straightforward: After a π% reduction in GDP, $y_{i,P}^{A_i} = y_{i,B}^{A_i}(1-(\frac{\pi}{100}))$ for individuals in all cohorts.

D. Equivalent and Compensating Variations

Using the utility model set forth in Appendix B., we can define equivalent variations (EV_i) and compensating variations (CV_i). $EV_i(\mathbf{p}_i, \mathbf{y}_i) = \Delta y^*$ such that: $U_i((p_i^{A_i},...,p_i^T),(y_i^1,...,y_i^T)) = U_i((p_{i,B}^{A_i},...,p_{i,B}^T),(y_{i,B}^1,...,y_{i,B}^{A_i-1}, y_{i,B}^{A_i}+\Delta y^*,...,y_{i,B}^T))$. $CV_i(\mathbf{p}_i, \mathbf{y}_i) = \Delta y^+$ such that: $U_i((p_{i,B}^{A_i},...,p_{i,B}^T),(y_{i,B}^1,...,y_{i,B}^T)) = U_i((p_i^{A_i},...,p_i^T),(y_i^1,...,y_i^{A_i-1}, y_i^{A_i}-\Delta y^+,...,y_i^T))$.

Assume, more specifically, that $(\mathbf{p}_i, \mathbf{y}_i)$ differs from baseline only with respect to current survival probability and/or current income, if at all. As a short hand, let U_i denote $U_i(\mathbf{p}_i, \mathbf{y}_i)$ and $U_{i,B}$ denote $U_i(\mathbf{p}_{i,B}, \mathbf{y}_{i,B})$. Then (1) $U_{i,B} = \sum_{t=1}^{A_i-1} u(y_{i,B}^t) + \pi_{i,B}(A_i;A_i)u(y_{i,B}^{A_i}) + \sum_{t=A_i+1}^{T} \pi_{i,B}(t;A_i)u(y_{i,B}^t)$.
Also, the compensating variation Δy^+ is s.t. (2) $U_{i,B} =$
$\sum_{t=1}^{A_i-1} u(y_{i,B}^t) + \pi_i(A_i;A_i)u(y_i^{A_i}-\Delta y^+) + \sum_{t=A_i+1}^{T} \pi_i(t;A_i)u(y_{i,B}^t)$

(3) $U_i = \sum_{t=1}^{A_i-1} u(y_{i,B}^t) + \pi_i(A_i;A_i)u(y_i^{A_i}) + \sum_{t=A_i+1}^{T} \pi_i(t;A_i)u(y_{i,B}^t)$. Also, the equivalent variation Δy^* is s.t. (4) $U_i = \sum_{t=1}^{A_i-1} u(y_{i,B}^t) + \pi_{i,B}(A_i;A_i)u(y_{i,B}^{A_i}+\Delta y^*) + \sum_{t=A_i+1}^{T} \pi_{i,B}(t;A_i)u(y_{i,B}^t)$.

Subtracting the fourth equation from the first, we have that $U_{i,B} - U_i =$ $\pi_{i,B}(A_i;A_i)\left(u(y_{i,B}^{A_i}) - u(y_{i,B}^{A_i}+\Delta y^*)\right)$. Subtracting the third from the second, we have that $U_{i,B} - U_i = \pi_i(A_i;A_i)\left(u(y_i^{A_i}-\Delta y^+) - u(y_i^{A_i})\right)$. Observe also that $\pi_{i,B}(A_i;A_i) = p_{i,B}^{A_i}$ and that $\pi_i(A_i;A_i) = p_i^{A_i}$.

Covid Economics 33, 30 June 2020: 1-45

With the logarithmic specification of $u(\cdot)$, $u(y) = \log(y) - \log(1000)$, we can use these last equations to solve explicitly for Δy^* and Δy^+. $\Delta y^* = y_{i,B}^{A_i}\left(\exp\left(\frac{U_i - U_{i,B}}{p_{i,B}^{A_i}}\right) - 1\right)$.

$$\Delta y^+ = y_i^{A_i}\left(1 - \exp\left(\frac{U_{i,B} - U_i}{p_i^{A_i}}\right)\right).$$

The gender gap in mental well-being during the Covid-19 outbreak: Evidence from the UK[1]

Ben Etheridge[2] and Lisa Spantig[3]

Date submitted: 24 June 2020; Date accepted: 25 June 2020

We document a decline in mental well-being after the onset of the Covid-19 pandemic in the UK. This decline is more than twice as large for women as for men. We seek to explain this gender gap by exploring gender differences in: family and caring responsibilities; financial and work situation; social engagement; health situation, and health behaviours, including exercise. We discuss two dimensions of gender differences, the extent to which particular circumstances relate to well-being and the share of individuals facing a given circumstance. Overall, we find that differences in family and caring responsibilities can explain a part of the gender gap, but the bulk is explained by social factors such as loneliness. Other factors such as financial difficulties or age are similarly distributed across genders and thus play little role in explaining the gap.

1 Thanks to James Banks for comments. Further comments welcome. All errors remain the responsibility of the authors. This paper draws on data from Understanding Society, distributed by the UK Data Service. Understanding Society is an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public.
2 Department of Economics, University of Essex.
3 Department of Economics, University of Essex.

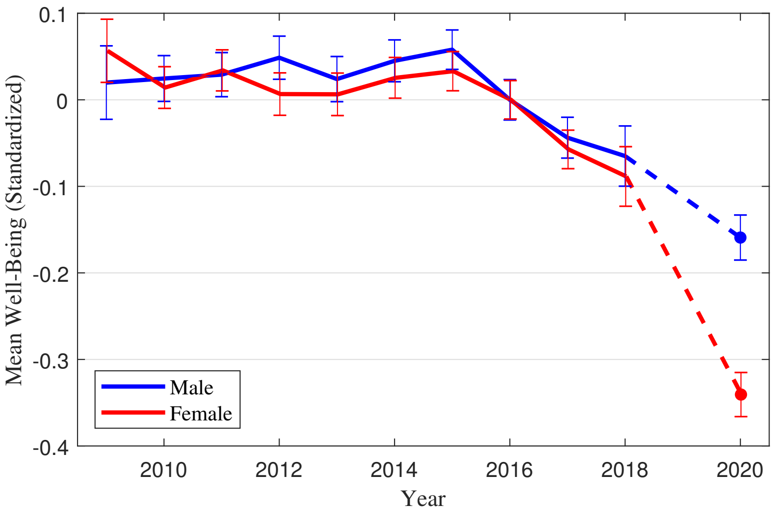
Copyright: Ben Etheridge and Lisa Spantig

Covid Economics 33, 30 June 2020: 46-72

1 Introduction

The Covid-19 pandemic has caused large disruption to much of the population across the globe, and along many dimensions. Early evidence suggests this disruption has negatively and substantially affected mental well-being (Adams-Prassl et al. 2020b; Banks and Xu 2020; Davillas and Jones 2020). This same early evidence suggests that the effects on well-being have been felt unequally, and have been noticeably born by women.[1]

We document a large decline in mental well-being after the Covid outbreak in the UK. To illustrate, Figure 1 displays average well-being by gender over time, including after the pandemic's onset. It shows a large drop after onset, and, consistently with the existing evidence, a disproportionate effect on women, for whom the impact appears to be over twice as large.[2] In this paper, we use



Notes: Data from UKHLS waves 1-9 and Covid module. Data for 2019 are collected mainly in wave 10, which is not yet publicly available. Figure shows standardized, seasonally-adjusted and inverted Likert score, obtained from 12 questions in the General Health Questionnaire. See section 2 for more details. Profiles indexed to 0 in 2016. Sample consists of all those responding to the Covid module, whether or not they responded to previous waves. Error bars are 95% confidence intervals.

rich representative data from the UK Household Longitudinal Survey to explore potential reasons

[1] In addition to the papers mentioned above see further the dedicated literature review below.

[2] Given the low frequency of the data, we cannot causally attribute the decline to the pandemic alone. Banks and Xu (2020) focus on the issue of time trends and show that, after accounting for these, a large portion of the decline remains unexplained. Given largely similar profiles for men and women before 2020, controlling for trends has a negligible effect on the gender gap. We thus refrain from making additional assumptions that would be required for the analysis of trends and focus on the gender gap in the 'raw' data.

Covid Economics 33, 30 June 2020: 46-72

for this differential impact. We first relate changes in well-being at the individual level (measured by the General Health Questionnaire (GHQ)) with various changes in circumstance and discuss to which extent women and men face differential exposure to these circumstances. In particular, we document how well-being of women and men is related to a variety of factors that have been shown to be affected by Covid, such as the economic situation and time use within the household. Building on the literature in psychology (see e.g. Holt-Lunstad et al. 2015, for a recent review), we additionally consider social circumstances such as friendships and loneliness. We then examine to what extent differential exposure to these factors can help explain the widened gender gap. Accordingly we shed light on which of the particular dimensions of disruption might be affecting well-being to the largest degree.

As of today, the UK has been one of the countries most affected by the Covid pandemic. At its peak in mid-April, the 7-day moving average of official daily deaths was 950 (14 per million per day), among the highest rates in the world. Meanwhile, the data we examine were collected only a little afterwards, when the death rate was around 800 per day.[3] At this time the 'lockdown' was in full force, including strict social distancing measures.[4] Indicators of economic activity were sharply negative.[5] At the same time, the main policy tools relating to the economy, such as the UK Job Retention ('furloughing') Scheme, were already well established.[6]

Regarding time use and economic factors that have been shown to be affected by the pandemic, we find we find that those with high childcare duties have shown noticeable deteriorations in well-being. A higher number of women report these duties and women are more affected by them than men. Similarly we find large declines in well-being reported by those in a tough financial situation, with similar numbers of women and men facing such circumstances. Consistently with the literature (L. Winkelmann and R. Winkelmann 1998), large declines in well-being are reported by those who have entirely lost their job. However, the size of this group is dwarfed by the number of workers who've seen a reduction in hours, yet remain in contact with their employer, such as through furloughing. For these workers, declines in well-being are small, which indicates the beneficial effects of the unusual labour market policies in place at the time.

In addition, we document a strong correlation between declines in well-being and social factors. The declines in well-being are particularly large for those who report often feeling lonely, and similarly, those who report an increase in loneliness since their last pre-Covid interview. These correlations are larger for women and women are more likely than men to report higher levels of loneliness. To investigate this further, we make use of the background data collected before 2020. We find

[3]The data were collected from April 24th. See Section 2. Death rates obtained from https://www.worldometers.info/coronavirus/country/uk/, accessed on June 4th 2020.

[4]See, for example, the cross-country tracker of policy responses in Hale et al. (2020).

[5]For example, over the weekend during which most of our data were collected the FTSE 100 stock market index stood at 5750 points, 25% below its level at the beginning of the year.

[6]The UK Job Retention Scheme was introduced on March 20, 2020. All the other main schemes were introduced at a similar time, including the Self-Employed Income Support and Mortgage Relief Schemes, among others.

Covid Economics 33, 30 June 2020: 46-72

that those who previously reported fewer friends are less affected by the pandemic, presumably because they are impacted less by the social distancing policies imposed. In terms of demographics, we document that those aged between 16 and 30, both men and women, have been much more negatively affected than older individuals.

Our main empirical exercise is to examine which of these factors help explain the gender gap in well-being on aggregate. Here we focus on *differential exposures* across women and men. For example, even though the young have been more affected than the old, the gender composition across age categories is balanced. Therefore, gender differences in well-being cannot be explained by age. Likewise, women and men have born adverse financial outcomes similarly, and so the gender gap is not explained by this factor. To explain the gap, we find some role for family time use and childcare. Women spend more time with their children than men. However, the large majority of adults do not have young children, and so the role of this factor overall is limited. In fact, most important appear to be social factors. This is likely explained by the two observations mentioned above. First, a larger number of women reports high levels of loneliness. Second, higher levels of loneliness are strongly related to declines in well-being. Thus, social factors appear to be important in explaining differential impacts by gender.

In terms of implications for policy, our results are suggestive of the strongly adverse and unequal effect of social distancing. Our results can therefore inform debates at a time when policy makers must weigh up many competing objectives. In particular, our results suggest that lockdown is impacting mental well-being of women less through its effect on the labour market or wider economy, and more through the direct loss of social interaction. Further evidence of an adverse effect of policies rather than, say, the virus itself, is the fact that younger individuals see larger well-being declines, despite the widespread reporting that the young are not as affected by the disease.

After a review of the literature the paper proceeds as follows. We describe the data in Section 2. In Section 3 we show the gender-specific distributions of well-being scores and present gender-specific correlations of variables that have been suggested or shown to affect mental well-being. We also discuss the distribution of gender within these variables. We then evaluate to what extent these variables can help explain the gender gap in mental well-being that has arisen in the UK during the pandemic. Section 4 concludes.

Related Literature

In general, it has been established that mental well-being is related to both economic and social factors. In particular, mental well-being can be negatively affected by bad economic outcomes. For example, using quarterly British data from 2002-2016, (Janke et al. 2020) estimate that a one percent increase in the employment rate leads to a 4.2% reduction in mental health conditions.

Evidence from previous economic downturns suggests that such events affect mental health of men and women differently (Chang et al. 2013; Dagher et al. 2015).

Regarding social factors, both objective isolation and its subjective perception (loneliness) can negatively impact mental health (Cacioppo et al. 2015, 2011; Holt-Lunstad et al. 2015; House et al. 1988).[7] Focusing on subjective well-being specifically, loneliness has been shown to have a moderately strong association (VanderWeele et al. 2012). Even though closely related and similarly impacting well-being, loneliness needs to be conceptually distinguished from social support networks (Golden et al. 2009). We thus include both the number of close friends as a proxy for social support networks and loneliness in our analyses.

Shortly after the outbreak of Covid-19, international organizations and researchers started to warn about not only the immediate physical, but also the psychological effects of the pandemic (Holmes et al. 2020; World Health Organization 2020). Based on a review of existing evidence Brooks et al. (2020) conclude that quarantine is negatively related to a variety of psychological factors, including (post-traumatic) stress symptoms and anxiety. Indeed, Fetzer et al. (2020a) causally show that higher perceived mortality and contagiousness increases anxiety related to economic outcomes.

Early Covid-related studies have started to examine the channels through which well-being and mental health are being affected. Results from google trends analyses suggest that lockdowns can have severe mental health implications with search terms loneliness, worry and sadness increasing under lockdowns in different countries (Brodeur et al. 2020; Knipe et al. 2020; Tubadji et al. 2020). Other research suggests that a more nuanced view is necessary: Fetzer et al. (2020b) show that mental well-being temporarily *increases* with lockdowns and relate this to the fact that many individuals perceived the early government responses to be inefficient and are thus relieved when lockdowns are imposed. Evidence from Germany suggests that mental health (measured by calls to the largest helpline) worsened just after the lockdown and started easing with the third week of lockdown (Armbruster and Klotzbuecher 2020). The authors discuss that the decline in mental health is not driven by financial worries or fear of the disease, but is due to higher levels of loneliness and anxiety. Contrary to this, and using Swiss helpline data, (Brülhart and Lalive 2020) find an increase only in calls directly related to the pandemic (i.e. elderly persons calling and calls related to fear of infection). Early results from a UK stakeholder survey at the very beginning of the pandemic (Cowan 2020) reveal that many concerns regarding mental health revolve around anxiety, isolation and access to support. Further worries concern the impact of the pandemic on family and relationships. The report hints at differential concerns: women report being more worried about isolation, social distancing and mental health.

Even early in the pandemic, it became clear that the crisis would have differential economic and health impacts on different socio-economic groups (e.g. Alon et al. 2020; Dingel and Neiman 2020). Adams-Prassl et al. (2020a) collect two independent waves of survey data in late March and early

[7] In general, women are more likely to be affected by anxiety (Remes et al. 2016).

Covid Economics 33, 30 June 2020: 46-72

April in the UK, US and Germany. They find that women in the UK and the US (though not in Germany) are 5 percentage points more likely to have lost their jobs than men and are 5 percentage points less likely to be able to work from home. Additionally, women spend significantly more time taking care of children and homeschooling, even if they are still employed and able to work from home. For those who still work, there is no gender difference in the likelihood of experiencing a fall in income. A closer look within the household is offered by Andrew et al. (2020) who conducted a survey of UK families around early May with time use data collected from both parents. They document that mothers and fathers in two opposite-gender households are differentially affected by the lockdown. These differences in economic outcomes and time use during the lockdown appear to be natural candidates to explain the gender gap in mental well-being. In terms of mental well-being, Adams-Prassl et al. (2020b) find in a cross-sectional survey in the US that women tend to be worse off during lockdown.

The main drawback of many of these early studies is that they rely on cross-sectional surveys with limited background characteristics of respondents. We use rich data from an established longitudinal survey (University of Essex 2019) that allows the analysis of 1) within-person changes, 2) background characteristics before the outbreak of the pandemic and 3) different individuals in the same household. As such, our work is most similar to Banks and Xu (2020), Daly et al. (2020), Davillas and Jones (2020) and Zhou et al. (2020), who use the same data to us.[8] While all these studies document larger declines in mental well-being effects during the pandemic for women, they do not investigate in detail the contributory factors to these differential impacts.

More broadly, our work relates to an established literature that addresses gender gaps in well-being over time. For example, (Stevenson and Wolfers 2009) address the 'paradox' of declining women's happiness in the U.S. against the background of increasing success of women across a range of economic and social spheres. In fact we similarly find a persistent gender gap in well-being scores across all waves of UKHLS. It should be noted, however, that in our analysis we difference out all individual heterogeneity in reporting and focus on the *extra* effect of the Covid outbreak. Overall, and in the context of this literature, our work is therefore informative about the role of differences in social needs and social engagement in the production of mental well-being across genders.

2 Data

We use the Covid-19 module from the UK Household Longitudinal Survey (UKHLS), administered monthly from April 2020. The current dataset uses the Covid module's first wave and will be updated once new waves become available. The interviews used here were conducted in the 7 days from Friday April 24, with 75% of interviews completed by Sunday April 26. We merge these

[8]Benzeval et al. (2020) use the same data and document large heterogeneity in *economic* impacts.

Covid Economics 33, 30 June 2020: 46-72

data with waves 1-9 of the 'parent' UKHLS (also known as 'Understanding Society'), a large-scale national survey administered yearly from 2009.

The UKHLS Covid April wave was conducted entirely over the internet. The underlying sampling frame consists of all those who participated in the UKHLS main survey's last two waves. To conduct the fieldwork, the sample was initially contacted using a combination of email, telephone, postal and SMS requests. From the underlying sampling frame, the response rate was a little under 40%. To adjust our analysis for non-response, we use the survey weights provided. In addition, to allow for the fact that many respondents are related either through primary residence or through the extended family, we cluster all regressions at the primary sampling unit level. For a further discussion of the Covid module and underlying UKHLS design see Social and Research (2020), ISER (2020).

The main variable of interest is mental well-being. Our measure is derived from the Likert index that sums 12 questions from the General Health Questionnaire (GHQ-12). The GHQ battery asks questions regarding, for example, the ability to concentrate, loss of sleep and enjoyment of day-to-day activities. Importantly, the questionnaire asks participants to evaluate their well-being with respect to 'usual' and thus induces a reference point against which respondents evaluate their current feelings. This feature distinguishes our measure from other measures of mental well-being such as the WHO 5-question module (used e.g. in Adams-Prassl et al. 2020b) or the PHQ9 depression questionnaire (adopted e.g. in Fetzer et al. 2020b) that ask about occurrence of specific feelings or behaviors over the last two weeks. While the latter measures have been shown to reflect the cognitive dimension of well-being, our measure captures affective well-being (see e.g. Diener et al. 1985). The GHQ-12 from this survey has been widely used, both in psychological (e.g. Bridger and Daly 2019) and other social sciences research (e.g. Clark et al. 2019; Davillas et al. 2016; Davillas and Jones 2020; Powdthavee et al. 2019). Importantly the GHQ questionnaire has been administered in all waves of UKHLS in exactly the same form. For precise details on the GHQ questionnaire see Appendix B.

Each component of the GHQ can be scaled from 0 (least distressed) to 3 (most distressed). The Likert score is obtained by summing these scores to yield a total score between 0 and 36. We standardize this score across all waves and invert it so that, in our analysis, lower scores indicate lower well-being. To remove seasonal effects in mood, we take account for month effects, adjusting all pre-Covid data to 'April equivalents'. To remove individual factors in reporting style, we typically use differences of the Covid-modules measures from wave 9. It should be noted that wave 9 data were collected mainly in 2017, but also, to a lesser extent, in 2018 and 2019. We treat all these data as uniformly 'pre-Covid' and, other than by the seasonal adjustment, do not adjust for differences in interview timing.

We make use of the extensive background information collected in the Covid April wave, as well as the prior UKHLS surveys. In the Covid module, participants were asked a battery of questions about their current experiences. These include questions on employment, on health, on caring

Covid Economics 33, 30 June 2020: 46-72

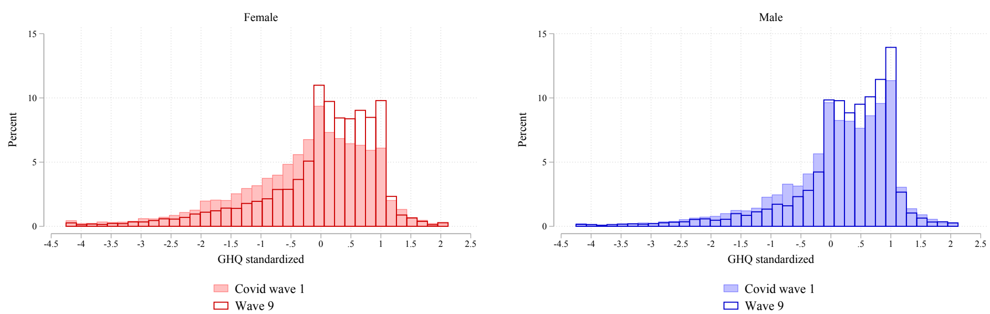
responsibilities, on time use and childcare, as well as self-assessments of financial situation and feelings of loneliness. From previous waves of UKHLS, we make particular use of a specific module conducted in wave 9 on social networks. This module contains detailed self-reports on the quantity, intensity and nature of friendships. Our measure of gender is self-reported sex, copied over from UKHLS wave 9. The adjusted number of interviews for which full information is available on all measures, including relevant measures from wave 9, is $12,250$.

3 Results

3.1 Distribution of Well-Being by Gender

We focus on wave 9 (last wave before 2020) and the first wave during the pandemic, to take a closer look at the distribution of well-being by gender. Figure 2 shows the distribution of standardized well-being scores by gender.[9] The left panel shows scores for women, where solid bars indicate values in 2020 and transparent bars show values in wave 9 (mostly in 2017). A solid bar being larger than a transparent one implies that a larger fraction of participants scored this value in 2020 than in wave 9. For both women and men, the left tail has become fatter, suggesting that a wide spectrum of individuals has been affected.

Figure 2: Distributions of Mental Well-Being Before and During the Pandemic, by Gender



Notes: Data from UKHLS wave 9 and Covid module. Figure shows standardized and inverted Likert score, obtained from 12 questions in the General Health Questionnaire.

Examining within-individual changes, we find that about 54 percent of respondents have worse mental well-being in 2020 than in wave 9. Sixty-one percent of those are women. Eleven percent of respondents do not see a change in their well-being and 34 percent have better well-being in 2020

[9]For the ease of exposition, this figure is not seasonally adjusted. In all of our regressions reported below, we control for seasonality.

Covid Economics 33, 30 June 2020: 46-72

than in wave 9. The composition of these latter two groups across genders is roughly balanced.[10] In the following, we will explore the gap more rigorously and discuss how factors mentioned in the introduction, such as economic concerns, differences in time use and changes in social interactions contribute to these patterns.

3.2 Declines in Well-Being by Gender and by Salient Factors

We now present correlations of a variety of background characteristics/circumstances with the change in subjective well-being. Here, we make use of the panel structure of the data to calculate within-individual changes in subjective well-being by taking the first difference in measures (using wave 9 and the first Covid-19 wave). This holds constant any individual differences in reporting style. We present correlations separately for women and men and in light of potential differences in exposure to the circumstances that men and women face.

Time Use

We start with factors that relate to the situation within the household. Alon et al. (2020) discuss that the closure of schools and daycare facilities is likely to affect women more than men and that these effects are likely to be stronger than effects relating to employment. For the UK, Andrew et al. (2020) show that mothers in households with two opposite-gender parents bear a disproportionate share of household responsibilities. We therefore examine whether changes in well-being are related to caring duties, child care and time spent doing housework. Accordingly, Table 1 shows the change in well-being by gender and when individuals are grouped according to their current time use. Importantly, Table 1 includes all respondents, with and without children.

Columns 1 and 4 show that when splitting respondents by the time they are currently spending on child care, all of the different groups face on average worse well-being in 2020. Notably, there appears to be no significant difference in average well-being changes between those with currently moderate childcare duties (1 to 15 hours per week), and those with no childcare duties. Only those with substantial child care duties of over 15 hours appear to be significantly worse off, and women more so than men.

We show the proportions of the sample making up each category in Figure A.1, which we use extensively to discuss differential exposure. The figure indicates that the majority of those with substantial child care responsibilities are women.[11] However, it also shows clearly that the overwhelming majority of both women and men do not perform any childcare at all. This is due to the fact that most adults do not have young children. These proportions should be born in mind when

[10] Fifty-two percent of those without a change and 44 percent of those with better well-being are men.
[11] 14 percent of all women fall in this category, as do 8 percent of all men.

Covid Economics 33, 30 June 2020: 46-72

Table 1: Well-Being by Gender: Family, Time Use and Caring

	Female	Female	Female	Male	Male	Male
Child Care: 0 hrs	-0.23***			-0.11***		
	(0.02)			(0.02)		
1 to 15 hrs	-0.21***			-0.12**		
	(0.06)			(0.05)		
> 15 hrs	-0.43***			-0.25***		
	(0.05)			(0.06)		
House Work: < 6 hrs		-0.12***			-0.11***	
		(0.04)			(0.03)	
6 to 10 hrs		-0.26***			-0.12***	
		(0.03)			(0.03)	
> 10 hrs		-0.31***			-0.13***	
		(0.02)			(0.03)	
Caring Duties: No			-0.24***			-0.12***
			(0.03)			(0.02)
Yes			-0.28***			-0.13***
			(0.02)			(0.02)
Observations	7235	7237	7237	5243	5242	5242
Adjusted R^2	0.056	0.056	0.053	0.021	0.019	0.019

Notes: Data from UKHLS wave 9 and Covid module. Table reports grouped means of outcome variable, which is the individual change in standardized, seasonally-adjusted and inverted GHQ Likert score. Standard errors clustered at the primary sampling unit and presented in parentheses. Covid survey weights used in all computations. Child care and house work hours from self-reported time use in the previous week. Caring duties variable is from self-report on responsibilities of caring for somebody outside the current residence in the previous 4 weeks.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

attempting to explain the gender gap in well-being in the aggregate. Even if well-being is negatively impacted by substantial childcare duties, and women perform a larger share of childcare duties, it seems unlikely that childcare alone can explain the gender gap overall.

Columns 2 and 5 examine the relationship between changes in well-being and time currently spent on housework. Again, all three groups face on average worse mental health in 2020 and this is true for both genders. However, the decrease in well-being seems to be declining further with more house work for women, but not for men. We also note imbalances in how many men and women fall into each of the three categories: 39 percent of men report doing housework for less than six hours, whereas only 20 percent of women are in the same situation (top center graph in Figure A.1). Over half of all surveyed women report substantial housework (more than ten hours per week); for men this number is 29 percent.

Columns 3 and 6 show the correlation of changes in well-being with caring responsibilities for others outside of the household. Both groups, those who care for others and those who don't,

Covid Economics 33, 30 June 2020: 46-72

face on average similar well-being declines in 2020. Again, we note a gender imbalance in caring responsibilities: 51 percent of women report caring for others while this number is 45 percent for men (top right graph in Figure A.1).

Economic Impacts

Much of the current literature on consequences of the pandemic has focused on economic impacts such as hours worked or facing financial difficulties. Table 2 shows the relationship of changes in well-being by gender and various indicators of economic position. Columns 1 and 4 show mean group correlation for a subjective measure asking how well respondents are getting by.[12] We use this measure as a summary of the complex impacts of loss of earnings and other incomes, as well as changes in expenditure patterns induced by the pandemic. We note a similar pattern as in Table 1, where all subgroups face a decline in well-being on average, irrespective of gender. Not surprisingly, we see a stronger average decline for those who report a worse subjective financial situation. The majority of respondents are 'living comfortably' or 'doing alright' (76 percent of women and 79 percent of men) and only a small fraction finds their situation 'quite' or 'very' difficult (six percent of each gender; see also middle left graph in Figure A.1).

Columns 2 and 4 show similar patterns for bill payments, a more objective measure of financial situation. Here, the average decline in well-being for those who are behind with all bills is not statistically significant for either women nor men, likely because the numbers affected are very small ($N = 18$ and 19, respectively).

In Columns 3 and 6 we turn to furloughing and job loss, the latter of which is usually a strong predictor of subjective well-being (L. Winkelmann and R. Winkelmann 1998). Those who have lost their jobs fully have seen large declines in well-being, with women being affected worse than men. However, only less than one percent of the sample falls into this category, which implies that the explanatory power of job loss for the gender gap is likely to be limited. More usually, hours have been cut or employees have been furloughed. The decline in well-being is not significantly different for these individuals than for those who did not experience a reduction in working hours. Examining the fraction of women and men who lost their job or were furloughed, we do not see a difference between gender (15 percent for both genders; middle right graph in Figure A.1). Overall, we note that in terms of the financial and employment measures used here, women and men do not appear to be very differently affected. It is thus unlikely that financial measures can help explain a large share of the gender gap documented in Figure 1.[13]

[12] Respondents are asked 'How well would you say you yourself are managing financially these days? Would you say you are...', and then given 5 options: 'Living comfortably'; 'Doing alright'; 'Just about getting by'; 'Finding it quite difficult', and 'Finding it very difficult'.

[13] Note however, that particular subgroups of women, such as mothers in households with two opposite-gender parents or single mothers appear to be economically worse affected by the pandemic (Andrew et al. 2020; Benzeval et al. 2020).

Covid Economics 33, 30 June 2020: 46-72

Covid Economics 33, 30 June 2020: 46-72

Table 2: Well-Being by Gender: Finances and Work

	Female	Female	Female	Male	Male	Male
Finances: Comfortable	-0.17***			-0.04*		
	(0.02)			(0.02)		
Doing alright	-0.25***			-0.12***		
	(0.03)			(0.02)		
Just about getting by	-0.38***			-0.18***		
	(0.05)			(0.05)		
Quite difficult	-0.36***			-0.36***		
	(0.12)			(0.12)		
Very difficult	-0.61***			-0.69***		
	(0.20)			(0.26)		
Bills: Up to date		-0.25***			-0.11***	
		(0.02)			(0.02)	
Behind with some		-0.40***			-0.35***	
		(0.09)			(0.12)	
Behind with all		-0.42			-0.20	
		(0.71)			(0.29)	
No hours reduction			-0.24***			-0.12***
			(0.02)			(0.02)
Hours cut or furlough			-0.31***			-0.13**
			(0.05)			(0.05)
Job loss			-1.22***			-0.69***
			(0.27)			(0.24)
Observations	7233	7209	7238	5239	5223	5243
Adjusted R^2	0.058	0.054	0.053	0.030	0.021	0.019

Notes: Data from UKHLS wave 9 and Covid module. Table reports grouped means of outcome variable, which is the individual change in standardized, seasonally-adjusted, and inverted GHQ Likert score. Standard errors clustered at the primary sampling unit and presented in parentheses. Covid survey weights used in all computations. Finances score is from self-report of present financial situation: variable 'finnow'. Lateness of bill payments is from self-report: variable 'xpbills'.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Social Factors

Perhaps the most immediate consequence of the pandemic has been social distancing induced by the lockdown policy. We therefore examine the role of social relationships and loneliness, which have been associated with subjective well-being in a predominantly psychological literature (for a review see e.g. Cacioppo et al. 2015). Table 3 shows social factors and their correlation with changes in well-being.

In contrast to the previous tables, in which we document a decline in well-being across all groups, different levels of social factors are differently related to changes in well-being. For respondents who

Table 3: Well-Being by Gender: Social Factors

	Female	Female	Female	Male	Male	Male
Lonely: Never	-0.02			0.01		
	(0.02)			(0.02)		
Sometimes	-0.41***			-0.35***		
	(0.03)			(0.04)		
Often	-1.01***			-0.82***		
	(0.08)			(0.12)		
Friends: 0		0.11			0.05	
		(0.11)			(0.08)	
1 to 3		-0.22***			-0.07**	
		(0.03)			(0.03)	
4 to 6		-0.29***			-0.16***	
		(0.03)			(0.03)	
> 6		-0.31***			-0.16***	
		(0.03)			(0.03)	
Lonely: Less			0.26***			0.24***
			(0.04)			(0.05)
Same			-0.19***			-0.07***
			(0.02)			(0.02)
More			-0.89***			-0.77***
			(0.04)			(0.06)
Observations	7237	7160	7234	5242	5161	5236
Adjusted R^2	0.120	0.052	0.150	0.074	0.018	0.107

Notes: Data from UKHLS wave 9 and Covid module. Table reports grouped means of outcome variable, which is the individual change in standardized, seasonally-adjusted and inverted GHQ Likert score. Standard errors clustered at the primary sampling unit and presented in parentheses. Covid survey weights used in all computations. Loneliness score is from self-report in Covid module: variable 'sclonely'. Number of close friends is from the social networks module conducted in wave 9, with total count of close friends grouped into bins. Third and sixth columns use change in loneliness variable (sclonely) from wave 9.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

currently state they never feel lonely (71 percent of men and 55 percent of women), the loneliness variable is not correlated with well-being (see Columns 1 and 4).[14] However, for those who report being lonely, the correlation between loneliness and well-being changes is strongly increasing in the level of reported loneliness. While this pattern holds for both genders, the fraction of affected women is higher: Thirty-four percent of women sometimes feel lonely and eleven percent often feel lonely, while these numbers are lower for men (23 and 6 percent, respectively; see also bottom left graph in Figure A.1).

[14]Note this is not explained by a lack of power as nearly two thirds of the sample fall into this category.

Covid Economics 33, 30 June 2020: 46-72

A useful feature of UKHLS is that we can examine background characteristics from before the onset of the pandemic. To examine the role of social connectedness in more detail, we make use of a special module conducted in wave 9 that elicits the number of close friends. *A priori*, it is not clear how the number of close friends would relate to subjective well-being during the pandemic and in particular the lockdown. On the one hand, one might hypothesise that a strong social network can help coping with such a difficult situation, thus leading to a positive correlation between number of friends and well-being changes. On the other hand, related to the above discussion of loneliness, being more connected might lead to increased feelings of loneliness during physical distance measures and lockdowns. Columns 2 and 4 suggest that the latter explanation applies: individuals with more close friends face larger declines in well-being. The pattern is similar for women and men, but the correlation between the number of close friends and changes in well-being appear to be stronger for women. Interestingly, the proportions of women and men who fall into the respective categories is largely similar: 64 percent of women report more than three close friends while only 2 percent report no friends at all; for men the respective proportions are 61 percent and 4 percent.

Finally, returning to loneliness, we check whether the declines in well-being are associated with some persistent loneliness trait or reported changes in this variable. The results are shown in Columns 3 and 6. Those who report less loneliness (in the Covid wave as compared to wave 9) show substantially *higher* well-being, while, consistently with columns 1 and 4, those who report an increase in loneliness are substantially less happy. While 21 percent of women report more loneliness, the respective fraction of men amounts to only 14 percent.

Other Factors

We investigate additional correlations of well-being changes with medical and health factors, health behaviors and key demographics. These are presented in Appendix tables A.1, A.2 and A.3, respectively. In sum, we find negative average changes in well-being for all groups and largely similar patterns for both genders.

Regarding medical factors, we see that those who experienced Covid symptoms (six percent of the sample) and those who are receiving help from outside the house (32 percent) experience larger declines in well-being. The only category that is not statistically significantly correlated with changes in well-being is being 'vulnerable', i.e. having received a letter to stay at home or being pregnant.[15]

Regarding health behaviors (shown in Table A.2), the most notable finding is that not being able to eat healthy meals is associated with significant declines in well-being, especially for women. In contrast, we do not see large differences in well-being changes for those who report to exercise as compared to those who do not. Likewise, those who consume alcohol appear to be similarly affected as those who abstain.

[15]Only four percent fall in this category and the coefficient is less precisely estimated.

Covid Economics 33, 30 June 2020: 46-72

Covid Economics 33, 30 June 2020: 46-72

We present correlations with key demographics in Table A.3. Interestingly, being in a couple seems only to have comparatively benefitted men. Those with children face slightly larger declines in well-being than those without. Most notable, however, are results by age (in line with Banks and Xu 2020; Davillas and Jones 2020). Youths of both genders face a substantially larger decline in well-being than older individuals. This is particularly interesting as overall, the young appear to be comparatively more affected by the policy response to Covid than by Covid itself.

3.3 Explaining the Gender Gap

We discussed in the previous subsection that women's well-being might be more affected by the pandemic when facing a given circumstance. This would be reflected in stronger correlations of being in a given category and well-being changes for women (shown in Tables 1, 2 and 3). This is the case for variables such as hours spent on child care (all three correlations are stronger for women than for men), job loss (the correlation is nearly twice as large for women) or having more than three friends. Additionally, we discussed that, while women might be similarly affected as men by a given circumstance (reflected in similar correlations as for e.g. spending less than six hours on house work), they might have a different likelihood of facing the circumstance (as shown in Figure A.1; in the example of house work, the fraction of women falling into the less than six hours category is 20 percent, while the fraction is nearly twice as large for men (39 percent)). We now explore the extent to which these differential exposures help explain the gender gap in well-being shown in Figures 1 and 2.

Table 4 shows the results of a linear regression to which we add different sets of controls to explore potential drivers of the gender gap. Column 1 presents the raw gap in subjective well-being changes, where we regress the dependent variable on a gender dummy only. The negative coefficient shows that women experienced a 0.14 standard deviations larger decline in mental well-being during the pandemic than men. This is comparable to Adams-Prassl et al. (2020b) who calculate the gender gap in mental health in US states with and without the lockdown but without access to a panel dimension.[16]

Columns 2, 3 and 4 include as controls the variables discussed in Tables 1, 2 and 3, respectively. Column 2 controls for factors related to time use within the household, as well as caring responsibilities. These factors narrow the initial estimate of the gender gap by 0.04 standard deviations, or roughly 25 percent. This finding is in contrast to Adams-Prassl et al. (2020b) who conclude that caring responsibilities do not play a large role.[17] Controlling only for financial and work-related factors in Column 3 does not reduce the gap significantly. This is in line with our discussion above

[16]More precisely, Adams-Prassl et al. (2020b) do not present an estimate of the gender gap without additional control variables. The gender gap in well-being in states without a lockdown is estimated to be 0.21 standard deviations, whereas the gender gap is 0.14 standard deviations larger in states with a lockdown.

[17]They measure caring responsibilities coarsely by a binary indicator 'Have to change your work patterns to care for others' and find that this variable is not correlated with well-being for either gender.

Table 4: Gender Gap in Mental Well-Being

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.144***	-0.108***	-0.140***	-0.048**	-0.131***	-0.035	-0.148***
	(0.023)	(0.025)	(0.023)	(0.023)	(0.023)	(0.024)	(0.021)
Controls:							
Family & Caring		X				X	
Financial & Work			X			X	
Social				X		X	
Medical					X	X	
Health Behaviour					X	X	
Demographic					X	X	
HH Fixed Effects							X
Observations	12252	12250	12252	12252	12252	12250	8758
Adjusted R^2	0.005	0.010	0.013	0.079	0.019	0.094	0.420

Notes: Data from UKHLS wave 9 and Covid module. Dependent variable is individual change in standardized, seasonally-adjusted and inverted GHQ Likert score. Sample in columns (1) to (6) is all individuals aged 16 and over with full interviews. Sample in column (7) comprises individuals in households with multiple interviews only. Covid survey weights used in all regressions. Standard errors clustered at the primary sampling unit and presented in parentheses. Column (1) includes a gender dummy and a constant, but no other controls. Family controls are hours of housework, hours of childcare, and a dummy for caring responsibilities. Financial controls are categorical variables for: self-reported financial situation; whether having trouble paying bills; having experienced job loss, or having been furloughed. Social controls are categorical variables for the number of close friends declared in wave 9, and self-reported loneliness. Medical controls are categorical variables for: having Covid symptoms either in the past or present; receiving external care from outside the house; being shielded or pregnant. Health behaviour controls are indicators for: drinking alcohol regularly; exercising vigorously or moderately at least 3 times a week; being able to eat nutritious meals. Demographic controls are age categories, indicators for the presence of children; being in a couple.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

that the fraction of women and men experiencing financial problems is very similar. This is also consistent with Adams-Prassl et al. (2020b) who discuss that the gender gap is not explained by realized economic impacts. Next, we explore the role of social factors. Loneliness and the number of close friends close the gap by about 67 percent, leaving only 0.05 standard deviations of the initial gap unexplained.[18] This finding accords with the discussion above, in which we documented that more women are experiencing loneliness more often, and that the correlation between loneliness and changes in well-being is large.

Lastly, we include medical factors, health behaviors and demographic characteristics in combination. As documented above, some of these factors, notably age, are strongly related to declines in well-being. However, these factors are unlikely to explain the gender gap because they are similarly distributed across men and women. In total we find these factors explain about 10 percent (Column 5). When including all controls in Column 6, we can explain the gender gap to the extent

[18] In this specification, we can only reject the non-existence of the gender gap at the 5 percent level, instead of the 1 percent level in the previous specifications. Note that this is not driven by an increase in imprecision, but a reduction of the coefficient.

Covid Economics 33, 30 June 2020: 46-72

CEPR PRESS

that the coefficient on the female dummy becomes statistically indistinguishable from zero. Overall, including all controls reduces the gender gap by around 76 percent.

In the final column we exploit a useful feature of the data, whereby all individuals in a residence are interviewed. Here we regress the dependent variable on the gender dummy and a set of household fixed effects. We do this only for households with more than one full interview.[19] Column 7 shows that the household fixed effect has a large explanatory power for well-being, indicated by the high R^2. However, the fixed effects do not affect the gender coefficient. This finding indicates that the gender gap is not explained by household effects: women are not overly represented in households that are doing worse during the pandemic.

Table 4 uses a fairly rich set of controls and shows that these can explain most of the gender gap. As a robustness check, we consider what happens to the gender gap when the set of controls is extended further, and includes relevant interactions. In Appendix table A.4, we allow for polynomials in childcare and hours of housework, both interacted with the presence of children. Additionally, we include indicators for being a keyworker, working from home in February 2020 (before the onset of the pandemic) and receiving formal care from outside the home. We also add indicators for the number of over-70s in the household, an indicator for holding a degree and interactions of these and basic demographic variables (the presence of children and being in a couple) with age. The estimates presented in Table A.4 do not differ from our main results in Table 4.

4 Conclusion

Early studies produced since the onset of the Covid-19 pandemic have found deteriorating mental well-being, particularly apparent for women. Using rich data we document a similar finding for the UK. We add to these existing studies by showing that much of the gender gap in well-being can be explained by gender differences in social factors and feeling lonely more often. We also show that gender differences in family-related time use and caring responsibilities appear to play a role.

The implications for policy are noteworthy. Our results are suggestive of the strongly adverse and unequal effect of social distancing. Our results can therefore inform debates at a time when policy makers must weigh up many competing objectives. In particular, our results suggest that lockdown is impacting mental well-being less through its effect on the labour market or wider economy, and more through the direct loss of social interaction.

[19]This reduces the same to 8758 observations. Running the baseline regression of Column 1 again, we find a similar gender dummy for this reduced sample.

Covid Economics 33, 30 June 2020: 46-72

Covid Economics 33, 30 June 2020: 46-72

References

Adams-Prassl, Abi, Teodora Boneva, Marta Golin and Christopher Rauh (2020a): 'Inequality in the Impact of the Coronavirus Shock: Evidence from Real Time Surveys'. *IZA Discussion Paper* 13183.

— (2020b): 'The Impact of the Coronavirus Lockdown on Mental Health: Evidence from the US'. *HCEO Working Paper* 2020-030.

Alon, Titan, Matthias Doepke, Jane Olmstead-Rumsey and Michele Tertilt (2020): 'The impact of Coivd-19 on gender equality'. *Covid Economics* 4, pp. 62–85.

Andrew, Alison et al. (2020): 'How are mothers and fathers balancing work and family under lockdown?' *IFS Briefing Note* BN290.

Armbruster, Stephanie and Valentin Klotzbuecher (2020): 'Lost in lockdown? Covid-19, social distancing, and mental health in Germany'. *Covid Economics* 22, pp. 117–153.

Banks, James and Xiaowei Xu (2020): 'The mental health effects of the first two months of lockdown and social distancing during the Covid-19 pandemic in the UK'. *Covid Economics* 28, pp. 91–118.

Benzeval, Micheala et al. (2020): 'The Idiosyncratic Impact of an Aggregate Shock: The Distributional Consequences of COVID-19'. *IFS Working Paper* W20/15.

Bridger, Emma and Michael Daly (2019): 'Cognitive ability as a moderator of the association between social disadvantage and psychological distress: Evidence from a population-based sample'. *Psychological Medicine* 49 (9), pp. 1545–1554.

Brodeur, Abel, Andrew Clark, Sarah Fleche and Nattavudh Powdthavee (2020): 'Assessing the impact of the coronavirus lockdown on unhappiness, loneliness, and boredom using Google Trends'. *arXiv* 2004.12129.

Brooks, Samantha K. et al. (2020): 'The psychological impact of quarantine and how to reduce it: rapid review of the evidence'. *The Lancet* 395 (10227), pp. 912–920.

Brülhart, Marius and Rafael Lalive (2020): 'Daily Suffering: Helpline Calls during the Covid-19 Crisis 1'. *Covid Economics* 19, pp. 143–158.

Cacioppo, John, Stephanie Cacioppo, John Capitanio and Steven Cole (2015): 'The Neuroendocrinology of Social Isolation'. *Annual Review of Psychology* 66 (1), pp. 733–767.

Cacioppo, John, Louise Hawkley, Greg Norman and Gary Berntson (2011): 'Social isolation'. *Annals of the New York Academy of Sciences* 1231 (1), pp. 17–22.

Chang, Shu Sen, David Stuckler, Paul Yip and David Gunnell (2013): 'Impact of 2008 global economic crisis on suicide: Time trend study in 54 countries'. *BMJ* 347, pp. 1–15.

Clark, Ben, Kiron Chatterjee, Adam Martin and Adrian Davis (2019): 'How commuting affects subjective wellbeing'. *Transportation*.

Cowan, Katherine (2020): *Survey results: Understanding people's concerns about the mental health impacts of the COVID-19 pandemic.* Tech. rep.

Dagher, Rada K., Jie Chen and Stephen B. Thomas (2015): 'Gender differences in mental health outcomes before, during, and after the great recession'. *PLoS ONE* 10 (5), pp. 1–16.

Daly, Michael, Angelina Sutin and Eric Robinson (2020): 'Longitudinal changes in mental health and the COVID-19 pandemic: Evidence from the UK Household Longitudinal Study'. *PhyArXiv*.

Davillas, Apostolos, Michaela Benzeval and Meena Kumari (2016): 'Association of adiposity and mental health functioning across the lifespan: Findings from understanding society (The UK household longitudinal study)'. *PLoS ONE* 11 (2), pp. 1–19.

Davillas, Apostolos and Andrew M. Jones (2020): ‘The Covid-19 pandemic and its impact on inequality of opportunity in psychological distress in the UK’. *ISER Working Paper* 2020-07.

Diener, Ed, Robert A Emmons, Randy J Larsen and Sharon Griffin (1985): ‘The Satisfaction With Life Scale’. *Journal of Personality Assessment* 49 (1), pp. 71–75.

Dingel, Jonathan and Brent Neiman (2020): ‘How Many Jobs Can be Done at Home?’ *NBER Working Paper* 26948.

Fetzer, Thiemo, Lukas Hensel, Johannes Hermle and Christopher Roth (2020a): ‘Coronavirus Perceptions And Economic Anxiety’. *Review of Economics and Statistics* forthcoming.

Fetzer, Thiemo et al. (2020b): ‘Perceptions of an Insufficient Government Response at the Onset of the COVID-19 Pandemic are Associated with Lower Mental Well-Being’. *PsyArXiv*.

Golden, Jeannette et al. (2009): ‘Loneliness, social support networks, mood and wellbeing in community-dwelling elderly’. *International Journal of Geriatric Psychiatry* 24, pp. 694–700.

Hale, Thomas, Anna Petherick, Toby Phillips and Samuel Webster (2020): ‘Oxford COVID-19 Government Response Tracker’. *Blavatnik School of Government Working Paper* 31.

Holmes, Emily A. et al. (2020): ‘Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science’. *The Lancet Psychiatry* 7, pp. 547–560.

Holt-Lunstad, Julianne, Timothy Smith, Mark Baker, Tyler Harris and David Stephenson (2015): ‘Loneliness and Social Isolation as Risk Factors for Mortality: A Meta-Analytic Review’. *Perspectives on Psychological Science* 10 (2), pp. 227–237.

House, James, Karl Landis and Debra Umberson (1988): ‘Social relationships and health’. *Science* 241, pp. 540–545.

Janke, Katharina, Kevin Lee, Carol Propper, Kalvinder Shields and Michael A Shields (2020): ‘Macroeconomic Conditions and Health in Britain: Aggregation, Dynamics and Local Area Heterogeneity’. *IZA Discussion Paper* 13091.

Knipe, Duleeka, Hannah Evans, Amanda Marchant, David Gunnell and Ann John (2020): ‘Mapping population mental health concerns related to COVID-19 and the consequences of physical distancing: a Google trends analysis’. *Wellcome Open Research* 5 (May).

Powdthavee, Nattavudh, Anke Plagnol, Paul Frijters and Andrew Clark (2019): ‘Who Got the Brexit Blues? The Effect of Brexit on Subjective Wellbeing in the UK’. *Economica* 86 (343), pp. 471–494.

Remes, Olivia, Carol Brayne, Rianne Van Der Linde and Louise Lafortune (2016): ‘A systematic review of reviews on the prevalence of anxiety disorders in adult populations’. *Brain and behavior* 6 (7), e00497.

Social, Institute for and Economic Research (2020): ‘Understanding Society: COVID-19 Study’, SN: 8644, 10.5255/UKDA-SN-8644–1.

Stevenson, Betsey and Justin Wolfers (2009): ‘The paradox of declining female happiness’. *American Economic Journal: Economic Policy* 1 (2), pp. 190–225.

Tubadji, Annie, Frederic Boy and Don J. Webber (2020): ‘Narrative economics, public policy and mental health’. *Covid Economics* 20, pp. 109–131.

University of Essex, Institute for Social and Economic Research and NatCen Social Research, Kantar Public (2019): ‘Understanding Society: Waves 1-9, 2009-2018’.

VanderWeele, Tyler, Louise Hawkley and John Cacioppo (2012): ‘On the reciprocal association between loneliness and subjective well-being’. *American Journal of Epidemiology* 176 (9), pp. 777–784.

Covid Economics 33, 30 June 2020: 46-72

Winkelmann, Liliana and Rainer Winkelmann (1998): 'Why are the unemployed so unhappy? Evidence from panel data'. *Economica* 65 (257), pp. 1–15.

World Health Organization (2020): *Mental Health and Psychosocial Considerations During COVID-19 Outbreak*. Tech. rep.

Zhou, Muzhi, Ekaterina Hertog, Kamila Kolpashnikova and Man-yee Kan (2020): 'Gender inequalities: Changes in income, time use and well-being before and during the UK COVID-19 lockdown'. *SocArXiv*.

Covid Economics 33, 30 June 2020: 46-72

CEPR PRESS

Covid Economics 33, 30 June 2020: 46-72

Appendix

A Additional Tables and Figures

Table A.1: Well-Being by Gender: Medical/Health Factors

	Female	Female	Female	Male	Male	Male
Symptoms: No	-0.24***			-0.11***		
	(0.02)			(0.02)		
Yes	-0.36***			-0.22***		
	(0.05)			(0.05)		
Vulnerable: No		-0.27***			-0.12***	
		(0.02)			(0.02)	
Yes		-0.12			-0.12	
		(0.09)			(0.08)	
Receiving help: No			-0.23***			-0.08***
			(0.02)			(0.02)
Yes			-0.30***			-0.21***
			(0.03)			(0.03)
Observations	7238	7238	7238	5243	5243	5242
Adjusted R^2	0.049	0.049	0.049	0.017	0.016	0.020

Notes: Data from UKHLS wave 9 and CoViD module. Table reports grouped means of outcome variable, which is the individual change in standardized, seasonally adjusted and inverted GHQ Likert score. Standard errors clustered at the primary sampling unit and presented in parentheses. CoViD survey weights used in all computations. 'Symptoms' comes from self-reported presence of symptoms since the onset of the pandemic. 'Vulnerable' takes value "yes" either if the individual has received an NHS letter requesting they should stay at home ('shielded') or the individual is pregnant. 'Help' is a self-report of whether the individual has received care from outside the house, either from family or not.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Well-Being by Gender: Health Behaviours

	Female	Female	Female	Male	Male	Male
Alcohol: No	-0.22***			-0.13***		
	(0.03)			(0.04)		
Yes	-0.28***			-0.12***		
	(0.02)			(0.02)		
Exercise: No		-0.23***			-0.16***	
		(0.03)			(0.03)	
Yes		-0.28***			-0.09***	
		(0.02)			(0.02)	
Healthy Meals: No			-0.24***			-0.12***
			(0.02)			(0.02)
Yes			-0.68***			-0.32*
			(0.16)			(0.18)
Observations	7238	7238	7236	5243	5243	5238
Adjusted R^2	0.049	0.049	0.053	0.016	0.017	0.017

Notes: Data from UKHLS wave 9 and Covid module. Table reports grouped means of outcome variable, which is the individual change in standardized, seasonally-adjusted and inverted GHQ Likert score. Standard errors clustered at the primary sampling unit and presented in parentheses. Covid survey weights used in all computations. Alcohol variable uses self-report of whether respondent has drunk any alcohol in previous 4 weeks. Exercise variable is based self-report of whether the individual has either done moderate exercise or vigorous exercise on three days in the previous week. Healthy meals variable is based on self-report of whether the individual was able to eat health food in the previous week: variable 'lacknutr'.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Covid Economics 33, 30 June 2020: 46-72

CEPR PRESS

Table A.3: Well-Being by Gender: Demographic Factors

	Female	Female	Female	Male	Male	Male
Age: 16 to 29	-0.46***			-0.31***		
	(0.06)			(0.06)		
30 to 49	-0.29***			-0.11***		
	(0.03)			(0.03)		
50 to 69	-0.17***			-0.06***		
	(0.02)			(0.02)		
Over 70	-0.20***			-0.12***		
	(0.03)			(0.03)		
Couple: Yes		-0.25***			-0.10***	
		(0.02)			(0.02)	
No		-0.27***			-0.18***	
		(0.03)			(0.04)	
Children: No			-0.23***			-0.10***
			(0.02)			(0.02)
Yes			-0.32***			-0.18***
			(0.04)			(0.04)
Observations	7238	7238	7238	5243	5243	5243
Adjusted R^2	0.055	0.048	0.049	0.023	0.018	0.017

Notes: Data from UKHLS wave 9 and Covid module. Table reports grouped means of outcome variable, which is the individual change in standardized, seasonally-adjusted and inverted GHQ Likert score. Standard errors clustered at the primary sampling unit and presented in parentheses. Covid survey weights used in all computations.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Covid Economics 33, 30 June 2020: 46-72

Table A.4: Gender Gap in Mental Well-Being: Richer Controls

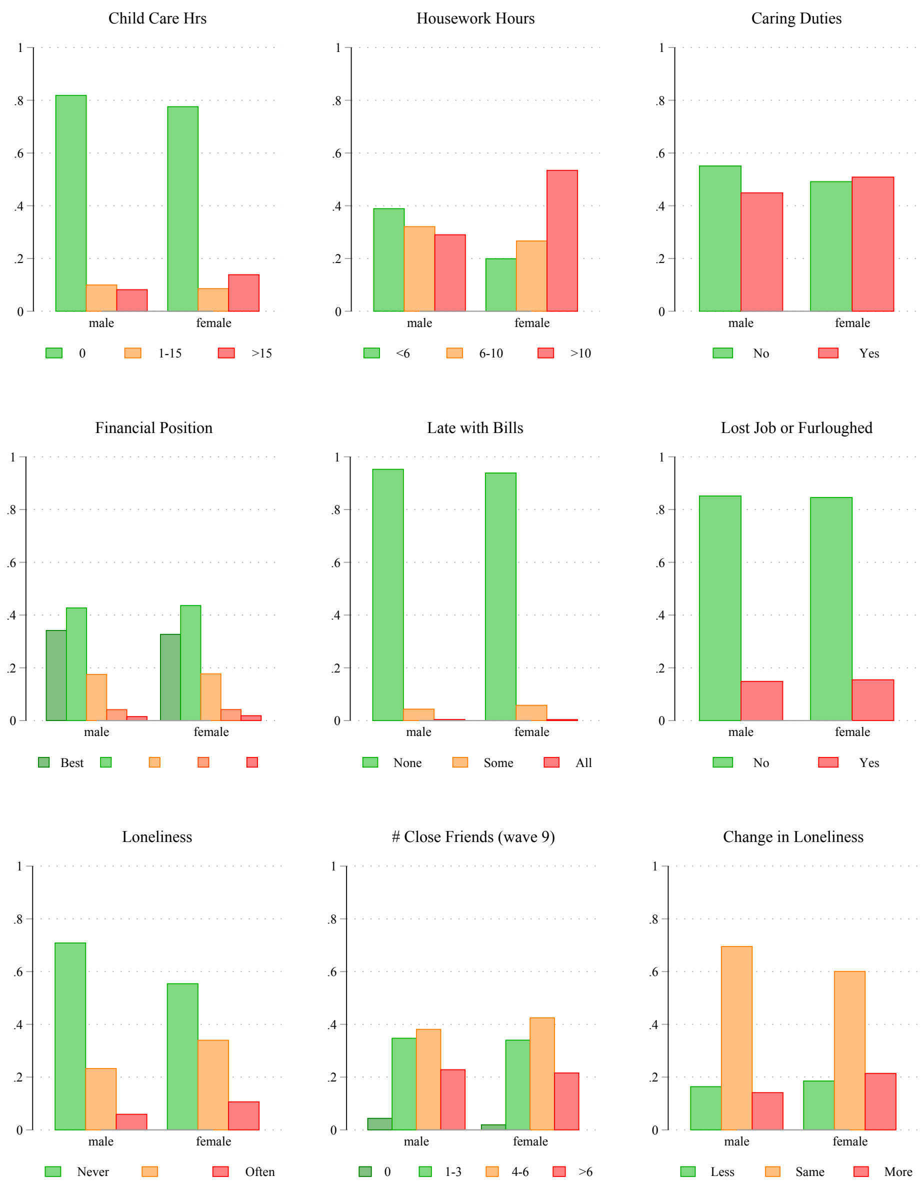
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.144***	-0.112***	-0.142***	-0.048**	-0.133***	-0.039	-0.148***
	(0.023)	(0.025)	(0.023)	(0.023)	(0.024)	(0.025)	(0.021)
Controls:							
Family & Caring		X				X	
Financial & Work			X			X	
Social				X		X	
Medical					X	X	
Health Behaviour					X	X	
Demographic					X	X	
HH Fixed Effects							X
Observations	12252	12250	12252	12252	12246	12244	8758
Adjusted R^2	0.005	0.011	0.014	0.079	0.028	0.105	0.420

Notes: Similar to Table 4, with some modifications to controls. Full list of controls as follows. Family controls are an interaction of the presence of children with: a cubic polynomial in hours of housework; cubic polynomial in hours of childcare; a dummy for caring responsibilities. Financial controls are categorical variables for: self-reported financial situation; whether having trouble paying bills; having experienced job loss, or having been furloughed; whether worked from home in February 2020. Social controls are categorical variables for the number of close friends declared in wave 9, and self-reported loneliness. Medical controls are categorical variables for: having Covid symptoms either in the past or present; receiving external care from outside the house; being shielded or pregnant; use of a formal carer; being a keyworker. Health behaviour controls are indicators for: being a smoker; drinking alcohol regularly; exercising vigorously or moderately at least 3 times a week; being able to eat nutritious meals. Demographic controls are age categories interacted with: presence of children; being in a couple; having elderly people in the home; having a degree.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Covid Economics 33, 30 June 2020: 46-72

CEPR PRESS

Figure A.1: Proportions by Type and Gender



Notes: Data from UKHLS wave 9 and Covid module, including survey weights. Variables are the same as those presented in tables 1, 2 and 3. For each variable, figure reports proportions of the sample taking each value, by gender.

Covid Economics 33, 30 June 2020: 46-72

CEPR PRESS

B Description of GHQ-12 Questionnaire

As discussed in section 2 our measure of mental well-being comes from the Likert scale derived from the 12-question GHQ questionnaire. The GHQ questions are listed below. The Likert scale is obtained by recoding so that the scale for individual variables runs from 0 to 3 instead of 1 to 4, and then summing, giving a scale running from 0 (the least distressed) to 36 (the most distressed). The questionnaire is administered to everyone.

In our analysis we standardize this variable across gender and wave to have a mean of zero and a standard deviation of one. We then multiply by −1 to obtain a scale that runs from negative (more distressed) to positive (less distressed).

Covid Economics 33, 30 June 2020: 46-72

Wording of the questions:

ghqa [GHQ: concentration]: The next questions are about how you have been feeling over the last few weeks. Have you recently been able to concentrate on whatever you're doing?
1. Better than usual 2. Same as usual 3. Less than usual 4. Much less than usual

ghqb [GHQ: loss of sleep]: Have you recently lost much sleep over worry?
1. Not at all 2. No more than usual 3. Rather more than usual 4. Much more than usual

ghqc [GHQ: playing a useful role]: Have you recently felt that you were playing a useful part in things?
1. More so than usual 2. Same as usual 3. Less so than usual 4. Much less than usual

ghqd [GHQ: capable of making decisions]: Have you recently felt capable of making decisions about things?
1. More so than usual 2. Same as usual 3. Less so than usual 4. Much less capable

ghqe [GHQ: constantly under strain]: Have you recently felt constantly under strain?
1. Not at all 2. No more than usual 3. Rather more than usual 4. Much more than usual

ghqf [GHQ: problem overcoming difficulties]: Have you recently felt you couldn't overcome your difficulties?
1. Not at all 2. No more than usual 3. Rather more than usual 4. Much more than usual

Covid Economics 33, 30 June 2020: 46-72

ghqg [GHQ: enjoy day-to-day activities]: Have you recently been able to enjoy your normal day-to-day activities?
1. More so than usual 2. Same as usual 3. Less so than usual 4. Much less than usual

ghqh [GHQ: ability to face problems]: Have you recently been able to face up to problems?
1. More so than usual 2. Same as usual 3. Less able than usual 4. Much less able

ghqi [GHQ: unhappy or depressed]: Have you recently been feeling unhappy or depressed?
1. Not at all 2. No more than usual 3. Rather more than usual 4. Much more than usual

ghqj [GHQ: losing confidence]: Have you recently been losing confidence in yourself?
1. Not at all 2. No more than usual 3. Rather more than usual 4. Much more than usual

ghqk [GHQ: believe worthless]: Have you recently been thinking of yourself as a worthless person?
1. Not at all 2. No more than usual 3. Rather more than usual 4. Much more than usual

ghql [GHQ: general happiness]: Have you recently been feeling reasonably happy, all things considered?
1. More so than usual 2. About the same as usual 3. Less so than usual 4. Much less than usual

Covid Economics **Issue 33, 30 June 2020**

The effect of the 1918 influenza pandemic on income inequality: Evidence from Italy[1]

Sergio Galletta[2] and Tommaso Giommoni[3]

Date submitted: 25 June 2020; Date accepted: 28 June 2020

Covid Economics 33, 30 June 2020: 73-109

In this paper, we estimate the effect of the 1918 influenza pandemic on income inequality in Italian municipalities. Our identification strategy exploits the exogenous diffusion of influenza across municipalities by infected soldiers on leave from World War I operations at the peak of the pandemic. Our measures of income inequality come from newly digitized historical administrative records on Italian taxpayer incomes. We show that in the short-/medium-run (i.e., after five years), income inequality is higher in Italian municipalities more afflicted by the pandemic. The effect is mostly explained by a reduction in the share of income held by poorer people. Finally, we provide initial evidence that these differences in income inequality persist even after a century.

1 We are grateful to Elliott Ash, Francesco Cutolo, Alessio Fornasin and Mario Jametti for their valuable comments and suggestions. We thank Matteo Pinna and David Cai for their helpful research assistance.

2 Assistant Professor, University of Bergamo.
3 Postdoctoral Researcher, ETH Zurich.

Copyright: Sergio Galletta and Tommaso Giommoni

CEPR PRESS

Covid Economics 33, 30 June 2020: 73-109

1. Introduction

The recent COVID-19 outbreak is expected by many to increase income inequality. While there are widespread negative economic impacts of a pandemic, the poorer are likely to be hit harder. For example, the probability of being infected during the epidemic is potentially higher for poorer than for richer individuals. Moreover, the lower-income strata of the population may not have accumulated wealth to smooth the severity of the crisis and better recover once the pandemic has passed.[1]

To date, there is very little evidence regarding whether and to what extent global pandemics can affect the income distribution (Karlsson et al., 2014; Alfani and Ammannati, 2017; Furceri et al., 2020). The limited number of empirical results is mostly due to the relative rarity of pandemics in recent history, as well as to the scant availability of economic data for less recent periods. The main objective of this paper is to address this gap by providing empirical evidence of the redistributive effect of the 1918 influenza pandemic, also known as the Spanish Flu, in Italian municipalities.

Our empirical analysis employs unique data and exploits the heterogeneous diffusion of the influenza across localities in a cross-sectional framework. The identification and estimation of the pandemic's effect on inequality are challenging due to potential reverse causality and correlated omitted variables: for example, one might expect the pandemic to have a greater impact in areas with higher inequality. In addition, there may be unobservable municipal characteristics associated with both inequality and the severity of the pandemic.

To address such identification issues, we rely on a unique natural experiment in Italy during World War I (WWI). At that time, pandemic diffusion was linked with the movement of Italian troops in the national territory. According to historical accounts, the relocation of contagious soldiers played an important role in the diffusion of the flu from the war front to Italian communities (Tognotti, 2015; Cutolo, 2020). Drawing on these accounts, we exploit the plausibly exogenous variation in the number of WWI soldiers who returned to their hometown on leaves of absence during the epidemic's peak.

To construct a proxy for soldier-related flu exposure, we use the Italian "Honor Roll of World War I Dead" ("*Albo dei Caduti Italiani della Grande Guerra*"). This compre-

[1] See, for example, some recent studies suggesting that the COVID-19 pandemic will have stronger negative effects on the more vulnerable categories of individuals (Bell et al., 2020; Adams-Prassl et al., 2020; Alon et al., 2020).

CEPR PRESS

hensive publication reports information on place of birth, place of death, day of death, and reason of death for the 540,401 Italian soldiers who died from 1915 to 1920.[2] We use this data to create a proxy for pandemic exposure. Specifically, our treatment variable is the number of soldiers returning from the front who died of illness in their hometown during the peak of the epidemic.

We validate the pandemic treatment variable in two ways. First, we provide evidence that at the regional level, the per capita number of soldiers who died of illness in their hometown during the year of the epidemic (1918) is significantly and positively correlated with the number of deaths due to influenza. This correlation does not hold for other years of the war, or for the per capita soldier deaths from the broader region outside the hometown. Thus, we can argue that we are picking up variation in the severity of the disease rather than the general local effects of WWI. Second, we show that the national monthly variation in excess mortality of civilians is highly correlated with the monthly variation of soldiers who passed away because of illness. At least during the flu peak, soldiers who died of illness are most likely victims of the Spanish Flu. Therefore, our treatment captures both geographic and time variation in the severity of the disease.

Our identification assumption is that the number of infected soldiers who returned to their hometown and eventually died of illness is exogenous, conditional on controls. We report evidence in support of this assumption. First, we show that our proxy is not correlated with pre-determined municipal characteristics such as demographic features, geography, or local policies that might be correlated with both local inequality and influenza outcomes. Second, we show using the method from Oster (2019) that omitted variable bias is likely to be limited, as the addition of pre-determined municipal characteristics and a broad set of fixed effects marginally alter the estimated coefficients while sensibly increasing the R^2.

Besides this new identification strategy for pandemic exposure, our second empirical contribution is a new historical measure of income inequality from Italian municipalities. We collected and digitized income declaration reports published by the Italian Ministry of Finance in 1924 for the main tax on income. In the early 1900s, this tax was the single largest source of government tax revenue.[3] These publications include individual income

[2]Fornasin et al. (2018) use the same kind of information for a sample of Italian regions to provide new estimate on the number of victims of Spanish flu in Italy.

[3]In the period considered in the analysis, it accounted for more than half of the revenues collected by the Italian Treasury from direct taxes.

Covid Economics 33, 30 June 2020: 73-109

Covid Economics 33, 30 June 2020: 73-109

data on the universe of taxpayers with income from commercial, industrial, or other professional activities. Using these income data, we calculate several income inequality measures for nearly 2,000 municipalities across Italy.

Our main result is that a one-standard-deviation increase in our proxy for pandemic exposure in 1918 caused a $2-3.4\%$ increase in inequality as measured by the 1924 Gini index. The effect is mostly driven by reduced income at the bottom of the distribution, rather than increased income at the top (where there is no effect). These effects appear to be persistent, as the most severely afflicted municipalities still have more unequal incomes even a century later (2018 data). Moreover, placebo checks suggest that this effect only depends on the intensity of the disease as we find no effect when using variables measuring the local severity of WWI.

These results contribute to a number of literatures. First, we add to the research on the effect of pandemics on inequality. Karlsson et al. (2014) study the impact of the 1918 influenza pandemic on economic outcomes in Sweden. In that setting, using a difference-in-difference strategy on Swedish counties, they show that the epidemic increased the share of poor people, as measured by the fraction of inhabitants living in public poorhouses. Alfani (2015) and Alfani and Ammannati (2017) focus on the effect of the fourteenth-century Black Death on income inequality in a set of Italian regions. They show that the plague reduced inequality over-time. Furceri et al. (2020) apply a panel cross-country analysis of the few epidemics that occurred since 2000 and find an increase in inequality in those countries that were more severely affected. Finally, Galasso (2020) uses survey evidence to show that the COVID-19 epidemic worsens the labor market outcomes of low-income individuals immediately after the introduction of lockdown measures.

More generally, our paper complements the rapidly growing body of research aiming to understand the economic and social consequences of pandemics. Carillo and Jappelli (2020), is the closest study to ours also focusing on the economic consequences of the 1918 influenza in Italy. Their analysis shows that regions with the highest mortality rates reported a significant decrease in GDP compared to less affected regions. They also show that the effect did not last long, disappearing four years after the pandemic. The presence of a negative effect on GDP is also emphasized by other studies (Correia et al., 2020; Karlsson et al., 2014; Dahl et al., 2020). In particular, Correia et al. (2020) exploit variation in the mortality rate due to the the 1918 influenza across major U.S. cities to analyze output. Aside from the negative effect of the influenza on GDP, they found that non-pharmaceutical interventions did not harm the economy.

Finally, we contribute to the literature on the long-term consequences of pandemics. There is already evidence hinting at the long term effect of the 1918 pandemic on individuals' outcomes such as health and economic status as well as attitudes and trust (Almond, 2006; Guimbeau et al., 2020; Beach et al., 2018; Lin and Liu, 2014; Percoco, 2016; Aassve et al., 2020). Yet, the only papers linking pandemics and inequality in the long-run are Alfani (2015) and Alfani and Ammannati (2017), who study the effect of the Black Death and found that inequality decreased in the following centuries.

To the best of our knowledge, our paper is one of the first to study the causal effects of pandemics on income inequality focusing on local jurisdictions. Moreover, we go beyond the analysis of GDP by showing how the 1918 Influenza Pandemic affected the income distribution. Finally, we provide novel evidence on the long-term effects of the Spanish Flu, showing that the rise in inequality has persisted over time, for over a century.

The remainder of the paper is structured as follow. Section 2 discusses the history of the 1918 influenza in Italy. Section 3 discusses the data used for the analysis. Section 4 describes the identification strategy. Section 5 presents the results. Section 6 concludes.

2. Historical background

The 1918 flu is considered one of the deadliest pandemics experienced in modern history. Estimates suggest that 500 million people were infected worldwide and that between 20 million and 50 million died as a consequence of the disease (Johnson and Mueller, 2002).

Italy reported one of the highest mortality rates in Europe, with 600,000 people falling victim of the flu.[4] Three different waves occurred between spring 1918 and early 1919. The first and third waves were of moderate intensity, while most casualties were a result of the second wave, in fall 1918, particularly in October and November 1918; Figure 1 reports the total number of deaths from influenza over time, and it clearly shows the outbreak of 1918.

The severity of the epidemic was a result of several factors. First, the standard of living was very low for most of the population. Hygienic conditions were generally

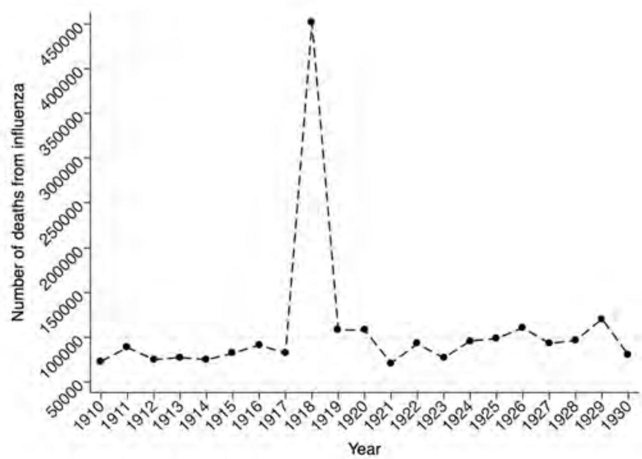
[4] Tognotti (2015) reports an accurate description of the Italian experience with the influenza epidemic and its interplay with WWI. What we described in this section is based on this source, unless noted otherwise. Cutolo (2020) and Alfani and Melegaro (2010) provide additional evidence on the Spanish Flu in Italy.

Covid Economics 33, 30 June 2020: 73-109

CEPR PRESS

Covid Economics 33, 30 June 2020: 73-109

Figure 1: Number of deaths from Influenza in Italy by Year



Notes: The plot shows the numbers deaths from Influenza over time, according to the publication *Cause di Morte (1887-1955)* from Istituto Centrale di Statistica (1958). The count includes the number of victims for flu and pneumonia.

inadequate to help oppose the spread of the virus. For example, just a quarter of the population had access to running water. Similarly, not everyone lived in houses with private toilets or access to sewer lines.

Second, public authorities produced ineffective non-pharmaceutical interventions. They were implemented too late and with a lack of coordination between the responsible offices and were generally not enforced. Some argue that the initial lack of involvement was part of a strategy that aimed not to further demoralize citizens and soldiers in a crucial moment for the resolution of WWI. This is also revealed in the few mentions of the influenza by elected representatives in the parliament during the worst months of the crisis. By inspecting parliamentary speeches, we find only one request to introduce social-distancing measures in late November.[5]

Finally, it is generally recognized that WWI played a major role the proliferation of the virus (Crosby, 1989; Winter, 2010; Tognotti, 2015). This is true when considering the spread of the disease to a global scale as well as for the Italian case. It has been suggested that the disease was brought to Europe during the spring of 1918 by the U.S. army when 200,000 American soldiers crossed the Atlantic to join the Allies in the

[5]Francesco Rota, 21st of November 1918 - https://storia.camera.it/regno/lavori/PDF/RI_LEG24/unica/17306.pdf

battlefields. Moreover, it has been documented that the end of the war and the following demobilization in the autumn of 1918 caused the situation to deteriorate further, helping the disease to spread (Oxford et al., 2005; Taubenberger and Morens, 2006; Herring and Sattenspiel, 2011).[6] Similarly, several sources of historical evidence report the presence of the disease among Italian soldiers in military camps and trenches and document the central role played by the troopers in spreading the flu in the internal front. Italy had an active internal front in the national territory, which made it impossible to avoid interactions between civilians and soldiers. Indeed, the frequent relocation of troops, as well as the presence of soldiers on leave, have been considered to be important actors in the propagation of the disease across the Italian regions (Tognotti, 2015).[7] In this setting, the virus was able to spread prior to the cessation of warfare between Italy and Austria-Hungary, on November 4th 1918.[8]

For instance, considering the second wave, the first outbreak among soldiers had been reported as early as in mid-August in a military camp nearby Parma. The health inspector of the camp indicated that this outbreak was associated with soldiers on leave returning to Italy from northern countries. A local newspaper reported that in the week from August 19 to 25, 77 people died because of the flu, 37 of whom were soldiers. Inspecting the data we gathered about military casualties, we believe these to be conservative numbers. We found that in August 1918, in Parma, 95 soldiers died because of a disease, 90 of them between August 16 and 31.[9]

Such episodes warned the military administration of the health risks and persuaded it to intervene to limit the spread of the flu among soldiers, unfortunately without much success. Cutolo (2020) reports that general measures of controls decided by the

[6]Following Oxford et al. (2005) "Demobilisation in the autumn of 1918 would have provided ideal opportunities for further intimate person-to-person spread and wide dispersion as soldiers returned home by sea and rail to countries around the globe. Family parties organised to celebrate a soldier's homecoming would have further exacerbated the situation."

[7]Soldiers could ask for a period of leave for a variety of reasons. In addition to the ordinary leave to which each soldier was entitled to receive after a certain period at the front, it was also possible, for example, to ask for sick leave as well as for work-related leave (e.g., to support the agricultural sector when needed). The length of the leave was typically between ten and fifteen days. The management of leaves was modified during the years of the war. These were allowed very rarely when the Italian army was under the command of General Luigi Cadorna. When General Armando Diaz became Commander in Chief, at the end of 1917, the policy changed in order to grant soldiers to get their rightful time of leave. The scope of the policy was to increase the morale of troops after the Caporetto disaster.

[8]Furthermore, the actual demobilization of the Royal Italian Army took place mostly in 1919.

[9]Around two-thirds belonged to the Brigata Sicilia – 61st e 62nd reggimento fanteria.

Covid Economics 33, 30 June 2020: 73-109

ministry of war had been implemented starting in early September, requiring strict medical controls of soldiers returning to the front after a period of leave. Initially, some military units attempted to stop soldiers from going on leave, but the complaints raised by soldiers and the lack of coordination among the various commands halted any nationwide limitation on leave. Indeed, soldiers were able to request sick leave to recover from the flu.[10] Civilians, meanwhile, were also worried about the possible spread of the virus from soldiers.[11]

To provide an insightful anecdote about the situation, we include an extract of a letter of the American philanthropist, resident in the Italian municipality of Bagni di Lucca, Evangeline Whipple sent on October 28, 1918, to the Rector of New York's Grace Church:

> *There is a mountain village of this Commune, about ten miles up the valley, from which point one must climb on foot for about an hour and a half before reaching.... Most of the men, all of fighting age, are at the Front. But they brought the Spanish fever in its most virulent form to this remote place, on their ten days' leave of absence. In an isolated place like this sky village, with no water except rain water, of course the contagious sickness has full sway. The priest, the only one in authority, felt the sickness coming on, and ran away....*

Overall, anecdotal and historical evidence points to a causal link between infected soldiers and civilians. This evidence becomes of primary relevance in motivating the estimation strategy detailed in Section 4.

3. Data

This paper relies on a variety of sources of information. First, we have data on the victims of influenza and casualties in World War I, and second, we have income data at the individual level from the income declaration of 1924. Finally, we complement these data with additional variables from a variety of sources. Descriptive statistics are reported in Appendix Table A.1.

[10] See, for example, Arturo Radici Valenti's description of his experience with the flu (Capodarca, 1991).

[11] Again, Cutolo (2020) reports an article from the newspaper *Il Tempo*, describing a popular unrest in a city in the south of Italy in which two soldiers, one in leave, were killed because they where considered to be virus spreaders. The article was supposed to be published at the end of October 1918 but was banned.

Covid Economics 33, 30 June 2020: 73-109

CEPR PRESS

Covid Economics 33, 30 June 2020: 73-109

3.1. Influenza and general mortality

We gather yearly information about *casualties from influenza* in the period 1915 to 1919 at the regional level from the publication *"Cause di morte:* 1887 − 1955*"* issued by the national statistical office. Following the literature, we count deaths related to influenza and pneumonia together. The lowest level of geographic aggregation of such data is the region. The average number of casualties for the period 1915-1920 is 9,617 (Table A.1, Panel A). As Figure 1 shows there is a large jump in 1918, corresponding of the peak of pandemic.

Next, we gathered national-level data with the *monthly excess of mortality* in the period 1915 − 1920 from Mortara (1925). This variable reports the ratio between the number of deaths in each month and the average number of deaths in the same month for the triennium 1911 − 1913.

3.2. Income declaration data

The main dependent variable of the analysis consists of a measure of inequality at the municipality level. To generate such an indicator, we rely on a unique historical source of information on Italian incomes. Specifically, we collected and digitized a series of publications listing individual income declarations for 24 Italian provinces composed of around 2,000 municipalities in 1924.[12] These tabulations were issued by the Italian tax authority in application of a law enacted in 1922 (*Decreto Reale 16 dicembre 1922 n. 1631*) and contain information about taxable income that was borne only by taxpayers who declared revenues from commercial and industrial activities, category B, or the profession of liberal arts, category C.[13] Categories A and D include, respectively, incomes deriving from capital and from salaries and pensions.

To provide information about how representative our sample of taxpayers is, in Table 1 we report the number of taxpayers for each income category for the year 1924: It is clear that the groups under consideration, categories B and C, represent the majority of the Italian taxpayers (67.3%), while the other two groups represent a smaller fraction.

[12]It is worth noting that the selection of such provinces is not due to any specific reason. These are the publications that we have been able to collect and digitize thus far as part of a larger data collection effort with the final goal of digitizing this publication for all Italian provinces. For our analysis, is important to note that these provinces are spread across Italian regions.

[13]The majority of taxpayers in the dataset were owners of small businesses, mostly individual companies, or small stores and artisans.

Similarly, if we focus on the aggregate declared income, the categories under consideration account for more than 9 billions Liras, equal to the 75.6% of all declared income in that year. Our final dataset includes 221,139 taxpayers from a total of 1,219,192 (categories B and C), therefore accounting for nearly 18% of the national sample.[14]

Table 1: Tax payers category, 1924

Category	N. Tax payers	Sum income decleared	Average Income
A	574,025	1,296,994,396	2,259.47
B	**1,016,139**	**6,086,711,600**	**5,990.04**
C	**203,053**	**3,233,377,995**	**15,923.81**
D	26,075	1,717,108,803	65,852.69
Total	1,812,213	12,334,192,794	6,806.15

Notes: This table reports information about the different categories of tax-payers that were subject to the income tax (Italian Ministry of Finance).

For the purpose of our analysis, we generate a series of indicators in order to measure the income inequality of Italian municipalities in 1924. First, we compute the municipal *Gini index*: this variable has an average value of 0.45 (Table A.1, Panel B). Figure 2, Plot A, shows the distribution of this variable: This indicator peaks in correspondence to the average value, with a second, smaller, peak between 0.7 and 0.8. Moreover, Figure 3, Plot A, shows the distribution of this indicator over the Italian territory. Two remarks can be made: First, as already mentioned, the income data at our disposal cover provinces spread across regions. Second, municipalities in the sample show large variations in the level of income inequality, even within the same province.

Next, we compute the fraction of income owned by the top 20% and the bottom 20% taxpayers, with means of, respectively, 0.53 and 0.05. The distributions of these variables are shown in Figure 2, Plots B and C: The top 20% shows a bell distribution, while the bottom 20% has a long right tail. Finally, we generate a series of additional inequality indicators such as the income standard deviation and the fraction of income owned by the top 10% and the bottom 10% of the population.

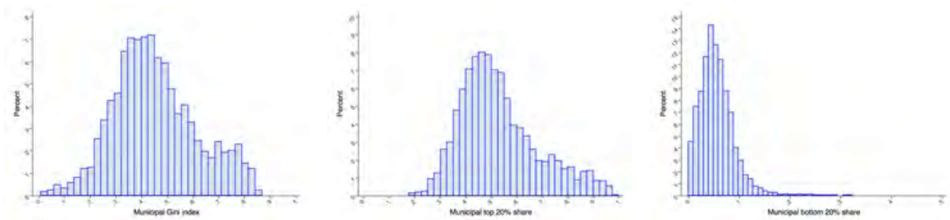
[14]We are aware that our data might be affected by taxpayers' false declaration. This should not be an issue as far as there is not a systematic correlation among a municipality's tendency to evade or elude taxes, pre-existing levels of inequality and our treatment variable. As all individuals within a district are subject to the same tax-enforcement office they face the same probability of audit and detection, the inclusion of district fixed effects should therefore account for most of potential concerns.

Covid Economics 33, 30 June 2020: 73-109

Covid Economics 33, 30 June 2020: 73-109

Figure 2: Distributions of Inequality Indices by Municipality (1924)

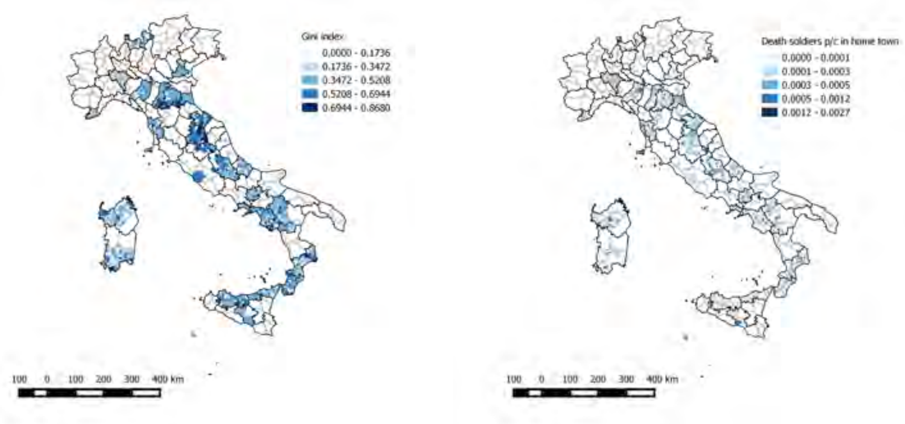
(A) Gini index (B) Top 20% income share (C) Bottom 20% income share



Notes: The figure shows the distribution of the inequality measure: municipal Gini index (Plot A), fraction of income owned by the top 20% taxpayers (Plot B) and the fraction of income owned by the bottom 20% taxpayers (Plot C).

Figure 3: Distribution of Gini index and war-related casualties by illness (hometown)

(A) Gini index (B) Death by Illness (hometown)



Notes: Panel A shows the geographic distribution of the Gini index across Italian municipalities. These data are only available for 2,038 municipalities. Instead, Panel B shows the number of soldiers who died by illness in their hometown (per-capita).

CEPR PRESS

Covid Economics 33, 30 June 2020: 73-109

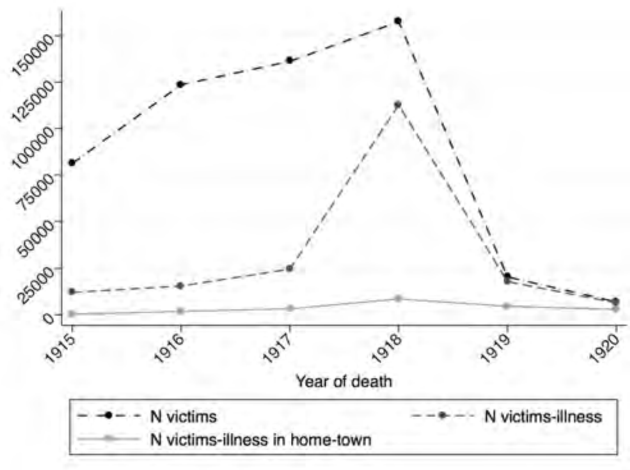
3.3. World War I casualties

To construct the explanatory variable of the analysis, we collected detailed information on the universe of Italian military victims in WWI. These data come from the *"Albo dei Caduti Italiani della Grande Guerra"*, a publication containing details on all the Italian military victims that have been identified: The dataset contains information on 540,401 victims from 1915 to 1920. The number of Italian military victims in WWI has been estimated at around 651,000 (Mortara, 1925), suggesting that the coverage of the Albo dei Caduti is rather high.

As reported by General Fulvio Zugaro, head of the Royal Army's Statistics Office and chief scientific advisor of the operation, the publication contains the information on soldiers who: a) died in combat or due to war injuries, b) went missing due to war-related causes, c) died or went missing in captivity (except deserters), d) died to illness related to war service, e) died of an accidental cause related to war service, and f) died for suicide whose cause was related to war service (Zugaro, 1926).

The dataset contains information on the demographics of the deceased soldiers, such as the name, the date of birth, and the city of origin. Moreover, information on the circumstances of the death such as the date and the place of death is also available. Finally, it also reports the cause of the decease. For instance, we know whether a soldier died while fighting, due to injuries sustained during the military combat, or due to illness. We use this information to create variables that proxy the severity of the Spanish Flu at the municipality level, in light of the historical evidence suggesting that an important cause of the diffusion of the disease was the movement of soldiers. First, we create a municipal variable counting the *number of soldiers who died* and the *number of soldiers who died because of some illness* – that is, point d) of the previous list. Second, we create a variable identifying the *number of soldiers who died of illness in their hometown*, which measures those victims of illness who came back and passed away in their hometown: This is the subset of victims whose place of death coincides with their city of origin. As argued in the historical background section, these soldiers were mostly on leave. To refine this variable as a proxy for the severity of the influenza, we focus only on those casualties that occurred during the months of the peak of the Spanish Flu, that is, August-December 1918. This approach makes it very likely that the death cause is the Spanish Flu and not something else. This latter variable is the main regressor in our central analysis and measures the extent to which troopers bring the Spanish Flu back to their hometown.[15]

Figure 4: Italian victims of WWI - over time



Notes: The plots shows the numbers of Italian victims in WWI according to the *Albo dei Caduti Italiani della Grande Guerra* over time.

Figure 4 shows the number of Italian casualties in WWI over time. The total number of victims (black dashed-dotted line) shows a positive trend in the years 1915 – 1918 and a peak in 1918, with around 160,000 casualties, before decreasing drastically in 1919 and 1920.[16] The number of victims due to illness (grey, shaded line) shows a different pattern, as this is low until 1917 thereafter sharply increasing in 1918 by around 350% compared to the previous year, with around 113,000 victims. It is important to note that in 1918 more than two-thirds of the war deaths were due to illness and, as already discussed, this peak is mostly due to the diffusion of the Spanish Flu in the trenches. Finally, the number of those who died of illness in their hometowns (light grey, solid line) in 1918 is around 8,500. This measure had a substantial increase from 1917 to 1918 of around 160%. During the months of the peak of the epidemic (i.e., August–December

[15]We cannot categorically exclude the possibility that, for at least some soldiers, the direction of contagion runs opposite, meaning they were infected while at home. However, given the rich historical evidence, we believe this represents a small fraction of our sample of victims.

[16]Italy entered the war on May 24 1915 and the conflict ended on November 11 1918. Nevertheless, a limited number of casualties in the dataset, only 5%, is attributed to the years 1919 and 1920: These are soldiers who passed away after the end of the war for war-related reasons, 88% of whom due to illness.

Covid Economics 33, 30 June 2020: 73-109

Covid Economics 33, 30 June 2020: 73-109

1918) we count a total of 3,113 soldiers who died of illness in their hometowns. For the sample of cities included in our analysis (24 provinces in which fiscal data are available), it is equal to 1,504. In the analysis we use these variables in per capita terms. Figure 3, Panel B, displays the spatial distribution of this variable, in per capita terms, for the sample of municipalities used in the analysis. Interestingly, we find relevant variation both within and between provinces across Italian regions.

3.4. Other data

To conduct our analysis, we collected a set of additional municipality characteristics. First, we track *municipal population* size making use of the decennial census. In particular, we recover these data from the censuses conducted in 1901 and in 1911. From the censuses, we also recover important information on Italian municipalities such as the *area in square kilometer*, the administrative importance (i.e., whether the *city is capital of a province*), and the *geographic features* (i.e., whether the city is in a mountain region, on the hill, or on a plain, as well as whether it is on the coast or not). We then compute municipality *population density* as the ratio between city population (in 1911) and city area. The average population of Italian cities in 1911 is 4,308 inhabitants, with an average growth rate of 5.8% between 1901 and 1911. Finally, the average population density is 1.64 inhabitants per square kilometer (Table A.1, Panel D).

Second, we collected relevant local fiscal policy measures from the 1912 municipal budgets by digitizing the volume "*Bilanci Comunali per l'anno 1912*", published by the Italian Statistical Office. Specifically, we collected information about *total expenditure*, *total surplus*, and various categories of spending: *police and sanitation/hygiene services*, *justice and security*, and *education*. These are important characteristics to help to understand the potential differences among municipalities characterized by different inflows of ill soldiers, our main explanatory variable. Total expenditures were around 23.9 Liras per inhabitant. On average, spending in police and sanitation/hygiene services represented 20.5% of total expenditures, spending in justice and security accounts for 0.9% and education spending represented 17.8% (Table A.1, Panel D).

Finally, we also measured the value of the *Gini index in 2018*, exploiting the municipal data on income declarations from the Ministry of Economy and Finance.[17] This indicator has an average value of 0.40 in the sample under analysis and of 0.39 in the entire sample

[17] See Giommoni (2019) for the use of this dataset to construct other types of income inequality indicators.

of Italian municipalities (Table A.1, Panel B). This dataset does not contain individual incomes but only information about grouped income, we therefore rely on this group-specific data to construct the Gini index.[18]

4. Empirical approach

4.1. Estimation strategy

Our primary aim is to identify the short-/medium-run effect of the severity of the 1918 influenza on income inequality in Italian local jurisdictions. In an ideal setting in which the geographic spread of influenza was orthogonal to ex ante local inequality, we could simply estimate the following equation via OLS:

$$Inequality_{idp} = \beta \; Influenza \; severity_{idp} + \gamma \mathbf{X}_{idp} + \eta_{idp} \tag{1}$$

where $Inequality_{idp}$ is a measure of income inequality in a municipality i belonging to a district d and a province p, while $Influenza\ severity_{idp}$ denotes the harshness of the influenza at the municipality level. To improve the precision of estimate, we can also add a set of fixed effects and pre-determined municipal characteristics as controls $\mathbf{X}_{idp}$, while η_{ip} denotes the error term. β would be the coefficient of interest, identifying the nature of the relationship between inequality and influenza severity.

Unfortunately, we face data limitations when assessing the actual level of the influenza severity. In fact, as already reported in Section 3, while our data collection allows us to compute the level of income inequality at the local level, the only available information about deaths related to the influenza pandemic is reported at a higher level of aggregation, that is, the regional level. Moreover, even assuming that we could avail ourselves of influenza casualties at the municipal level, the use of an OLS estimator without accounting for potential endogeneity concerns will most likely produce biased coefficients. Indeed, there are a number of reasons one can expect this variable to be, either directly or indirectly, related to pre-existing local economic conditions. For instance, in areas with a high share of poor individuals, it was common for more families to share the same home unit, therefore limiting the possibility of social-distancing. Similarly, in poorer

[18] This is a common practice when individual-level data are missing. See, for example, Sala-i-Martin (2006) and Davies et al. (2011).

Covid Economics 33, 30 June 2020: 73-109

Covid Economics 33, 30 June 2020: 73-109

localities the absence of minimal hygienic conditions might have worsened the effect of the disease compared to richer areas.

To overcome these issues, we decide to focus on a different variable that would work as a proxy for the severity of the influenza, and that we argue is less subject to endogeneity concerns. We take advantage of the well-documented role played by soldiers from WWI in the diffusion of the 1918 Influenza, as we discussed in Section 2. Specifically, we link inequality with the municipal exposure to infected soldiers returning home from the war front.

Therefore, we base our main results on the following equation:

$$Inequality_{idp} = \beta \, Victims\ WWI\ for\ illness\ hometown_{idp} + \gamma \mathbf{X}_{idp} + \eta_{idp} \qquad (2)$$

where all terms are as previously defined, while the main regressor *Victims WWI for illness hometown* is equal to the number of soldiers who returned to their hometowns from the war frontline and died due to a disease complication during the peak of the epidemic (August–December 1918). This variable is expressed in per-capita terms (with 1911 population). Again, β is our coefficient of interest. $\mathbf{X}_{idp}$ includes a set of fixed effects and controls in order to improve the precision of the estimates. In particular, the full specification includes district fixed effects, quartiles of taxpayers fixed effects (capturing the number of taxpayers) and geography fixed effects. Moreover, the set of municipal-specific controls includes a province capital dummy, population density (ratio of 1911 population to city area), and a set of budget variables, in per capita terms, from the 1912 municipal budget: total expenditures, budget surplus, spending on police and sanitation/hygiene services, spending on justice and security, spending on education. Finally, standard errors η_{ip} are robust and clustered at the district level.

Overall, our strategy hinges on two assumptions. First, our proxy is indeed correlated with the number of deaths due to influenza. Second, the proportion of ill soldiers returning to their hometown is plausibly exogenous, conditional on controls.[19]

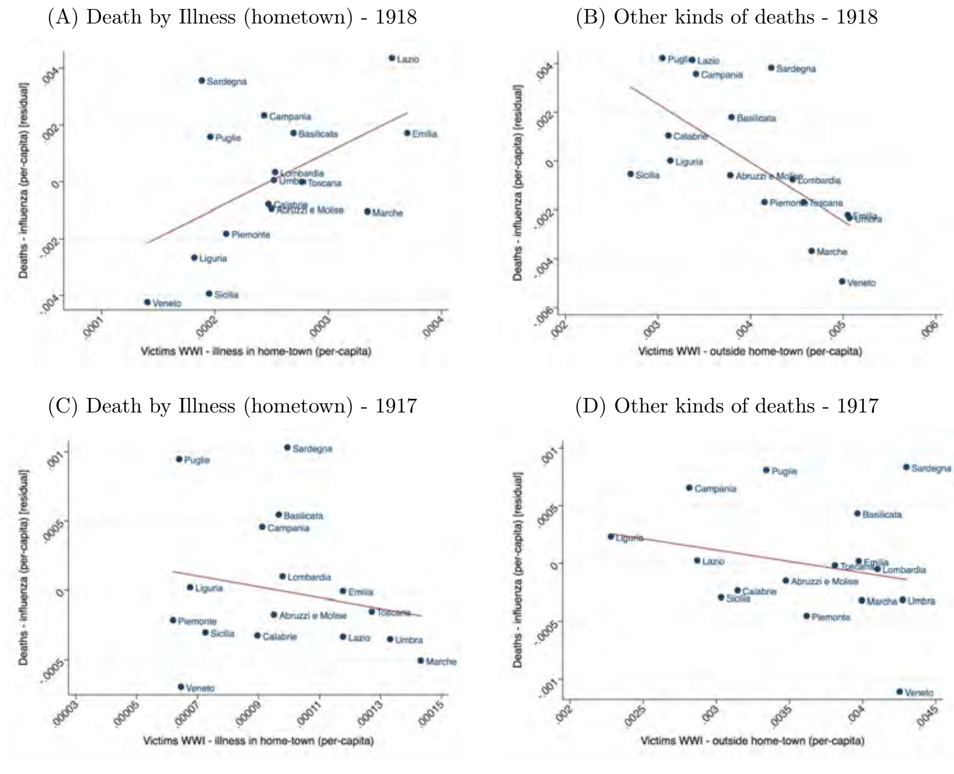
[19] To put it differently, one can consider the estimates from Equation 2 to be similar to a *reduced form* effect of an instrumental variables strategy, where the number of soldiers returning in their hometown who died due to a disease complication is the instrument for the endogenous variable, the count of influenza casualties in a municipality.

Covid Economics 33, 30 June 2020: 73-109

4.2. Assessing the quality of the proxy

In this section, we test whether our main explanatory variable is a good measure to capture the severity of the epidemics at the municipal level. In particular, we want to show that the number of soldiers who died in their hometown during the pandemic as a result of some disease is positively correlated with the number of deaths due to influenza, both variables are expressed in per-capita terms. To this end, we exploit geographic and yearly variation in the regional-level data and monthly variation in the national level data.

Figure 5: WWI victims and influenza



Notes: The Panel A shows the correlation between the number of soldiers who died of illness in their home town and the number of victims for influenza for Italian regions in 1918. The Panel B shows the correlation between the number of military victims outside their hometown and the number of victims for influenza for Italian regions in 1918. The Panel C shows the correlation between the number of soldiers who died of illness in their home town and the number of victims for influenza for Italian regions in 1917. The Panel D shows the correlation between the number of military victims outside their hometown and the number of victims for influenza for Italian regions in 1917. All the variables are expressed in per-capita terms, with 1911 population as the benchmark. The number of victims for influenza has been residualized.

First, we focus on regional data, as only at this level of aggregation we have real

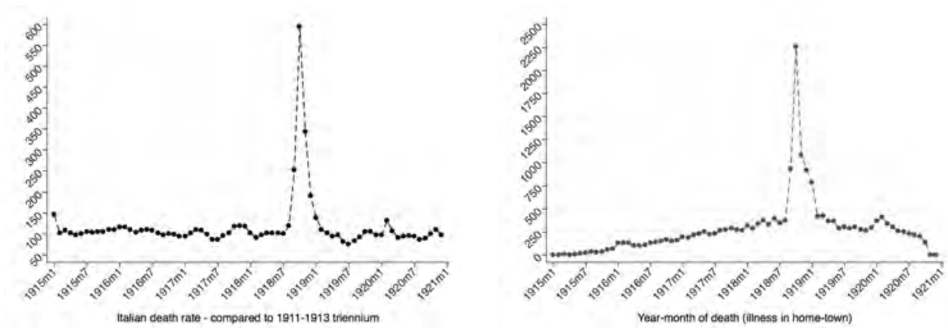
Covid Economics 33, 30 June 2020: 73-109

numbers for the casualties of influenza. Figure 5 and Table A.2 present the results graphically and numerically, respectively. Panel A of Figure 5 shows a positive relationship between the number of soldiers who died of a disease in 1918 -in their hometown- and the total number of deaths due to influenza. Therefore, we confirm the expected relationship between the two variables of interest. One potential concern is that our proxy is informative not just of the link between casualties of infected soldiers that came back and the influenza severity in a specific locality, but more generally to the connection between the number of WWI victims and the influenza severity. If that were the case, our proxy would be likely to also detect variation across localities in the severity of WWI. To account for this aspect, we display in Figure 5, Panel B, the relationship between the number of deaths due to influenza and the total number of military victims outside their hometown. Reassuringly, we see that the sign of the relationship is the opposite of that of our interest: It seems that regions with more victims are less affected by the pandemic. An additional possible issue is that the evidence shown so far could be mostly explained by pre-existing relationships to the variable of interest and therefore unrelated to the 1918 influenza outbreak. This seems unlikely as, in Panels C and D of Figure 5 and in Column 2 of Table A.2, we show that there is not a significant relationship between the number of WWI victims (both by illness in hometown and total outside of hometown) and influenza casualties in 1917.

Figure 6: Death rate vs Italian victims of WWI - monthly data

(A) Death rate Italian population

(B) Victims WWI-illness (hometown)



Notes: The left plot shows the Italian death rate over time, according to Mortara (1925). The average in the triennium 1911-1913 is the benchmark, with value of 100. The right plot shows the number of victims per illness who died in their hometown from "Albo caduti". **Correlation= 0,874**.

Table A.2 reports the actual coefficient estimates of the scatter plots: Columns 1 and

2 show the analysis conducted for 1918 and 1917, while Columns 3 and 4 confirm that our evidence also holds when considering post-pandemic ($1918-1920$) and pre-pandemic ($1915-1917$) years.[20]

Next, we exploit time variation. Figure 6 displays the Italian excess of mortality from the beginning of 1915 to the end of 1920 compared to some month for the period $1911-1913$ (Panel A), and the monthly number of deaths of ill soldiers in their home town (Panel B). The two time-series appear to follow a very similar trend. It is comforting to see that both peak in October 1918 and that overall the two variables have a high level of correlation $0,874$).[21]

Overall, our evidence tends to support the quality of our proxy variable.

4.3. Exogeneity of the proxy variable

In this section, we discuss the exogeneity of our proxy variable and provide initial supporting evidence.

Our causal identification rests on the idea that the probability for a soldier to return back to his municipality of origin and eventually die there because of a disease, during the peak of the pandemic, is orthogonal with respect to the local level of income inequality. This condition may be violated, for instance, when soldiers from a more/less unequal municipality are more/less likely to get infected by the virus while at the front; if there are local conditions that correlate with local inequality, which might favor/disfavor the return of soldiers; or if municipal policies of more/less unequal municipalities affect the survival rate of returning sick soldiers.

About the first possibility, it is worth mentioning that the spread of the virus among soldiers was mainly occured within the same battalion. However, soldiers belonging to a given battalion were from municipalities located in different areas of Italy, which implies that clusters of virus diffusion for soldiers from the same municipality while at the front were unlikely. Therefore we expected the individual probability of infection to be uncorrelated with the features of the city of origin. For instance, considering the 1918 August outbreak in Parma, the 90 soldiers who died were from the same unit, but were from 83 different municipalities.

To account for the second and the third point, we do two things: First, we rely on a

[20]In Table A.2 we use standardized regressor variables.

[21]This also suggests that the majority of soldiers who died of illness during the peak died largely due to the Influenza.

Covid Economics 33, 30 June 2020: 73-109

CEPR PRESS

Covid Economics 33, 30 June 2020: 73-109

set of balance checks to determine whether cities with different inflows of infected soldiers showed systematic differences in initial characteristics. In particular, we perform a set of regressions to check whether the proxy is correlated with pre-1918 or time-invariant municipal features. Second, we are comforted by the fact that the inclusion of additional covariates and fixed effects, marginally affect the size of the effect of the treatment variable on inequality, hinting at the presence of a limited omitted variable bias. (more details in Section 5).

Table 2: Correlation with municipal characteristics

	Population growth rate (1901-1911) (1)	Population density 1911 (2)	Geographical zone (3)	N taxpayers quartile (4)	Province capital (5)
Panel A. Municipal demographic and geographical features					
Victims WWI - illness (hometown)	-0.002	-0.030	-0.005	0.006	0.003
	(0.002)	(0.029)	(0.026)	(0.026)	(0.003)
N	1795	1810	1821	1821	1821
R^2	0.290	0.262	0.568	0.314	0.098
Geography FE	Yes	Yes	No	Yes	Yes
N taxpayers quartile FE	Yes	Yes	Yes	No	Yes
District FE	Yes	Yes	Yes	Yes	Yes
	Total expenditures (1)	Budget surplus (2)	Police, Hygiene expenditure (3)	Justice expenditure (4)	Education expenditure (5)
Panel B. Municipal budget (1912)					
Victims WWI illness (hometown)	-1.134	-0.030	-0.061	-0.033	-0.265
	(0.733)	(0.035)	(0.193)	(0.020)	(0.227)
N	1807	1807	1807	1807	1807
R^2	0.210	0.106	0.102	0.125	0.144
Geography FE	Yes	Yes	Yes	Yes	Yes
N taxpayers quartile FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is reported in the column head. *Victims WWI - illness (hometown)* captures the number of soldiers who died of illness in their home town between August and December 1918, in per-capita terms (with population of 1911) and standardized. Robust standard errors are clustered at the district level are in parentheses: * p < 0.1, ** p < 0.05 and *** p < 0.01.

Table 2 reports the results of the balance checks. The main regressor is the number of ill soldiers who died in their home town during the peak of the pandemic in per capita terms, while the dependent variable is a different covariate each time. Panel A focuses on demographic and geographic features while Panel B accounts for local policies. All regressions include district fixed effects, and when possible geographic and number of contributors (quartile) fixed effects to be as close as possible to the most saturated specification that we use in the main analysis. It is reassuring to see that none of the estimated coefficients are statistically different from 0 at the conventional level. The first set of regressions suggests that our treatment is not correlated with *population growth*, *population density*, the *geographic characteristics* of a municipality or the *number of contributors*. In addition, the administrative status of a city does not matter (*Province*

CEPR PRESS

Covid Economics 33, 30 June 2020: 73-109

capital). Notably, the number of ill soldiers who died in their hometown is not correlated with *population density*, a variable that we should expect to have an independent effect on the number of influenza casualties. In the second panel, we show that local fiscal policies are also not correlated with our main regressor. We account for *total expenditure*, *budget surplus/deficit*, *expenditure on police and sanitation/hygiene services*, *expenditure on justice* and *expenditure on education*. The focus is mainly on the expenditure side of the budget as we want to address concerns about a systematic relationship between the provision of a specific public good and the probability of having soldiers returning and eventually dying in their hometown. Importantly, our proxy variable is not associated with expenditure on sanitation and hygiene services.

5. Results

5.1. Main results

We now turn to the estimates of our central results as delineated above. Table 3 reports the estimate of our main analysis. The explanatory variable is the number of soldiers who returned ill in their hometown and died there during the peak of the epidemic, August–December 1918, which captures the local severity of influenza. The measure is expressed in per-capita terms (with the 1911 population as a benchmark) and standardized.

Panel A considers the impact on the Gini index, computed from the municipal tax declaration in 1924, and expressed in logarithm.[22] In each of the proposed specifications, the estimates suggest that cities hit more severely by the epidemics display higher levels of income inequality five years later, as measured by the Gini index. In Column (1) we show the effect without including any covariates, while from Column (2) to (4) we begin progressively adding a rich set of fixed effects and municipal controls. In particular, an increase by one standard deviation of our proxy variable raises the Gini index by between 2% and 2.7%. The size of this effect is not negligible: for instance, it is higher than that found by Furceri et al. (2020) who estimate an average effect of the presence of an epidemic on the Gini index to be between 0.75% and 1.25%. It is reassuring for the causal interpretation of our results that the coefficient remains rather stable as we gradually include covariates. More specifically, in the spirit of Oster (2019), we estimate

[22] The main results emerge also if we use the Gini index not expressed in logarithm.

CEPR PRESS

Covid Economics 33, 30 June 2020: 73-109

Table 3: Impact of WWI victims on inequality indicators.

	(1)	(2)	(3)	(4)
Panel A: Municipal Gini index (log)				
Victims WWI - illness (hometown)	0.027***	0.021**	0.021**	0.020**
	(0.010)	(0.009)	(0.009)	(0.008)
N	1804	1804	1804	1804
R^2	0.004	0.225	0.276	0.279
Panel B: Income share of top 20%				
Victims WWI - illness (hometown)	0.006	0.003	0.004	0.003
	(0.004)	(0.004)	(0.004)	(0.004)
N	1774	1774	1774	1774
R^2	0.001	0.152	0.222	0.225
Panel C: Income share of bottom 20%				
Victims WWI - illness (hometown)	-0.002**	-0.002*	-0.002*	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
N	1774	1774	1774	1774
R^2	0.004	0.199	0.258	0.262
Province FE	No	Yes	No	No
Geography FE	No	Yes	Yes	Yes
N contributors quartile FE	No	Yes	Yes	Yes
District FE	No	No	Yes	Yes
Municipal controls	No	No	No	Yes

Notes: The dependent variable is municipal Gini index in 1924 (Panel A), the income share of top 20% in 1924 (Panel B) and the income share of bottom 20% in 1924 (Panel C). The variable *Victims WWI-illness (hometown)* captures the number of soldiers who died of illness in their home town between August and December 1918, in per-capita terms (with population of 1911) and standardized. Municipal controls include a dummy whether the city is a province capital, population density of the city (defined as the ration between 1911 population and city area), total expenditure, budget surplus and municipal spending in police, health service, justice and education (all budget variables are in per-capita and refers to the year 1912). Fixed effects are described in section 4. Robust standard errors are clustered at the district level are in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

Covid Economics 33, 30 June 2020: 73-109

how large the unobservables must be in comparison to the observables, to nullify the effect that we find. To put it differently, this helps evaluate how likely the presence of omitted variable bias is in the estimates. Larger values are associated with smaller chances that OVB would substantially affect our findings. Moving from the estimate of Column (1) to that of Column (4) there is a decrease in the effect of 0.7% and an increase in the R^2 from 0.004 to 0.279($= \tilde{R}$), which implies that effect of unobservables must be no less than 7.3($= \delta$) times larger than that of observables to cancel out the effect.[23]

In Panel B and C of Table 3, we explore which part of the income distribution is more affected by the epidemics in order to understand the effect on income inequality. With this aim in mind, we focus on the income of resources held by the top 20% and by the bottom 20% of the population. These estimates demonstrate that in municipalities in which the epidemic was harsher, (1) the fraction of resources held by the top 20% is higher, though the effect is not statistically significant, and (2) the share of resources detained by the bottom 20% significantly decreases. The latter effect is large, as one-standard-deviation increase in the treatment reduces the income share of the bottom 20% by 0.2%.[24] These findings suggest that the increase in inequality is mostly driven by a significant impoverishment of the poor groups and a weak increase in the income of the rich.[25] Furthermore, we conducted a set of additional analyses, displayed in Table A.3 in the Appendix, confirming the overall evidence. In particular, we show that the occurrence of the influenza leads to a significant increase in municipal income dispersion, measured by income standard deviation, and to an increase, though not significant, in the income share of the top 10% and to a decrease, also not significant, in the income share of the bottom 10%.

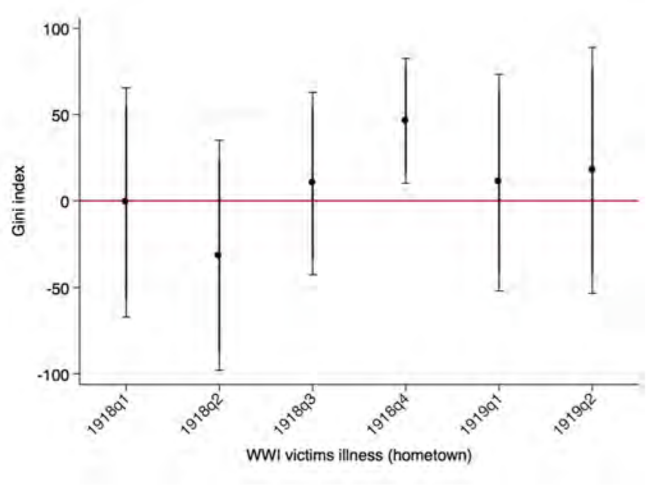
[23]This is well above the cutoff value of $\delta = 1$ (Altonji et al., 2005; Oster, 2019). When $\delta = 1$, the observables are as important as the unobservables. To compute these measures, we follow Oster (2019) and use an $R_{max} = \tilde{R} \times 1.3$. We also tested with alternative values of R_{max}. For instance, when we use an $R_{max} = \tilde{R} \times 2$ the effect of unobservables must be no less than 2.2 times larger than that of observables to cancel out the effect.

[24]We replicate the Oster (2019) test when using the share of the bottom 20% as a dependent variable, and we find very similar results (i.e., $\delta = 6.7$ when $R_{max} = \tilde{R} \times 1.3$ and $\delta = 2.0$ when $R_{max} = \tilde{R} \times 2$).

[25]One potential concern regarding these findings may be if, in the cities most affected by the Spanish Flu, a fraction of the poor taxpayers went out of business and no longer appeared in the tabulations in 1924. Even if it that was the case, this would work against our findings as this would mechanically increase, ceteris paribus, the resources of the poor in the cities heavily affected by the flu. Importantly, we show in Section 5.2 that the number of taxpayers was not affected by the severity of the flu at the local level. Overall, we find no evidence that our sample is censored.

Covid Economics 33, 30 June 2020: 73-109

Figure 7: Effect of pandemic exposure on the Gini index



Notes: The plot shows the estimates according to Equation 3. For each coefficient, 95% (delimited by horizontal bars) and 90% (bold line) confidence intervals are included. The dependent variable is municipal Gini index in 1924 and the explanatory variable is the number of soldiers that returned -infected- and died in their hometown in the corresponding quarter (per-capita). The specification is the same as in column (1) of Table 3.

CEPR PRESS

Covid Economics 33, 30 June 2020: 73-109

Finally, we conduct a series of placebo tests to show that our treatment do not capture factors other than the severity of the disease at the municipal level. First, we show that what matters is the number of soldiers who died in their hometown during the peak of the flu. We are interested in studying how the effect of the inflow of ill soldiers varies over time. In Figure 7 we test the impact of the main treatment by quarter on the Gini index focusing on the period January 1918–June 1919, in which the epidemic started and developed.[26] The graph shows that the inflow of ill soldiers led to an increase in the Gini index only when the pandemic was particularly severe (i.e., in the fourth quarter of 1918), while the impact was zero in the other quarters. More formally, in Appendix Table A.5 Column 1, we find that the number of soldiers dying at home from a disease before the peak of the flu, in January–July 1918, do not correlate with local income inequality (Appendix Table A.5 Column 1). These results further confirm that our main explanatory variable is a good proxy for the local incidence of the Spanish Flu and that the channel is the spread of the disease in the localities that received more ill soldiers in the months of the peak.

Second, we show that what really define the effect on local inequality is the number of soldiers who died because of illness in their hometown rather than the overall number of victims of illness. Thus, we evaluate the impact of the number of soldiers who died of illness in the period of the peak outside that locality, rather than in their hometown (Appendix Table A.5 column 2).

Third, we show that the effect does not depend on the overall exposure of the locality to the war. We report that inequality is not affected by the total number of WWI victims from a specific municipality, excluding those victims considered in our main treatment (Appendix Table A.5 Column 3).[27]

[26]For this analysis we estimate the following model:

$$Inequality_{idp} = \sum_{q=1918q1}^{1919q2} \beta_q \, Victims\ WWI\ for\ illness\ hometown_{idpq} + \gamma \mathbf{X}_{idp} + \eta_{idp} \tag{3}$$

where *Victims WWI for illness hometown* is equal to the number of reported soldiers returning from the war frontline who died because of a disease complication in municipality i (district d, province p) during the quarter q, and all other terms are defined as in Equation 2.

[27]Coherently, in unreported estimates, we find that our main results are not affected if we include this variable as a control. Moreover, we find that inequality is not affected when considering as treatment, either the total number of WWI victims or the number of victims once excluding those who died of illness.

Covid Economics 33, 30 June 2020: 73-109

Overall, these results show that the diffusion of the Spanish Flu across Italian municipalities increased income inequality and that this effect was mostly driven by a reduction of the resources of the poor strata of the population.

5.2. Additional results

In this section, we discuss a set of additional analysis and robustness checks. These results are shown in Table 4.

Table 4: Additional analysis

	Dep. var.: Gini index				Dep. var.: Economic outcomes	
	Entire sample (1)	Entire sample (2)	Without province capitals (3)	Without provinces at the front (4)	Total declared income (per-capita) (5)	Number of taxpayers (per-capita) (6)
Victims WWI - illness (hometown) - binary	0.034**					
	(0.015)					
Victims WWI - illness (place of death)		0.019*				
		(0.011)				
Victims WWI - illness (hometown)			0.021**	0.015*	1.281	0.000
			(0.008)	(0.008)	(0.894)	(0.000)
N	1804	1804	1775	1744	1805	1805
R^2	0.278	0.277	0.278	0.275	0.375	0.372
Geography FE	Yes	Yes	Yes	Yes	Yes	Yes
N contributors decile FE	Yes	Yes	Yes	Yes	No	No
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Municipal controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is municipal Gini index in 1924 (columns 1-4), the municipal total declared income in thousands of Liras (column 5) and the number of taxpayers (column 6). The variable *Victims WWI-illness (hometown) - binary* is a binary variable equal to one if there is at least one soldier who died of illness in his hometown between August and December 1918. The variable *Victims WWI-illness (place of death)* captures the number of soldiers who died of illness and focuses on the municipality of death (rather than their birthplace) between August and December 1918. The variable *Victims WWI-illness (hometown)* captures the number of soldiers who died of illness in their home town between August and December 1918. Both these variables are in per-capita terms (with population of 1911) and standardized. Municipal controls include a dummy whether the city is a province capital, population density of the city (defined as the ration between 1911 population and city area), total expenditure, budget surplus and municipal spending in police, health service, justice and education (all budget variables are in per-capita and refers to the year 1912). Fixed effects are described in section 4. Robust standard errors are clustered at the district level are in parentheses:* p<0.10, ** p<0.05, *** p<0.01.

First, we show that the main result with respect to the Gini index is robust to modifications in how the explanatory variable is constructed and in the sample of municipalities included. In Column (1) of Table 4 we use an explanatory variable expressed in binary term: In particular, we define the variable *Victims WWI illness (hometown) - binary* as equal to one if a municipality has a positive inflow of ill soldiers who died there and zero otherwise. The impact of this treatment is larger than that of the main variable: Treated cities experienced an average increase in the Gini index of 3.4%. Moreover, in Column (2) of Table 4, we construct the variable *Victims WWI illness (place of death)*: in which we treat a municipality with the number of soldiers who died there by illness (in the months August-December 1918), regardless of whether that locality was their hometown. This variable is larger, on average, than our main explanatory variable, which only focuses on the soldiers who died in their hometown. The scope of this test is, indeed, studying whether there is an effect on inequality if we include in our treatment those

Covid Economics 33, 30 June 2020: 73-109

soldiers who died of illness in a city of passage. This alternative treatment has a positive and significant effect on the Gini index, but the magnitude is slightly smaller than the main one. This is reasonable, a soldier on leave is more likely to have multiple interactions with the community and this may spread the contagion more compared to infected troopers who arrived in a non-hometown city. Finally, we show that the main results hold if we exclude from the sample municipalities that are province capitals (Column 3) and those located in provinces at the Italian war front (Column 4).

Next, we study whether the Spanish Flu had an effect on economic activities. Column (5) of Table 4 shows the impact on the total income declared, expressed in per capita terms; no significant effects emerge. Moreover, Column (6) of Table 4 shows that the Spanish Flu does not affect the number of taxpayers, also in per capita terms. These results suggest that the epidemics did not have a substantial effect on economic activities as proxied by declared income, at least in the medium term. This finding is not surprising and is in line with the results of Carillo and Jappelli (2020) on Italian regions. Indeed, the authors show that the 1918 influenza had an initial negative effect on GDP that vanished after four years. This suggests that, while the contraction of the GDP vanished relatively soon, the impact on income inequality appears to have persisted over time, an issue that next section explores further.

5.3. Long term effects

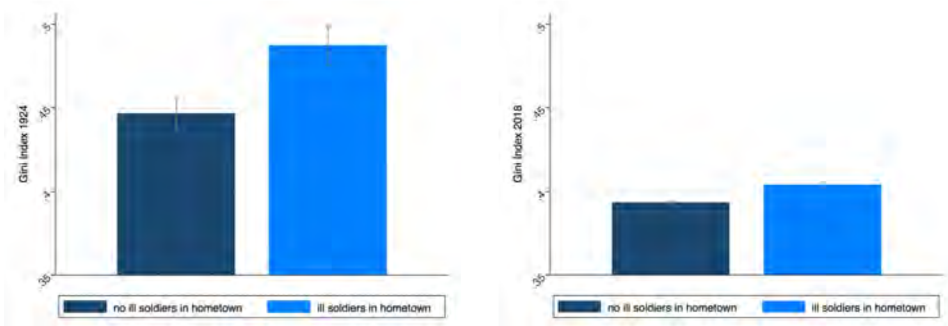
While the main analysis addresses the short-/medium-term effect of the Spanish Flu on income inequality, here, we provide some initial evidence of its long-term effects.

To conduct this analysis we constructed a measure of municipal inequality in 2018 allowing us to study the effect of Spanish Flu severity on inequality after 100 years. Specifically, we replicate the main analysis, this time using the Gini index in 2018 as the dependent variable and, alternatively, the continuous and the binary version of our treatment variable.[28] This allows us to measure inequality for each municipality in Italy. Figure 8 displays a simplified version of the results graphically, while the actual estimates are presented in Appendix Table A.4. Panel A of Figure 8 shows the difference in the average Gini index in 1924 between municipalities that had at least one soldier who died of illness in the hometown during the pandemic, and those who did not, while Panel B shows the same relationship using Gini index in 2018. In both cases, there is a significant

[28]As mentioned, in 2018, we do not avail ourselves of individual income declaration; therefore, the inequality index come from grouped data.

Figure 8: Inequality in the long-run

(A) Gini index in 1924 (B) Gini index in 2018



Notes: The plots show the Gini index in 1924 (left plot) and in 2018 (right plot) for Italian municipalities dividing between cities where no ill soldiers came back from the front and those where a positive number of ill soldier that returned in the period August-December 1918.

difference between the two groups, with the municipality in which the treatment is 0 reporting a lower level of income inequality. From the estimated coefficients, when using the saturated specification, we find that treated municipalities have a Gini index in 2018 that is 1.5% (significant at the 1% level) higher using the whole sample of Italian municipalities, and 1.2% (significant at the 5% level) with the sample of municipalities used in the main analysis.[29] Instead, when using the continuous treatment variable, the coefficient is still positive, but is not always statistically significant.

Altogether, these results provide initial evidence of the long-lasting effect of the 1918 Influenza pandemic on shaping inequality.

6. Conclusion

The recent development of the COVID-19 epidemic is expected to, at least in the short term, negatively affect the economy. However, the potential distributive effect remains unclear.

In this paper, we provide some initial evidence for the effect of pandemics on income inequality using the Italian experience with the 1918 influenza as a case study. Our results, based on geographic variation in the severity of the disease, induced by the

[29] The long-term effect is very similar when considering the broader sample of Italian municipalities and our restricted one. Indeed, this result helps us also in discouraging the worries that our main results are merely driven by the restricted sample of municipalities for which we have income data.

Covid Economics 33, 30 June 2020: 73-109

CEPR PRESS

Covid Economics 33, 30 June 2020: 73-109

likely exogenous presence of ill soldiers in local jurisdictions, suggests that pandemics increase income inequality. In the short/medium term this effect comes mainly from a reduction in the share of income generated by the poorer side of the population, while top earners do not seem to be affected. Finally, the general effect on inequality appears to have long-lasting consequences, as after 100 years we still find that municipalities that experienced the most damage from the flu report a less equal distribution of income.

We are clearly aware that our study is focused on a specific historical setting and it is not directly comparable with the pandemics of previous centuries, nor with what the world is experiencing today. This implies that our results are difficult to generalize and that we face external validity issues. Nevertheless, our findings are in line with existing evidence focusing on more recent epidemics (Furceri et al., 2020), supporting the possibility that a similar pattern will occur as a result of the COVID-19 pandemic. Therefore, also in light of potential long-term effects, this study's outcomes would call for interventions aimed at attenuating potential distributive consequences that may increase inequalities also for future generations.

There are many possible lines of research that future studies may follow. First, the main channels at the basis of our results require further exploration: Understanding under which conditions pandemics may trigger income inequality is important, from both a positive and a normative perspective. Second, the study of which policy interventions are more successful in limiting the distributive effects of an epidemic is of vital importance, especially today, to avoid a surge in future levels of inequality.

References

Aassve, A., Alfani, G., Gandolfi, F., Le Moglie, M., 2020. Epidemics and trust: the case of the spanish flu. IGIER Working Paper No. 661 .

Adams-Prassl, A., Boneva, T., Golin, M., Rauh., C., 2020. Inequality in the impact of the corona- virus shock: Evidence from real time surveys. .

Alfani, G., 2015. Economic inequality in northwestern italy: A long-term view (fourteenth to eighteenth centuries). The Journal of Economic History 75, 1058–1096.

Alfani, G., Ammannati, F., 2017. Long-term trends in economic inequality: the case of the florentine state, c. 1300–1800. The Economic history review 70, 1072–1102.

Alfani, G., Melegaro, A., 2010. Pandemie d'Italia: dalla peste nera all'influenza suina : l'impatto sulla società. Cultura e società, Egea.

Almond, D., 2006. Is the 1918 influenza pandemic over? long-term effects of in utero influenza exposure in the post-1940 us population. Journal of political Economy 114, 672–712.

Alon, T., Doepke, M., Olmstead-Rumsey, J., Tertilt., M., 2020. The impact of the coronavirus pan- demic on gender equality. Covid Economics Vetted and Real-Time Papers, Issue 4 .

Altonji, J.G., Elder, T.E., Taber, C.R., 2005. An evaluation of instrumental variable strategies for estimating the effects of catholic schooling. Journal of Human resources 40, 791–821.

Beach, B., Ferrie, J.P., Saavedra, M.H., 2018. Fetal shock or selection? The 1918 influenza pandemic and human capital development. Technical Report. National Bureau of Economic Research.

Bell, B., Bloom, N., Blundell, J., Pistaferri, L., 2020. Prepare for large wage cuts if you are younger and work in a small firm. VoxEU.org 06/04/2020 .

Capodarca, V., 1991. Le ultime voci della grande guerra. Tra le righe.

Carillo, M.F., Jappelli, T., 2020. Pandemic and local economic growth: Evidence from the Great Influenza in Italy. Technical Report. Covid Economics 10: 1-23.

Correia, S., Luck, S., Verner, E., 2020. Pandemics depress the economy, public health interventions do not: Evidence from the 1918 flu. Technical Report.

Crosby, A., 1989. America's forgotten pandemic: The influenza of 1918. Cambridge University Press, Cambridge .

Cutolo, F., 2020. L' influenza spagnola del 1918-1919. La dimensione globale, il quadro nazionale e un caso locale. I.S.R.Pt Editore.

Dahl, C.M., Hansen, C.W., Jense, P.S., 2020. The 1918 epidemic and a V-shaped recession: Evidence from municipal income data. Technical Report.

Davies, J.B., Sandström, S., Shorrocks, A., Wolff, E.N., 2011. The level and distribution of global household wealth. The Economic Journal 121, 223–254.

Fornasin, A., Breschi, M., Manfredini, M., 2018. Spanish flu in italy: new data, new questions. Le infezioni in medicina 26, 97.

Furceri, D., Loungani, P., Ostry, J.D., Pizzuto, P., 2020. Will Covid-19 affect inequality? Evidence from past pandemics. Technical Report. Covid Economics 12: 138-57.

Galasso, V., 2020. Covid: not a great equaliser. Covid Economics 19 (2020): 241-265 .

Giommoni, T., 2019. Does progressivity always lead to progress? the impact of local redistribution on tax manipulation. CESifo Working Paper Series No. 7588 .

Guimbeau, A., Menon, N., Musacchio, A., 2020. The brazilian bombshell? the long-term impact of the 1918 influenza pandemic the south american way. Technical Report. National Bureau of Economic Research.

Herring, D.A., Sattenspiel, L., 2011. Death in winter: Spanish flu in the canadian subarctic, in: Killingray, D., Phillips, H. (Eds.), The Spanish Influenza Pandemic of 1918-1919: New Perspectives. Routledge Studies. Routledge Studies in the Social History of Medicine. chapter 10.

Istituto Centrale di Statistica, 1958. Istituto Centrale di Statistica: Cause di Morte (1887-1955).

Johnson, N.P., Mueller, J., 2002. Updating the accounts: global mortality of the 1918-1920" spanish" influenza pandemic. Bulletin of the History of Medicine , 105–115.

Karlsson, M., Nilsson, T., Pichler, S., 2014. The impact of the 1918 spanish flu epidemic on economic performance in sweden: An investigation into the consequences of an extraordinary mortality shock. Journal of Health Economics 36, 1 – 19.

Lin, M.J., Liu, E.M., 2014. Does in utero exposure to illness matter? the 1918 influenza epidemic in taiwan as a natural experiment. Journal of health economics 37, 152–163.

Mortara, G., 1925. La salute pubblica in Italia durante e dopo la guerra. v. 9, G. Laterza & figli.

Oster, E., 2019. Unobservable selection and coefficient stability: Theory and evidence. Journal of Business & Economic Statistics 37, 187–204.

Covid Economics 33, 30 June 2020: 73-109

Oxford, J.S., Lambkin, R., Sefton, A., Daniels, R., Elliot, A., Brown, R., Gill, D., 2005. A hypothesis: the conjunction of soldiers, gas, pigs, ducks, geese and horses in northern france during the great war provided the conditions for the emergence of the "spanish" influenza pandemic of 1918–1919. Vaccine 23, 940–945.

Percoco, M., 2016. Health shocks and human capital accumulation: the case of spanish flu in italian regions. Regional Studies 50, 1496–1508.

Sala-i-Martin, X., 2006. The world distribution of income: falling poverty and... convergence, period. The Quarterly Journal of Economics 121, 351–397.

Taubenberger, J.K., Morens, D.M., 2006. 1918 influenza: the mother of all pandemics. Emerging infectious diseases 12(1), 15—-22.

Tognotti, E., 2015. La" spagnola" in Italia. Storia dell'influenza che fece temere la fine del mondo (1918-1919): Storia dell'influenza che fece temere la fine del mondo (1918-1919). FrancoAngeli.

Winter, J., 2010. L'inuenza spagnola. in: La prima guerra mondiale, a cura di S. Audoin-Rouzeau, J. Becker, Einaudi, Torino 2010, pp. 288 .

Zugaro, F., 1926. Bollettino dell' ufficio storico. l'albo d'oro dei caduti per l'iitalia nella guerra mondiale .

Covid Economics 33, 30 June 2020: 73-109

CEPR PRESS

Covid Economics 33, 30 June 2020: 73-109

A. Appendix - Additional Figures and Tables

Table A.1: Descriptive statistics

	Average value	Standard deviation	N
Panel A: Influenza (region)			
Casualties Influenza 1915-1920	9,617.8	11,055.4	96
Panel B: Inequality indicators (municipality)			
Gini index (1924)	.459	.162	1,804
Top 20% share	.535	.152	1,774
Bottom 20% share	.055	.034	1,774
Top 10% share	.395	.170	1,716
Bottom 10% share	.022	.016	1,716
Income standard deviation	4,326.9	5,692.0	1,804
Gini index (2018)	.401	.034	1,781
Gini index (2018) [entire sample]	.395	.040	6,747
Panel C: Victims WWI in per capita terms (municipality)			
Casualties 1915-1920	.015	.0060	1,804
Casualties illness 1915-1920	.0054	.0029	1,804
Casualties illness (hometown) 1915-1920	.0006	.0007	1,804
Casualties illness (hometown) Aug-Dec 1918	.0001	.0002	1,804
Casualties illness (hometown) Jan-Jul 1918	.00007	.0002	1,804
Casualties illness (front) Aug-Dec 1918	.0016	.0012	1,804
Casualties illness (place of death) Aug-Dec 1918	.0003	.0009	1,804
Casualties illness hometown (binary) Aug-Dec 1918	.3314	.4708	1,804
Casualties illness (hometown) 1918q1	.00003	.0001	1,804
Casualties illness (hometown) 1918q2	.00003	.0001	1,804
Casualties illness (hometown) 1918q3	.00004	.0001	1,804
Casualties illness (hometown) 1918q4	.00009	.0002	1,804
Casualties illness (hometown) 1919q1	.00005	.0001	1,804
Casualties illness (hometown) 1919q2	.00002	.0001	1,804
Panel D: Additional information (municipality)			
Population 1911	5,040.0	11,223.5	1,804
Population density 1911	1.64	4.23	1,804
Population growth rate 01-11	.058	.105	1,787
Mountain city	.768	.667	1,804
Coastal city	.253	.547	1,804
Province capital	.016	.125	1,804
Total expenditures	23.90	24.60	1,804
Budget surplus	.64	1.99	1,804
Expenditures in police and sanitation/hygiene services	4.91	8.42	1,804
Expenditures in justice and security	.22	1.07	1,804
Expenditures in education	4.27	6.21	1,804

Notes:Casualties Influenza measures the regional and yearly number of victims for Influenza and Pneumonia. *Gini index (1924)* measures the municipal Gini index in 1924, it is constructed with individual incomes of taxpayers in the categories B and C. *Top 20% share*, *Top 10% share*, *Bottom 20% share*, *Bottom 10% share* measure the fraction of wealth of, respectively, the top 20%, the top 10%, the bottom 20% and the bottom 10% taxpayers. *Gini index (2018)* measures the municipal Gini index in 2018, it is constructed with bracket-specific income data.*Casualties 1915-1920*, *Casualties illness 1915-1920* and *Casualties illness (hometown) 1915-1920* measure, respectively, the number of military victims, the number of military victims by illness, the number of military victims by illness (in their hometown). The variables *Casualties illness (hometown) Aug-Dec 1918/Jan-Jul 1918* measures the number of military victims by illness (in their hometown) in the reference period. The variable *Casualties illness (front) Aug-Dec 1918* measures the number of military victims by illness (outside their hometown) in the reference period. The variable *Casualties illness (place of death) Aug-Dec 1918* measures the number of military victims by illness in the reference period, attributed to the municipality where they passed away. The variable *Casualties illness hometown (binary) Aug-Dec 1918* is the binary version of the variable *Casualties illness (hometown) Aug-Dec 1918*. *Population 1911* is municipal population in 1911, *Population density 1911* is the ration between municipal population and city area in 1911.*Population growth rate 01-11* is the growth rate from 1901 to 1911 census. *Mountain city*, *Coastal city* and *Province capital*, are dummy variable capturing whether the city, respectively, is in a mountain area, is on the coast and is a province capital. The variables *Total expenditures*, *Budget surplus*, *Expenditures in police and health services*, *Expenditures in justice and security* and *Expenditures in education* are expressed in per-capita terms.

CEPR PRESS

Covid Economics 33, 30 June 2020: 73-109

Table A.2: Soldiers casualties and influenza deaths

	N victims of Influenza			
	Year 1918	Year 1917	Post-pandemic 1918-1920	Pre-pandemic 1915-1917
	(1)	(2)	(3)	(4)
Victims WWI - illness (hometown)	0.418**	-0.0890	0.296**	-0.131
	(0.174)	(0.126)	(0.132)	(0.112)
Victims WWI - others	-0.988***	-0.0902	-0.962***	-0.0811
	(0.289)	(0.109)	(0.307)	(0.0732)
Year FE	No	No	Yes	Yes
N	16	16	48	48
R^2	0.513	0.163	0.935	0.195

Notes: Each column represents a single regression. All the variables are in per-capita terms (with population of 1911) and standardized. The unit of observation are Italian regions. Robust standard errors are clustered at the region level are in parentheses: * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

CEPR PRESS

Covid Economics 33, 30 June 2020: 73-109

Table A.3: Impact of WWI victims on inequality indicators (additional variables).

	(1)	(2)	(3)	(4)
Panel A: Income standard deviation (log)				
Victims WWI - illness (hometown)	0.085***	0.061**	0.058**	0.055**
	(0.030)	(0.027)	(0.027)	(0.026)
N	1804	1804	1804	1804
R^2	0.005	0.208	0.262	0.269
Panel B: Income share of top 10%				
Victims WWI - illness (hometown)	0.003	0.001	0.001	0.001
	(0.004)	(0.004)	(0.004)	(0.004)
N	1716	1716	1716	1716
R^2	0.000	0.131	0.197	0.200
Panel C: Income share of bottom 10%				
Victims WWI - illness (hometown)	-0.001	-0.001	-0.001	-0.001
	(0.000)	(0.001)	(0.001)	(0.001)
N	1716	1716	1716	1716
R^2	0.002	0.209	0.263	0.270
Province FE	No	Yes	No	No
Geography FE	No	Yes	Yes	Yes
N contributors decile FE	No	Yes	Yes	Yes
District FE	No	No	Yes	Yes
Municipal controls	No	No	No	Yes

Notes: The dependent variable is income standard deviation in 1924 (Panel A), the income share of top 10% in 1924 (Panel B) and the income share of bottom 10% in 1924 (Panel C). The variable *Victims WWI-illness (hometown)* captures the number of soldiers who died of illness in their home town between August and December 1918, in per-capita terms (with population of 1911) and standardized. Municipal controls include a dummy whether the city is a province capital, population density of the city (defined as the ration between 1911 population and city area), total expenditure, budget surplus and municipal spending in police, health service, justice and education (all budget variables are in per-capita and refers to the year 1912). Fixed effects are described in section 4. Robust standard errors are clustered at the district level are in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

CEPR PRESS

Covid Economics 33, 30 June 2020: 73-109

Table A.4: Long-term effect on income inequality

	All municipalities				Municipalities in 1924 sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Continuous treatment								
Victims WWI - illness (hometown)	0.004***	0.003**	0.002	0.001	0.003	0.002	0.002	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)
N	6747	6747	6747	6747	1781	1781	1781	1781
R^2	0.002	0.253	0.323	0.339	0.001	0.292	0.356	0.371
Panel B: Binary treatment								
Victims WWI - illness (hometown)	0.029***	0.018***	0.018***	0.013***	0.019***	0.015***	0.015***	0.012**
	(0.004)	(0.003)	(0.003)	(0.003)	(0.006)	(0.005)	(0.005)	(0.005)
N	6747	6747	6747	6747	1781	1781	1781	1781
R^2	0.017	0.259	0.328	0.342	0.012	0.298	0.361	0.374
Province FE	No	Yes	No	No	No	Yes	No	No
Geography FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
N contributors decile FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
District FE	No	No	Yes	Yes	No	No	Yes	Yes
Municipal controls	No	No	No	Yes	No	No	No	Yes

Notes: The dependent variable is municipal Gini index in 2018. The sample includes all Italian municipalities in columns (1-4) and only those used in the main analysis in columns (5-9). The variable *Victims WWI-illness (hometown)* captures the number of soldiers who died of illness in their home town between August and December 1918, in per-capita terms (with population of 1911) and standardized. The variable is continuous on Panel A and it is binary -0 in cities with no returning soldiers and 1 in those with a positive number- in Panel B. Municipal controls include a dummy whether the city is a province capital and population density of the city (defined as the ration between 1911 population and city area). Fixed effects are described in section 4. Robust standard errors are clustered at the district level are in parentheses: * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

CEPR PRESS

Covid Economics 33, 30 June 2020: 73-109

Table A.5: Impact of WWI victims on inequality indicators - Placebo tests

	(1)	(2)	(3)
Panel A: Municipal Gini index (log)			
Victims WWI illness hometown (January-July 1918)	-0.005		
	(0.009)		
Victims WWI illness others (August-December 1918)		0.018	
		(0.019)	
Victims WWI all (1915-1920)			0.008
− illness hometown (August-December 1918)			(0.128)
N	1804	1804	1804
R^2	0.277	0.277	0.277
Panel B: Income share of top 20%			
Victims WWI illness hometown (January-July 1918)	-0.003		
	(0.004)		
Victims WWI illness others (August-December 1918)		0.004	
		((0.005)	
Victims WWI all (1915-1920)			0.004
− illness hometown (August-December 1918)			(0.005)
N	1774	1774	1774
R^2	0.226	0.225	0.225
Panel C: Income share of bottom 20%			
Victims WWI illness hometown (January-July 1918)	0.001		
	(0.001)		
Victims WWI illness others (August-December 1918)		-0.001	
		(0.001)	
Victims WWI all (1915-1920)			-0.001
− illness hometown (August-December 1918)			(0.001)
N	1774	1774	1774
R^2	0.261	0.261	0.261
Province FE	No	No	No
Geography FE	Yes	Yes	Yes
N contributors quartile FE	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Municipal controls	Yes	Yes	Yes

Notes: The dependent variable is municipal Gini index in 1924 (Panel A), the income share of top 20% in 1924 (Panel B) and the income share of bottom 20% in 1924 (Panel C). The variable *Victims WWI-illness hometown (January-July 1918)* captures the number of soldiers who died of illness in their home town between January and July 1918. The variable *Victims WWI-illness others (August-December 1918)* captures the number of soldiers who died of illness outside their home town between August and December 1918. *Victims WWI all (1915-1920)* is the total number of soldiers who died during WWI. *Victims WWI-illness hometowns (August-December 1918)* captures the number of soldiers who died of illness in their home town between August and December 1918. All these variables are in per-capita terms (with population of 1911) and standardized. Municipal controls include a dummy whether the city is a province capital, population density of the city (defined as the ration between 1911 population and city area), total expenditure, budget surplus and municipal spending in police, health service, justice and education (all budget variables are in per-capita and refers to the year 1912). Fixed effects are described in section 4. Robust standard errors are clustered at the district level are in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

Valuing goods online and offline: The impact of Covid-19[1]

Diane Coyle[2] and David Nguyen[3]

Date submitted: 26 June 2020; Date accepted: 27 June 2020

This paper uses a survey representative of the UK online population to assess the willingness to accept loss of certain goods. We had conducted an initial survey in February, focusing on 'free' online goods and some potential substitutes and comparators. Consistent with other contingent valuation studies, consumers on average assigned valuations to many of these goods, particularly when benchmarked against revenue figures for the services. Our pilot studies, discussed in a forthcoming paper, also suggested that the actual valuations are not well anchored, but the methodology can give consistent rankings among goods. It is also a useful way to assess changes in valuations. Repeating the survey in May, during the UK, lockdown, we observed significant changes in the valuations of different goods and services, with some large differences by age and gender. In this sense the lockdown has acted as a natural experiment testing for the extent to which digital goods and physical goods are substitutes. These valuation changes may indicate which services are most valuable in a post-pandemic world where more activity takes place online. They also provide important, policy-relevant insights into distributional questions.

1 We thank Avi Collis, Joel Rogers & YouGov, and to Richard Heys & colleagues at the Office of National Statistics for their help and comments. This work was funded by the Office for National Statistics through the Economic Statistics Centre of Excellence.

2 University of Cambridge and ESCoE.
3 National Institute of Economic and Social Research and ESCoE.

Copyright: Diane Coyle and David Nguyen

CEPR PRESS

1. Introduction

Contingent valuation methods have been widely used in environmental economics to estimate the value of ecosystem services, either with no market price or characterised by externalities that drive a wedge between the market and social values. We have been exploring the applicability of these methods to 'free' digital goods, following Brynjolfsson et al. (2018), who found high values for services such as Facebook and Google search for which users do not have to pay a monetary price. Specifically, we wanted to calibrate these valuations by comparing them to other free goods, such as access to parks, and to paid-for substitutes, such as newspapers.

We therefore ran a survey (using YouGov's online panel) at the end of February and beginning of March (ahead of the UK lockdown), following a series of pilots, to elicit respondents' valuations for 30 goods. Following the lockdown, we repeated the survey on 14th May (before easing measures had begun) to see how the values had changed, and what this might tell us about changes in the economic welfare provided by different goods during periods of lockdown, given people's preferences and their consumption opportunities. We consider gender and age differences. Our forthcoming paper will consider the wider methodological issues and results.

The Covid19 lockdown has acted as a natural experiment providing insight into the changes in valuations, varying by age and gender, for a range of digital goods which are free to use and therefore do not provide market signals. We found there were large changes in both usage and willingness-to-accept valuations of some of the goods and services, such as online grocery shopping, online learning, WhatsApp and Netflix. The valuation of public parks also increased. There were declines in the case of some of the goods and services, such as LinkedIn and personal emails.

Younger age groups valued some digital services, such as online learning and Instagram and Twitter, far more highly than did older age groups, and the valuations they stated increased by more between February and May. Older people increased their valuation of online grocery shopping by substantially more than did young people, however. Women's stated values for online grocery shopping also increased by more than men's.

As in other work using contingent valuation methods, we found high willingness-to-accept valuations for the goods and services considered. These might be considered implausible when compared to benchmark figures such as average revenue per user, but our results between February and May suggest there is useful insight to be gained from changes in valuations, in rankings of the stated values attached to different goods and services, and in the comparison of valuations as between different groups.

2. Previous literature

There is a large literature on the use of contingent valuation methods in environmental economics and cultural economics (see Carson, Flores & Meade, 2001 and McFadden & Train, 2017 for surveys). Recently the approach has been applied in the context of digital goods and services, for which there is no direct market price, or where there are likely to be significant externalities including network effects, as part of a debate

CEPR PRESS

Covid Economics 33, 30 June 2020: 110-123

about whether and how these should be accounted for in aggregate economic measurement.

In one of the first of these contributions applying contingent valuation methods, Brynjolfsson, Collis & Eggers (2019) used large-scale online choice experiments to elicit valuations of consumer surplus and concluded that the welfare value was large. For instance, in the incentive-compatible discrete choice experiments, the median US Facebook user needed around $48 to give it up for a month, and $322 for one year. Others have reported a range of median valuations – a lower figure of $59 willingness-to-accept and a median $1 willingness-to-pay in Sunstein (2018) to over $1000 in Corrigan et al (2018).

The method was extended by Brynjolfsson et al (2019b) to calculate an extended GDP, GDP-B, who used estimates of consumer welfare elicited from online discrete choice experiments for a number of goods to calculate growth in the wider measure compared to conventional GDP, concluding that it would add 0.05 to 0.11 percentage points a year to US growth.

In a recent study Allcott et al (2020) found median annual valuations for Facebook of around $100 using similar methods, but queried aspects of the methodology. For example, some studies did not require users to actually deactivate their social media accounts. In particular, though, the paper finds that willingness-to-accept valuations are not firmly anchored, and changed after users in their experiment had gone without Facebook: "We find that four weeks without Facebook improves subjective well-being and substantially reduces post- experiment demand."

In our ongoing work, we are seeking to understand such findings in the context of growing interest in wider measures of economic growth and welfare (Heys, Martin & Mkandwire 2019). We used a UK survey of willingness-to-accept measures for a range of digital goods, other free goods, and marketed substitutes for digital goods.

3. The Surveys

The main survey with a sample size of 10,500 adults was conducted by YouGov between 27 February and 3 March and a smaller follow-up with a sample size of 1,600 ran from 14 -15 May.[4]

Survey participants were asked about their willingness to accept giving up access to 30 different goods for a period of time. Specifically, participants were asked to choose a sum of money based on pre-determined valuation bands as shown in Figure 1. For the main survey, half of participants were asked to give up access for a period of one month, while the other half was asked to consider twelve months. In the follow-up survey, we asked only about the twelve months.

[4] The first Covid-19 death in the UK occurred on the 5th March and the country officially went into lockdown on the 23rd March. The first steps in easing lockdown restrictions in the UK occurred on 13th May. Italy went into lockdown on the 9th March, the first country in Europe to do so. More details on the survey, including summary statistics can be found in Appendix 1.

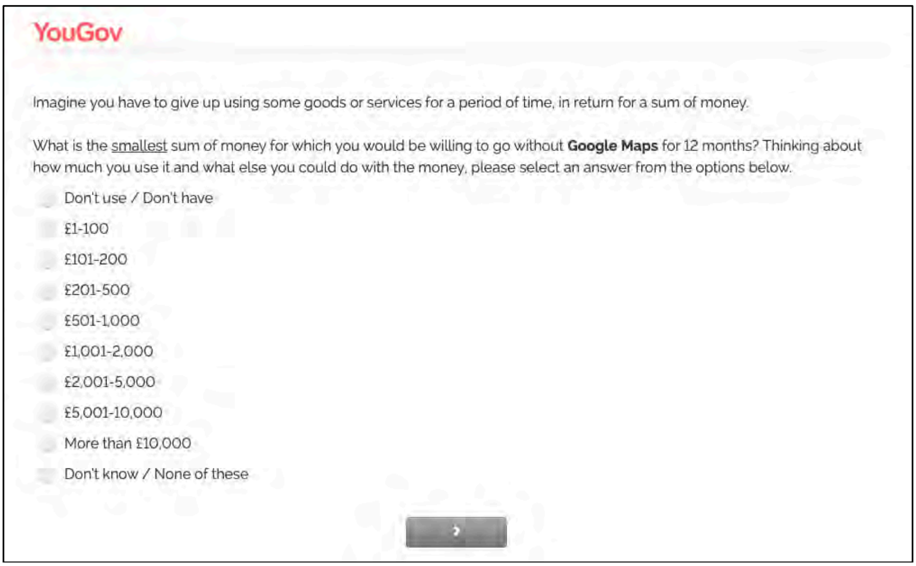
Covid Economics 33, 30 June 2020: 110-123

Both surveys were representative of the population in terms socio-economic factors including age, gender, income, education and region. However, considering that 13% of households in the UK do not have access to the internet (Ofcom, 2019) our findings likely only hold for the UK's online population.[5]

In addition, we had previously run a number of pilots to test questions and valuation bands. A total of four pilots of approximately 1,600 adults each ran between March and November 2019. We tested:

- Open box questions versus using bands
- Different time periods (week, month, 3 months, 6 months, 12 months)
- Categories versus specific goods (e.g. "All social media" versus "Facebook")
- Usage intensity

Figure 1. Typical survey question



YouGov

Imagine you have to give up using some goods or services for a period of time, in return for a sum of money.

What is the smallest sum of money for which you would be willing to go without **Google Maps** for 12 months? Thinking about how much you use it and what else you could do with the money, please select an answer from the options below.

- Don't use / Don't have
- £1-100
- £101-200
- £201-500
- £501-1,000
- £1,001-2,000
- £2,001-5,000
- £5,001-10,000
- More than £10,000
- Don't know / None of these

We selected 30 goods for the survey, based on 1) number of users and time spent on them; 2) goods used in the previous literature, to allow us to make some comparisons; 3) a wider coverage of categories than prior studies (for example including banking, gaming, news, and some non-digital goods that are potential substitutes). Where specific goods have high usage rates among the population (e.g. Facebook) we opted to asking about them specifically rather than at the category level (i.e. all social media). Asking about categories is more useful where there are many competing providers (e.g. ride-hailing services); however it is possible that people might not consider the full ramifications of giving up access (i.e. no substitutes) when compared to individual goods.

---

[5] More details on the UK's online population can be found in Appendix 2.

4. The effects of lockdown on valuations: results

Usage change, February to May

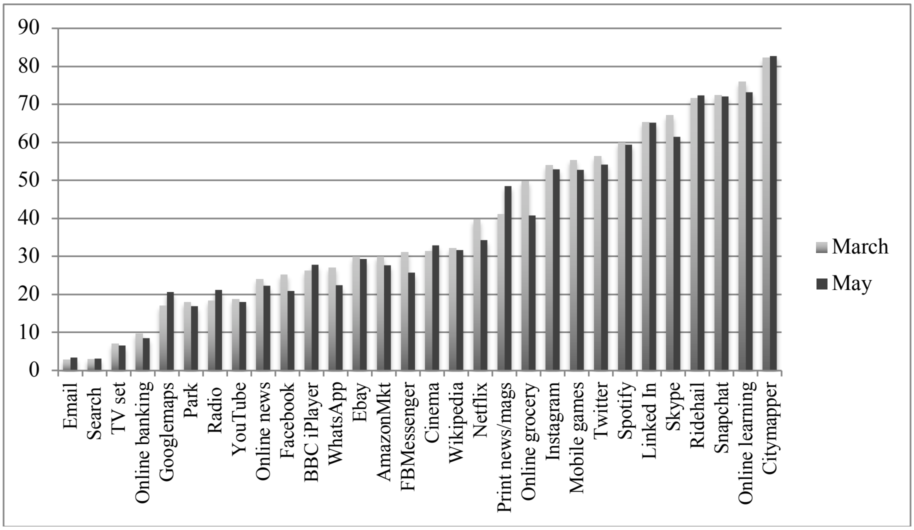
There is a substantial variation in the level of usage of the different goods included in the survey (Figure 2). Personal email, search, TV and online banking are almost universally used among the online population, while some specific digital services are used by fewer people. Between the end of February and mid-May, there were some significant changes for some goods in the proportion of respondents saying they used or did not use them (Table 1).

As the two survey waves were only 10 weeks apart and people were asked to consider the next 12 months, one might not usually expect large changes in usage rates. However, due to the Covid-10 'shock' we observe significant differences, which move in the expected direction.

While in February around 50% reported that they do not shop online for groceries, this had declined to 40% in the second survey wave. The share of people that do not use Skype, Facebook Messenger, Netflix and WhatsApp also decreased by around 5 percentage points. Other goods that saw a decline in non-usage rates are Facebook, online learning, mobile games, Amazon and Twitter.

On the other hand, the usage of various other goods has declined. For example, while in February around 41% reported they do not use consume newspapers this has increased to 48.5% by mid-May. In addition, the reported use of Google Maps, Radio, BBC iPlayer and cinemas has decreased somewhat.

Figure 2: Proportion (%) who *do not use*



Covid Economics 33, 30 June 2020: 110-123

Table 1: Changes in usage, measured in % who *do not use*, ranked by February-May change (*in percentage points*)

Item	February	May	Change
Online groceries	49.84	40.76	-9.1%
Skype	67.10	61.47	-5.6%
Messenger	31.10	25.71	-5.4%
Netflix	39.69	34.37	-5.3%
WhatsApp	27.09	22.48	-4.6%
Facebook	25.26	20.89	-4.4%
Online learning	75.92	73.19	-2.7%
Mobile games	55.36	52.76	-2.6%
Amazon	29.99	27.64	-2.4%
Twitter	56.37	54.20	-2.2%
Online news	24.04	22.27	-1.8%
Online banking	9.80	8.53	-1.3%
Instagram	54.01	52.90	-1.1%
Public park	17.99	16.94	-1.1%
YouTube	18.81	18.03	-0.8%
Spotify	59.97	59.38	-0.6%
eBay	29.89	29.38	-0.5%
TV set	7.08	6.66	-0.4%
Wikipedia	32.14	31.75	-0.4%
Snapchat	72.41	72.09	-0.3%
LinkedIn	65.34	65.19	-0.2%
Online search	3.11	3.16	+0.1%
Citymapper	82.29	82.67	+0.4%
Personal email	2.94	3.37	+0.4%
Online ride hailing	71.67	72.38	+0.7%
Cinema	31.44	32.93	+1.5%
BBC iPlayer	26.28	27.78	+1.5%
Radio	18.43	21.24	+2.8%
Google maps	17.07	20.62	+3.6%
Printed newspapers	41.05	48.50	+7.5%

Changes in valuations, February to May

While the average valuations people report differ considerably across the 30 goods, we focus on the change in valuations over the period from end of February to mid-May. Table 2 shows the annual valuations in February and May as well as the percentage change. The magnitudes of the valuations themselves are largely consistent with the literature; our forthcoming paper will discuss in more detail on how to interpret them.[6]

The valuations changed in line with expectations and have increased the most for online grocery shopping (+47%), online learning (+25%), WhatsApp (+12%) and Netflix (+8%). They also increased for Facebook, public parks, physical TV sets at home, eBay, Facebook Messenger, mobile games, and Instagram.

At the same time valuations have decreased for Twitter (-27%), printed newspapers (-24%), Google Maps (-21%), and cinemas (-18%). There were also substantial

[6] These are minimum valuations, as we are using the lower end of the valuation bands offered, shown in Figure 1 (i.e. £1, £101, £201, etc). While there is no obvious way to decide which value to choose, we err on the side of caution.

CEPR PRESS

decreases in valuations for Citymapper, online ride hailing (e.g. Uber), Spotify, any radio (often listened to in cars), LinkedIn, personal email, and Wikipedia.

We also calculated the upper and lower 95% confidence bands for the annual valuations in both periods (see Table 2). This allows us to see whether the changes in valuations between February and May are statistically significant. In the case of increasing valuations, they are significant (i.e. outside the upper confidence band of February) in the case of online groceries, online learning, WhatsApp, Netflix, Facebook, public parks, and TV sets. For most of the decreasing valuations, the May valuations are outside the lower confidence band of the February figures, with the exceptions of online ride hailing, LinkedIn, Wikipedia and Snapchat.

Table 2. Annual valuations, ranked by February-May % change

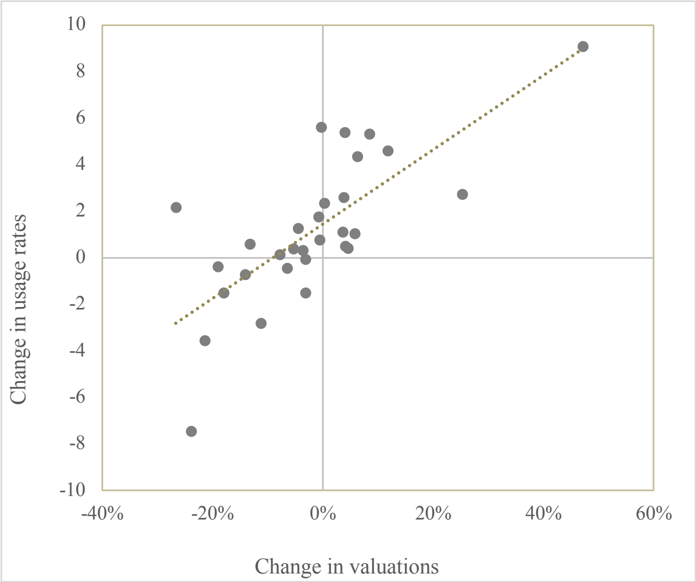
Item	Feb	Conf. interval Feb		May	Conf. interval May		Change
	(in £)	(low)	(high)	(in £)	(low)	(high)	
Online groceries	720	656	783	1059	918	1199	+47.1%
Online learning	247	208	285	309	230	388	+25.3%
WhatsApp	1588	1496	1680	1774	1599	1950	+11.8%
Netflix	1267	1185	1349	1373	1219	1528	+8.4%
Facebook	1278	1195	1360	1358	1202	1514	+6.3%
Public park	1951	1848	2053	2063	1869	2258	+5.8%
TV set	3300	3182	3419	3449	3226	3673	+4.5%
eBay	787	723	852	819	696	943	+4.0%
Messenger	1088	1011	1166	1131	987	1275	+3.9%
Mobile games	592	534	650	614	506	721	+3.7%
Instagram	657	597	717	681	569	792	+3.6%
Amazon	1044	968	1119	1046	908	1185	+0.3%
Skype	335	290	379	334	251	416	-0.3%
YouTube	1399	1313	1485	1390	1234	1546	-0.6%
Online news	1253	1170	1336	1243	1092	1393	-0.8%
Online search	3095	2977	3214	2998	2777	3219	-3.1%
BBC iPlayer	821	757	885	795	677	914	-3.2%
Snapchat	350	304	396	337	255	419	-3.6%
Online banking	2790	2674	2906	2664	2455	2874	-4.5%
Wikipedia	694	633	756	657	547	767	-5.4%
Personal email	3402	3282	3522	3181	2958	3404	-6.5%
LinkedIn	238	201	275	220	155	284	-7.8%
Radio	1713	1617	1809	1520	1350	1689	-11.3%
Spotify	696	633	759	604	499	709	-13.3%
Online ride hailing	240	204	277	206	147	266	-14.1%
Cinema	719	657	780	589	486	693	-18.0%
Citymapper	174	142	207	141	90	192	-19.0%
Google maps	1307	1224	1390	1027	889	1164	-21.4%
Printed newspapers	566	510	621	430	342	518	-23.9%
Twitter	556	501	612	408	322	494	-26.7%

These changes in valuations were strongly positively correlated with changes in usage (Figure 3), with a correlation coefficient of 0.74.

Covid Economics 33, 30 June 2020: 110-123

CEPR PRESS

Figure 3. Correlation between change in usage rates and change in valuations



The contingent valuation results can be compared to revenues per user for the service providers. Ofcom (2019) estimates various per capita revenues for online services in the UK in 2018.[7] For example, online search was estimated to be £101, £45 for social media, £27 for free video streaming, £11 for online news, £1,094 for online shopping, £47 for online entertainment, and £63 for online gaming. For almost all of these (the exception being online grocery shopping in February), the stated valuations in our survey exceed these revenue per user figures by a large margin. This is consistent with findings of a large gap between willingness-to-accept and willingness-to-pay valuations (Sunstein 2019).

Age group differences, May

In order to compare the relative valuations of different age groups across goods, we computed the ratio of valuations reported by those aged under 50 to those aged 50 or above (Figure 4). On average, younger people tend to place a higher valuation on most of the goods included.

The relative difference is even more pronounced looking at the ratio between those aged 18-24 to those aged 65 or over. The difference in valuations in this case is most pronounced in the case of Snapchat (by 50 times), Instagram and Spotify (15 times), online learning and Twitter (10 times). The differences are less pronounced but still large when splitting respondents at the age of 50.

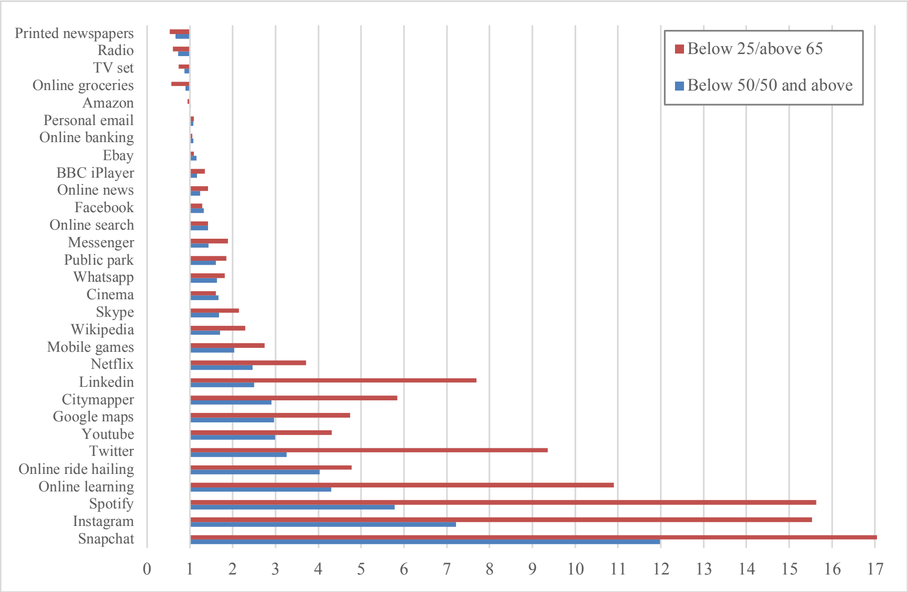
[7] Based on estimations of UK market share in total global revenues of large businesses, averaged across population in the UK rather than actual users.

Covid Economics 33, 30 June 2020: 110-123

As expected, however, older people tend to value non-digital services more than the younger people. For instance, valuations of printed newspapers, radio, and a physical TV set were twice as high for those above 65 than for those aged 18-24. Older people also valued online grocery shopping more highly, which is perhaps also not a surprise given the age incidence of serious covid19 illness.

In the case of Amazon, personal email, online banking, eBay and BBC iPlayer there appear to be no significant difference in valuations between younger and older age groups.

Figure 3. Annual valuations: young compared to old



Notes: The ratio for Snapchat in the case of "Below 25/above 65" is off the scale with a value of 53. Category "Below 25/above 65" shows ratio of average valuations of those aged below 25 to those aged 65 and above. Category "Below/above 50" shows ratio of average valuations of those aged below 50 to those aged 50 and above.

Changes in valuations by age group, February to May

Changes in valuations differed across age groups. For example, while valuations of Facebook decreased by 2% to 4% for those aged 25-65, it increased by 26% for those aged 18-24 and by 38% for those aged above 65.

There were also stark differences in the case of online grocery shopping, which increased in value for all age groups apart from those aged 18-24. The value that people aged 65 and above attached to this increased by 127%, while for people between 25-64 it increased by 37%.

Wikipedia was another interesting contrast. For those aged 18-49 the value decreased by 13-16%, while for those aged 50 and above valuations increased by 14%.

Turning to the smaller changes, the valuation of mobile games decreased for the youngest and oldest age groups but increased for the two groups in the middle.

Online learning increased in stated value by between 20-40% for all age groups below 65, while its value decreased by 20% for those 65 and above.

The value that different age groups attached to public parks increased markedly for those aged 18-24 (+25%) and 25-49 (+13%), while it appeared to have decreased for the groups of 50-64 (-3%) and over 65 (-13%).

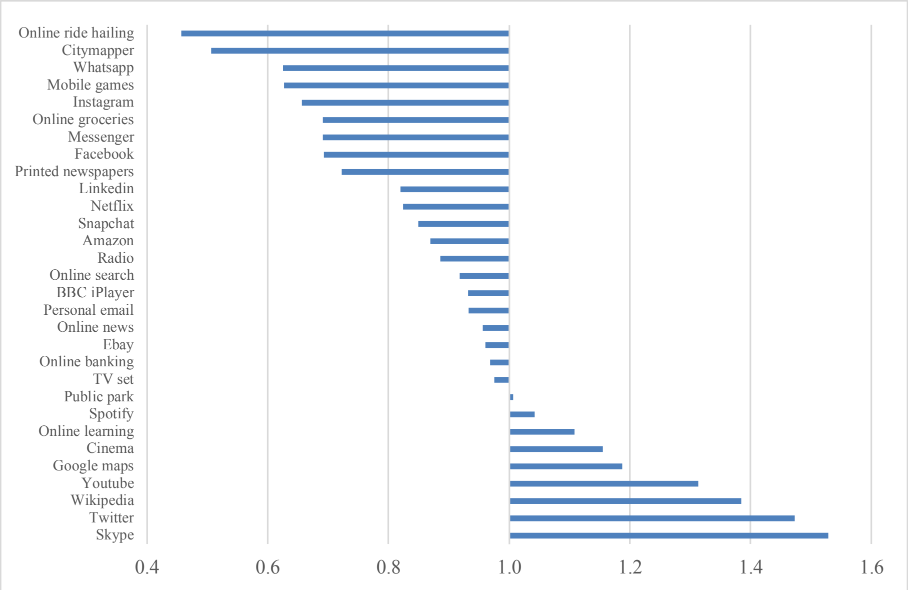
In another case, the value of the BBC iPlayer appears to have decreased for all age groups (between -2% to -9%), apart from those aged 18-24 (+16%). A similar pattern holds for radio.

Gender differences

Differences in valuations by men and women are also striking (see Figure 5). In the May results, for most of the goods, men appear to have a lower willingness to accept than women. The differences are most pronounced in the case of Skype and Twitter (male valuations around 50% higher) and online ride hailing and Citymapper (female valuations around 50% higher).

Valuations of women also were around a third higher for WhatsApp, mobile gaming, Instagram, online grocery shopping, Facebook Messenger, and Facebook. Men on the other hand attach higher valuations to Wikipedia (+38%), YouTube (+31%), and Google Maps (+19%).

Figure 5. Annual valuations: ratio of male to female, May 2020



Covid Economics 33, 30 June 2020: 110-123

CEPR PRESS

There were some large changes between February and May. In February around 51% of men and 49% of women did not use online grocery shopping. Over the lockdown period these proportions decreased to 44% for men and 38% for women. Valuation for online grocery shopping thus increased relatively more for women (+51%) than for men (+41.5%).

We also found large difference in changes in valuations in the case of Skype (women -15%, men +14%), online news (women +15%, men -14%), online learning (women +62%, men +4%), LinkedIn (women +21%, men -30%), online ride hailing (women +25%, men -50%), mobile gaming (women +16%, men -12%), printed newspapers (women -2%, men -43%), and WhatsApp (women +18%, men +2%).

Changes in valuations by social group, February to May

Next we look at the changes in valuations by six socio-economic grades and we find stark differences across groups.[8] In most cases the changes in valuations have different signs across the groups. The valuations increased for all groups (online grocery shopping) or decreased for all (cinema) for only a few goods.

Interestingly, the valuations for LinkedIn and online learning decreased for grades A to C2, but increased considerably for grades D and E. For example, in the case of online learning the WTA loss of access increased by more than 400% for semi-, unskilled and manual workers (grade D). This group also stands out as having the largest increase in valuations for Wikipedia, and various social media sites (excluding Facebook where their valuations were already the highest among all the social grades).

For people in grade A (*High managerial, administrative or professional*) valuations decreased for most goods and by the most for mobility apps (online ride hailing, Google Maps, and Citymapper) along with eBay, Twitter, printed newspapers and cinemas. Apart from online grocery shopping valuations of this group only markedly increased for YouTube (+17%).

Valuations for people in grade B (*Intermediate managerial, administrative or professional*) decreased the most for printed news, Skype and Snapchat, but increased for eBay, Facebook and online grocery shopping.

Valuations for grade C1 (*Supervisory, clerical and junior managerial, administrative or professional*) decreased considerably for Twitter, and printed newspapers, and increased significantly only for online grocery shopping.

For C2 (*Skilled manual workers*) valuations did not increase much, apart from online groceries and Amazon and to a lesser degree Netflix, Facebook, WhatsApp and a TV set at home. Their valuations went down considerably for Skype and cinemas and also Twitter, LinkedIn, Spotify and YouTube.

[8] The NRS six social grades are: A-High managerial, administrative or professional; B-Intermediate managerial, administrative or professional; C1-Supervisory, clerical and junior managerial, administrative or professional; C2-Skilled manual workers; D-Semi and unskilled manual workers; E-State pensioners, casual or lowest grade workers, unemployed with state benefits only. The social grades refer to the chief income earner in a household.

Covid Economics 33, 30 June 2020: 110-123

Covid Economics 33, 30 June 2020: 110-123

People in grade E (*State pensioners, casual or lowest grade workers, unemployed with state benefits only*) reported a large decrease in valuations for Spotify, Google maps and Snapchat as well as Wikipedia and Twitter. Apart from online groceries they saw an increase in the value of LinkedIn (+32%).

Conclusions

Our ongoing work looking at contingent valuation methods is assessing whether the methodology could be a useful way of accounting for the value of 'free' digital goods, either in economic welfare terms or particularly in terms of macroeconomic statistics. We used a survey representative of the UK's online population, rather than incentive-compatible laboratory experiments, because the use of the method would need to be scalable for statistical production.

Consistent with other studies, we found that consumers on average assign high willingness-to-accept valuations to many of these goods, particularly when benchmarked against revenue figures for the services. Our pilot studies, discussed in our forthcoming paper, also suggested that the actual valuations are not well anchored, but the methodology can give consistent rankings among goods.

It is also a useful way to assess changes in valuations. During the lockdown, we observed rapid changes in the contributions different goods and services make to consumer welfare, with some significant differences by age group and gender. In this sense the lockdown has acted as a natural experiment testing for the extent to which digital goods and physical goods are substitutes. As many of the goods we considered are free to use, these valuation changes give useful insights into economic welfare and activity that are not captured by changes in prices. They act as a forward-look at which services are most valuable in a post-pandemic world where more activity takes place online, compared with the recent past. They also provide important, policy-relevant insights into distributional questions as between men and women and different age and socio-economic groups.

References

Brynjolfsson, E., Collis, A., and Eggers, F. (2019a) Using massive online choice experiments to measure changes in well-being, Proceedings of the National Academy of Sciences 116 (15): 7250-7255.

Brynjolfsson, E., Collis, A., Diewert, W. E., Eggers, F., and Fox, K. J. (2019b) *GDP-B: Accounting for the Value of New and Free Goods in the Digital Economy*. NBER Working Paper No. 25695, March 2019.

Carson, R.T., Flores, N.E. & Meade, N.F. (2001) Contingent Valuation: Controversies and Evidence. *Environmental and Resource Economics* 19, 173–210.

Corrigan JR, Alhabash S, Rousu M, Cash SB (2018) How much is social media worth? Estimating the value of Facebook by paying users to stop using it. PLOS ONE 13(12): e0207101.

Heys, R. Martin, J. and Mkandawire, W. (2019) *GDP and Welfare: A Spectrum of Opportunity*, ESCoE Discussion Paper 2019-16.

McFadden, D., & Train, K. (2017) *Contingent Valuation of Environmental Goods: A Comprehensive Critique*, Edward Elgar.

Ofcom (2019) *Online Nation*, May 2019.

Sunstein, C. (2019). Valuing Facebook. *Behavioural Public Policy,* 1-12.

CEPR PRESS

Covid Economics 33, 30 June 2020: 110-123

Appendix

A1. Description of survey sample

The average age in our sample is 49 years of which just under 49% are male and 51% female. Around 30% hold a degree, 16% A-Levels, 14% GCSEs and 7% report no formal qualification (around one third of respondents do not report their highest degree). The median annual household income (of those reporting their income) is between £25-35k and hence the official ONS estimate of £29.6k for 2019 falls just in the middle.[9] The regional distribution of our survey sample also closely matches the latest subnational population estimates by the ONS from mid-2019 (see Table A1).[10] Finally, 60% of respondents took the survey on a phone or tablet while the remaining 40% used a laptop or desktop computer.

Table A1. Regional breakdown of sample and official population estimates

Region	Our sample	ONS estimates
North East	4.1%	4.2%
North West	11.4%	11.5%
Yorkshire and the Humber	8.6%	8.6%
East Midlands	7.8%	7.6%
West Midlands	8.8%	9.3%
East of England	9.5%	9.8%
London	12.0%	14.1%
South East	14.1%	14.4%
South West	10.0%	8.8%
Wales	5.1%	5.0%
Scotland	8.6%	8.6%

A2. The UK's 'online' population.

Following Ofcom (2019) around 13% of adults in the UK reported in 2019 that they do not use the internet. This share is higher for those aged 55-64 (19%), 65-74 (33%), and 75+ (48%). Further, 14% of people in social grades D (i.e. semi-skilled and unskilled manual workers) and E (i.e. state pensioners, casual and lowest grade workers, unemployed with state benefits only) do not use the internet, as compared to 4% of the rest. 47% of those not using the internet report that do not need it and 12% say that someone else is going online for them. Eurostat figures for 2019 show that 96% of people in the UK have used the internet in the past 3 months.[11] This share decreases to 85% for those aged 65-74.

[9] Average household income, UK: financial year ending 2019, Office for National Statistics: https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/householddisposableincomeandinequality/financialyearending2019

[10] We use data from the following ONS population released in May 2020 and recalculate the regional distribution by excluding Wales as our sample only covers Great Britain. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2019#population-age-structures-of-uk-countries-and-english-regions

[11] Eurostat, *ICT usage in households and by individuals (isoc_i), 2019.*

Covid Economics Issue 33, 30 June 2020

Investor rewards to environmental responsibility: Evidence from the COVID-19 crisis[1]

Alexandre Garel[2] and Arthur Petit-Romec[3]

Date submitted: 23 June 2020; Date accepted: 23 June 2020

The COVID-19 shock and its unprecedented financial consequences have brought about vast uncertainty concerning the future of climate actions. We study the cross-section of stock returns during the COVID-19 shock to explore investors' views and expectations about environmental issues. The results show that firms with responsible strategies on environmental issues experience better stock returns. This effect is mainly driven by initiatives addressing climate change (e.g., reduction of environmental emissions and energy use), is more pronounced for firms with greater ownership by investors with long-term orientation and is not observed prior to the COVID-19 crisis. Overall, the results indicate that the COVID-19 shock has not distracted investors' attention away from environmental issues but on the contrary led them to reward climate responsibility to a larger extent.

1 We thank Thomas Bourveau, François Derrien, Gunther Capelle-Blancard, Alberta Di Giuli, José Martin-Flores, and David Stolin for helpful discussions.

2 Audencia Business School.
3 SKEMA Business School, Université Côte d'Azur.

Copyright: Alexandre Garel and Arthur Petit-Romec

Covid Economics 33, 30 June 2020: 124-162

CEPR PRESS
COVID ECONOMICS
VETTED AND REAL-TIME PAPERS

1. Introduction

Climate change and the large financial risks it imposes on companies represent a key economic challenge (e.g., Carney 2015; Litterman 2017; Nordhaus 2019). Survey evidence suggests that investors believe that climate risks have financial implications for their portfolio firms and increasingly call for climate risk reporting (e.g., Ilhan, Krueger, Sautner, and Starks 2019; Krueger, Sautner, and Starks 2020). While investors, corporate leaders, policy makers, and other stakeholders were focusing on environmental issues, the COVID-19 crisis came as a complete surprise since pandemic risk was not receiving a lot of attention.[1] Ramelli and Wagner (2020) note that the five risks listed as being most likely to materialize in the World Economic Forum's Global Risk Report 2020 concern environmental issues whereas infectious diseases were ranked 10 and quite unlikely.

Has the COVID-19 outbreak distracted investors' attention away from environmental issues, which may have suddenly been perceived as less urgent? On the contrary, has the COVID-19 outbreak reinforced the value that investors attach to responsible initiatives on environmental and climate issues? We provide insights on these questions by studying the cross-section of stock price reactions to the COVID-19 shock.[2] Specifically, using data from Thomson Reuters Asset4 ESG database for a sample of large U.S. listed companies, we explore whether a firm's environmental score affects its stock price reaction to the COVID-19 shock. The environmental score measures a firm's commitment and effectiveness toward adopting responsible initiatives and strategies on environmental issues including the reduction of environmental emissions (e.g., greenhouse gases, ozone-depleting substances) and the efficient use of natural resources in the production process.

As highlighted by Ramelli and Wagner (2020), the cross-section of returns to the COVID-19 shock provides an opportunity to observe how investors responded to the crisis and what they expect for

[1] Schoenfeld (2020) finds that managers systematically underestimated their exposure to pandemics in their SEC-mandated risk factors. Loughran and McDonald (2020) document that about 80% of companies had no pandemic risk disclosure in 2018.
[2] Although the COVID-19 shock led to one of the fastest declines in market valuations with the S&P index taking 16 days in February 2020 to fall by 20%, there is substantial cross-sectional variation in stock price reactions to the COVID-19 shock.

Covid Economics 33, 30 June 2020: 124-162

CEPR PRESS

Covid Economics 33, 30 June 2020: 124-162

the future. If, as a result of the COVID-19 shock, investors update their perceptions about climate risk and consider that environmental efforts and investments in climate responsibility have been done in vain, companies with responsible initiatives on environmental issues may end up losing more value. By the same token, if investors anticipate that the COVID-19 shock may weaken or delay environmental regulations and climate actions, companies with responsible strategies on climate change may experience a stronger loss of value.[3] On the contrary, anecdotal evidence suggests that the COVID-19 has actually strengthened the need for climate actions.[4] If investors consider that environmental and climate responsibility will become more important following the COVID-19 shock, firms that had already adopted responsible initiatives on environmental issues should have an edge and experience better returns. Whether a firm's environmental score affects its stock price reaction to the COVID-19 shock is therefore an empirical question.

We find that firms with good environmental scores perform significantly better during the COVID-19 crisis (i.e., the period from February 20 to March 20).[5] The results are robust to controlling for industry fixed effects and various firm characteristics known to be related to stock returns. The effect of the environmental score is economically sizeable. A one-standard deviation higher environmental score is associated with stock returns that 1.4 percentage points during the COVID-19 shock. The economic significance of the environmental score in explaining the cross-section of returns is almost of the same order of magnitude as the economic significance of cash holdings or long-term debt. These results are inconsistent with the COVID-19 shock distracting investors' attention away from environmental and

[3] Several global business sectors have called for suspension of environmental protections (https://www.theguardian.com/environment/2020/apr/17/polluter-bailouts-and-lobbying-during-covid-19-pandemic)

[4] The front page of the Economist states that the COVID-19 shock creates a chance to flatten the climate curve (https://www.economist.com/leaders/2020/05/21/countries-should-seize-the-moment-to-flatten-the-climate-curve). Related anecdotal evidence indicates that despite the COVID-19 crisis, there is a surge in climate actions and rebellions in 2020 (https://www.ft.com/content/c10056af-306f-4d9d-8e97-5ffa112ddf49).

[5] Although it is difficult to set a clear starting date for the COVID-19 crisis, we retain the date of February 20 because we observe graphically that a prolonged decline in major U.S. indices (i.e., Russell, S&P, and Dow Jones) started on February 20 (see Figure 1).

climate issues. On the contrary, they indicate that the COVID-19 shock leads investors to reward companies with responsible initiatives on environmental issues.

Next, we explore whether our results reflect investors penalizing firms with poor environmental performance or investors rewarding firms with superior environmental performance. Specifically, we assess whether the results are more pronounced at very high or very low levels of environmental score. We find that the difference in returns between firms in the best and worst quartile in terms of environmental score is statistically significant and equal to about 3.7 percentage points. On the contrary, we find that the differences in returns between firms in the second or third quartile and worst quartile are positive but not statistically different. Our results therefore indicate that investors rewarded firms with superior environmental score during the COVID-19 crisis.

It could be that investors rewarded firms with corporate social responsibility (CSR) in general beyond environmental responsibility. Lins, Servaes, and Tamayo (2017) show firms with high CSR intensity experience better stock returns during the 2008-2009 financial crisis, consistent with the trust between a firm and its stakeholders, built through investments in CSR, paying off when the level of trust in corporations suffers a negative shock. The causes and scope of COVID-19 crisis are fundamentally different from the ones of 2008 global financial crisis (GFC).[6] In particular, the COVID-19 crisis starts outside the financial sector and was not produced by excessive risk-taking. Given its unprecedented nature, it is not clear whether the COVID-19 crisis represents a negative shock in the level of trust in corporations and whether all investments in CSR would pay off. We find that responsible strategies on social issues have a positive, though not statistically significant, association stock returns during the COVID-19 crisis. These results highlight that investors specifically reward environmental responsibility are consistent with investor awareness on environmental and climate issues being much more pronounced today than it was in 2008.[7]

[6] https://www.project-syndicate.org/commentary/covid19-crisis-has-no-economic-precedent-by-carmen-reinhart-2020-03

[7] Albuquerque et al. (2019, 2020) develop a model predicting that by increasing product differentiation and customer loyalty, CSR investments decrease systemic risk and strengthen firm resilience in shocks and provide empirical evidence supporting their predictions. Our results show that only environmental responsibility and not social responsibility is associated with higher stock returns in the COVID-19 crisis. This may be due to the fact that customer loyalty plays a

We then examine whether specific components of the environmental score are more important for stock returns. Given the emphasis placed on climate change and, in particular, on concerns over global warming linked to CO2 emissions, we expect that responsible initiatives specifically addressing firms' climate performance are more likely to be rewarded by investors. Consistent with this prediction, we find that the subcomponent related to the reduction of environmental emissions strongly affects stock returns during the COVID-19 shock. This result complements prior studies documenting that carbon emissions increase downside risk (e.g., Ilhan et al. 2019). The other subcomponent that mainly influences stock returns during the COVID-19 shock is a firm's resource use, which also captures, to some extent, initiatives addressing climate responsibility (e.g., water use and energy use in the supply chain). The decomposition of the environmental score therefore indicates that investors have mainly rewarded initiatives that specifically address climate responsibility.

Investors differ in the importance they place on environmental responsibility. It is therefore unlikely that all investors will reward environmental responsibility to the same extent. Prior evidence suggests that investors with long-term orientation are more concerned with environmental responsibility (e.g., Gibson and Krueger 2018; Ramelli et al. 2018). We therefore explore whether responsible strategies on environmental issues are rewarded to a larger extent when investors have a long-term orientation. Using two standard measures of investor horizons (based on investors' portfolio turnover or on long-term orientation associated with specific investor types) at the firm level, we find that the positive effect of the environmental score on stock price reactions to the COVID-19 crisis is significantly more pronounced for firms with greater long-term investor ownership. This result is consistent with the evidence in Ramelli et al. (2018) that although Donald Trump's election and his nomination of Scott Pruitt, a climate skeptic, to lead the Environmental Protection Agency drastically downshifted expectations on U.S. climate-change policy, long-term investors actually rewarded climate-responsible companies.

less important role in the COVID-19 crisis that has far-more dramatic immediate real effects than previous crises as many firms experience a sudden stop in their ability to produce and customers are no longer able to consume.

The evidence that firms with responsible initiatives on environmental issues experience better stock returns during the COVID-19 crisis indicates that investors expect them to do better in the long-run. It could however be that investors only value environmental responsibility to the extent that it allows firms to generate short-term profits or to access financial resources. We provide two additional tests to assess this possibility. First, we do not find evidence that environmental score is related to financial analysts' forecast revisions of one-year ahead earnings per share (EPS) during the COVID-19 crisis. Second, if firms with high environmental score had better returns during the COVID-19 crisis because they have more financial flexibility or better access to finance, one would expect those firms to benefit less from the news concerning stimulus package and policy responses to the COVID-19 crisis on March 24. We do not find evidence that it is the case.

Overall, our analyses indicate that although the COVID-19 shock and its severe consequences took virtually all investors and companies by surprise, those with high environmental score were substantially less affected. Investors rewarded environmental responsibility such as the reduction of environmental emissions during the COVID-19 crisis. These results suggest that the unprecedented and novel risk posed by the COVID-19 crisis has not led investors to distract their attention away from environmental issues.

Our paper contributes to different streams of the literature. First, our results are related to a stream of research analyzing stock market reaction to different events related to environmental issues. Prior studies have investigated sock market reaction to industrial disasters (Capelle-Blancard and Laguna 2010), the inclusion in a sustainability stock index (Oberndorfer et al. 2013), the disclosure of positive environmental information (Wang et al. 2019), the announcement of corporate news related to environment (Flammer 2013), the announcement of membership in a program targeting reductions in greenhouse gas emissions (Fisher-Vanden and Thorburn 2011), Donald Trump's election and his nomination of Scott Pruitt, a climate skeptic, to lead the Environmental Protection Agency (Ramelli et al. 2018), the Paris agreement (Monasterolo and De

Angelis 2020), global climate strikes (Ramelli et al. 2020), and extreme temperatures (Choi et al. 2020). Although the COVID-19 crisis was not directly caused by the materialization of environmental risks, it has brought about substantial questions regarding the future of environmental actions. From this perspective, studying the stock market reaction of firms with greater environmental responsibility is informative about the views and anticipations of investors regarding the value attached to environmental responsibility. The reward to environmental responsibility during the COVID-19 shock indicates that investors expect that environmental responsibility will be value-enhancing in the future.

Second, our results are related to prior studies analyzing the links between environmental responsibility and financial performance at the investors' portfolio or firm level. Prior studies show that decarbonization or more broadly the integration of environmental criteria does not impair portfolio performance (e.g., Trinks, Scholtens, Mulder, and Dam 2018; Boermars and Galema 2019). At the firm level, prior studies examine the links between environmental performance and financial performance or risk (e.g., Lioui and Sharma 2012; Lyons et al. 2013; Gonenc and Scholtens 2017; Ilhan et al. 2019). Other studies find that environmental performance and green investing affects the cost of capital (e.g., Chava 2014; Sharfman and Fernando 2008; De Angelis, Tankov, and Zerbib 2020). We expand these studies by examining the association between environmental responsibility and stock performance in the COVID-19 shock that was largely unexpected and hit firms with unprecedented force. Our results indicate that the COVID-19 shock has led investors to revise the value attached to environmental responsibility and may mark a turning point in the relationship between environmental responsibility and financial performance.

Third, our paper contributes to the burgeoning literature on the COVID-19 shock and its consequences on firms. Fahlenbrach et al. (2020) and Ramelli and Wagner (2020) show a negative (positive) relation between stock returns during the COVID-19 shock and leverage (cash holdings). Albuquerque et al. (2020) show that firms with high social capital have better stock returns during the COVID-19 crisis, confirming the results from Lins, Servaes, and Tamayo (2017) for the 2008

Covid Economics 33, 30 June 2020: 124-162

financial crisis. Using an international sample, Ding et al. (2020) examine how different firm characteristics (financial flexibility, global supply chains, corporate social responsibility, and corporate governance) accentuate or mitigate the effect of the number of COVID-19 cases on weekly stock returns. Using stress-tests under different scenarios of revenue shortfalls, De Vito and Gómez (2020) investigate how the COVID-19 shock could affect the liquidity of listed firms. Other papers provide evidence that managers systematically underestimated their exposure to pandemics and that the large majority of companies had no pandemic risk disclosure (Loughran and McDonald 2020; Schoenfeld 2020).[8]

The remainder of the paper is organized as follows. The next section presents the sample construction and the main variables used in the empirical analysis. Section 3 presents the results and robustness tests. We conclude in Section 4.

2. Data, Sample and Summary Statistics

2.1 Data and Sample Construction

We obtain data on environmental responsibility from Thomson Reuters Asset4 ESG database, which offers one of the most comprehensive ESG databases covering over 70% of global market capitalization, across more than 450 different ESG metrics, with history going back to 2002. Thomson Reuters collects information from a wide variety of data sources (e.g., corporate annual reports, stock exchange filings, corporate socially responsible reports, non-profit organizations, the news media) to produce three pillar scores: Environmental, Social, and Corporate Governance. The pillar scores are based on a large number of individual indicators capturing different aspects of environmental or social responsibility. Many recent studies rely on Asset4 ESG database because

[8] Schoenfeld (2020) finds that an indicator variable for whether that firm includes pandemics as a risk factor in their risk disclosure is not associated with stock returns during the COVID-19 crisis. Our results are therefore unlikely to be driven by firms with high environmental score being more concerned about pandemic risk.

Covid Economics 33, 30 June 2020: 124-162

Covid Economics 33, 30 June 2020: 124-162

it has significantly increased its coverage and contrary to MSCI does not face a major structural break in the series (e.g., Cheng et al. 2014; Dyck et al. 2019; Gonenc and Scholtens 2017).

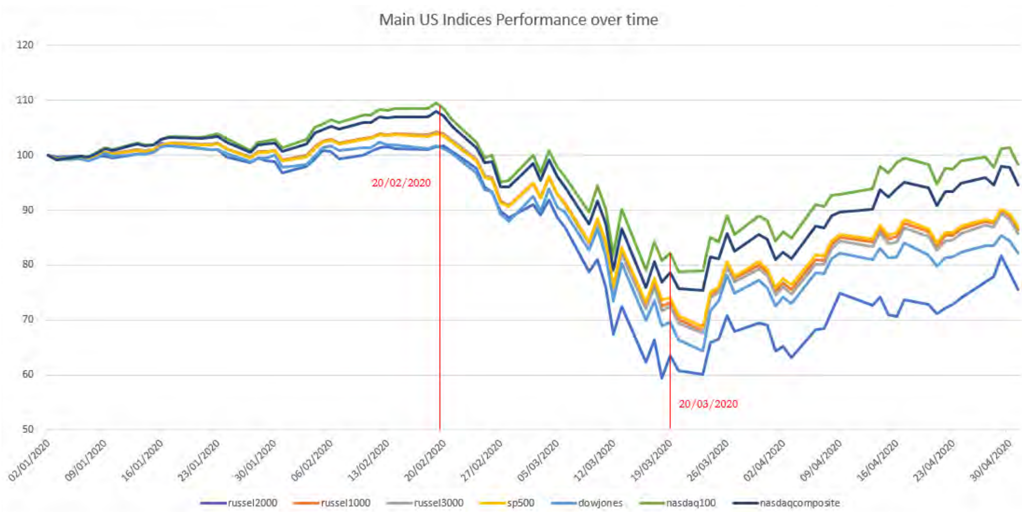
We obtain stock market data and accounting data from Thomson Reuters database (EIKON). Accounting data are based on the last quarter ending at or before the end of 2019. We drop financial firms, utilities, not for profit and governmental firms, and firms with non-U.S. headquarters. We also remove stocks with prices of less than $1 (Fahlenbrach et al. 2020). Finally, we drop firms for which information on key control variables, which are described in the next section, is missing. These restrictions result in a final sample of 1,626 large U.S. listed firms for which all key variables are available for the COVID-19 crisis period.[9]

2.2 Main Variables

Our main dependent variable is the buy-and-hold stock return during the COVID-19 crisis. Although it is difficult to set a clear starting date for the COVID-19 crisis, we initially focus on the period from February 20 to March 20 based on what we observe for the time series of index returns. Figure 1 shows the returns of major U.S. indices (i.e., Russell, S&P, and Dow Jones) since January 2020 and indicates that a prolonged decline in major indices started on February 20. This period is very close to the *fever* period (i.e., February 24 to March 20) used in Ramelli and Wagner (2020) and starting after the strict lockdown that Italy imposed in Lombardy. In robustness tests, we show that our main results are unchanged if we use the *fever* period or the period from February 3 to March 23, which corresponds to the collapse period in Fahlenbrach et al. (2020).

[9] In the online appendix, we provide the list of companies' Thomson identifiers included in our analysis to ease comparability and replication.

Figure 1: U.S. stock market performance in the COVID crisis



Main U.S. indices over 01/2020-04/2020, based at 100 as of the 01/01/2020.

Our main independent variable is a firm's environmental score measured as of 2018, the last available year prior to the COVID-19 outbreak. The environmental score is composed of information on three subcomponents: i) Resource Use, ii) Emission Reduction, and iii) Green Innovation. The resource use subcomponent reflects a company's performance and capacity to reduce the use of materials, energy or water, and to find more eco-efficient solutions by improving supply chain management. The Emission Reduction subcomponent measures a company's commitment and effectiveness towards reducing environmental emissions in the production and operational processes. The Green Innovation subcomponent reflects a company's capacity to reduce the environmental costs and burdens for its customers, thereby creating new market opportunities through new environmental technologies and processes or eco-designed products.[10] A key feature of the Asset4 ESG's scoring methodology is that the scores are based on relative performance with the company's sector.[11] Moreover, the scoring methodology accounts for the fact that the importance of ESG factors differ industries by using ESG magnitude

[10]These three subcomponents aggregate data from numerous individual indicators: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/esg-scores-methodology.pdf
[11] Asset4 ESG's scoring methodology defines a company's sector based on TRBC industry groups.

Covid Economics 33, 30 June 2020: 124-162

Covid Economics 33, 30 June 2020: 124-162

(materiality). In our empirical analysis, we use either the overall environmental score or the three subcomponents.

Our set of control variables include three proxies for financial flexibility: cash over assets, short-term debt over assets, and long-term debt over assets. Fahlenbrach, Rageth, and Stulz (2020) and Ramelli and Wagner (2020) show that companies with higher financial flexibility (i.e., more cash and less debt) did significantly better during the COVID-19 crisis. We further control for firm characteristics known to be related to stock returns. Specifically, we control for the firm's factor loadings based on the Fama-French three-factor model, momentum (the firm's raw return over the period November 2019 to January 2020), firm size (the logarithm of total assets), and profitability (the ratio of EBITDA to total assets).

2.3 Summary Statistics

Table 1, Panel A provides descriptive statistics for our main variables. All continuous variables are winsorized at the 1st and 99th percentiles. Unsurprisingly, returns during the COVID-19 crisis are negative and large with a mean of -40.5% and a median of -39.7%. There is however substantial cross-sectional variation in stock price reactions to the COVID-19 crisis as the interquartile difference is 23.3% and the standard deviation of stock returns is 18.11%. These figures are in line with other studies investigating the determinants of the cross-section of returns during the COVID-19 crisis (e.g., Fahlenbrach et al. 2020; Ramelli and Wagner 2020). The environmental score has a mean value of 20.91 with a standard deviation of 25.7, indicating substantial cross-sectional variation. A significant portion of companies has an environmental score of 0, indicating that several firms have not taken any responsible initiatives on environmental issues. A closer look indicates that most companies with environmental score of 0 are concentrated in industries in which reporting on environmental issues remains scarce (e.g., Healthcare Equipments & Providers, Software & IT).[12] The average firm in our sample has total assets of $8.85 billion, an equity beta of 1.12, a ratio of long-term debt to total assets of 26%, a profitability of 3%, and a ratio of cash to total assets of 22%.[13]

Table 1. Descriptive statistics

	Summary Statistics					
Variables	#Obs.	Mean	SD.	P25	P50	P75
COVID-19 Crisis Return (%)	1,626	-40.55	18.11	-52.43	-39.69	-29.16
Environmental Score	1,626	20.91	25.73	0.00	8.57	35.48
Beta MKT	1,626	1.12	0.62	0.73	1.08	1.47
Beta HML	1,626	0.81	0.96	0.20	0.67	1.31
Beta SMB	1,626	0.07	0.98	-0.44	0.10	0.59
Total Assets ($Billion)	1,626	8.85	30.25	0.59	1.69	5.37
Long-Term Debt	1,626	0.26	0.22	0.06	0.23	0.39
Short-Term Debt	1,626	0.01	0.02	0.00	0.00	0.00
Profitability	1,626	0.03	0.19	0.02	0.07	0.11
Cash Holdings	1,626	0.22	0.26	0.03	0.10	0.29
Momentum	1,626	5.65	20.83	-5.58	4.42	13.65

This table presents descriptive statistics for the main variables used in the empirical analysis. The sample consists of 1,626 firms. *Environmental Score* is the Thomson Reuters' environmental score. It aggregates information on a company's performance and capacity to reduce the use of materials, energy or water, and to find more eco-efficient solutions by improving supply chain management (*Resource Use Score*), a company's commitment and effectiveness towards reducing environmental emission in the production and operational processes (*Emission Reduction Score*), and a company's capacity to reduce the environmental costs and burdens for its customers, thereby creating new market opportunities through new environmental technologies and processes or eco-designed products (*Green Innovation Score*). *COVID-19 Crisis Return* is the buy-and-hold return computed over the period February 20-March 20. Accounting data are based on the last quarter ending at or before the end of 2019. *Beta MKT*, *Beta HML*, and *Beta SMB* are the factor loadings of the three-factor Fama-French model estimated in the 60 months prior to the crisis (we require at least 12 months of data available). *Total Assets* is in billions of dollars. *Long-Term Debt* is computed as long-term debt divided by assets. *Short-Term Debt* is computed as debt in current liabilities divided by assets. *Cash Holdings* is computed as cash and marketable securities divided by assets. *Profitability* is computed as operating income divided by assets. *Momentum* is the buy-and-hold return over the period November 2019 – January 2020. All variables are winsorized at the 1st and 99th percentile.

3. Results

3.1 Environmental score and stock returns during the COVID-19 crisis

We start by providing graphical evidence on the relationship between environmental scores and stock returns during the COVID-19 shock. Figure 2 shows the average industry-adjusted

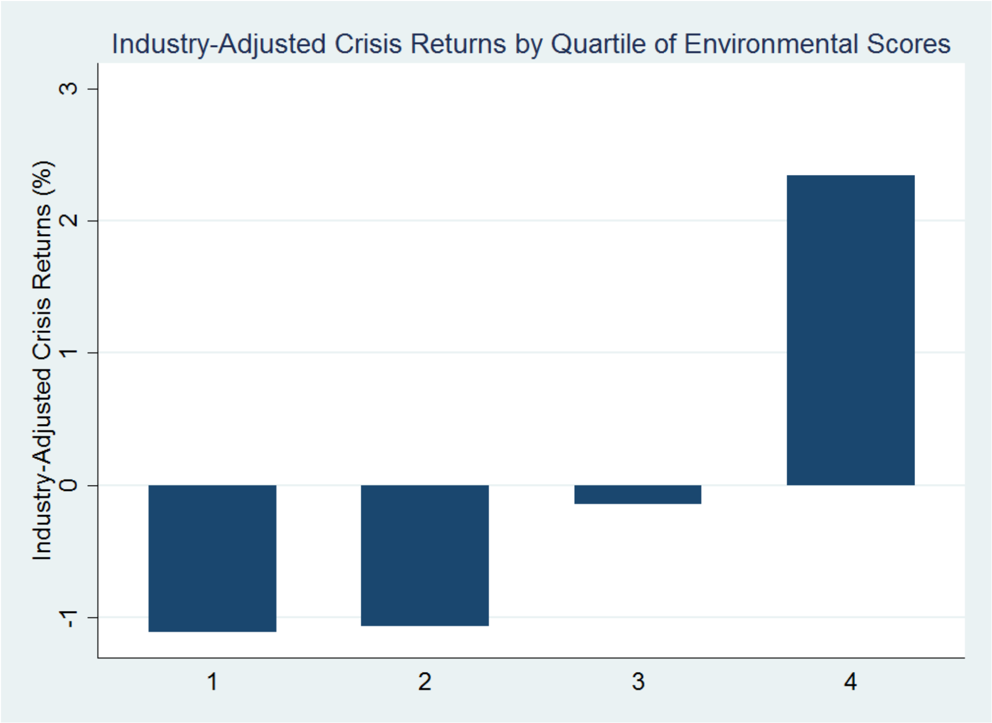
[12] In unreported tests, we find that all results hold if we exclude firms with an environmental score equal to zero.
[13] Correlations among the variables we use in our regressions are reported in the Appendix.

Covid Economics 33, 30 June 2020: 124-162

CEPR PRESS

returns during the COVID-19 shocks by quartiles of environmental scores. It suggests that firms with high environmental scores performed better during the crisis.

Figure 2: Industry-Adjusted Crisis Returns by Quartile of Environmental Score



This figure shows average crisis returns by quartile of environmental scores. Crisis returns are computed over the February-20 to March-20 period. Returns are computed in excess of companies' industry returns (TRBC industry groups).

We then formally investigate whether environmental score explains the cross-section of returns to the COVID-19 shock. In Table 2, we estimate regressions of stock returns during the COVID-19 crisis (i.e., over the period February 20-March 20) on pre-crisis environmental score and control variables. All regressions include industry fixed effects[14] to account for the fact that firms in some industries may have been differentially affected by the COVID-19 crisis. Columns 1 and 2 show that, both with

[14] We include industry fixed effects based on TRBC industry groups because they are the ones used by Asset4 to calculate the environmental score of companies relative to their industry peers. In robustness tests, we check that our results hold when we use industry fixed effects based on other industry classifications.

CEPR PRESS

Covid Economics 33, 30 June 2020: 124-162

and without controls, the coefficient on *environmental score* is positive and statistically significant at the 1% level, indicating that companies with responsible strategies on environmental issues experience better returns during the COVID-19 shock. The effect of environmental score is economically sizeable. According to Column 2, a one-standard deviation higher environmental score is associated with a 1.41% (=25.7×0.055) higher stock returns during the COVID-19 crisis. By comparison, a one-standard-deviation increase in cash holdings (long-term debt) is associated with a 2.7% (1.9%) higher returns. The economic significance of the environmental score in explaining the cross-section of returns represents respectively 52% of the economic significance of cash holdings and 75% of the economic significance of long-term debt, which are two well-established determinants of the cross-section of returns during the COVID-19 crisis (Fahlenbrach et al. 2020; Ramelli and Wagner 2020).

In Columns 3 and 4, we reproduce the regression of Column 2 using two alternative dependent variables: stock returns over the collapse period (i.e., February 3 to March 23) used by Fahlenbrach et al. (2020) and the fever period (i.e., February 24 to March 20) used by Ramelli and Wagner (2020). The results show that the environmental score has a positive and significant effect on stock returns over these two alternative periods, indicating that our results are not driven by our choice to use February 20, the day in which major stock indices started to decline, as the start of the COVID-19 shock. Overall, these results are inconsistent with the COVID-19 shock distracting investors' attention away from environmental issues.

Covid Economics 33, 30 June 2020: 124-162

Table 2. Environmental score and stock returns during the COVID-19 crisis

Panel A: Environmental score and stock returns in the COVID-19 crisis

	(1) COVID-19 crisis returns (February 20-March 20)	(2) COVID-19 crisis returns (February 20-March 20)	(3) Fahlenbrach et al.'s (2020) collapse period (February 3-March 23)	(4) Ramelli and Wagner's (2020) fever period (February 24-March 20)
Environmental Score	0.068***	0.055***	0.045**	0.053***
	(0.016)	(0.019)	(0.020)	(0.019)
Beta MKT		-3.788***	-3.535***	-3.231***
		(0.755)	(0.808)	(0.755)
Beta HML		-2.175***	-1.631***	-2.090***
		(0.565)	(0.611)	(0.563)
Beta SMB		-1.163**	-1.310**	-1.241**
		(0.590)	(0.640)	(0.581)
Ln(Total Assets)		-0.158	-0.158	-0.014
		(0.385)	(0.401)	(0.382)
Long-Term Debt		-8.630***	-9.565***	-9.454***
		(2.471)	(2.508)	(2.491)
Short-Term Debt		-22.469	-23.489	-27.123*
		(14.677)	(14.803)	(14.995)
Cash Holdings		10.401***	13.181***	12.131***
		(2.780)	(2.896)	(2.696)
Profitability		19.373***	19.487***	20.788***
		(3.549)	(3.703)	(3.430)
Momentum		-0.018	0.011	-0.013
		(0.025)	(0.025)	(0.024)
Observations	1,626	1,626	1,626	1,626
Industry Dummies	Yes	Yes	Yes	Yes
Adjusted R-squared	0.195	0.266	0.276	0.262

Panel A presents regression estimates of stock returns during the COVID-19 crisis on environmental score and control variables. We use different definitions of the COVID-19 crisis period. In columns 1 and 2, we use the February-20 to March-20 period (based on the observation of stock market index returns in Figure 1). In columns 3 and 4, we use the crisis period definitions of Fahlenbrach et al. (2020) and Ramelli and Wagner (2020), respectively. Variables are winsorized at the 1st and 99th percentiles. All regressions include industry fixed effects based on TRBC industry group classifications. Variables are defined in Appendix A. Constants are not reported. Heteroskedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

Panel B: Environmental score quartiles and stock returns in the COVID-19 crisis				
	(1) COVID-19 crisis returns (February 20-March 20)	(2) COVID-19 crisis returns (February 20-March 20)	(3) Fahlenbrach et al.'s (2020) collapse period (February 3-March 23)	(4) Ramelli and Wagner's (2020) fever period (February 24-March 20)
Env. Score 2	0.727	0.713	0.006	0.782
	(1.470)	(1.408)	(1.433)	(1.434)
Env. Score 3	1.635	1.598	1.245	1.686
	(1.224)	(1.193)	(1.274)	(1.210)
Env. Score 4	4.306***	3.313**	2.580*	3.278**
	(1.171)	(1.332)	(1.409)	(1.335)
Beta MKT		-3.829***	-3.561***	-3.272***
		(0.755)	(0.809)	(0.756)
Beta HML		-2.191***	-1.642***	-2.105***
		(0.566)	(0.612)	(0.564)
Beta SMB		-1.157*	-1.306**	-1.234**
		(0.590)	(0.641)	(0.581)
Ln(Total Assets)		-0.047	-0.072	0.092
		(0.370)	(0.386)	(0.367)
Long-Term Debt		-8.614***	-9.599***	-9.432***
		(2.494)	(2.528)	(2.515)
Short-Term Debt		-20.908	-22.408	-25.509*
		(14.757)	(14.858)	(15.063)
Cash Holdings		10.818***	13.399***	12.583***
		(2.787)	(2.920)	(2.715)
Profitability		19.250***	19.433***	20.658***
		(3.559)	(3.712)	(3.443)
Momentum		-0.018	0.011	-0.013
		(0.025)	(0.025)	(0.024)
Observations	1,626	1,626	1,626	1,626
Industry Dummies	Yes	Yes	Yes	Yes
Adjusted R-squared	0.193	0.264	0.274	0.261

Panel B presents regression estimates of stock returns during the COVID-19 crisis on dummy variables denoting whether a firm's environmental score is in the second, third, or top quartile and control variables. We use different definitions of the COVID-19 crisis period. In columns 1 and 2, we use the February-20 to March-20 period (based on the observation of stock market index returns in Figure 1). In columns 3 and 4, we use the crisis period definitions of Fahlenbrach et al. (2020) and Ramelli and Wagner (2020), respectively. Variables are winsorized at the 1st and 99th percentiles. All regressions include industry fixed effects based on TRBC industry group classifications. Variables are defined in Appendix A. Constants are not reported. Heteroskedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

Our results from Table 2, Panel A are equally consistent with investors penalizing companies with poor environmental performance and with investors rewarding firms with superior

Covid Economics 33, 30 June 2020: 124-162

environmental performance during the COVID-19 crisis. In Panel B, therefore, we analyze whether there are asymmetries in the relation between environmental score and returns during the COVID-19 crisis. We split firms into quartiles based on their environmental score and create indicator variables for each of the four groups. *Env score 1* contains all observations whose environmental score is among the 25% lowest. Panel B reports results of regressions in which we replace the environmental score with the quartile indicators. The omitted group is *Env score 1,* the quartile of firms that have the lowest environmental scores. The results from Column 1, which does not include control variables, show that only the coefficient on *Env score 4* is statistically significant, and it is much larger than the other coefficients. The difference in returns during the COVID-19 crisis between firms in the best and worst quartile in terms of environmental score is statistically significant and equal to about 4.3 percentage points. Adding control variables attenuates the effect to a certain extent. However, in Column 2, which includes the same control variables as in Panel A, the coefficient on *Env score 4* is still statistically significant at the 1% level and equal to 3.3. In Columns 3 and 4, we reproduce the same analysis for stock returns computed over the collapse period (i.e., February 3 to March 23) used by Fahlenbrach et al. (2020) and the fever period (i.e., February 24 to March 20) used by Ramelli and Wagner (2020). In both cases, the results show that firms with superior environmental score experience higher stock returns during the COVID-19 crisis.

3.2 Environmental subcomponents, social responsibility, and corporate governance

So far, our analysis has focused on environmental responsibility because we are interested in understanding whether the COVID-19 outbreak distracted investors' attention away from environmental issues. While our results indicate the COVID-19 shock has rather reinforced the extent to which investors reward responsible strategies on environmental issues, it could be that investors rewarded firms with socially responsible strategies in general beyond environmental issues. Lins, Servaes, and Tamayo (2017) show firms with high CSR intensity experience better stock returns during the 2008 global financial crisis. As they explain, their evidence suggests that the trust between a firm and its stakeholders, built through

Covid Economics 33, 30 June 2020: 124-162

investments in CSR, pays off when the overall level of trust in corporations suffers a negative shock. In Table 3, we reproduce our regressions from Table 2 replacing the environmental score by the social score. Column 2 shows that the coefficient on *social score* is positive but statistically not significant. In Column 3, we find similarly that the coefficient on *governance score* is positive but statistically not significant. The results from Column 4 show that the environmental score has a positive effect on stock returns during the COVID-19 shock after controlling for the social score and the governance score. Therefore, our results indicate that investors rewarded to a larger extent environmental responsibility than social responsibility, consistent with investor awareness on environmental and climate issues being much more pronounced today than it was in 2008. In particular, the COP 21 Paris agreement in 2015 played an important part in raising awareness on climate issues (Andersson, Bolton, Samama 2016).

Another non-mutually exclusive explanation for the fact that only environmental responsibility plays a role in the COVID-19 crisis is that, as stressed by Carmen Reinhart among others, the causes and scope of COVID-19 crisis are fundamentally different from the ones of the 2008 global financial crisis (GFC).[15] Contrary to the GFC, the COVID-19 crisis started outside the financial sector and was not produced by financial imbalances and risks growing over many years. Moreover, while the GFC affected firms through its impact on financial intermediaries and credit markets, the COVID-19 crisis has much more severe direct consequences as many firms experience a sudden stop in their ability to generate revenues. From this perspective, it is not clear whether the COVID-19 crisis represents a negative shock in the level of trust in corporations which investments in social capital have not paid off in this shock.

[15] Source: https://www.project-syndicate.org/commentary/covid19-crisis-has-no-economic-precedent-by-carmen-reinhart-2020-03

Table 3. ESG pillars and stock returns during the COVID-19 crisis

Crisis Return	(1)	(2)	(3)	(4)
Environmental Score	0.055***			0.060**
	(0.019)			(0.025)
Social Score		0.035		-0.011
		(0.024)		(0.032)
Governance Score			0.016	0.001
			(0.021)	(0.022)
Beta MKT	-3.788***	-3.842***	-3.874***	-3.789***
	(0.755)	(0.755)	(0.756)	(0.755)
Beta HML	-2.175***	-2.257***	-2.289***	-2.175***
	(0.565)	(0.562)	(0.563)	(0.566)
Beta SMB	-1.163**	-1.147*	-1.153*	-1.164**
	(0.590)	(0.590)	(0.592)	(0.591)
Ln(Total Assets)	-0.158	0.131	0.324	-0.132
	(0.385)	(0.395)	(0.357)	(0.410)
Long-Term Debt	-8.630***	-8.886***	-8.950***	-8.630***
	(2.471)	(2.473)	(2.481)	(2.478)
Short-Term Debt	-22.469	-21.617	-21.268	-22.459
	(14.677)	(14.578)	(14.662)	(14.714)
Cash Holdings	10.401***	10.763***	10.793***	10.363***
	(2.780)	(2.775)	(2.784)	(2.781)
Profitability	19.373***	19.273***	19.036***	19.367***
	(3.549)	(3.558)	(3.570)	(3.559)
Momentum	-0.018	-0.017	-0.017	-0.018
	(0.025)	(0.025)	(0.025)	(0.025)
Observations	1,626	1,626	1,626	1,626
Industry Dummies	Yes	Yes	Yes	Yes
Adjusted R-squared	0.266	0.263	0.263	0.265

This table presents regression estimates of stock returns during the COVID-19 crisis (i.e., February 20-March 20, 2020) on environmental score, social score, governance score, and control variables. All variables are winsorized at the 1st and 99th percentiles. All regressions include industry fixed effects based on TRBC industry group classifications. Variables are defined in Appendix A. Appendix B provides summary statistics on the social score and the governance score. Constants are not reported. Heteroskedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

Next, we examine whether specific components of the environmental score are more important for stock returns. Given the emphasis placed on climate change and the large financial risks it imposes on companies, we expect that responsible initiatives specifically addressing firms' climate performance are more likely to be rewarded by investors. As explained in section 2, the environmental score is composed of three subcomponents: i) Resource Use, ii) Emission Reduction, and iii) Green Innovation. Given that concerns over global warming linked to CO2 emissions have become salient, the Emission

Covid Economics 33, 30 June 2020: 124-162

Reduction is the subcomponent that most specifically addresses climate responsibility. To a lower extent, the Resource Use subcomponent may also capture initiatives addressing climate responsibility (e.g., water use and energy use in the supply chain).

In Table 4, we estimate regressions of stock returns during the COVID-19 crisis (i.e., over the period February 20-March 20) on the different subcomponents of the environmental score and control variables. The results from Column 1 show that the coefficient on *Emission Reduction* is positive and statistically significant at the 5% level, suggesting a firm's capacity to reduce the use of materials, energy or water, and to find more eco-efficient solutions by improving supply chain management is associated with better stock returns during the COVID-19 crisis. The results from Column 2 show that the coefficient on *Emission Reduction* is positive and statistically significant at the 1% level, suggesting that companies with responsible strategies to reduce emissions experience better stock returns during the COVID-19 crisis. This result is consistent with prior studies documenting that carbon emissions increase downside risk (Ilhan et al. 2019). Finally, the results in Column 3 show that the green innovation subcomponent is not statistically associated with stock returns during the COVID-19 crisis. The green innovation subcomponent captures a firm's ability to create new market opportunities through new environmental technologies or eco-designed, dematerialized products with extended durability. Contrary to the other two subcomponents, it does not capture a firm's efforts to limit its impact on the environment.

Overall, the results from Table 4 indicate that the effect of the environmental score on the stock returns during the COVID-19 crisis is mainly driven by initiatives that specifically address climate responsibility. These results further indicate that the COVID-19 outbreak has not distracted investors' attention away from environmental and climate issues and that, on the contrary they have started to reward environmental responsibility to a larger extent. This suggests that investors anticipate that issues related to climate change will be more important than ever with the COVID-19 crisis.

Covid Economics 33, 30 June 2020: 124-162

CEPR PRESS

Covid Economics 33, 30 June 2020: 124-162

Table 4. Subdimensions of environmental score stock returns during the COVID-19 crisis

	(1)	(2)	(3)
Resource Use Score	0.037**		
	(0.015)		
Emission Reduction Score		0.046***	
		(0.017)	
Green Innovation Score			0.012
			(0.016)
Beta MKT	-3.810***	-3.795***	-4.713***
	(0.754)	(0.753)	(0.850)
Beta HML	-2.209***	-2.185***	-2.803***
	(0.565)	(0.565)	(0.623)
Beta SMB	-1.166**	-1.175**	-1.036
	(0.591)	(0.590)	(0.680)
Ln(Total Assets)	-0.044	-0.116	0.260
	(0.379)	(0.388)	(0.348)
Long-Term Debt	-8.724***	-8.630***	-9.998***
	(2.472)	(2.468)	(2.449)
Short-Term Debt	-21.164	-22.019	-26.523*
	(14.697)	(14.568)	(14.565)
Cash Holdings	10.506***	10.344***	12.303***
	(2.783)	(2.782)	(2.953)
Profitability	19.403***	19.335***	21.061***
	(3.555)	(3.546)	(4.665)
Momentum	-0.017	-0.018	-0.013
	(0.025)	(0.025)	(0.030)
Observations	1,626	1,626	1,418
Industry Dummies	Yes	Yes	Yes
Adjusted R-squared	0.265	0.266	0.293

This table presents regression estimates of stock returns during the COVID-19 crisis (i.e., February 20-March 20, 2020) on the subdimensions of the environmental score (i.e., resource use, emission reduction, and green innovation) and control variables. All variables are winsorized at the 1st and 99th percentiles. All regressions include industry fixed effects based on TRBC industry group classifications. Variables are defined in Appendix A. Appendix B provides summary statistics on the subcomponents of the environmental score. Constants are not reported. Heteroskedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

3.3 Environmental score and stock returns before the COVID-19 crisis

Our main argument revolves around investors changing their perception of the value of environmental responsibly as a result of the COVID-19 crisis. We thus do not expect to find a similar surge in the relative value of companies with strong environmental scores in January 2020, when attention to the COVID-19 was still limited. To explore this issue, in Table 5, we examine whether environmental scores affect stock returns in January 2020 as well as during the periods from January 2 to January 17 and

CEPR PRESS

Covid Economics 33, 30 June 2020: 124-162

from January 20 to February 21, corresponding respectively to the *incubation* period and the *outbreak* period in Ramelli and Wagner (2020). During these periods, attention to the COVID-19 among U.S. firms was still very low. Ramelli and Wagner (2020) highlight that the first conference call discussing the COVID-19 took place on January 22 (i.e., the one of United Airlines).

Table 5. Environmental score and stock returns before the COVID-19 crisis

	(1) January 2020	(2) Ramelli and Wagner (2020): Incubation (02/01/2020 – 17/01/2020)	(3) Ramelli and Wagner (2020): Outbreak (20/01/2020 – 21/02/2020)
Environmental Score	0.007 (0.010)	0.008 (0.008)	0.010 (0.014)
Observations	1,626	1,626	1,626
Control Variables	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes
Adjusted R-squared	0.410	0.259	0.132

This table presents regression estimates of stock returns during the pre-COVID-19 crisis period on environmental score and control variables. In Column 1, the dependent variable is the stock return over January 2020. In Column 2, the dependent variable is the stock return over the period January 2 to January 17, which corresponds to the incubation period in Ramelli and Wagner (2020). In Column 3, the dependent variable is the stock return over the period January 20 to February 21, which corresponds to the outbreak period in Ramelli and Wagner (2020). All variables are winsorized at the 1st and 99th percentiles. All regressions include industry fixed effects based on TRBC industry group classifications. Variables are defined in Appendix A. Appendix B provides summary statistics for the stock returns in the pre-crisis periods. Constants are not reported. Heteroskedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

The results show that a firm's environmental score is not statistically associated with its stock returns during these three periods, suggesting that the COVID-19 shock has reinforced the extent to which investors reward companies with responsible strategies on environmental issues. Ramelli and Wagner's (2020) find that cash holdings and debt only started to affect stock returns during the fever period, suggesting that the health crisis has morphed into a possible financial crisis. The results therefore indicate that investors started to reward firms with good environmental score only when it became clear that the COVID-19 crisis would have severe financial consequences.

3.4 Do all investors reward environmental responsibility during the COVID-19 crisis?

Investors differ in the importance they place on environmental responsibility. It is therefore unlikely that all investors will reward environmental responsibility to the same extent. In this section, we explore whether some investors are more likely than others to reward environmental responsibility. Prior evidence suggests that investors with a long-term orientation are more concerned with climate responsibility (e.g., Gibson and Krueger 2018; Ramelli et al. 2018). Moreover, prior studies show that investors with short-term horizons sell their stockholdings to a larger extent that investors with long-term horizons during episodes of market turmoil (Cella et al. 2013). From this perspective, we expect that investors with long-term horizons are more likely to reward environmental responsibility.

In Table 6, we examine whether the effect of environmental score on stock returns during the COVID-19 crisis is more pronounced for firms with greater ownership by investors with long-term horizons. We use two alternative proxies for investor horizons. First, following common approach in the literature (e.g., Gaspar et al. 2005; Derrien et al. 2013), we capture an institutional investor's investment horizon through its portfolio turnover. Although investor horizons are not directly observable, the rationale behind this approach is that an investor which changes very frequently the composition of its portfolio is more likely to have a shorter investment horizon. In line with existing literature, we compute measures of investor horizons only for institutional investors covered by the 13F Thomson Files, for which data on stock portfolio composition is available over time. Following Derrien et al. (2013), we start by computing the portfolio turnover of each institutional investor. At the firm level, we then compute the weighted average of the portfolio turnover ratios of a firm's investors. Higher values of the average turnover therefore indicate shorter investor horizons. Second, as an alternative measure of long-term investor ownership, we sum the ownership by institutional investors which are likely to have a long-term orientation (i.e., pension funds, endowment funds, foundation funds, insurance companies,

Covid Economics 33, 30 June 2020: 124-162

CEPR PRESS

Covid Economics 33, 30 June 2020: 124-162

sovereign funds, venture capital funds, and private equity funds).[16] The rationale for this alternative measure is that investor horizons may be better captured by their orientation than their turnover (Edmans and Holderness 2017).

Table 6 reports regressions of stock returns during the COVID-19 crisis on pre-crisis environmental score and control variables estimated separately for subsamples sorted by our two proxies for long-term investor ownership. The results from Columns 1 and 2 indicate that the positive and significant association between environmental score and stock returns during the COVID-19 crisis is driven by the group of firms with below-median average investor turnover (i.e., with high long-term investor ownership). The Wald test of coefficient equality shows that the difference between coefficients is statistically significant. Similarly, the results from Columns 3 and 4 show that the effect of environmental score on stock returns during the COVID-19 crisis is more pronounced for the group of firms with above-median ownership by institutional investors with a long-term orientation. The Wald test of coefficient equality shows that the difference between coefficients is again statistically significant.

Table 6. Environmental score and stock returns during the COVID-19 crisis: cross-sectional tests based on long-term investor ownership

	(1) Low Investor Turnover	(2) High Investor Turnover	(3) Low % Investors with long-term orientation	(4) High % Investors with long-term orientation
Environmental Score	0.074***	0.002	0.010	0.075***
	(0.023)	(0.034)	(0.031)	(0.025)
Observations	813	813	813	813
Control Variables	Yes	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes	Yes
Adjusted R-squared	0.346	0.202	0.257	0.284

[16] The universe of investor types as identified by Thomson Reuters consists of : *Bank and Trust, Corporation, Endowment Fund, Foundation, Hedge Fund, Holding Company, Individual Investor, Insurance Company, Investment Advisor, Investment Advisor/Hedge Fund, Other Insider Investor, Pension Fund, Private Equity, Research Firm, Sovereign Wealth Fund, and Venture Capital.*

CEPR PRESS

P-value of Wald test (1) > (2)	0.040	
P-value of Wald test (4) > (3)		0.045

This table presents regression estimates of stock returns during the COVID-19 crisis (i.e., February 20-March 20, 2020) on environmental score and control variables. In columns 1 and 2, we split the sample according to median share-weighted average investor turnover. In column 3 and 4, we split the sample according to the median of ownership by investors with a long-term orientation. The group of investors with a long-term orientation consists of pension funds, endowment, foundation, insurance companies, and sovereign funds. Investor variables are measured as of end of 2019. All variables are winsorized at the 1st and 99th percentiles. All regressions include industry fixed effects based on TRBC industry group classifications. Variables are defined in Appendix A. Appendix B provides summary statistics for the ownership variables. Constants are not reported. Heteroskedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

Overall, the results from Table 6 indicate that responsible strategies on environmental issues are rewarded to a larger extent when investors have a long-term orientation. This result is consistent with the evidence in Ramelli et al. (2018) that although Donald Trump's election and his nomination of Scott Pruitt, a climate skeptic, to lead the Environmental Protection Agency drastically downshifted expectations on U.S. climate-change policy, long-term investors actually rewarded climate-responsible companies. While both the climate policy shock and the COVID-19 outbreak could have resulted in environmental responsibility being penalized, long-term institutional investors reward firms with responsible strategies on environmental issues presumably because they are anticipating the long-term consequences of environmental responsibility (Ramelli et al. 2018).

3.5 Environmental score, EPS revisions during the COVID-19 crisis, and stock reaction to policy responses

The evidence that firms with responsible strategies on climate change experience better stock returns during the COVID-19 crisis indicates that investors' attention has not been distracted away from environmental issues. Rather, the results suggest that investors, in particular those with a long-term orientation, have started to reward to larger extent firms with responsible initiatives on environmental issues and except them to do better in the long run. It could however be that the greater stock returns of firms with responsible initiatives on environmental issues reflects investors

Covid Economics 33, 30 June 2020: 124-162

valuing the possibility that these firms are better able to generate short-term profits or to access to financial resources.[17]

In this section, we provide two additional tests suggesting that firms with responsible initiatives on environmental issues are rewarded during the COVID-19 crisis because investors expect them to do better in the long-run rather than because they are in a better position to improve their immediate financial conditions. First, we examine whether environmental score is related to financial analysts' forecast revisions of one-year ahead earnings per share (EPS). In Table 7, we estimate regressions of the analysts' median forecast revision of one-year ahead EPS between February 2020 and March 2020 on pre-crisis environmental score and control variables. More specifically, we compute the average median EPS forecast for March 2020 and the average media EPS forecast for February 2020 and consider the raw difference, the percentage difference, and the difference scaled either by forecast dispersion (as of February 2020) or share price (at beginning of February 2020). In all columns, the results show that the coefficient on *Environmental score* is very small and not statistically significant, indicating that financial analysts do not expect firms with responsible strategies on environmental issues to generate higher earnings at a one-year horizon.

Table 7. Environmental score and one-year-ahead EPS consensus forecast revision

	(1) Median EPS Forecast Revision	(2) Pct. Median EPS Forecast Revision	(3) Median EPS Forecast Revision scaled by Forecast Dispersion	(4) Median EPS Forecast Revision scaled by Share Price
Environmental Score	0.001	0.044	-0.001	0.000
	(0.001)	(0.056)	(0.001)	(0.000)
Observations	1,157	1,157	1,157	1,157
Control Variables	Yes	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes	Yes
Adjusted R-squared	0.378	0.0130	0.117	0.436

This table presents regression estimates of one-year-ahead EPS consensus forecast revision between February and March 2020 on environmental score and control variables. *Median Forecast Revision* is the difference between average

[17] Cheng, Ioannou, and Serafeim (2013) show that superior performance on corporate social responsibility (CSR) strategies leads to better access to finance.

median forecast for one-year-ahead EPS in February and March. In Column 1, the dependent variable is the EPS median forecast revision between February and March. In Column 2, our main dependent variable is the EPS median forecast revision between February and March expressed as percentage of the average EPS median forecast of February 2020. In Columns 3 and 4, the dependent variable is EPS median forecast revision scaled by forecast dispersion (in February) and the share price (beginning of February 2020), respectively. We require at least two analysts to cover the firms and two estimates to calculate the median forecast for a given month. All variables are winsorized at the 1st and 99th percentiles. All regressions include industry fixed effects based on TRBC industry group classifications. Variables are defined in Appendix A. Appendix B provides summary statistics for EPS variables. Constants are not reported. Heteroskedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

Second, we investigate the possibility that investors may have rewarded firms with high environmental score for their superior access to finance. To do so, we examine whether firms with high environmental score benefit from the news concerning policy responses to the COVID-19 crisis on March 24. As explained by Fahlenbrach et al. (2020), on March 24th, once the approval of a large-scale stimulus became likely, the stock market responded positively with the best one-day performance since 2008.[18] If firms with high environmental score had better returns during the COVID-19 crisis because they have more financial flexibility or better access to finance, one would expect those firms to benefit less from the policy changes associated with the stimulus package. Fahlenbrach et al. (2020) show that firms with more cash and less debt experience higher returns during the COVID-19 crisis and benefit less from the policy responses on March 24.

In Table 8, we estimate regressions of the stock return from March 23 to March 24, 2020 on pre-crisis environmental score and control variables. The results show that both with and without controls, the coefficient on *Environmental score* is positive but not statistically significant, indicating that companies with responsible strategies on environmental issues do not benefit from the news concerning policy responses to the COVID-19 crisis on March 24. This finding does not support the notion that firms with high environmental score experience better returns during the COVID-19 crisis because they have more financial flexibility or better access to finance.

[18] The Fed stated that it would buy Treasuries "in the amounts needed to support smooth market functioning and effective transmission of monetary policy to broader financial conditions and the economy" and opened new facilities designed to provide credit to employers and to support the corporate bond market (Fahlenbrach et al. 2020).

Covid Economics 33, 30 June 2020: 124-162

CEPR PRESS

Covid Economics 33, 30 June 2020: 124-162

Table 8. Environmental score and stock reaction to policy responses to the COVID-19 crisis.

Return	Aid Plan 24 March 2020 (1)	Aid Plan 24 March 2020 (2)
Environmental Score	0.013	0.005
	(0.008)	(0.010)
Beta MKT		0.081
		(0.363)
Beta HML		0.458*
		(0.235)
Beta SMB		-0.545**
		(0.255)
Ln (Total Assets)		0.335*
		(0.184)
Long-Term Debt		3.773***
		(1.110)
Short-Term Debt		3.960
		(8.472)
Cash Holdings		-0.370
		(1.204)
Profitability		-3.470**
		(1.407)
Momentum		0.003
		(0.010)
Observations	1,626	1,626
Industry Dummies	Yes	Yes
Adjusted R-squared	0.114	0.130

This table presents regression estimates of stock returns from March 23 to March 24, 2020 on environmental score and control variables. All variables are winsorized at the 1st and 99th percentiles. All regressions include industry fixed effects based on TRBC industry group classifications. Variables are defined in Appendix A. Appendix B provides summary statistics for the stock returns from March 23 to March 24, 2020. Constants are not reported. Heteroskedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

Overall, the results from this section provide evidence to suggest that investors reward firms with responsible initiatives on environmental issues because they except them to do better in the long-run rather than because these firms are able to generate short-term profits or enjoy a better access to finance.

3.6 Robustness tests

In this section, we present robustness tests we have conducted to demonstrate that the results of our main analysis are not driven by the set of control variables or the empirical specification.

First, all regressions include industry fixed effects based on TRBC industry groups because they are the ones used by Asset 4 to calculate the environmental score of companies relative to their sector. In Table 9, Panel A, we reproduce our main regression from Table 2, Panel A including industry fixed effects based on different classifications. Specifically, we consider SIC industry dummies, GICS industry dummies, and NAICS industry dummies. The results show that the coefficient on *Environmental score* is positive and statistically significant in all specifications, indicating that our results are not sensitive to the choice of industry classifications.

Second, the results from Table 3 show that the effect of environmental score on stock returns during the COVID-19 crisis is robust to controlling for the governance score. Previous studies show that a firm's ownership and corporate governance structure affect stock returns in periods of market turmoil (e.g., Cella et al. 2013; Erkens et al. 2012; Ding et al. 2020). Table 9, Panel B presents the results of estimating our main regression from Table 2, Panel A with additional governance variables. Specifically, we control for several governance variables including institutional ownership, antitakeover devices, board size, board independence, and CEO duality dummy. The results indicate that institutional ownership (CEO duality) is negatively (positively) associated with stock returns during the COVID-19 crisis. Most importantly for our purpose, the coefficient on *Environmental score* remains statistically significant at the 1% level in all specifications.

Third, we further control for product market competition to alleviate the concern that firms operating in less competitive environments may be better able to both adopt responsible strategies on environmental issues and to resist the COVID-19 shock. Table 9, Panel C presents the results of estimating our main regression from Table 2, Panel A including Hoberg and Phillips' (2010) text-based measures of product

Covid Economics 33, 30 June 2020: 124-162

CEPR PRESS

market competition as additional control variables.[19] These measures are based on web crawling and text parsing algorithms that process the text in the business descriptions of 10-K annual filings. *HP HHI* and *HP SIM* are measures of market structure and market power respectively. The results show that these two measures of product market competition are not statistically associated with stock returns during the COVID-19 crisis. Most importantly for our purpose, the coefficient on *Environmental score* remains statistically significant at the 1% level in all specifications.

Finally, we check that our results are not driven by the inclusion of some sectors that are likely to have been particularly affected by the COVID-19 shock[20]. In Panel D, we estimate our baseline regression excluding companies in the energy, transportation, automobile, and personal services sector, respectively. The results show that our main results hold in all subsamples.

Table 9. Robustness tests

Panel A. Alternative definitions of industry fixed effects

Crisis Return	SIC Industry Dummies (1)	GICS Industry Dummies (2)	NAICS Industry Dummies (3)
Environmental Score	0.053***	0.040**	0.055***
	(0.020)	(0.019)	(0.020)
Observations	1,626	1,626	1,626
Control Variables	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes
Adjusted R-squared	0.277	0.277	0.256

This table presents regression estimates of stock returns during the COVID-19 crisis (i.e., February 20-March 20, 2020) on environmental score and control variables (Table 2, Panel A, Column 2) using different industry classifications. Columns 1 to 3 reports results of our baseline regression with industry dummies based on alternative industry group classifications: 2-digit SIC codes, NAICS industry groups and GICS industry groups, respectively. Variables are defined in Appendix A. Heteroskedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively. Constants are not reported.

[19] The latest available year for data on product market competition is 2017.
[20] The environment scores provided by Asset 4 are based on relative performance to the company's industry group, which mitigates the concern that some industries that were severely hit by the COVID-19 shock drive our results.

Covid Economics 33, 30 June 2020: 124-162

Panel B. Additional controls for corporate governance

Crisis Return	(1)	(2)	(3)	(4)	(5)	(6)
Environmental Score	0.052***	0.053***	0.061***	0.057***	0.057***	0.057***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.020)
Institutional Ownership	-0.045**					-0.035*
	(0.020)					(0.021)
Nb. Antitakeover Devices		-0.309				-0.242
		(0.206)				(0.209)
Ln(Board Size)			-2.462			-2.003
			(2.091)			(2.082)
% Independent Board Members				-0.031		0.011
				(0.034)		(0.036)
CEO Duality Dummy					1.616**	1.405*
					(0.802)	(0.804)
Observations	1,626	1,621	1,615	1,626	1,618	1,608
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.268	0.266	0.266	0.266	0.267	0.268

This table presents regression estimates of stock returns during the COVID-19 crisis on environmental score and control variables (Table 2, Panel A, Column 2) including additional governance variables. All regressions include industry fixed effects based on TRBC industry group classifications. Variables are defined in Appendix A. Appendix B provides summary statistics for governance variables. Constants are not reported. Heteroskedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

Panel C. Controlling for product market competition

Crisis Return	(1)	(2)	(3)
Environmental Score	0.052***	0.053***	0.061***
	(0.019)	(0.019)	(0.019)
HP HHI	0.922		0.585
	(1.789)		(1.827)
HP SIM		-0.060	-0.056
		(0.065)	(0.067)
Observations	1,449	1,449	1,449
Control Variables	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes
Adjusted R-squared	0.244	0.244	0.244

This table presents regression estimates of stock returns during the COVID-19 crisis on environmental score and control variables (Table 2, Panel A, Column 2) including product market competition variables from Hoberg and Phillips (2010). All regressions include industry fixed effects based on TRBC industry group classifications. Variables are defined in Appendix A. Appendix B provides summary statistics for product market competition variables. Constants are not reported. Heteroskedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

Covid Economics 33, 30 June 2020: 124-162

Covid Economics 33, 30 June 2020: 124-162

Panel D. Removing the influence of specific sectors

Crisis Return (%)	(1) Excluding Energy Sector	(2) Excluding Transportation Sector	(3) Excluding Automobile Sector	(4) Excluding Personal Services Sector
Environmental Score	0.049**	0.055***	0.055***	0.045**
	(0.019)	(0.019)	(0.019)	(0.020)
Observations	1,512	1,575	1,594	1,484
Control Variables	Yes	Yes	Yes	Yes
Industry Dummies	Yes	Yes	Yes	Yes
Adjusted R-squared	0.243	0.260	0.265	0.249

This table presents regression estimates of stock returns during the COVID-19 crisis on environmental score and control variables (Table 2, Panel A, Column 2) for subsamples excluding the energy, transportation, automobile, and personals services related companies, in columns 1, 2, 3, and 4, respectively. The energy sector corresponds to the TRBC sector "Energy - Fossil Fuels" and "Renewable Energy", the transportation sector corresponds to the TRBC sector "Transportation", the automobile sector corresponds to the TRBC sector "Automobiles & Auto Parts", and the personals services sector corresponds to the TRBC sector "Cyclical Consumer Services" and "Personal & Household Products & Services". All regressions include industry fixed effects based on TRBC industry group classifications. Variables are defined in Appendix A. Appendix B provides summary statistics for product market competition variables. Constants are not reported. Heteroskedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

4. Conclusion

In recent years, environmental issues and climate change have increasingly become a daily preoccupation for many companies. While climate change and the large financial risks it imposes were receiving growing attention, the COVID-19 outbreak came as a complete surprise for investors, corporate leaders, and policy makers. Although the near-economic standstill induced by the COVID-19 has immediate positive effects on the environment with levels of pollution and carbon emission dropping everywhere, the need for economic recovery potentially threatens climate actions as governments may be tempted to postpone stricter environmental actions.[21]

While there is still vast uncertainty concerning the financial consequences and the future of climate actions, our work looks at stock price reactions to the COVID-19 crisis, which capture the views and

[21] https://www.nytimes.com/2020/05/09/world/europe/will-the-coronavirus-crisis-trump-the-climate-crisis.html

expectations regarding future consequences of the virus. The evidence that firms with good environmental scores have significantly higher returns during the COVID-19 crisis indicates that investors have started to reward companies with responsible strategies on climate change to a larger extent. The view provided by the stock market on what investors expect for the future therefore suggests that companies with responsible strategies on climate issues will do better in the long run.

Future research should examine the realized consequences of the COVID-19 shock regarding environmental issues: Will there be a surge in shareholder activism on environmental issues? Will companies react to the COVID-19 crisis by fostering responsible initiatives on environmental and climate issues? Will customers favor environment-friendly products? Empirical analysis of these questions will be informative on the commitment of different stakeholders to tackle environmental issues and climate change.

Covid Economics 33, 30 June 2020: 124-162

References

Albuquerque, R., Koskinen, Y., & Zhang, C. (2019). Corporate social responsibility and firm risk: Theory and empirical evidence. *Management Science*, *65*(10), 4451-4469.

Albuquerque, R. A., Koskinen, Y., Yang, S., & Zhang, C. (2020). Love in the time of covid-19: The resiliency of environmental and social stocks.

Andersson, M., Bolton, P., & Samama, F. (2016). Governance and climate change: A success story in mobilizing investor support for corporate responses to climate change. *Journal of Applied Corporate Finance*, *28*(2), 29-33.

Boermans, M. A., & Galema, R. (2019). Are pension funds actively decarbonizing their portfolios?. *Ecological economics*, *161*, 50-60.

Capelle-Blancard, G., & Laguna, M. A. (2010). How does the stock market respond to chemical disasters?. *Journal of Environmental Economics and Management*, *59*(2), 192-205.

Carney, M. (2015). Breaking the Tragedy of the Horizon–climate change and financial stability. *Speech given at Lloyd's of London*, *29*, 220-230.

Cella, C., Ellul, A., & Giannetti, M. (2013). Investors' horizons and the amplification of market shocks. *The Review of Financial Studies*, *26*(7), 1607-1648.

Chava, S. (2014). Environmental externalities and cost of capital. *Management Science*, *60*(9), 2223-2247.

Cheng, B., Ioannou, I., & Serafeim, G. (2014). Corporate social responsibility and access to finance. *Strategic management journal*, *35*(1), 1-23.

Choi, D., Gao, Z., & Jiang, W. (2020). Attention to global warming. *The Review of Financial Studies*, *33*(3), 1112-1145.

De Angelis, T., Tankov, P., & Zerbib, O. D. (2020). Environmental Impact Investing. *Available at SSRN 3562534*.

De Vito, A., & Gómez, J. P. (2020). Estimating the COVID-19 cash crunch: Global evidence and policy. *Journal of Accounting and Public Policy*, 106741.

Derrien, F., Kecskés, A., & Thesmar, D. (2013). Investor horizons and corporate policies. *Journal of Financial and Quantitative Analysis*, *48*(6), 1755-1780.

Ding, W., Levine, R., Lin, C., & Xie, W. (2020). *Corporate immunity to the covid-19 pandemic* (No. w27055). National Bureau of Economic Research.

Dyck, A., Lins, K. V., Roth, L., & Wagner, H. F. (2019). Do institutional investors drive corporate social responsibility? International evidence. *Journal of Financial Economics*, *131*(3), 693-714.

Edmans, A., & Holderness, C. G. (2017). Blockholders: A survey of theory and evidence. In *The handbook of the economics of corporate governance* (Vol. 1, pp. 541-636). North-Holland.

Erkens, D. H., Hung, M., & Matos, P. (2012). Corporate governance in the 2007–2008 financial crisis: Evidence from financial institutions worldwide. *Journal of corporate finance*, *18*(2), 389-411.

Fahlenbrach, R., Rageth, K., & Stulz, R. M. (2020). *How valuable is financial flexibility when revenue stops? Evidence from the COVID-19 crisis* (No. w27106). National Bureau of Economic Research.

Fisher-Vanden, K., & Thorburn, K. S. (2011). Voluntary corporate environmental initiatives and shareholder wealth. *Journal of Environmental Economics and management*, *62*(3), 430-445.

Flammer, C. (2013). Corporate social responsibility and shareholder reaction: The environmental awareness of investors. *Academy of Management Journal*, *56*(3), 758-781.

Gaspar, J. M., Massa, M., & Matos, P. (2005). Shareholder investment horizons and the market for corporate control. *Journal of financial economics*, *76*(1), 135-165.

Gibson, R., & Krueger, P. (2018). The sustainability footprint of institutional investors. *Swiss Finance Institute Research Paper*, (17-05).

Gonenc, H., & Scholtens, B. (2017). Environmental and financial performance of fossil fuel firms: A closer inspection of their interaction. *Ecological Economics*, *132*, 307-328.

Hoberg, G., & Phillips, G. (2010). Product market synergies and competition in mergers and acquisitions: A text-based analysis. *The Review of Financial Studies, 23(10),* 3773-3811.

Ilhan, E., Krueger, P., Sautner, Z., & Starks, L. T. (2019). Institutional investors' views and preferences on climate risk disclosure. *Available at SSRN 3437178*.

Ilhan, E., Sautner, Z., & Vilkov, G. (2019). Carbon tail risk. *Available at SSRN 3204420*.

Krueger, P., Sautner, Z., & Starks, L. T. (2020). The importance of climate risks for institutional investors. *The Review of Financial Studies*, *33*(3), 1067-1111.

Lins, K. V., Servaes, H., & Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance*, *72*(4), 1785-1824.

Lioui, A., & Sharma, Z. (2012). Environmental corporate social responsibility and financial performance: Disentangling direct and indirect effects. *Ecological Economics*, *78*, 100-111.

Litterman, B. (2017). *Climate Risk: Tail Risk and the Price of Carbon Emissions-Answers to the Risk Management Puzzle*. John Wiley & Sons.

Loughran, T., & McDonald, B. (2020). Management disclosure of risk factors and COVID-19. *Available at SSRN 3575157*.

Lyon, T., Lu, Y., Shi, X., & Yin, Q. (2013). How do investors respond to Green Company Awards in China?. *Ecological Economics*, *94*, 1-8.

Monasterolo, I., & De Angelis, L. (2020). Blind to carbon risk? An analysis of stock market reaction to the Paris Agreement. *Ecological Economics*, *170*, 106571.

Nordhaus, W. (2019). Climate change: The ultimate challenge for Economics. *American Economic Review*, *109*(6), 1991-2014.

Oberndorfer, U., Schmidt, P., Wagner, M., & Ziegler, A. (2013). Does the stock market value the inclusion in a sustainability stock index? An event study analysis for German firms. *Journal of*

Appendix A. Variable definitions

Variables	Definition	Source
% Independent Board Members	Proportion of independent directors sitting on the board	TR Asset 4
Beta HML	Loading on the HML premium based on the Fama-French three factors model estimated over 60 months prior to the crisis. We require at least 12 months of data available.	TR and Kenneth French's Website
Beta MKT	Loading on the market premium based on the Fama-French three factors model estimated over 60 months prior to the crisis. We require at least 12 months of data available.	TR and Kenneth French's Website
Beta SMB	Loading on the SMB premium based on the Fama-French three factors model estimated over 60 months prior to the crisis. We require at least 12 months of data available.	TR and Kenneth French's Website
Cash Holdings	as cash and marketable securities scaled by total assets	TR Fundamentals
CEO Duality Dummy	Dummy variable coding for whether the CEO is also the chairman of the board.	TR Asset 4
COVID-19 Crisis Return	Buy-and-hold stock return over the period February 20-March 20	TR
Emission Reduction	The Emission Reduction Score measures a company's commitment and effectiveness towards reducing environmental emission in the production and operational processes.	TR Asset 4
Environmental Score	*Environmental Score* aggregates information on a company's performance and capacity to reduce the use of materials, energy or water, and to find more eco-efficient solutions by improving supply chain management (*Resource Use Score*), a company's commitment and effectiveness towards reducing environmental emission in the production and operational processes (*Emission Reduction Score*), and a company's capacity to reduce the environmental costs and burdens for its customers, thereby creating new market opportunities through new environmental technologies and processes or eco-designed products (*Green Innovation Score).*	TR Asset 4
Governance Score	*Governance Score* aggregates information a company's commitment and effectiveness towards following best practice corporate governance principles (*Management Score*), a company's effectiveness towards equal treatment of shareholders and the use of anti-takeover devices (*Shareholders Score*), and a company's practices to communicate that it integrates the economic (financial), social and environmental dimensions into its day-to-day decision-making processes (*CSR Strategy Score*).	TR Asset 4
Green Innovation	The Green innovation Score reflects a company's capacity to reduce the environmental costs and burdens for its customers, and thereby creating new market opportunities through new environmental technologies and processes or eco-designed products	TR Asset 4
HP HHI	This data is based on web crawling and text parsing algorithms that process the text in the business descriptions of 10-K annual filings on the SEC Edgar website from 2017. HHI is a measure of concentration within TNIC industry classification.	Hoberg's website
HP SIM	This data is based on web crawling and text parsing algorithms that process the text in the business descriptions of 10-K annual filings on the SEC Edgar website from 2017. SIM is a measure of product differentiation within TNIC industry classification.	

Covid Economics 33, 30 June 2020: 124-162

Institutional Ownership	Percentage of a firm's common shares outstanding held by institutional investors.	TR
Investor Turnover	The shares-weighted average portfolio turnover of a firm's investor, whereby portfolio turnover is the average investor holding period calculated based on the previous 12 quarters (36 months) of portfolio holdings.	TR
Ln(Board Size)	Natural logarithm of the number of board members	TR Asset 4
Ln(Total Assets)	Natural logarithm of total assets	TR Fundamentals
Long-Term Debt	Long-term debt scaled by total assets	TR Fundamentals
Median EPS Forecast Revision	*Median EPS Forecast Revision* is the difference between the average median forecast of one-year-ahead EPS in March and the one in February 2020.	TR IBES
Momentum	Buy-and-hold return over the period November 2019 – January 2020	TR
Nb. Antitakeover Devices	Number of anti-takeover devices in place	TR Asset 4
Pct. Long-term Investors	Ownership by investors with a long-term orientation, defined based on shareholder types (pension funds, endowment, foundation, insurance companies, sovereign funds, private equity, venture capital).	
Pct. Median EPS Forecast Revision	The difference between the average median forecast of one-year-ahead EPS in March and the one in February scaled by the average median forecast of one-year-ahead EPS in February.	TR IBES
Profitability	EBITDA scaled by total assets	TR Fundamentals
Resource Use	The Resource Use Score reflects a company's performance and capacity to reduce the use of materials, energy or water, and to find more eco-efficient solutions by improving supply chain management.	TR Asset 4
Short-Term Debt	Short-term debt scaled by total assets	TR Fundamentals
Social Score	*Social Score* aggregates information on the extent to which firms enhance employee welfare (*Workforce Score*), promote human rights (*Human Rights Score*), engage in community development (*Community Score*), and fulfill their responsibilities to consumers (*Product Responsibility Score*).	TR Asset 4

Appendix B. Extra summary statistics

Variables	#Obs	Mean	S.D.	.25	Mdn	.75
Institutional Ownership	1,626	83.78	21.56	72.94	89.29	98.92
Nb. Antitakeover Devices	1,621	6.67	1.98	5.00	7.00	8.00
Ln(Board Size)	1,615	2.16	0.24	1.95	2.20	2.30
% Independent Board Members	1,626	79.61	12.15	75.00	83.33	88.89
CEO Duality Dummy	1,618	0.54	0.50	0.00	1.00	1.00
Governance Score	1,626	45.48	22.97	26.47	44.93	63.83
Social Score	1,626	39.93	20.56	23.57	36.09	52.63
Resource Use Score	1,626	24.38	31.21	0.00	4.82	46.69
Emission Score	1,626	20.81	28.75	0.00	3.88	35.29
Green Innovation Score	1,418	17.89	27.68	0.00	0.00	35.29
Crisis Return (Fahlenbrach et al., 2020)	1,626	-39.06	19.24	-52.06	-38.67	-27.27
Incubation (Ramelli and Wagner, 2020)	1,626	2.29	8.85	-2.05	1.70	5.95
Outbreak (Ramelli and Wagner, 2020)	1,626	-1.89	12.76	-9.09	-2.14	5.30
Fever (Ramelli and Wagner, 2020)	1,626	-37.95	18.34	-49.90	-36.75	-26.12
Aid Plan	1,626	10.70	8.18	5.43	9.56	14.56
Investor Turnover (%)	1,626	40.03	10.79	32.49	37.74	45.38
Pct. LT Ownership (%)	1,626	5.74	7.18	2.76	4.10	5.57
Median EPS Forecast Revision	1,157	-0.18	0.43	-0.18	-0.05	-0.01
Pct. EPS Median Forecast Revision (%)	1,157	-5.33	36.00	-9.27	-2.97	-0.34
HP HHI	1,449	0.28	0.25	0.10	0.19	0.37
HP SIM	1,449	7.25	14.80	1.12	1.56	3.31

Institutional Ownership is the percentage of a firm's common shares outstanding held by institutional investors. *Nb. Antitakeover Devices* is the number of anti-takeover devices in place. *Ln(Board Size)* is the natural logarithm of the number of board members. *% Independent Board Members* is the proportion of independent directors sitting on the board. *CEO Duality Dummy* codes for whether the CEO is also the chairman of the board. *Environmental Score* aggregates information on a company's performance and capacity to reduce the use of materials, energy or water, and to find more eco-efficient solutions by improving supply chain management (*Resource Use Score*), a company's commitment and effectiveness towards reducing environmental emission in the production and operational processes (*Emission Reduction Score*), and a company's capacity to reduce the environmental costs and burdens for its customers, thereby creating new market opportunities through new environmental technologies and processes or eco-designed products (*Green Innovation Score*). *Social Score* aggregates information on the extent to which firms enhance employee welfare (*Workforce Score*), promote human rights (*Human Rights Score*), engage in community development (*Community Score*), and fulfill their responsibilities to consumers (*Product Responsibility Score*). *Governance Score* aggregates information a company's commitment and effectiveness towards following best practice corporate governance principles (*Management Score*), a company's effectiveness towards equal treatment of shareholders and the use of anti-takeover devices (*Shareholders Score*), and a company's practices to communicate that it integrates the economic (financial), social and environmental dimensions into its day-to-day decision-making processes (*CSR Strategy Score*). *Investor Turnover* is the shares-weighted average portfolio turnover of a firm's investor, whereby portfolio turnover is the average investor holding period calculated based on the previous 12 quarters (36 months) of portfolio holdings. *Pct. LT Ownership* is the ownership of investors with a long-term orientation, defined based on shareholder types (pension funds, endowment, foundation, insurance companies, sovereign funds, private equity, venture capital). *Median EPS Forecast Revision* is the difference between the average median forecast of one-year-ahead EPS in March and the one in February 2020. *Pct. Median EPS Forecast Revision* is the difference between the average median forecast of one-year-ahead EPS in March and the one in February scaled by the average median forecast of one-year-ahead EPS in February.

Covid Economics 33, 30 June 2020: 124-162

Covid Economics 33, 30 June 2020: 124-162

Appendix C. Correlation Matrix

This table presents the correlation matrix for the main dependent and independent variables we use in our empirical analysis. All variables are defined in Appendix A.

Correlation Matrix											
	1	2	3	4	5	6	7	8	9	10	11
1. Crisis Return (%)	1.00										
2. Environmental Score	0.0630	1.0000									
3. Ln(Total Assets)	-0.1793	0.0139	1.0000								
4. Long-Term Debt	-0.1294	-0.3330	-0.0373	1.0000							
5. Short-Term Debt	-0.2042	0.1478	0.1346	-0.0030	1.0000						
6. Profitability	0.0393	0.4122	-0.0254	-0.2160	0.0254	1.0000					
7. Cash Holdings	-0.1882	0.1348	0.0403	-0.0960	0.1008	0.0572	1.0000				
8. Momentum	-0.0148	0.1074	-0.0392	-0.0629	0.0230	0.0946	-0.0372	1.0000			
9. Beta MKT	0.0310	0.2975	-0.1668	-0.2971	0.3079	0.0982	0.1770	0.0479	1.0000		
10. Beta HML	0.1669	-0.3243	0.0627	0.2673	-0.4190	-0.1185	-0.3450	-0.1115	-0.6324	1.0000	
11. Beta SMB	0.0206	-0.0701	0.1166	0.0482	-0.3184	-0.0225	-0.0372	-0.0465	-0.1919	0.2014	1.0000

CEPR PRESS

Covid Economics **Issue 33, 30 June 2020**

Cultural differences in COVID-19 spread and policy compliance: Evidence from Switzerland

Fabrizio Mazzonna[1]

Date submitted: 24 June 2020; Date accepted: 25 June 2020

Despite the COVID-19 pandemic is currently spreading all over the world, we still observe dramatic variation between and, even within, countries in the speed of the infection, in the observed fatality rates and in the effectiveness of the containment measures put in place by most countries. This paper sheds light on the role of culture exploiting the large cultural variation between German and Latin (French and Italian) speaking regions in Switzerland. Consistently with the large difference in social contacts across generations between these two distinct cultural groups, it shows that the disease affected disproportionately elderly people only in Latin regions. Then, it shows that cultural differences are also associated with different levels of compliance with the containment measures put in place by the Swiss government. Mobility data by Google and Apple clearly show that people living in Latin-speaking regions started reducing their movements a week before the lockdown and then complied more strictly than their German counterparts with the policy. This differential compliance across language regions clearly affected the epidemic curves. Using an event study design, we reveal that Latin regions experiencing a faster decline in the growth rate of new cases, hospitalizations and deaths than their German counterpart.

1 Associate Professor, Department of Economics (IdEP), Università della Svizzera Italiana.

Copyright: Fabrizio Mazzonna

Covid Economics 33, 30 June 2020: 163-185

CEPR PRESS

Covid Economics 33, 30 June 2020: 163-185

1 Introduction

The first wave of COVID-19 is almost over in Europe and in most Asian countries, but the pandemic is not yet under control. Part of the USA, Latin America and India are still recording thousands of new infections every day and new waves are expected all over the world. Unfortunately, we still lack a clear understanding of the routes of transmission of the virus. For instance, some countries or regions were affected more severely than others and case fatality rates vary substantially across different areas. Differences in testing criteria undoubtedly play an important role, but other crucial factors, which have not been sufficiently investigated, include cultural differences in social contacts, especially across generations, and citizens' compliance with social distancing measures implemented worldwide to limit the outbreak. Starting from China in late January, massive non-pharmaceutical interventions (NPIs) have been implemented worldwide to limit the outbreak but, as it stands, it is not completely clear how and to what extent the containment measures are effective and what is the role of people's compliance with these measures. Policy makers need to know how different social and cultural environments react to these measures to obtain the best behavioral response to mitigate the impact of the pandemic (Van Bavel et al., 2020).

This paper provides three main pieces of evidence which should improve our understanding of the complex interaction between cultural factors, spread of the disease and policy effectiveness. First, by exploiting the well-documented within country cultural divide in Switzerland between cantons where a German or Latin (French and Italian) language is spoken (Eugster et al., 2011), we show that the spread of the virus disproportionally affected elderly people only in Latin cantons. After accounting for several potential confounders (e.g., differential timing of the epidemic, population density and age structure), the large difference in the incidence of the disease between these two language groups remains unaffected only among the elderly. On the opposite, differences between language regions become small and often not statistically significant when we focus on people below age 65. The cultural explanation is further supported by an investigation of the social network of the elderly, which shows large differences in social contacts between the two linguistic regions, especially across generations. Alternative explanations, like the proximity of some Latin cantons to hot clusters in neighboring countries (which are also at the border with some German cantons) can hardly explain the presence of such large difference across age groups.

Second, we document a substantial heterogeneity in the compliance with the "soft" lockdown policy implemented on March 16, when the Swiss federal government closed non-essential high street shops and services and banned any gathering of more than 5 people, but did not order to stay at home. Using mobility reports provided by Google and Apple, we show that people living in Latin cantons anticipated the lockdown by reducing their movements one week before and then complied

CEPR PRESS

Covid Economics 33, 30 June 2020: 163-185

more strictly with the policy than their German counterparts.

Third, using an event study design meant to evaluate the effect of the lockdown, we show that the heterogeneity in compliance clearly influences the effectiveness of the policy. In particular, the estimated decline in the growth rate of confirmed cases, hospitalizations and deaths starts earlier in Latin cantons and remains substantially larger even several weeks after the implementation of the policy.

Given the larger initial incidence of the disease in Latin cantons, our results might suggest that risk perception and salience have a key role in driving the level of compliance. However, when we account for this channel a substantial part of the estimated cultural divide remains unexplained. One plausible explanation is the different cultural attitude towards government mandates (and more generally towards the role of the government) between the two language regions. While individual responsibility and freedom are crucial in the Swiss-German culture, in fact, people in Latin regions are more in favor of government intervention, as in France and several countries in Southern Europe (Eugster et al., 2011).[1]

This paper contributes to the current literature about the COVID-19 pandemic. First, it has been argued that intergenerational contacts and co-residence may have been a key driver of the observed variation in the spread of disease and in fatality rates, especially in Mediterranean countries like Italy and Spain (Dowd et al., 2020). However, early cross-country evidence about the positive correlation between family ties and case fatality rates (Kuhn and Bayer, 2020) has been highly criticized, especially its cross-country design and the use of the case fatality rate as a proxy for the spread of the disease among the elderly (Belloc et al., 2020). This paper, instead, focuses on within country cultural variation and measures the spread of the disease using the incidence of confirmed cases, hospitalizations and deaths by language regions.

Second, we contribute to the ongoing literature about the determinants of people's compliance with NPIs. Political beliefs (Painter and Qiu, 2020), income and age structure (Engle et al., 2020), social capital (Bartscher et al., 2020) and political regime (Frey et al., 2020) have been recently pointed out as important drivers of citizens' compliance with NPIs. None of the existing papers, however, considers the role of risk perception and culture. More generally, we provide additional evidence of the importance of culture in shaping socio-economic outcomes (see Alesina and Giuliano, 2015 for a review).

The rest of this paper is organized as follows. Section 2 describes the Swiss cultural background and the epidemic. Section 3 describes our data. Section 4 investigates the heterogeneity in the incidence of the disease. Section 5 shows the estimated effects of the lockdown. Finally, Section 6 concludes.

[1] See for instance the recent anti-lockdown rallies on May 16 in Zurich, Bern and Basel: https://www.swissinfo.ch/eng/police-intervention_anti-lockdown-rallies-stopped-at-early-stage/45765040.

Covid Economics 33, 30 June 2020: 163-185

2 Background

2.1 Language regions and culture

Switzerland is a confederation of 26 states called cantons and has four different official languages, namely German (62.6%), French (22.9%), Italian (8.2%) and Romansh (0.5%). Language borders are historically determined and geographically delimited (see Figure A.1), so the discontinuity in the probability of speaking a given language is quite sharp at such borders.[2] These language groups correspond to two main cultural groups, German-speaking communities and communities speaking a language of Latin origin (French, Italian and Romansh). Several papers have already exploited this discontinuity at the language border to show how the large cultural difference between these two broad language groups affects a variety of outcomes, namely preferences for social insurance (Eugster et al., 2011, 2017), taxation (Eugster and Parchet, 2019), international trade (Egger and Lassmann, 2015) and fertility decisions (Steinhauer, 2018).

Closely related to this study is the work by Gentili et al. (2017) on the relationship between culture and long-term care arrangements in Switzerland, which shows the presence of stronger family ties among Swiss-Latin individuals leads them to enter nursing homes later than their German-speaking neighbors and to use more home-care services. Differences in family ties also imply different levels of intergenerational contacts, which may represent one of the reasons for the differential spread of the infection across linguistic regions, especially among the elderly, as described in this paper.

This hypothesis is further investigated exploiting the social network module in the sixth wave (2015) of the Survey of Health Aging and Retirement in Europe (SHARE).[3] As expected, Figure 1 shows the presence of huge differences in the social network of people aged 65+ across language groups. In particular, old age respondents in cantons where a Latin language is spoken report both a significantly larger social network and a higher level of social connectedness (in both cases, the difference corresponds to a quarter of a standard deviation).[4] This is partly driven by family contacts, especially intergenerational contacts with children and grandchildren. Latin speaking respondents are significantly more likely to have weekly contacts with their children, to have children living within one Km and to do baby-sitting for their grandchildren. In addition, these individuals are also more likely to have friends outside their family network and this suggests that differences in social contacts between language groups are not exclusively driven by family ties.

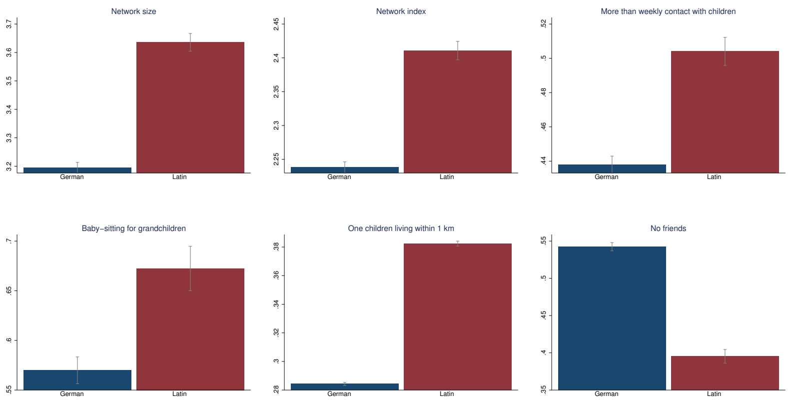
[2] Language borders correspond to canton borders except for three bilingual cantons (Bern, Fribourg and Valais).

[3] SHARE is a multidisciplinary and cross-national panel database of micro data on health, socio-economic status and social and family networks of about 140,000 individuals aged 50 or older in 27 European countries (Börsch-Supan, 2019). More information on the data are provided in the online Appendix.

[4] The level of social connectedness is assessed using a summary scale of the social network based on Litwin and Stoeckel (2016).

Covid Economics 33, 30 June 2020: 163-185

Figure 1: Social network differences in Latin and German speaking areas (people 65+)



Notes: This figure shows 6 social network characteristics of old age people (65+) living in Switzerland testing for differences between the two main language group German and Latin (French and Italian) using data are drawn from the sixth wave of the Survey of Health Aging and Retirement in Europe (SHARE). The 6 dimensions reported are the following: the size of the social network; an overall index of social connectedness —a summary scale of the social network data (Litwin and Stoeckel, 2016); whether the respondents has more than weekly contacts with their children; whether the respondents does baby-sitting for their grand children; whether at least one chid lives with 1 km; whether there respondents has no friends (excluding family network). For each dimension we report the mean value for each language group and its 95% confidence interval.

2.2 The COVID-19 pandemic in Switzerland

The first COVID-19 case in Switzerland was detected on February 25 in Ticino, the southern canton at the border with Lombardy. Since that date, an increasing number of new cases started to be soon reported in all language regions, but during the following week the incidence grew rapidly especially in Ticino and in French-speaking cantons. As a response, the first containment measure introduced on February 28 was the ban of all events involving more than one thousand participants.

Given the increasing number of reported cases in the whole country during the subsequent weeks, on March 16 the Federal Government announced the implementation of a *soft lockdown*. These new measures, which became effective on March 17, included the closure of schools and childcare facilities as well as the interruption of all non-essential economic activities (i.e., shops, bars and restaurants). Furthermore, the Federal Government also banned any gathering of more than five people since March 20. Two of the most affected cantons, Ticino and Geneva anticipated some of these measures (i.e., school closure and banning of gathering of more than five people) by a couple of days. However, although recommended, a stricter lockdown, with an order to stay at home, has never been implemented.

Because of the remarkable decline in the number of new cases, these restrictions began to be partially loosened since April 27, when several shops were allowed to start their activities again. After this first phase, the opening of schools, museums and libraries took place on May 11, when also bars and restaurants had the possibility to open, although conditional upon the respect of strict social distancing measures. Finally, on June 8 a third phase started, since post-compulsory schools and universities reopened and events with a maximum of 300 participants started to be permitted.

3 Data

COVID-19 data

The main analysis is based on the data provided by the Swiss Federal Office for Public Health (FOPH) about the universe of people found positive to the COVID-19 virus. These data include the dates of test, symptoms onset (only for hospitalized cases) and death (when applicable), as well as socio-demographic information concerning age, sex and canton of residence. The geographical detail allows to broadly distinguish between two main cultural groups identified by their dominant language, namely Latin and German-speaking cantons, hereafter Latin and German cantons. While this broad distinction is used for descriptive purposes, the empirical investigation will be based on the share of German-speaking people within each canton, exploiting the substantial variations of this share across cantons, especially the bilingual ones.

Switzerland has massively tested its population (51 tests per thousand people) and, differently from other European countries (e.g., Italy or Spain), it has pursued this strategy since the very beginning of the COVID-19 crisis. This allowed the country to monitor the spread of the infection more closely and to intervene effectively to isolate infected people. In spite of these remarkable testing efforts, however, the majority of people who were either asymptomatic or with very mild symptoms were not traced: the government, in fact, decided to prioritize high-risk subjects and people with severe symptoms. As a consequence, elderly people were more likely to be tested with respect to young people. This should not represent a concern, unless testing criteria differ systematically across linguistic regions and age groups, and in a way it explains the higher incidence of cases detected among the elderly (relative to younger people) in Latin regions. As far as we know, testing criteria have been defined centrally by the FOPH, although the differential spread of the disease might have affected the testing policy across cantons. However, it is reassuring that in the data the median time between the onset of symptoms and the test is not significantly different across linguistic regions. This suggests that the largest incidence of the disease in Latin regions did not affect (delay) their testing capacity. Additionally, we find similar differences in the incidence of hospitalizations, which are unlikely to be driven by testing.

Social mobility data

People's compliance with the soft lockdown is measured using data from the developers of the two most popular mobile phone operating systems and application maps, namely Google and Apple. Google Community Mobility Reports chart movement trends over time by canton across different categories of

Covid Economics 33, 30 June 2020: 163-185

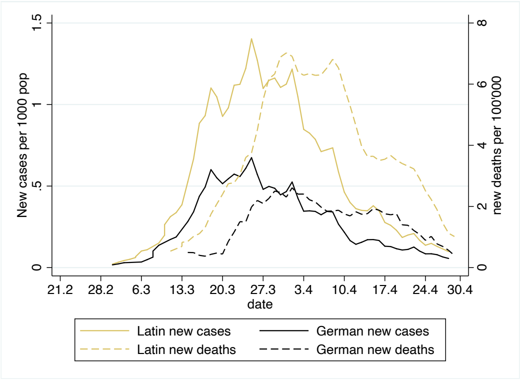
Covid Economics 33, 30 June 2020: 163-185

places such as retail and recreation, groceries and pharmacies, parks, transit stations, workplaces and residential areas. More specifically, they report the daily percentage changes in mobility with respect to the median value for the period Jan 3–Feb 6, 2020. Similarly, Apple Mobility Trends Reports reflect the requests for directions using Apple Maps: starting from January 13 (baseline day), they show how driving mobility changes daily and by canton.

4 Heterogeneity in COVID-19 incidence

Although the first COVID-19 cases were detected roughly at the same time in all language regions (German, French and Italian), Figure 2 shows that the spread of the disease was far more rapid in Latin cantons.

Figure 2: Incidence of new cases and deaths by language groups



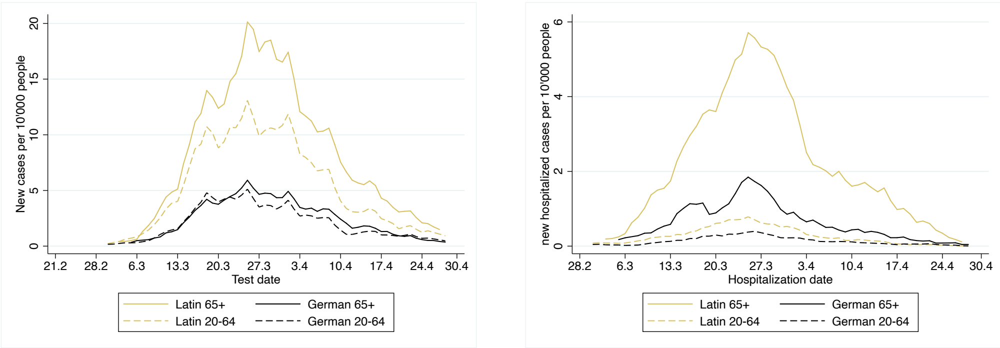
This might be partially explained by the close proximity to France and Italy, two of the most affected countries in Europe.[5] Such proximity, however, cannot explain why the observed differences across language groups are far more pronounced among the elderly, as reported in Figure 3. This is especially true for the incidence of new cases and hospitalizations, while deaths are actually reported almost exclusively among people older than 65. As previously mentioned, testing criteria might also play a role, but the higher number of hospitalizations and deaths in Latin cantons is clearly driven by the larger incidence of the disease among the elderly.

The descriptive evidence reported so far holds if we account for the differential timing of the epidemic across cantons (i.e., the time passed from the first 10 cases or an incidence of two cases per 100 thousand people), for the calendar date and for other potential confounders at canton level, namely population density, hospital beds and the share of people aged 65+ (Table A.1 and additional

[5] Actually, the French and Italian border regions were not severely affected by the virus. This is also true for the Italian provinces of Como and Varese that, differently from the rest of Lombardy, experienced infection and death rates which were significantly lower with respect to Ticino, the Southern Swiss canton.

Covid Economics 33, 30 June 2020: 163-185

Figure 3: Incidence of new cases and hospitalizations by language and age groups



(a) Positive cases

(b) Hospitalizations

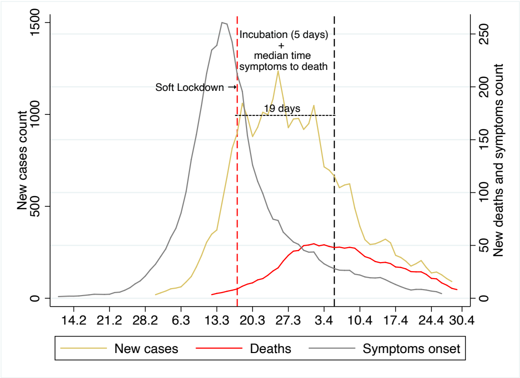
details about the empirical specification are provided in the Appendix). It is only worth noting that the cumulative number of confirmed cases and hospitalizations among people aged 65+ increases by over 150% when the share of Latin speaking people in a canton changes from 0 to 1, while this effect is much lower and often not statistically significant among people who are younger than 65. Considering the relatively low number of deaths before age 65 in Switzerland, this result also explains the largest incidence of deaths registered in Latin cantons. Although we already explained that the close proximity of some cantons to Italy and France can hardly explain our results, we find that our results hold even if we exclude Ticino and Geneva, the most affected cantons at the border with Italy and France.

5 Heterogeneity in the effect of the lockdown

Figure 4, which aggregates individual data about symptoms onset, test and death by date, shows the relationship between the containment measures put in place by the federal government and the epidemic curves. It is worth noting that the peak of symptoms onset anticipates the implementation of the policy, suggesting that the decline in the infection rate had already started a few days before, especially if we consider the incubation period, whose median value should be five days (WHO, 2020). Analogously, the actual peak in the number of new deaths seems to anticipate the predicted peak (summing up the median incubation period and the median number of days between symptoms onset and death from the FOPH data).

However, the evidence reported at national level masks substantial heterogeneity. A closer look at the previously mentioned Figure 2, in fact, suggests a sharper decline of the epidemic curves in Latin cantons with respect to German cantons. This is particularly clear when looking at the incidence of

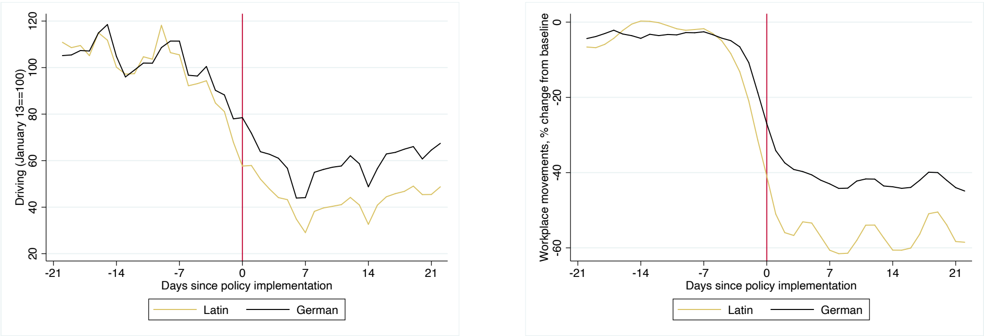
Figure 4: Epidemic curve, daily deaths and symptoms onset in Switzerland



new deaths. In Latin cantons, this incidence sharply declines after the peak before April 10, while in German cantons the peak is followed by a long plateau till April 20.

Such descriptive evidence might be the result of different levels of compliance with the policy between the two linguistic regions. Using Apple and Google mobility reports (Figure 5), indeed, we show that people living in Latin cantons are characterized by a higher level of compliance with the soft-lockdown implemented by the Swiss Government, which only recommended people to stay home, without any mandate as in Italy or Spain.

Figure 5: Mobility data by date and language group



(a) Apple

(b) Google

Notes: This figure shows mobility report data provided by Apple (left figure) and Google (right figure) by date and language group. Apple data track driving time while Google data movements towards the workplace.

Covid Economics 33, 30 June 2020: 163-185

CEPR PRESS

Covid Economics 33, 30 June 2020: 163-185

This figure also suggests that, especially in Latin cantons, people started reducing their movements roughly one week before the policy. Both driving times and movements towards work had already decreased by roughly 20% three days before the lockdown, whose implementation led to a further acceleration in the reduction of observed movements. Moreover, a substantial gap between Latin and German cantons is still observed three weeks after the policy.

Further evidence is provided in the Appendix (Table A.2), where we show the average percentage change in movements towards different venues (i.e., retail and recreation, parks, transit stations, workplaces and residential areas) by language, or selected cantons, using Google community movements data. With the only exception of parks, we observe large drops in movements in both language regions, but these negative changes are always larger in Latin cantons. Additionally, people living in German cantons increased their movements towards parks during the lockdown.

5.1 Econometric modeling of the effect of the lockdown

Estimating the heterogeneous impact of the policy across language groups has two main challenges. First, the effects can be observed only several days or weeks after because of the lags between exposure and recorded infections, hospitalizations or deaths. Second, we do not observe the counterfactual evolution of the pandemic in absence of intervention, which depends not only on the canton-specific timing of the epidemic (which we can control), but also on the "speed" of the disease, which is likely to be very different across language regions.

These issues are addressed using an "event-study" design which aims at tracing the dynamic effects of the policy and the potential anticipatory effects, taking into account the heterogeneity across language groups observed so far. Based on the classical models in epidemiology (e.g., SIR model) and the observed evolution of the pandemic over the world, we focus on growth rates to take into account the well-documented exponential growth of the disease, namely the fact that the number of new cases tends to be proportional to the number of already infected people.[6] More specifically, we estimate the following model:

$$G_{ct} = \beta_0 + \beta_1 SL_c + X_c\beta_2 + dow_t + \sum_{i=2}^{n} \lambda_{ict} + \sum_{s=-2}^{27} (\gamma_{1s} + \gamma_{2s} SL_c)\, 1[\tau_t = s] + \upsilon_{ct} \qquad (1)$$

where G_{ct} is the daily growth rate of confirmed COVID-19 cases, hospitalizations or deaths in canton c at time t; SL_c is the share of people speaking a Latin language in a canton (which varies across cantons and even within the German and the Latin language areas); X_c is a vector of controls at canton level that are likely to be correlated with the spread of diseases (e.g., population density, hospital beds per capita and the share of people aged 65+); dow_t is a vector of dummies for the days of the week, which

[6] An alternative would be to use of the log of new cases, hospitalizations and deaths, but the large number of zeros, especially for deaths and hospitalizations, prevents us from using it.

CEPR PRESS
Covid Economics 33, 30 June 2020: 163-185

is meant to take into account the presence of large variations in case reporting across the week; λ_{ict}, is a set of dummies for the days passed from the canton-specific start of the epidemic, which is measured as the occurrence of the first 10 cases (or, as robustness, to an incidence of at least two cases per 100'000 inhabitants). The dynamic effect of the policy across language regions is measured including a large set of dummies around the start of the policy on March 16 interacted with SL. Specifically, $1[A]$ is the indicator function of the event A and τ_t denotes the "event time", defined so that $\tau_t = 0$ on March 16.[7] Although this is a dynamic difference in difference specification, the coefficients of interest are both γ_{1s}, the sequence of event coefficients when the share of Latin speaking people is equal to zero, and γ_{2s}, the differential effect of these events when the share of Latin speaking people in a canton increases. In other words, given the potential differences in compliance across language regions observed in mobility data, the objective is to estimate the heterogeneity in the effect of the policy between language regions and the potential anticipation effects, especially in Latin cantons. Since it is difficult to predict precisely when the policy manifests its effects, the language specific average growth rate between three and five days before the policy implementation is used as reference period.

It is worth noting that the different language regions can also deviate —through the term SL— from the common growth rate of the disease relative to its (canton-specific) starting date, which is estimated using a full set of relative time fixed effects. This is meant to take into account that the incidence of the disease, and the speed at which it spreads, might vary across language groups. Point estimates are very similar if we include canton fixed effects or run separated estimates by language regions (German vs. Latin cantons) but, given the relatively small number of observations, we prefer to present the results of this model which is sufficiently rich to capture the relevant heterogeneity. As robustness check, we also estimate the model using the log of the cumulative number of cases, hospitalizations and deaths as dependent variables (results reported in the Appendix). Unfortunately, given the very different levels across language regions, we can run this model only by main language spoken in a canton, without fully exploiting the variation across cantons in the share of people speaking a Latin language.

5.2 Results

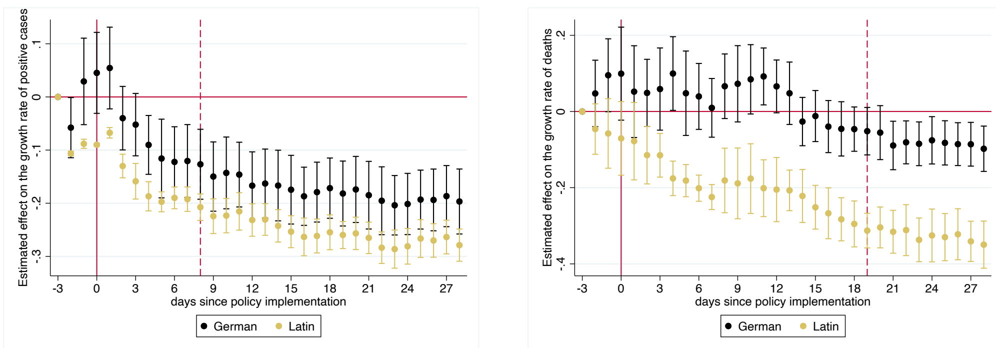
In Figure 6 we report the event study estimates of the effect of the lockdown on the growth rate of confirmed cases and deaths for each language group as described in equation (1), namely the series of γ_{1s} coefficients, which measure the dynamic effect of the policy for German cantons, and the sum of γ_{1s} and γ_{2s}, which measures the same effect for Latin cantons. We leave in the Appendix (Figure A.2) the results for the growth rate of hospitalizations, which are consistent with the other two outcomes. To ease the interpretation of the figure, for each outcome we draw a solid vertical red line at zero,

[7] The last event dummy, $\tau_t = 27$, includes all the calendar days in the observation window since the 27^{th} day after the policy implementation.

Covid Economics 33, 30 June 2020: 163-185

which represents the day of the lockdown implementation, and a vertical dashed line corresponding to the "expected lag effect" of the policy, namely the sum of the median incubation period (5 days) and the median time between symptoms onset and test, hospitalization or death. Since this is only an expected time, the lagged effects of the policy might show up earlier or later, but still represents a useful reference point.

Figure 6: Estimated effect of the policy by language group



(a) Confirmed cases

(b) Deaths

Notes: This figure shows the estimated impact of the policy on the daily growth rate of confirmed cases and deaths at canton level. For each language group, we report the estimated effect of a set of day dummies from 3 days before to 28 days after the policy announcement (and their 95% confidence intervals) with reference period the average growth rate 3–5 days before the policy implementation (March 12–14) by language group. The regressions also accounts for the canton-specific timing of the epidemic, including a full set of dummies for each day passed since the first 10 cases in a canton, dummies for the day of the week and the age distribution of cases in 10-year bins, population density, number of beds per capita, and dummies that account for those cantonal policies that anticipate some federal regulations or implement more stringent regulations. The solid vertical line indicates the day of the policy implementation while the dashed vertical line indicates when the policy announcement is expected to show its effects summing up the median incubation period (5 days) to the median time from symptoms to test or death.

Consistent with the evidence from mobile phone data, we find a quite clear anticipation effect for Latin cantons, namely a reduction in the growth rate of both outcomes that occurs already a few days after the lockdown. This is especially true for confirmed cases, where the expected time lag between the policy and its effect is relatively short. Specifically, at the time of the policy implementation the growth rate of confirmed cases is already 10% lower than the growth rate estimated for the reference period (3–5 days before the policy). However, the reduction in the growth rate substantially increases as the policy manifests its reasonable lag effect, achieving a total reduction that reaches almost 30%. For German cantons, there is no clear evidence of large anticipation effects, but we still estimate a substantial reduction in the growth rate of confirmed cases that goes somewhat in parallel with the effect estimated for Latin cantons reaching –20% two weeks after the policy implementation.

In the case of deaths, we again estimate an anticipation effect of the policy in Latin cantons, with

Covid Economics 33, 30 June 2020: 163-185

evidence of a reduction in the growth rate of deaths already 3–4 days after the policy, which is very unlikely considering a median expected lag of 19 days. On the opposite, for German cantons there is evidence of a significant reduction only 20 days after the policy. Differently from confirmed cases, the difference between the two language groups widens over time. More generally, the estimated reduction after three weeks is about 10% for German cantons while it is over 35% for Latin cantons. Even though the anticipation effect can explain part of the reduction in Latin cantons, we still observe a steeper reduction starting two weeks after the policy.

As robustness check, we show that our results are robust when we use the log of cumulative number of cases, hospitalizations or deaths as dependent variables (Figure A.3 in the Appendix). We also estimate the model including two time varying dummies that account for the fact that two cantons (Ticino and Geneva) anticipated by a few days the school closures and the ban on gatherings of more than five people but results are unchanged.

All in all, the empirical evidence reported so far shows a clear connection between the mobility data reported in Figure 5 and the evolution of the disease around the lockdown. People in Latin cantons anticipated the lockdown by reducing their movements, which explains part of the reduction in the spread of the disease. At the same time, the policy seems to accelerate this negative trend, explaining why the estimated reduction in the spread of the disease is not only earlier but also larger in Latin cantons.

Given the largest pre-policy incidence of the disease in Latin cantons, it would be natural to explain the large heterogeneity in policy compliance (and effectiveness) as the result of different risk perception between the two language groups. However, when we include the cumulated incidence of deaths at May 13 fully interacted with the daily dummies to account for risk perception, at least 50% of the language difference in the effect of the policy remains unexplained (Figure A.4 in the Appendix) and only the anticipation effect in Latin cantons seems to be largely reduced.

6 Conclusions

The vast majority of models used so far to predict the evolution of the disease, its fatality rate and the effect of social distancing measures are based on strong assumptions about the spread of the disease, such as constant attack rate across age groups and full policy compliance. This might lead to poor predictions and ineffective policies. A better understanding of the role of social interactions and an evaluation of the factors that contribute to the effectiveness of social distancing measures should help the policy makers to address the potential development of future waves of the infection and so its containment. The evidence from previous pandemics, especially the Spanish Flu, shows that the ability of virus containment has long term consequences on both population health (Almond, 2006) and economic activity (Barro et al., 2020; Jordà et al., 2020). This work demonstrates the importance of cultural differences in inter-generational contacts to explain the spread of the disease among the most vulnerable part of the population. It also shows the presence of a large cultural gradient in policy compliance that can be only partially explained by differences in risk perception. Different cultural attitudes towards government mandates might explain this gradient but it is not possible to exclude alternative explanations. Further research is therefore needed to better point out the mechanism which is needed to develop the optimal policy response.

Covid Economics 33, 30 June 2020: 163-185

CEPR PRESS

References

Alesina, A., Giuliano, P., 2015. Culture and institutions. Journal of Economic Literature 53 (4), 898–944.

Almond, D., 2006. Is the 1918 influenza pandemic over? long-term effects of in utero influenza exposure in the post-1940 us population. Journal of political Economy 114 (4), 672–712.

Barro, R. J., Ursúa, J. F., Weng, J., 2020. The coronavirus and the great influenza pandemic: Lessons from the spanish flu for the coronaviruss potential effects on mortality and economic activity. Tech. rep., National Bureau of Economic Research.

Bartscher, A. K., Seitz, S., Siegloch, S., Slotwinski, M., Wehrhöfer, N., 2020. Social capital and the spread of covid-19: Insights from european countries. Covid Economics 30, 63–82.

Belloc, M., Buonanno, P., Drago, F., Galbiati, R., Pinotti, P., 2020. Cross-country correlation analysis for research on covid-19.
URL https://voxeu.org/article/cross-country-correlation-analysis-research-covid-19

Börsch-Supan, A., 2019. Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6. Release version: 7.0.0. Munich: SHARE-ERIC. Data set.

Dowd, J. B., Andriano, L., Brazel, D. M., Rotondi, V., Block, P., Ding, X., Liu, Y., 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.

Egger, P. H., Lassmann, A., 2015. The causal impact of common native language on international trade: Evidence from a spatial regression discontinuity design. The Economic Journal 125 (584), 699–745.

Engle, S., Stromme, J., Zhou, A., 2020. Staying at home: mobility effects of covid-19. Available at SSRN 3565703.

Eugster, B., Lalive, R., Steinhauer, A., Zweimüller, J., 2011. The Demand for Social Insurance: Does Culture Matter? The Economic Journal 121 (556), F413–F448.

Eugster, B., Lalive, R., Steinhauer, A., Zweimüller, J., 2017. Culture, work attitudes, and job search: Evidence from the swiss language border. Journal of the European Economic Association 15 (5), 1056–1100.

Eugster, B., Parchet, R., 2019. Culture and taxes. Journal of Political Economy 127 (1), 296–337.

Frey, C. B., Chen, C., Presidente, G., 2020. Democracy, culture, and contagion: Political regimes and countries responsiveness to covid-19. Covid Economics 18, 1–20.

Gentili, E., Masiero, G., Mazzonna, F., 2017. The role of culture in long-term care arrangement decisions. Journal of Economic Behavior & Organization 143, 186–200.

Jordà, Ò., Singh, S. R., Taylor, A. M., 2020. Longer-run economic consequences of pandemics. Covid Economics 1, 1–15.

Kuhn, M., Bayer, C., 2020. Intergenerational ties and case fatality rates: A cross-country analysis. IZA Discussion Paper 13114.

Litwin, H., Stoeckel, K. J., 2016. Social network, activity participation, and cognition: A complex relationship. Research on aging 38 (1), 76–97.

Painter, M., Qiu, T., 2020. Political beliefs affect compliance with covid-19 social distancing orders. Available at SSRN 3569098.

Steinhauer, A., 2018. Working moms, childlessness, and female identity. CEPR Discussion Paper No. DP12929.

Van Bavel, J. J., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., Crockett, M. J., Crum, A. J., Douglas, K. M., Druckman, J. N., et al., 2020. Using social and behavioural science to support covid-19 pandemic response. Nature Human Behaviour, 1–12.

WHO, 2020. Report of the who-china joint mission on coronavirus disease 2019 (covid-19).

Covid Economics 33, 30 June 2020: 163-185

CEPR PRESS
Covid Economics 33, 30 June 2020: 163-185

A Appendix

Econometric modeling for the evolution of the relationship between language and the spread of the virus

To account for the differential timing of the epidemic across cantons, for the calendar date and for other potential confounders at canton level, namely population density, hospital beds and the share of people aged 65+, we use in Table A.1 an empirical model to systematically study the evolution of the relationship between language and the spread of the virus across Swiss cantons before the implementation on the federal lockdown on March 16 (to avoid the confounding effect of the lockdown measures). More specifically, we estimate the following model:

$$Log(Y_{ct}+1) = \alpha_0 + \alpha_1 SL_c + X_c\alpha_2 + \delta_t + \sum_{i=2}^{n} \lambda_{ict} + \epsilon_{ct} \tag{2}$$

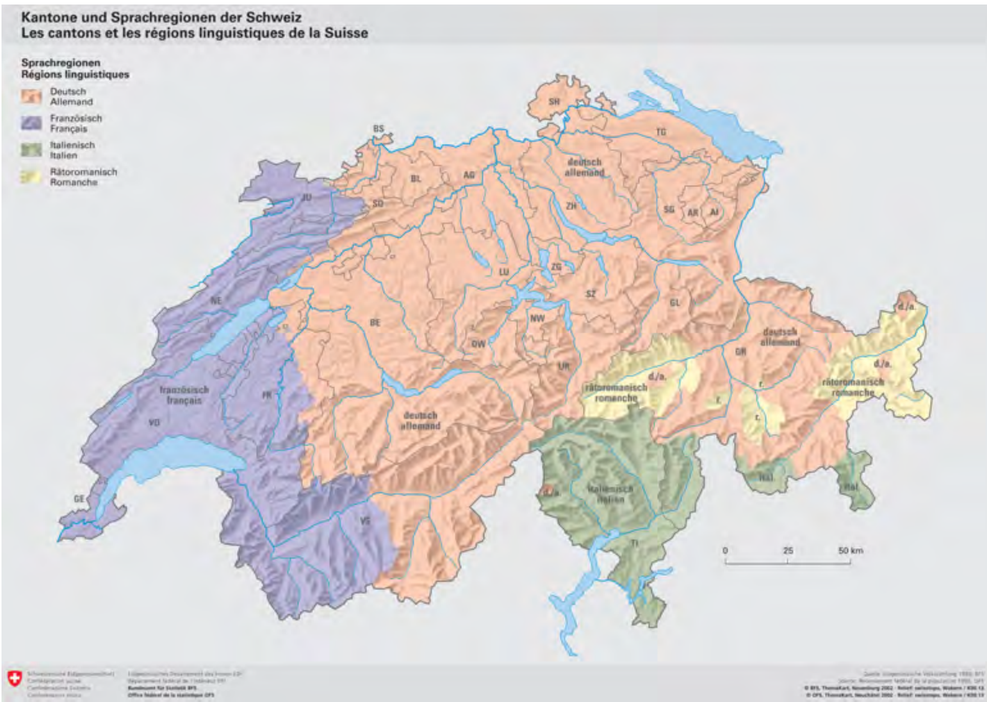
where Y_{ct} is the cumulative number of confirmed COVID-19 cases, hospitalizations or deaths in canton c at calendar time t; SL_c is the share of Latin speaking people in a canton, X_c is a vector of controls at canton level that are likely to be correlated with the spread of diseases (e.g., population density, hospital beds per capita and the share of people age 65+); δ_t is a calendar date fixed effect while λ_{ict}, is a set of dummies for the time passed in days from the canton-specific start of the epidemic, which is measured as the occurrence of the first 10 cases or as an incidence of at least two cases per 100'000 inhabitants.

Except for deaths, for which we have only a few fatalities under age 65, this model is estimated separately by age groups (20-64 vs. 65+) to explicitly test whether the differential spread of the disease between language groups is larger among the elderly. Given the large differences in population size across cantons (from 5'000 to over one million), all estimates are weighted using population weights especially to take into account the very different signal to noise ratio in the outcome variable.

CEPR PRESS

Additional figures and tables

Figure A.1: Linguistic areas across Switzerland



Notes: Colors correspond to different linguistic areas as in the legend in top left corner.
Sources: Swiss statistical office. ©OFS, ThemaKart

Covid Economics 33, 30 June 2020: 163-185

CEPR PRESS

Covid Economics 33, 30 June 2020: 163-185

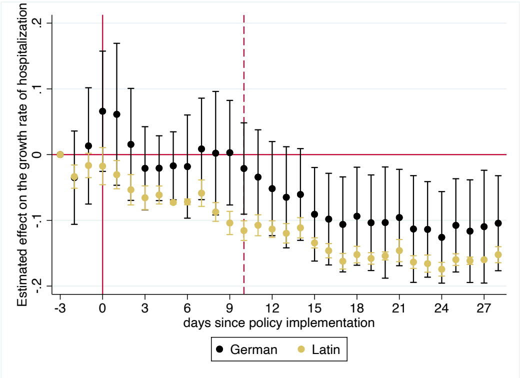
Table A.1: OLS regression of the log of confirmed cases, hospitalizations and deaths per 100'000 people on the share of people speaking a Latin language in a canton (up to May 16, 2020).

	Confirmed cases age 0–64			
Latin share	0.720	0.664***	0.652	0.407
	(0.461)	(0.112)	(0.481)	(0.458)
	Confirmed cases age 65+			
Latin share	1.536***	1.503***	1.481***	1.228***
	(0.350)	(0.251)	(0.361)	(0.323)
	Hospitalizations age 0–64			
Latin share	0.628	0.599**	0.528	0.339
	(0.510)	(0.242)	(0.518)	(0.502)
	Hospitalizations age 65+			
Latin share	1.757***	1.742***	1.711***	1.547***
	(0.369)	(0.285)	(0.374)	(0.340)
	Deaths			
Latin share	0.739***	0.756***	0.662***	0.266**
	(0.173)	(0.182)	(0.170)	(0.128)
Time control for the number of days since 10 cases:				
Exponential	Yes	No	No	No
fixed effects	No	Yes	No	No
Time control for the number of days since 2 cases per 10'000:				
Exponential	No	No	Yes	No
fixed effects	No	No	No	Yes
N	185	185	185	185

Notes: The table shows the estimated effect of the share of people speaking German in a canton on the incidence of positive cases, hospitalizations and deaths per 100'000 people. For positive cases and hospitalizations the incidence is split in two age groups (0–64, and 65+). The regression also includes controls for population density, the share of people 65+ in a canton, date fixed effects and a control for the differential timing of the epidemic across cantons which is different across columns. We use robust standard errors clustered at the canton level. Significance levels: *** < 0.01, ** < 0.05, * < 0.1.

Covid Economics 33, 30 June 2020: 163-185

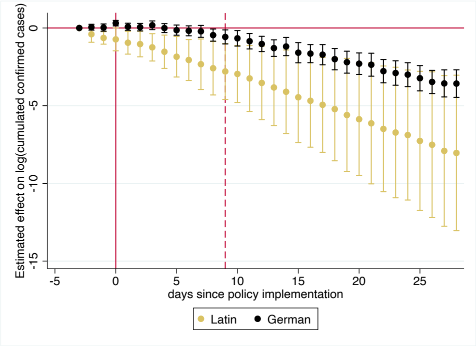
Figure A.2: Estimated effect of the policy on the growth rate of hospitalization by language group



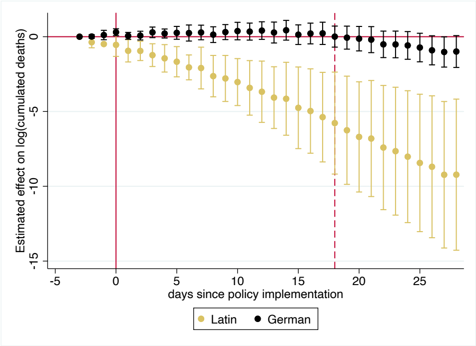
Notes: The figure shows the estimated impact of the policy on the daily growth rate of hospitalized cases at canton level. For each language group, we report the estimated effect of a set of day dummies from 3 days before to 28 days after the policy announcement (and their 95% confidence intervals) with reference period the average growth rate 3–5 days before the policy implementation (March 12–14) by language group. The solid vertical line represents the day of the policy implementation while the dashed vertical line indicates when the policy announcement is expected to show its effects summing up the median incubation period (5 days) to the median time from symptoms to hospitalization. The regression also accounts for differential timing of the epidemic across cantons, including a full set of dummies for each day passed since the first 10 cases in a canton, dummies for the day of the week and the age distribution of cases in 10-year bins, population density, number of beds per capita, and dummies that account for those cantonal policies that anticipate some federal regulations or implement more stringent regulation.

CEPR PRESS

Figure A.3: Estimated effect of the policy by language group, log specification



(a) Confirmed cases



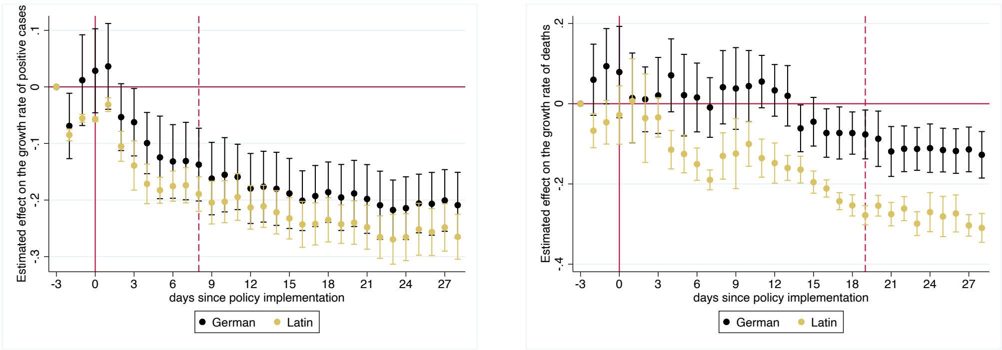
(b) Deaths

Notes: The two sub figures show the estimated impact of the policy on the daily log cumulated confirmed cases (left figure) and deaths (right figure) at canton level. For each language group, we report the estimated effect of a set of day dummies from 3 days before to 28 days after the policy announcement (and their 95% confidence intervals), with reference period the average growth rate 3–5 days before the policy implementation (March 12–14) by language group. The solid vertical line indicates the day of the policy implementation while the dashed vertical line indicates when the policy announcement is expected to show its effects summing up the median incubation period (5 days) and the median time from symptoms to test or death. The regression also accounts for differential timing of the epidemic across cantons, including a full set of dummies for each day passed since the first 10 cases in a canton, dummies for the day of the week and the age distribution of cases in 10-year bins, population density, number of beds per capita, and dummies that account for those cantonal policies that anticipate some federal regulations or implement more stringent regulation.

Covid Economics 33, 30 June 2020: 163-185

Covid Economics 33, 30 June 2020: 163-185

Figure A.4: Estimated effect of the policy by language group, controlling for pre-policy incidence of deaths



(a) Confirmed cases

(b) Deaths

Notes: This figure shows the estimated impact of the policy on the daily growth rate of confirmed cases and deaths at canton level controlling for pre-policy incidence of deaths. For each language group, we report the estimated effect of a set of day dummies from 3 days before to 28 days after the policy announcement (and their 95% confidence intervals) with reference period the average growth rate 3–5 days before the policy implementation (March 12–14) by language group. The regressions also accounts for the canton-specific timing of the epidemic, including a full set of dummies for each day passed since the first 10 cases in a canton, dummies for the day of the week and the age distribution of cases in 10-year bins, population density, number of beds per capita, and dummies that account for those cantonal policies that anticipate some federal regulations or implement more stringent regulation. Different from Figure 6 it also includes the cumulated incidence of deaths at cantonal level fully interacted with the 31 day dummies. The solid vertical line indicates the day of the policy implementation while the dashed vertical line indicates when the policy announcement is expected to show its effects summing up the median incubation period (5 days) to the median time from symptoms to test or death.

CEPR PRESS

Table A.2: Percentage change in Google community movements after lockdown

	retail & recreation	parks	work	transport	residential
Latin cantons	-72.22	-11.85	-50.06	-55.36	22.41
German cantons	-62.65	15.24	-37.34	-46.51	17.51
Selected Cantons					
Aargau (G)	-59.37	59.39	-35.70	-39.60	16.87
Basel City (G)	-73.90	-30.67	-44.33	-59.53	20.85
Bern (G)	-64.93	13.80	-37.20	-50.37	16.80
Geneva (F)	-73.00	-12.00	-58.17	-69.13	25.96
Lucerne (G)	-63.73	-3.17	-34.87	-47.53	17.33
St. Gallen (G)	-61.50	51.27	-32.00	-34.87	15.37
Ticino (I)	-80.67	-35.60	-55.03	-60.37	24.93
Valais (F)	-71.80	-33.37	-46.00	-54.83	19.07
Vaud (F)	-68.97	3.10	-52.37	-56.00	23.20
Zurich(G)	-66.73	43.03	-45.93	-55.07	20.70

Notes: The table shows the average percentage change in movements towards places such as retail and recreation, parks, transit stations, workplaces, and residential, by language or selected cantons Data are extracted by Google community movements data website. In parentheses the main language with G=German, F=French, I=Italian.

Covid Economics 33, 30 June 2020: 163-185