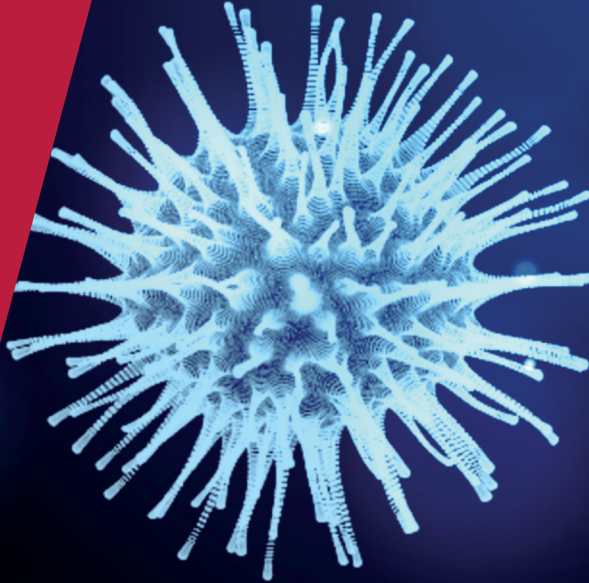


**CENTRE FOR
ECONOMIC
POLICY
RESEARCH**

CEPR PRESS



COVID ECONOMICS
VETTED AND REAL-TIME PAPERS

ISSUE 83
2 JULY 2021

**UNEMPLOYMENT INEQUALITY IN
SWEDEN**

Pamela Campa, Jesper Roine
and Svante Strömberg

ESG STOCKS RESILIENCE

Gianfranco Gianfrate, Tim Kievid
and Mathijs van Dijk

WHAT DRIVES BAILOUTS?

Jean-Marie Meier and Jake Smith

RETAIL INVESTORS

Katrin Tinn

CONSUMPTION AND INEQUALITY

Christopher D. Cotton, Vaishali Garga
and Justin Rohan

VARIANTS VS. VACCINATION

David Turner, Balázs Égert,
Yvan Guillemette and Jarmila Botev

RISK AND WILLINGNESS TO PAY

Sonja Warkulat, Sebastian Krull,
Regina Ortmann, Nina Klocke
and Matthias Pelster

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

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

Guido Alfani, Bocconi University and CEPR

Franklin Allen, Imperial College Business School and CEPR

Michele Belot, Cornell University and CEPR

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

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

Pierre-Yves Geoffard, Paris School of Economics and CEPR

Francesco Giavazzi, Bocconi University and CEPR

Christian Gollier, Toulouse School of Economics and CEPR

Timothy J. Hatton, University of Essex and CEPR

Ethan Ilzetzki, London School of Economics and CEPR

Beata Javorcik, EBRD and CEPR

Simon Johnson, MIT and CEPR

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

Tom Kompas, University of Melbourne and CEBRA

Miklós Koren, Central European University and CEPR

Anton Korinek, University of Virginia and CEPR

Michael Kuhn, International Institute for Applied Systems Analysis and Wittgenstein Centre

Maarten Lindeboom, Vrije Universiteit Amsterdam

Philippe Martin, Sciences Po and CEPR

Warwick McKibbin, ANU College of Asia and the Pacific

Kevin Hjortshøj O'Rourke, NYU Abu Dhabi and CEPR

Evi Pappa, European University Institute and CEPR

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

Richard Portes, London Business School and CEPR

Carol Propper, Imperial College London and CEPR

Lucrezia Reichlin, London Business School and CEPR

Ricardo Reis, London School of Economics and CEPR

Hélène Rey, London Business School and CEPR

Dominic Rohner, University of Lausanne and CEPR

Kjell G. Salvanes, Norwegian School of Economics 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.

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

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

Covid Economics

Vetted and Real-Time Papers

Issue 83, 2 July 2021

Contents

Editor's foreword <i>Charles Wyplosz</i>	v
Foreword from CEPR's President <i>Beatrice Weder di Mauro</i>	vii
Unemployment inequality in the pandemic: Evidence from Sweden <i>Pamela Campa, Jesper Roine and Svante Strömberg</i>	1
On the resilience of ESG stocks during COVID-19: Global evidence <i>Gianfranco Gianfrate, Tim Kievid and Mathijs van Dijk</i>	25
The COVID-19 bailouts <i>Jean-Marie Meier and Jake Smith</i>	54
Everyone is a stock trader now: Retail investors and Covid-19 <i>Katrin Tinn</i>	88
Consumption spending and inequality during the Covid-19 pandemic <i>Christopher D. Cotton, Vaishali Garga and Justin Rohan</i>	116
The tortoise and the hare: The race between vaccine rollout and new COVID variants <i>David Turner, Balázs Égert, Yvan Guillemette and Jarmila Botev</i>	154
COVID-19 reporting and willingness to pay for leisure activities <i>Sonja Warkulat, Sebastian Krull, Regina Ortmann, Nina Klocke and Matthias Pelster</i>	183

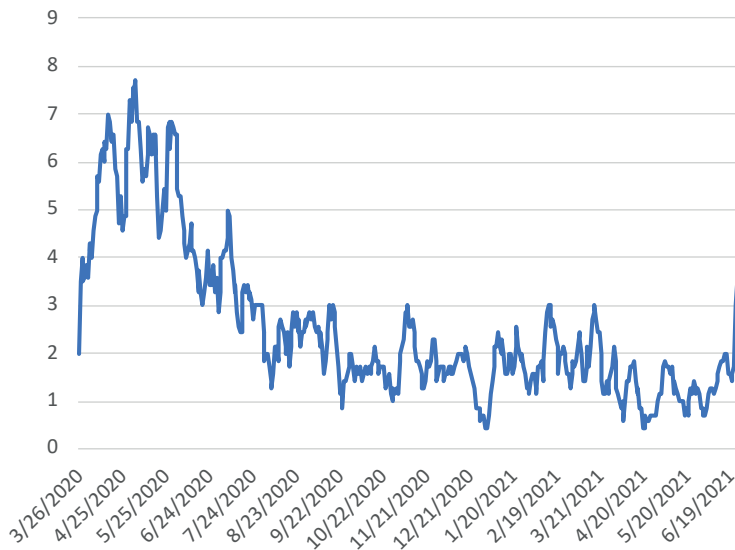
Editor's foreword

This is the last issue of *Covid Economics, Vetted and Real Times Papers* that I am publishing as Editor. Viewed from my desk, it has been a once-in-a-lifetime experience as economists from all over the world rushed to try and help understand a once-in-a-century (hopefully) pandemic. To be sure, the field of economic research on pandemics pre-existed the coronavirus, but it was relatively small. Within days of the onset of Covid, hundreds of economists dropped their ongoing work. They relied on well-established theories and techniques and on quickly expanding real-time data to fill a vacuum in the well-established field of epidemiology. Epidemiologists tracked viruses, economists looked at people behaviour and government responses.

The launch of *Covid Economics, Vetted and Real Times Papers* was intended to encourage prompt dissemination of this new research. There was a sense of urgency as the pandemic was spreading. There also was a sense of duty as better knowledge was bound to help policymakers faced with huge responsibilities in the midst of considerably uncertainty. Instructed by the experience of pre-prints in other scientific fields, CEPR decided to propose a place where sound papers were collected and made immediately available. Vetting replaced refereeing, with reviewers asked to quickly determine whether a paper was correctly executed, innovative and relevant.

Since the first issue, published on 3 April 2020, we have received 1,176 submissions, out of which 511 papers have been collected in 83 issues. Figure 1, which displays average daily submissions, shows a surge early on until the summer of 2020, followed by a slow decline and a stationary inflow afterwards. Submissions came from all over the world, prepared by seasoned economists from highly reputed centres to students from less well-known departments. The issues are freely available online.

The field of pandemic economics has now become mature. The low-hanging fruits have been harvested, few really new ideas and results emerge. The next step is to deepen this well of knowledge, which requires proper refereeing by specialists. This is not the time for emergency anymore, even though the pandemic is far from being contained around the world. I have concluded that *Covid Economics, Vetted and Real Times Papers* has served its purpose. CEPR is thinking about next steps, as explained in this issue by Beatrice Weder di Mauro.

Figure 1 Number of daily submissions (7-day moving average)

Nothing would have been possible without the Editorial Board. Some 50 colleagues committed to review papers within 48 hours, and they often have been asked to do so weekly. To my amazement, they have delivered. Beatrice Weder di Mauro, CEPR President, Richard Baldwin, the Editor of VoxEU and Tessa Ogden, the CEO of CEPR, had the initial idea and entrusted me to make it happen. CEPR's staff, especially Mariolina Ciccone, Anil Shamdasani and Sophie Roughton, created and managed the process from submissions to final publication – an incredible feat.

I am proud of being an economist. As a profession, we have responded to the challenge with ingenuity and imagination, drawing on our vast array of tools. I suspect that, collectively, we never created so much new knowledge in so short a time.

Charles Wyplosz
Editor

Foreword from CEPR's President

In recent months CEPR has published *Covid Economics: Vetted and Real Time Papers* less frequently as the pandemic has eased and researchers are returning to their original fields of study. Charles Wyplosz is stepping down as Editor and we have decided to cease publication in its present form.

Although this will be the last issue of *Covid Economics*, we recognise that the pandemic is not over and that research on these topics must go on. CEPR will continue to contribute to this field by creating a Research and Policy Network on Pandemic Economics later this year, and will also publish special issues of *Covid Economics* drawing the lessons from this crisis.

Charles and his board are owed an enormous debt of gratitude for this incredible service they have given to the economics profession and to the world.

Beatrice Weder di Mauro
President, CEPR

Unemployment inequality in the pandemic: Evidence from Sweden

Pamela Campa,¹ Jesper Roine² and Svante Strömberg³

Date submitted: 23 June 2021; Date accepted: 30 June 2021

Using the full population of registered unemployed individuals in Sweden, we study the unequal labor market impact of Covid-19 depending on gender, wage, age, and country of birth. Also, having very detailed data on the occupation of the unemployed, we can study inequalities both across and within occupations. We find that two demographic factors are associated with higher unemployment in the pandemic: being young and being foreign-born. Gender, however, does not seem to play a big role in the Swedish context, likely due to both institutional factors and labor market structure, as well as policy measures such as not closing schools and day-care facilities. We also find a clear wage gradient with lower-paying jobs having higher unemployment risk. Our results confirm previous findings on the most vulnerable being hit the hardest, but at the same time emphasize the importance of country specific studies to understand the economic impacts of the pandemic.

¹ Assistant Professor, Stockholm Institute of Transition Economics.

² Professor of Economics, Stockholm School of Economics and Deputy Director at Stockholm Institute of Transition Economics.

³ Research Assistant, Stockholm Institute of Transition Economics.

Copyright: Pamela Campa, Jesper Roine and Svante Strömberg

1 Introduction

The ongoing Covid-19 pandemic has affected the health of millions of people worldwide. It has also had an enormous impact on economic and living conditions through government policies aimed at containing the spread of the infection. While, at the onset of the pandemic, government officials, mainstream media, and even celebrities labeled COVID-19 “the great equalizer” (Mein, 2020), the reality has proven quite different, with the most vulnerable groups of the population appearing to be the most harmed by both the health and the economic crises.¹

Understanding which segments of society have borne the brunt of the economic crisis fuelled by the COVID-19 pandemic is crucial as governments worldwide discuss and implement eased restrictions and post-pandemic recovery plans. A broad policy discussion also exists around how national government and international organizations can best prepare for potential future pandemics; while the focus of such discussion is on the health crisis, a thorough understanding of the economic effects of the containment policies can help to solve more effectively the trade-offs implied by the goals of protecting health while preserving living standards. For instance, it is crucial to understand the role of school closures in, on one hand, the containment of the spread of the infection and kids’ learning outcomes, and on the other hand, the labor-market struggle of women throughout the pandemic, lamented by several sources.

In this paper we contribute to this discussion by leveraging new population-based monthly data on unemployment registrations, made available by the Swedish Public Employment Service, to study the impact of the pandemic on unemployment status in Sweden over the first wave of the health crisis (March 2020 to July 2020). The data distinguishes between 400 different occupations, allowing us to study the impact at a more detailed level than the sectoral one. We provide evidence on the extent to which the unemployment impact of the pandemic has been unequal across segments of the Swedish society, based on some characteristics that previous works have identified as indicators of vulnerability, namely gender, wage, age, and foreign-born status.

We construct a measure of unemployment impact of the pandemic, which compares the change in occupation-level unemployment risk during the first wave (March to July 2020) with respect

¹See, for instance, these articles edited by The World Economic Forum (<https://www.weforum.org/agenda/2020/10/covid-19-is-increasing-multiple-kinds-of-inequality-here-s-what-we-can-do-about-it/>), the IMF (<https://www.imf.org/external/pubs/ft/fandd/2020/09/COVID19-and-global-inequality-joseph-stiglitz.htm>).

to one month before the onset of the health crisis (February 2020) to the same change one year before, when the pandemic was neither occurring nor being anticipated. The occupation-level unemployment risk can be further disaggregated by gender, age, and foreign born status. We also complement these data with information on the average occupation-level wage in 2019, which is also disaggregated by gender. We then use the assembled dataset to answer the following questions: does the unemployment impact of the pandemic in Sweden differ systematically by gender, wage level, age-cohort, and foreign-born status? To what extent does sorting by demographics in different occupations explain differences in being affected by the pandemic impact? Our results contribute to the growing evidence that (a) the pandemic has not only exposed, but also broadened existing socio-economic inequalities in several countries worldwide, and (b) there is heterogeneity across countries in terms of which groups have been more severely affected, largely dependent on pre-existing vulnerabilities.

Our findings can be summarized as follows. First, there is large heterogeneity across occupations in the level of exposure to the pandemic. This heterogeneity mirrors the effects of the health crisis as well as the type of policy interventions that the Swedish government has adopted. Second, in a context characterized by high levels of pre-pandemic gender equality in the labor market, and by the policy choice not to close schools, we fail to find evidence of women paying a higher labor-market cost than men; if anything, the unemployment impact of the pandemic in Sweden appears to have been particularly large for some categories of male workers. Third, the wage gradient of unemployment risk, with lower-paying occupations typically suffering higher unemployment rates, has become steeper as a result of the economic crisis. Fourth, two demographics appear to be strong predictors of the unemployment impact of the pandemic in Sweden: age and foreign-born status; the increase in unemployment risk caused by the pandemic has been especially pronounced for young workers and workers born in non-EU countries, who were already more vulnerable before the pandemic.

Within the burgeoning literature on the pandemic's impact on different types and measures of inequality, a number of studies are especially related to ours, as they are focused on the labor market and study real-time data. Based on these studies, a number of patterns emerge.

First, the effect of the pandemic on increased probability of job loss appears stronger for low-skilled workers as proxied by education level (see e.g., [Adams-Prassl et al. \(2020\)](#), [Gaudecker et al.](#)

(2020), Casarico and Lattanzio (2020)). Gaudecker et al. (2020) also observe that in the Netherlands the negative education gradient has been mitigated by the government identifying some sectors of the economy as essential, since some of these sectors are characterized by a high concentration of low-educated workers. Our results on the wage-gradient are especially related to Cajner et al. (2020), who consider administrative data from the largest payroll processing company in the US, covering about 20% of total U.S. private employment, and find that employment losses in the first months of the pandemic have been disproportionately concentrated among low-wage workers.

Second, there is mixed evidence on the impact of the crisis on gender inequalities in the labor market. An early study by Alon et al. (2020), not based on real-time data, pointed to a number of reasons why the Covid-19 pandemic could be expected to have larger impacts on working women. They note that women are overrepresented in sectors of the economy most affected by social distancing measures and also point to the importance of factors such as opportunities for telecommuting and child care (closure of schools and day-care facilities leading to women taking more childcare responsibilities). Subsequent papers have painted a more mixed picture. While survey information from the UK and the US (Adams-Prassl et al., 2020) and administrative data from the U.S (Cajner et al., 2020) confirm that labor market outcomes for women have more severely deteriorated during the crisis, there is no evidence of unequal impacts by gender in Germany (Adams-Prassl et al., 2020) and Italy (Casarico and Lattanzio, 2020). Other papers find that the effect on labor-market outcomes by gender varies across contexts (see, e.g., Hupkau and Petrongolo, 2020 and Alon et al., 2021).

We contribute to these findings in a number of ways. We provide unique evidence from Sweden. As noted in previous studies (see e.g. Adams-Prassl et al., 2020), while the pandemic and the measures to contain its spread have reached virtually every country worldwide, the socio-economic impacts of the crisis, including its implications for inequality, are heterogeneous across countries. Studying inequalities during the pandemic in Sweden can provide unique insights, since the country is characterized by a relatively low level of income inequality (e.g. OECD, 2019), despite an upward trend over the past decades, as well as by high participation of women in the labor market, and high level of social inclusiveness (e.g. Gottfrieds, 2018 and OECD, 2016). Moreover, the Swedish government response to the Covid-19 crisis has been more “laissez-faire” than in many other countries (see e.g. Ellingsen and Roine, 2020 for an account of how Sweden responded to Covid-19

in comparison to the other Nordic countries). Sweden has not adopted stay-at-home orders that would have separated sectors of the economy between “essential” and “non-essential”; as a result, sectors that were typically shut down in other countries, for instance in the hospitality industry, have remained open during the first wave of the pandemic and have faced only partial limitations during the second wave. Our paper is therefore informative of whether the pandemic’s adverse socio-economic implications would have been avoided in the absence of lockdowns and similar policies, or are rather unavoidable in the presence of a massive health crisis. Finally, and crucially for the analysis of gender inequality, day-care facilities and schools below secondary level were never closed in Sweden. Recent studies, such as [Alon et al. \(2021\)](#), have shown that school closures might have been the main driver of the unequal impact of the pandemic by gender: the growing gender gap in labor market outcomes observed in some countries is indeed mostly driven by parents. Our paper complements this evidence by focusing on a context where the “childcare channel” is fixed, and we can thus separately study the importance of gender segregation across sectors of the economy, as well as of other factors that intersect with gender and determine gender-gaps in labor-market dynamics.

Our study is also unique in that we leverage a rich administrative dataset of weekly unemployment insurance (UI) claims grouped by a few demographics (gender, age, foreign-born status) and by occupational sector, defined in a very granular manner. This dataset is attractive because, as opposed to survey-based information relied upon in other papers ([Adams-Prassl et al., 2020](#)), it covers the universe of individuals who have lost their job during the pandemic. [Casarico and Lattanzio \(2020\)](#) also draw on population-level information, by accessing data on contracts, hiring and separations in Italy; since Italy has adopted a firing-freeze after the first two months of the pandemic, our study complements theirs by specifically analyzing a context where jobs were not protected by a similarly broad intervention. [Forsyth et al. \(2020\)](#) also study UI claims in the US labor market. As they note, UI claims are one of the few real-time indicators of the state of the labor market and as such have received increasing attention at the onset of the pandemic; however, information on UI claims from the US is highly aggregated, with only a few states reporting claims across industries. Instead, we observe UI claims at the occupational level, centrally aggregated for the whole national labor market. For instance, while the existing studies consider the hospitality

industry as a whole, we exploit detailed occupational information within this industry.²

Our study of the age and foreign-born status differences is also particularly interesting given the focus on the Swedish society, which is on the one hand characterized by high levels of inclusiveness, but on the other hand appears to face important challenges integrating young and foreign-born workers in the labor market; the difference in unemployment rates between native born and others, especially non-EU-born, is among the largest in the EU³ and youth unemployment is also relatively high in Sweden compared to the EU average⁴.

Overall, our contribution emphasizes how the inequality implications of the pandemic are largely dependent on both the pre-pandemic socio-economic conditions as well as the type of government response adopted. As a result, although the pandemic has reached virtually every corner of the world with similar sweeping power, academics and policy-makers should rely on context-specific studies to formulate policy proposals and actions that address the short and long-term societal consequences of the Covid-19 crisis.

The rest of the paper is organized as follows. Section 2 presents the data and our measure of unemployment impact of the pandemic. Section 3 shows the empirical evidence of impact inequalities based on a number of group classifications. Section 4 concludes.

2 Data

We use data from the registry of unemployed individuals kept by the Swedish Public Employment Service (Arbetsformedlingen), the government agency responsible for the functioning of the Swedish labor market, for the periods February to July 2019 and 2020. The incentives for laid-off individuals to register with the Employment Service are high, since the registration is directly connected to the right to claim various (relatively generous) unemployment benefits. As such, the data arguably includes a large share of employees who lost their job over the period studied. Based on the high incentives to register as unemployed, we also assume that the probability of registering does not differ substantially across the segments of the population that we compare. Nevertheless, the

²In concrete terms, we can, for instance, distinguish between chefs and waiters. Such a distinction might matter if there was a massive decline in the number of people who visited restaurants, but an increase in take-away orders.

³https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Migrant_integration_statistics_%E2%80%93_labour_market_indicators

⁴https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Participation_of_young_people_in_education_and_the_labour_market#Country_differences

probability of registering is expected to be lower for self-employed individuals.

The population-wide coverage is the main advantage of our data vis-à-vis the survey information used in many recent studies of the labor market throughout the pandemic (other studies using administrative data are Casarico and Lattanzio (2020) examining the Italian labor market, and Forsyth et al. (2020), who analyze the US case). Importantly, however, our data currently do not include furloughed workers, a significant group especially in the very early stages of the pandemic. Moreover, we do not observe information on number of hours worked, implying that we are unable to capture any labor-market adjustment on the intensive margin.

During the period from February to July 2019 the average number of individuals who registered monthly with the Swedish Public Employment Service was approximately 350,000, against an active population of roughly 5.5 million.^{5,6} We consider these monthly data, also for 2020 (when the corresponding number amounted to 447,00 individuals), grouped by 4-digit occupational classification (observing 400 occupations at this level). Each occupational group is further broken down by sex, age, and foreign-born status (specifically, *EU-born*, *EU no Sweden - born*, and *Outside EU - born*). We then divide the number of registered unemployed for the relevant group by the respective average number of individuals employed in 2017 and 2018.⁷ Throughout the analysis we refer to the ratio of registered unemployed in 2019 or 2020 over the group-level employment in the previous two years as the “unemployment risk”.

We also exploit information on the average wage by occupational group and gender in 2019, as reported by *Medlingsinstitutet* and publicly available at Statistics Sweden. This measure, although not being at the individual level, allows us to develop a relatively precise proxy of wages by occupation.

Measuring the unemployment impact of the pandemic With the data described above, we build the following occupation-level measure of change in unemployment risk between one month

⁵Source: Statistics Sweden, Population Statistics.

⁶Notice that, based on the 2019 Statistics Sweden Labor Force Survey, which is typically used to compute official statistics on unemployment rate, the number of unemployed persons in Sweden throughout 2019 was 373,000, which is largely comparable to the information from the Public Employment Agency that we rely on.

⁷Data on employed individuals by occupation and demographic groups are made available by Statistics Sweden.

before the pandemic onset (February 2020) and the five months after.

$$\Delta(u_{feb}^{mar-jul})_{i,2020} = \bar{u}_{i,mar-jul,2020} - u_{i,feb,2020} \quad (1)$$

where $u_{i,m,y}$ is the occupation-level monthly unemployment risk, i.e. the number of workers in 4-digit occupational sector i who registered as unemployed in month m of year y over the average number of employed in the same occupation in 2017 and 2018 (data available from Statistics Sweden). $\bar{u}_{i,mar-jul,2020}$ is the average unemployment risk during the months from March to July of 2020. Put it simply, $\Delta(u_{feb}^{mar-jul})_{i,2020}$ is an occupation-level indicator of the change in unemployment risk during the first five months of the pandemic as compared to one month before its onset.

Then, we consider the same measure for 2019, namely

$$\Delta(u_{feb}^{march-jul})_{i,2019} = \bar{u}_{i,mar-jul,2019} - u_{i,feb,2019} \quad (2)$$

and we use it to seasonally adjust the 2020 unemployment change. The resulting measure of *unemployment impact* of the pandemic is thus based on the change in unemployment at the occupational level between the months March to July *versus* February 2020, i.e. between five months after the start of the pandemic and the month before its onset, as compared to the equivalent change the year before:

$$Unemployment\ impact_i = \Delta(u_{feb}^{march-jul})_{i,2020} - \Delta(u_{feb}^{march-jul})_{i,2019}. \quad (3)$$

We thus account for seasonal factors by differencing out the unemployment change during the same months of 2019, when the pandemic was neither occurring nor being anticipated.⁸

The average unemployment impact was of 2.5 p.p. over the period studied. Figure A1 in the

⁸By accounting for seasonality, we aim to capture the change in unemployment risk *caused* by the pandemic. A concern is that the labor market conditions in some sectors would have evolved differently over March to July 2020 as compared to the same period of 2019, even in the absence of the pandemic, or that there were sectoral-specific shocks in February 2020, due to other determinants of unemployment incidence. We note, however, that the unprecedented shock caused by the pandemic is likely larger than many other labor-market shocks that could have happened during this period.

Appendix also shows the distribution of the unemployment impact across occupations, which is relatively skewed. The change in unemployment risk was minor for many occupations, but few occupations were severely negatively impacted by the pandemic, with increases in unemployment risk of up to 35 p.p. Table 1 below lists the ten most-affected occupations. We also report the number of employed in each occupation, as well as the share of female employment.

Table 1: Occupations by unemployment impact, ten most impacted

Occupation	Number of employed	Share female	Unemployment impact
Aircraft pilots and related associate professionals	1651	.076	.308
Choreographers and dancers	372	.570	.246
Masseurs and massage therapists	849	.812	.222
Cabin crew	2497	.785	.192
Travel guides	932	.580	.176
Bartenders	5531	.481	.168
Visual artists and related artists	595	.476	.160
Beauticians and related workers	1084	.967	.152
Chefs and sous-chefs	2570	.202	.151
Pizza makers and fast food preparers	3828	.179	.140
Average	10781	.49	.019

Note: This table shows the ten occupations with the highest unemployment impact of the pandemic. The unemployment impact is the change in unemployment risk during the first months of the pandemic (March to July 2020) versus one month before its onset (February 2020), as compared to the same change one year before (see definition in equation 3). We also report employment in 2017-18 in each occupation, and the respective share of women among the employed. The last row shows the mean of each variable in the sample. Many of the impacted occupations are relatively small.

The list of most-affected occupations reflects the evolution of the pandemic in Sweden. While the country did not adopt a regulated lockdown of the types observed in many other European countries, some activities were significantly reduced. Tourism from and to other countries was basically halted, due to a combination of both the Swedish government's recommendation to avoid unnecessary travel and border closures between European countries; this is reflected in workers in the aviation and tourism sector experiencing a dramatic increase in unemployment risk. Although in the first phase of the pandemic restaurants and bars were allowed to remain open, with only limited restrictions,⁹ data from mobile-phone users reveal that nonetheless mobility declined significantly (see [Dahlberg et al., 2020](#)), likely resulting in lower number of visits to bars and restaurants. In comparison, however, the cultural sector might have been even more affected, since several cultural

⁹Initially, the only restriction applied to restaurants and bars was the obligation to offer only table service. Starting from the Fall of 2020, however, more severe limitations regarding the maximum number of customers admitted and earlier closing times were adopted.

activities were virtually impeded due to the public order against even relatively small indoor events. The recommendation that individuals practice social distance has also likely reduced exposure to activities that involve physical proximity.

3 Empirical analysis: The unequal impact of the pandemic

In this Section we study whether the unemployment impact of the first months of the pandemic differs systematically by workers' gender, wage level, foreign-born status and age group, focusing on the Swedish labor market.

The unemployment impact of the pandemic by gender We estimate some variations of the following equation:

$$unemployment\ impact_{ig} = \alpha + \beta Female_g + \eta_{ig} \quad (4)$$

where the unit of observation is occupation-by-gender; *unemployment impact* is thus as described in (3), disaggregated by gender.¹⁰ *Female_g* indicates that the occupation-by-gender group considered is female (e.g., female bartenders as opposed to male bartenders). Since the occupation-by-gender groups differ substantially in size, we weight observations by the number of employed in the respective group, using the average employment between 2017 and 2018. We allow for arbitrary correlation of the error term η within 3-digit occupations. Practically speaking, the coefficient β describes the difference in the unemployment impact of the pandemic between women and men in the Swedish labor market, based on the occupation they work in.

We fail to find evidence for Sweden that women became relatively more at risk of being unemployed than men due to the pandemic (see Table 2, column (1), where we report the estimate of β from equation (4)). If anything, they were less impacted, although the difference is only possibly statistically significant (10%). Interestingly, once we account for occupation fixed effects, the estimated gender differ increases slightly and it is significant at conventional level (5%): in the months of March to July 2020 the risk of unemployment increased by roughly 2.5 p.p. more than over the

¹⁰We consider occupations-by-gender where in 2017-2018 there were at least 100 employees. Results are robust to increasing this threshold to 500.

same period of 2019 for men; the increase was 0.5 p.p., or 20% smaller, for women in the same occupation.

Early analysis at the onset of the pandemic has warned that the economic downturn following the Covid-19 outbreak would have a large impact on sectors with high female employment shares (Alon et al., 2020), although the evidence from empirical studies that emerged later on is more mixed. Media and civil-society advocates have also repeatedly warned that the Covid-19-induced economic crisis might spiral into a gender inequality crisis.

How can the evidence that we present here be reconciled with these concerns? First, we observe that conducting the analysis at the individual level (how many women and men have become unemployed?) rather than at the sectoral-level obviously captures different dynamics.¹¹ Some highly-feminized occupations (e.g. “masseurs and massage therapist” in the Swedish case) have been heavily impacted, but their relative size is small. Consistent with this observation, we also note that, based on our data, occupations with a larger share of female employees were indeed more impacted by the pandemic, although the difference is not significant at conventional level. We see this by regressing the overall sector-level unemployment impact against the respective share of female employment (see column (3) of Table 2). Since our data are more granular than those used in related work, for comparison we also estimate this regression using 2-digit occupation information (column 4), and the gender difference increases in size but remains statistically insignificant.

Second, the Swedish response to the pandemic was peculiar, namely (a) there was never a regulated lockdown that forced people to stay at home and businesses to close (with the exception of a few activities, most notably in the cultural sector), and, (b) crucially, day-care facilities and schools for kids aged less than 15 were never closed. The peculiarity of the Swedish response implies that the relative impact across sectors of the Swedish economy might have been different from elsewhere, and especially that Swedish families did not have to make significant adjustments to the time devoted to childcare. Since women taking on more childcare responsibilities has been one of the main reasons why researchers, policy-makers, media and civil-society advocates have described the pandemic as a gender-inequality crisis, it is perhaps not surprising that such crisis did not unfold in Sweden. Alon et al. (2021) find that the gender-unequal labor market impact in

¹¹Notice that, since we weigh occupation-by-gender groups by the number of employed pre-pandemic, we seemingly analyse the unemployment impact at the individual level.

the US is substantial especially among parents, suggesting that school closures might indeed have been a key factor behind women's struggles during the pandemic in other countries.

Third, we additionally note that the extent of occupational segregation differs among countries, and women in Sweden working predominantly in some occupations that were, if anything, more positively impacted by the pandemic, might have played a role. While a high female concentration in the health sector, one of the least impacted in terms of unemployment, is not unique to Sweden, the extent of the concentration is especially large in the Swedish labor market. Consider for instance the European Institute for Gender Equality (EIGE) Index, which ranks EU-28 countries based on a number of indicators of gender equality, including the extent of sectoral segregation by gender in the labor market.¹² In 2020, Sweden was the top-performer based on the overall EIGE index. Nonetheless, its score in terms of sectoral segregation was below the EU-28 average, denoting more gender inequality, due to the large concentration of women in sectors such as education, human health and social activities.¹³ According to our data, of the handful of occupations where unemployment risk decreased in relative terms, some of the largest are "Clinical and operations managers in health care", "Department managers in elderly care", "Pediatric nurses", and "Anaesthesia nurses", whose shares of female employment vary between 74% (anaesthesia nurses) and 97% (pediatric nurses). Job creation in these occupations presumably benefited more women than men. In terms of within-occupation differences, another important factor is the intersection between gender and other characteristics that affect unemployment risk, such as, for instance, tenure; in the following sections we will consider some of these characteristics. These last observations also highlight the importance of considering not only sectoral shocks, but also within sector dynamics of unemployment risk; the granularity of our data in defining occupations is particularly suitable to capture these dynamics.

The unemployment impact of the pandemic by labor income Some attention has also been devoted to the unequal economic impact of the pandemic based on skill-level, proxied by education, and the prevailing evidence is that less skilled workers have been more negatively affected (see, e.g.,

¹²See <https://eige.europa.eu/gender-equality-index/2020/compare-countries>.

¹³The EIGE index ranks the labor market as more segregated the larger is the difference between the shares of women and men working in "education, human health and social work activities". According to the data reported by EIGE, in Sweden 42% of all the women employed work in such sectors, compared to 12% of men, whereas these respective average shares among EU-28 countries are 31 and 8%.

Table 2: Gender and unemployment impact

Dep.variable:	Four-digit			Two-digit
	(1)	(2)	(3)	(4)
Unemployment impact				
Female	-0.004 (0.0022)	-0.005* (0.0024)		
Share female			0.003 (0.0098)	0.005 (0.0070)
Occupation FE	No	Yes	No	No
N	742	742	379	45
Mean y	0.025	0.025	0.024	0.060

Note: Unit of observation is occupation-by-gender group in columns (1) and (2), and occupation group in columns (3) and (4). Occupation groups are at 4-digit level in columns (1) to (3), and at 2-digit level in column (4). Observations are weighed by number of employed in 2017-2018 in columns (1) and (2). Standard errors are clustered by 3-digit occupational sector in columns 1-3, robust in column 4. Mean y is the mean of the dep. var. for the category *Male* in columns (1) and (2), and the mean of the dep. variable in the estimation sample in columns (3) and (4). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Adams-Prassl et al. (2020), Gaudecker et al. (2020), Casarico and Lattanzio (2020)). We enrich this strand of literature by considering wage inequality, similarly to Cajner et al. (2020) who study the US labor market; our study complements theirs by focusing on a country, Sweden, where the initial level of wage-inequality is substantially lower than in the US. Nevertheless, we reach similar conclusions to theirs, namely that the most-negatively impacted workers are those at the bottom of the wage distribution.

We first document this fact in Figure 1. We group occupations into wage deciles, using the occupation-level wage.¹⁴ Then, for each wage-decile, we consider the weighted (by the number of people employed in 2017 and 2018) average change in *unemployment impact*, as defined in (3), across occupations belonging to the decile. In Figure 1 we plot the unemployment impact against wage-decile and show that, while the risk of unemployment has increased due to the pandemic across all wage levels, the size of the increase is substantially larger for lower-wage occupations. In other words, the unemployment impact of the first months of the pandemic has been higher for the workers with the lowest wages in the Swedish labor market.

Since the bottom deciles were at higher risk of unemployment already before the pandemic (data

¹⁴Specifically, we consider the weighted (by number of people employed in 2017 and 2018) average of the occupation-by-gender wage for women and men.

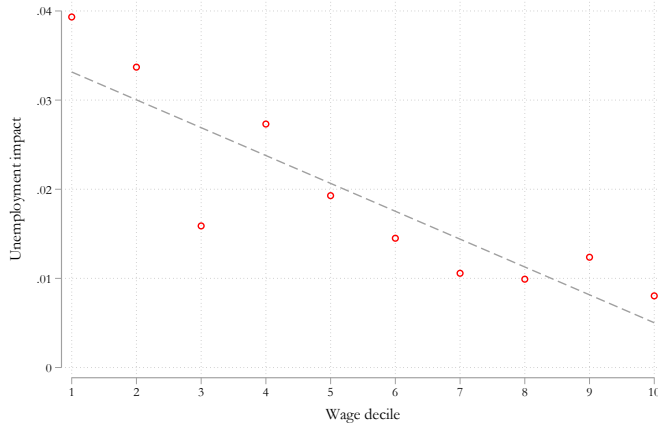


Figure 1: Unemployment impact by wage decile

Notes: This Figure shows the relationship between unemployment impact and wage decile. The unemployment impact is the change in unemployment risk during the first months of the pandemic (March to July 2020) versus one month before its onset (February 2020), as compared to the same change one year before (see definition in equation 3). The unemployment impact of the pandemic is higher for lower wage deciles.

not shown and available upon request), considering the unemployment impact *relative* to the pre-pandemic unemployment risk shows a reversed pattern: the unemployment cost of the pandemic is higher for the least vulnerable workers (see Appendix Figure A2), whose initial position has deteriorated to a larger extent.¹⁵

We further study the relationship between unemployment impact and wages estimating the following equation:

$$unemployment\ impact_{ig} = \alpha + \beta \log(wage)_{ig} + \eta_{ig} \quad (5)$$

where *unemployment impact* is as defined in (3), and $\log(wage)_{ig}$ is based on the respective average wage for occupation *i* and gender *g*. As before, we weight observations by group-size. We report the results in Table 3. Column (1), where we show estimates from the baseline specification in (5),

¹⁵This fact highlights that the researcher-chosen definition and metric of “impact” is consequential. We choose to focus on the absolute rather than the relative change in unemployment impact because we are interested in how unemployment inequality has evolved during the pandemic. As we document in Figure A3 in the Appendix, unemployment inequality in Sweden has increased in 2020 as compared to one year earlier. On the other hand, if we wanted to highlight which wage-group has experienced a more significant deterioration of their initial position, then we would use a different metric and reach opposite conclusions.

implies that a 20% wage increase, which corresponds to moving from the 1st to the 4th decile of the wage distribution, is associated with a decrease in unemployment impact of nearly 0.006, or 17% of one s.d. In column (2) we also account for gender, thus comparing workers of the same gender whose wages differ based on their respective occupation, and reach the same conclusion. Moreover, as compared to the estimates reported in Table 2, where we did not consider differences in wages, the estimated female advantage is larger. This is not surprising, since women, even in a relatively gender-equal country like Sweden, tend to be highly concentrated in lower-paying occupations (see Appendix Figure A4).¹⁶

Table 3: Wages and unemployment impact

Dep.variable:		
Unemployment impact	(1)	(2)
log(wage)	-0.031*** (0.0068)	-0.034*** (0.0069)
Female		-0.008*** (0.0023)
N	668	668

Note: Unit of observation is occupation-by-gender. Occupation groups are at 4-digit level. Observations are weighed by number of employed in 2017-2018. Standard errors are clustered by 3-digit occupational sector. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The unemployment impact of the pandemic by age-group The administrative data on unemployment registrations also report information disaggregated by six age-categories. In Table 4 we first show the pre-pandemic unemployment risk for these categories. Specifically, in column (1) we report estimates from a regression of the average unemployment risk during the months of March to July of 2019 for occupation i , gender g , and age group a , against age-group fixed-effects, omitting the age-group 40-49 (that is represented in the largest number of occupations) and weighting groups by the respective employment level in 2017-2018:

$$unemployment\ risk_{i,g,a}^{2019} = \alpha + \phi_a + \varepsilon \tag{6}$$

¹⁶Notice that the unequal impact by wage-level contributes to explain why the gender difference that we estimate in Table 2 is larger once we account for occupation fixed-effects: since women work predominantly in lower-paying occupations, which also are exposed to a larger unemployment impact, the estimate of women’s relative advantage during the first months of the pandemic increases once we condition on the occupation where they work.

where ϕ_a are the age-group fixed-effects. The coefficients on the age-group fixed-effects reveal an age-gradient for unemployment risk. Before the pandemic, younger cohorts were at higher risk of being unemployed than workers in the age-group 40-49, whereas there is no statistically significant difference for the older cohorts. The age-pattern is less clear once we account for occupation fixed-effects (column 2), suggesting that the main reason why younger cohorts are at higher risk of unemployment prior to the pandemic lies in the type of occupations where they work.^{17,18}

Occupational sorting does not seem to entirely explain the higher unemployment impact of the pandemic for younger cohorts. We show this in columns (3) to (5) of Table 4, where we report estimates from a few variations of the following equation:

$$\text{unemployment impact}_{i,g,a} = \alpha + \phi_a + \epsilon_{iga} \quad (7)$$

The results show that the unemployment impact of the pandemic is higher the younger is the age-cohort (column 3) and accounting for differences in occupation explains only partially this age-gradient (column 5). These findings confirm the evidence from Italy in [Casarico and Lattanzio \(2020\)](#), who also observe that younger workers were already suffering the consequences of the previous recession. The implications of the increased vulnerability of the relatively young cohorts can be far-reaching, given the high returns to experience in the labor market. Future work should study the long-term implications of Covid-19 for this segment of the population that was less affected by the health crisis, but seems to have paid the highest cost of the economic crisis, at least in terms of unemployment risk.

We also note that the gender difference is attenuated when the age-effects are included (column 4), suggesting that a different distribution of female and male workers across occupation-by-age groups explains in part the differential unemployment impact of the pandemic by gender.

The unemployment impact of the pandemic by foreign-born status In Table 5 we repeat the same analysis shown in Table 4 replacing the age-groups with foreign- status categories, namely

¹⁷Accounting for gender leaves the estimates virtually unchanged.

¹⁸However, for the age-group 30 to 39 the disadvantage persists even once differences in occupation are taken into account; this is not surprising, since the youngest cohorts likely include a large number of college students who work in part-time low-paying occupations; for the group aged 30 to 39 instead, the disadvantage with respect to the older cohorts is more likely determined by the type of contracts they are offered, as well as practices such as “last-in-first-out”, which companies tend to adopt when they downsize.

Table 4: Age group and unemployment impact

	(1)	(2)	(3)	(4)	(5)
Dep.variable	Unemployment risk 2019		Unemployment impact		
16-24	0.032*** (0.0067)	-0.020 (0.0135)	0.027*** (0.0034)	0.027*** (0.0034)	0.019*** (0.0035)
25-29	0.018*** (0.0037)	-0.003 (0.0048)	0.011*** (0.0015)	0.011*** (0.0015)	0.007*** (0.0015)
30-39	0.016*** (0.0022)	0.010*** (0.0023)	0.006*** (0.0008)	0.006*** (0.0008)	0.004*** (0.0007)
50-59	-0.000 (0.0023)	-0.001 (0.0025)	-0.001 (0.0007)	-0.001 (0.0007)	-0.001 (0.0007)
60-64	0.004 (0.0046)	0.003 (0.0047)	-0.003** (0.0010)	-0.003** (0.0010)	-0.001 (0.0011)
Female				-0.004* (0.0016)	-0.003 (0.0016)
Sector FE	No	Yes	No	No	Yes
Observations	4013	4013	3978	3978	3978
Reference mean	0.085	0.085	0.017	0.018	0.018

Note: Unit of observation is occupation-by-gender-by-age category. Omitted category is 40-49. Occupation groups are at 4-digit level. Observation are weighted by number of employed in 2017-2018. Standard errors are clustered by 3-digit occupational sector. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

EU-born, EU no Sweden - born, and Outside EU - born.

Table 5: Foreign-born status and unemployment impact

	(1)	(2)	(3)	(4)	(5)
Dep.variable	Unemployment risk 2019		Unemployment impact		
Europe outside of Sweden	0.064*** (0.0072)	0.056*** (0.0068)	0.006*** (0.0013)	0.006*** (0.0013)	0.004*** (0.0012)
Outside of Europe	0.242*** (0.0337)	0.230*** (0.0301)	0.029*** (0.0062)	0.029*** (0.0062)	0.026*** (0.0051)
Female				-0.004* (0.0016)	-0.002 (0.0015)
Sector FE	No	Yes	No	No	Yes
Observations	2055	2055	2043	2043	2043
Reference mean	0.051	0.051	0.022	0.022	0.022

Note: Unit of observation is occupation-by-gender-by-foreign born status. Omitted category is “Born in Sweden”. Occupation groups are at 4-digit level. Observation are weighted by number of employed in 2017-2018. Standard errors are clustered by 3-digit occupational sector. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

There is a large heterogeneity in the pre-pandemic risk of unemployment by foreign-born status. As compared to Sweden-born workers, the unemployment risk for workers born in other EU countries before the pandemic was 6 p.p. higher; this difference increases to a striking 25 p.p. for workers born outside of the EU (see column 1). These findings are consistent with reports documenting that Sweden has one of the largest gaps in unemployment rate between natives and foreign-born among OECD countries.¹⁹ Importantly, the foreign-born gap is not explained by occupational sorting (column 2) and is exacerbated by the pandemic (columns 3 to 5): based on our estimates, while the pandemic implied a 2.2 p.p. increase in the risk of unemployment for Sweden-born workers, the increase was 0.5 p.p. larger (8% of a s.d.) for other EU-born workers, and more than twice as large for workers born outside of the EU. The differences are statistically significant and virtually unchanged once gender and occupation are accounted for. Lastly, the gender difference is substantially attenuated in regressions that control for foreign-born status, confirming that female workers tend to differ from male workers on some characteristics that are important determinants of the unemployment impact of the pandemic, especially within occupations.²⁰

¹⁹https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Migrant_integration_statistics_%E2%80%93_labour_market_indicators.

²⁰Notice that we cannot study the relative importance of age-cohort and foreign-born status because the data are non disaggregated at occupation-by-gender, -age, and -foreign born status.

4 Conclusion

In this paper, we have studied differences in unemployment due to the Covid-19 pandemic across different groups of Swedish society. Using the full population of registered unemployed – around 400,000 individuals, grouped across some 400 occupations – we find important differences in unemployment risk across groups. Most adversely affected are the young and the foreign-born while gender does not seem to play any major role in the case of Sweden. Along the income dimension, the lower the wage the higher the unemployment.

These results both confirm what previous work has shown, namely that already vulnerable groups have been more affected by the economic impacts of Covid-19, but also remind us that country-specific institutions and other pre-pandemic factors, as well as policy responses, are likely to play an important role. In the case of Sweden, this is most obvious when it comes to the gender dimension, where we do not find that women fare worse in terms of unemployment due to the pandemic. If anything, it is the other way around. There are several plausible contributing factors for this. First, women in Sweden have high labor force participation and the welfare system has long focused on supporting possibilities to combine family and work-life. The one dimension where Sweden stands out as more gender segregated than many other countries is across sectors. But in the pandemic, this has, if anything, been to the advantage of women. Even though some of the most adversely affected jobs are dominated by women (such as cabin crew personnel) the size of these occupations is relatively small. At the same time, some of the sectors that instead have seen an increased demand for workers, such as many occupations in the health care sector, are also dominated by women and are much larger in size. Second, some aspects of the policy response to the pandemic have minimized the impact on women's labor market conditions compared to the situation in other countries. In particular, the fact that day-care facilities and schools basically remained open has largely taken away the channel identified as most important in explaining gender inequality in other countries, namely women taking a larger responsibility for childcare. The dimensions where we find the largest impact, age and being foreign-born (especially if born outside of Europe) are in line with what has been found in other studies but can also be related to weaknesses in the Swedish labor market prior to Covid-19. Prior to the pandemic, the inclusion of both these groups in the labor market stands out as an area where Sweden is doing poorly compared to most other

OECD countries. Importantly, our detailed occupational data allows us to study differences also within occupational groups so as to separate the effects of young as well as foreign-born potentially sorting into sectors which in turn have been most affected. We find that with respect to age and foreign-born status, the sorting is not the only factor explaining higher exposure to the economic downturn.

Finally, the wage gradient in unemployment suggests that there is indeed a possibility that the pandemic leads to larger inequality also in terms of disposable incomes. However, the steps from individual wages to living standards at the household level are numerous. Preliminary simulations made by the Swedish finance ministry in April 2020 indicate that overall inequality in terms of disposable incomes has not increased much in the pandemic. But this may be a consequence of transfers and other policy measures taken, again pointing to the importance of studying country-specific consequences for economic welfare.

References

- Adams-Prassl, A., T. Boneva, M. Golin, and C. Rauh (2020). Inequality in the impact of the coronavirus shock: Evidence from real time surveys. *Journal of Public Economics* 189, 104245.
- Alon, T., S. Coskun, M. Doepke, D. Koll, and M. Tertilt (2021). From Mancession to Shecession: Women's Employment in Regular and Pandemic Recessions. *National Bureau of Economic Research WP*.
- Alon, T. M., M. Doepke, J. Olmstead-Rumsey, and M. Tertilt (2020). The impact of COVID-19 on gender equality. *Covid Economics* 4, 62–85.
- Cajner, T., L. D. Crane, R. A. Decker, J. Grigsby, A. Hamins-Puertolas, E. Hurst, C. Kurz, and A. Yildirmaz (2020). The US labor market during the beginning of the pandemic recession. *National Bureau of Economic Research WP*.
- Casarico, A. and S. Lattanzio (2020). The heterogeneous effects of COVID-19 on labor market flows: Evidence from administrative data. *Covid Economics* 52, 152–174.
- Dahlberg, M., P.-A. Edin, E. Grönqvist, J. Lyhagen, J. Östh, A. Siretskiy, and M. Toger (2020). Effects of the COVID-19 pandemic on population mobility under mild policies: Causal evidence from Sweden. *Mimeo*.
- Ellingsen, T. and J. Roine (2020). Sweden and the Virus. In M. Carlsson-Wall, G. Lindqvist, S. Rosengren, J. Söderholm, A. Werr, and F. Wijkström (Eds.), *Sweden Through the Crisis*, pp. 36–51. Stockholm School of Economics.
- Forsyth, E., L. B. Kahn, F. Lange, and D. G. Wiczer (2020). Labor Demand in the time of COVID-19: Evidence from vacancy postings and UI claims. *National Bureau of Economic Research WP*.
- Gaudecker, H.-M., R. Holler, L. Janys, B. Siflinger, and C. Zimpelmann (2020). Labour supply in the early stages of the COVID-19 Pandemic: Empirical Evidence on hours, home office, and expectations.
- Gottfries, N. (2018). The labor market in Sweden since the 1990s. *IZA World of Labor* (411).
- Hupkau, C. and B. Petrongolo (2020). Work, care and gender during the Covid-19 crisis. *Fiscal studies* 41(3), 623–651.
- Mein, S. A. (2020). COVID-19 and health disparities: the reality of “the Great Equalizer”. *Journal of general internal medicine* 35, 2439–2440.
- OECD (2016). Promoting Well-being and Inclusiveness in Sweden.
- OECD (2019). OECD Economic Surveys: Sweden 2019.

Appendix

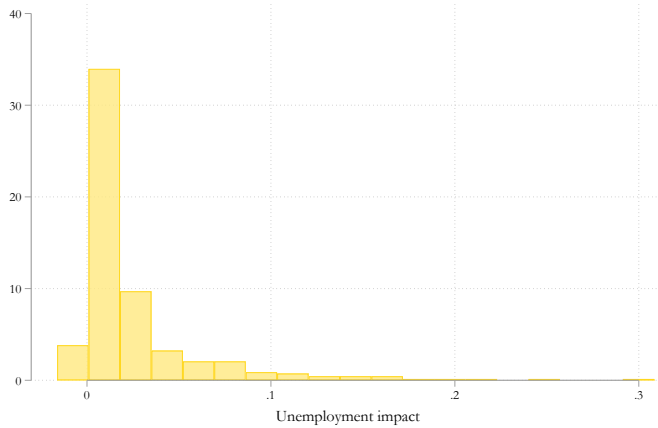


Figure A1: Distribution of the unemployment impact of the pandemic

Notes: This Figure shows the distribution of the unemployment impact of the pandemic at occupation-level. The unemployment impact is the change in unemployment risk during the first months of the pandemic (March to July 2020) versus one month before its onset (February 2020), as compared to the same change one year before (see definition in equation 3). The unemployment impact of the pandemic varies substantially across occupations.

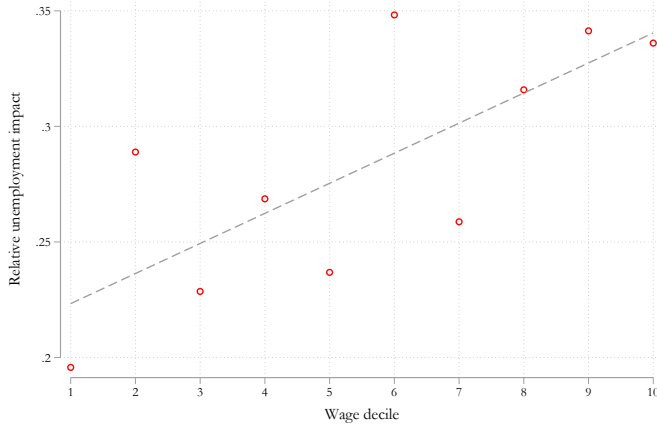


Figure A2: Relative unemployment impact by wage decile

Notes: This Figure shows the relationship between wage-decile and the occupation-level unemployment impact of the pandemic as a percentage of the pre-pandemic unemployment. The unemployment impact is the change in unemployment risk during the first months of the pandemic (March to July 2020) versus one month before its onset (February 2020), as compared to the same change one year before (see definition in equation 3). Relative to the pre-pandemic unemployment level, workers in the higher wage deciles experienced a larger increase in unemployment risk.

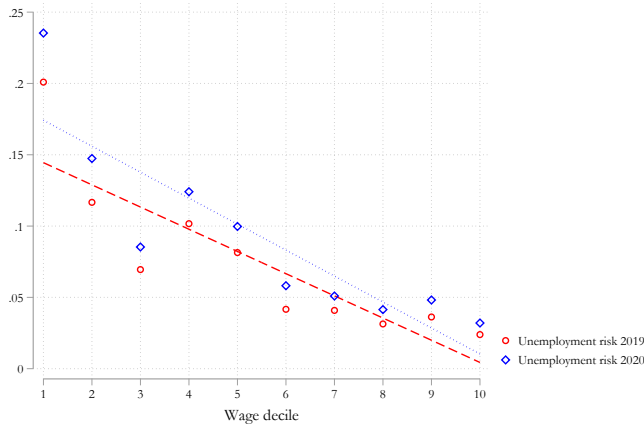
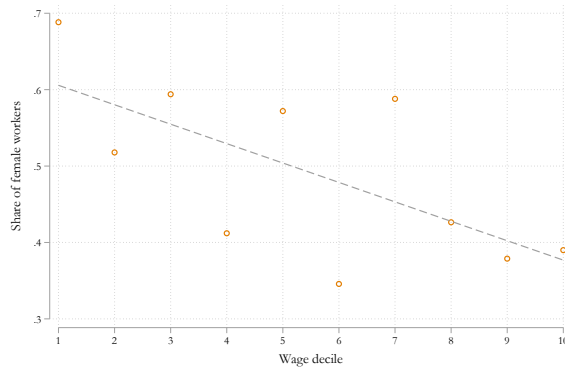


Figure A3: Unemployment inequality before and during the pandemic

Notes: This Figure shows the relationship between wage decile and unemployment risk in 2019 and 2020. Unemployment risk is unequal across wage-deciles: the lower the wage level, the higher the unemployment risk. The inequality increases during the first wave of the pandemic.

Figure A4: Share of female workers by wage decile



Notes: This Figure shows the relationship between wage decile and women as a share of total employment, based on occupation-level data. Women are more concentrated in lower paying occupations.

On the resilience of ESG stocks during COVID-19: Global evidence¹

Gianfranco Gianfrate,² Tim Kievid³ and Mathijs van Dijk⁴

Date submitted: 22 June 2021; Date accepted: 27 June 2021

Recent research on the U.S. stock market finds that the stocks of firms with high ESG (environmental, social, and corporate governance) ratings perform relatively well during crisis periods and thus serve as “rainy day assets” for investors. We assess this hypothesis in a global setting by assessing the relation between ESG and stock price performance during the COVID-19 crisis. In a sample of more than 6,000 stocks in 45 countries, we find little evidence that firms with higher ESG ratings had better stock market performance in the first quarter of 2020. The exception is North America, where stocks with higher ESG ratings have shown some degree of resilience during crises. Our findings indicate that the ability of socially responsible firms to deliver superior risk-adjusted stock market performance is still debatable and, at best, geography dependent.

1 We are grateful to Netspar for financial support through a topicality grant, to the Netspar Editorial Board for helpful comments, and to Jonas Elenskis for excellent research assistance.

2 Professor of Finance at EDHEC Business.

3 Research Affiliate, Rotterdam School of Management, Erasmus University.

4 Professor of Finance, Rotterdam School of Management, Erasmus University.

Copyright: Gianfranco Gianfrate, Tim Kievid and Mathijs van Dijk

1. Introduction

Recent studies present evidence suggesting that firms with higher environmental, social, and governance (ESG) ratings are relatively resilient to financial and economic crises. Lins, Servaes, and Tamayo (2017) show that, during the 2008-2009 financial crisis, stock returns for firms more active in ESG were significantly higher than those of firms that engaged less in ESG initiatives. Albuquerque, Koskinen, Yang, and Zhang (2020ab) find that U.S. firms with higher ESG ratings recorded higher stock returns during the COVID-19 crisis. Prior research considers a number of potential explanations for such resiliency. First, ESG can be considered as a product differentiating strategy that enhances customer loyalty (Albuquerque, Koskinen, and Zhang, 2019; Gantchev, Giannetti, and Li, 2019), thus reducing the stock sensitivity to general shocks. Second, firms with higher ESG ratings are held relatively more by socially-responsible investors, which are more resilient to shocks and less likely to engage in sell-offs (Heinkel, Kraus, and Zechner, 2001; Renneboog, Ter Horst, and Zhang, 2011; Ferriani and Natoli, 2020). Third, firms that consistently engage in ESG promote civic engagements and collaboration with all stakeholders, thus enhancing their social capital endowment (Lins et al., 2017) and allowing them to weather unexpected shocks relatively better.

In this paper, we examine the resilience of firms with high ESG ratings during the COVID-19 crisis in a global setting. We believe that such an analysis is not only relevant because it can provide out-of-sample evidence on the resilience of ESG stocks, but also because recent empirical studies show that financial markets are geographically segmented as far as responsible investment is concerned.¹ The COVID-19 crisis provides an opportunity to test whether responsible investment policies that assign more weight to ESG stocks fared relatively well for (institutional) investors in different regions.²

¹ For example, Gibson, Rajna, and Glossner (2020) show that global institutional investors publicly committed to integrate sustainability in their investment decisions exhibit higher ESG portfolio-level ratings, with the notable exception of U.S.-domiciled institutions. Consistently, Dyck, Lins, Roth, and Wagner (2019) show that U.S.-headquartered investors do not appear to affect the sustainability footprint of the firms that they invest in, but that European institutional investors, and especially those from countries whose culture strongly prizes sustainability, are able to eventually enhance the ESG ratings of the firms in their portfolios

² Because of its unexpected and dramatic impact, the COVID-19 pandemic represents an exogenous shock that allows researchers to shed light on the corporate characteristics that promoted financial resilience during crises. The early evidence suggests that access to liquidity, higher cash holdings, and lower financial leverage positively affected the resilience of stock prices at the height of the COVID-19 crisis (Acharya and Steffen, 2020; Fahlenbrach, Raged, and Stulz, 2020; Ramelli and Wagner, 2020).

For this purpose, we obtained detailed data on ESG ratings for a global sample of firms from Thomson Reuters' Refinitiv ESG database. Following Albuquerque et al. (2020ab), our main analyses omit the governance rating (G) and focus on the environmental and social (ES) rating. Albuquerque et al. (2020ab) argue that the superior stock market performance of U.S. firms during the COVID-19 crisis stems from their greater customer and investor loyalty, which are arguably more likely to be based on the E and S than the G rating.³ That said, financial media and industry reports on the outperformance of ESG stocks during the COVID-19 crisis also allude to their governance. For that reason, we also present results based on the overall ESG rating as well as on the separate E, S, and G sub-ratings. We collect data on stock returns and financial statement information from Datastream and Worldscope, respectively. Our final sample contains 6,824 firms from 45 countries.

We show that, without considering any control variables, the firm-level ES rating is positively associated with (abnormal) stock returns during the COVID-19 crisis in the first quarter of 2020. However, once we control for country-wide effects, this result disappears. In other words, in a global setting stronger ES firms did not show better stock returns because of their ES rating per se, but because they tend to be domiciled in countries whose stocks in general showed a relatively better (abnormal) stock return. Such country-wide effects may be due to a host of reasons why some countries performed better than others (including susceptibility and/or response to the COVID-19 pandemic) and can thus not be reliably attributed to the ES activities of individual firms.

When we rerun our baseline analyses for North America (Canada and the U.S.), we confirm the finding by Albuquerque et al. (2020ab) that stronger ES stocks in this region did show superior stock market performance, even after controlling for country-wide effects and a broader set of control variables than in their study – although the statistical significance of this effect is rather weak. However, for all other regions (Europe, Japan, Asia-Pacific, and Emerging Countries), the effect of ES on (abnormal) stock returns in 2020:Q1 is economically small (sometimes even negative) and statistically insignificant. When we zoom in on a number of ESG sub-ratings that are potentially more directly

³ See, e.g., AXA Investment Management (2020). Proactive measures to protect the workforce and supply chains in their ability to implement social distancing on workplaces have also positively contributed to the stock price performance of firms during the COVID-19 shock (Cheema-Fox, LaPerla, Serafeim, and Wang, 2020; Pagano, Wagner, and Zechner, 2020).

related to the resilience of different firms during the COVID-19 pandemic, we find that the S sub-rating for “Workforce” shows the strongest positive relation with (abnormal) stock returns, but only in North America. This finding suggests that the better working conditions of stronger ESG firms are a primary driver of the better performance of stronger ESG stocks during the COVID-19 crisis in that region, but not in other regions.

Taken together, our findings cast doubts on the view that better ESG is associated with enhanced resilience to significant economic shocks in a global setting. In particular, such a relation does not appear to hold for firms outside of North America. Cultural and institutional elements can weigh in the asymmetric role played by ESG ratings in the degree of resilience to the COVID-19 shock. Further empirical work should shed light on the heterogeneous response to the pandemic of socially responsible firms around the world.

We are able to dispel the view of Heinkel et al. (2001) and Renneboog et al. (2011) that, because high-ESG firms attract socially responsible investors and are assumed to be more resilient and less volatile, those firms are also the ones poised to weather well any exogenous shocks on the stock market. Firms with higher ESG ratings have been more resilient in the U.S, where investors are generally less committed to responsible investment strategies (Gibson et al., 2020). At the same time, European firms with higher ESG ratings have not recorded a rosier financial performance during the pandemic despite the fact that European investors are regarded as the champions of socially responsible investment (Dyck et al., 2019).

The remainder of the paper is structured as follows. Section 2 describes the data and sample construction. Section 3 discusses the methodology. Section 4 presents the results. Section 5 concludes and points out the limitations of our findings and future avenues of research.

2. Data and Sample Construction

We obtain measures of corporate ESG performance from Thomson Reuters’ Refinitiv ESG database. This dataset is frequently used in the sustainable finance literature (see, e.g., Dyck et al., 2019; Albuquerque et al., 2020ab)). Refinitiv analysts collect data from publicly available sources such as corporate reports, nongovernmental organizations, sustainability reports, and stock exchange filings.

Data coverage starts in 2002 and is updated on a yearly basis. As of 2020, the database includes over 9,000 publicly listed firms worldwide.

Refinitiv assesses ESG performance using the so-called E, S, and G pillars, which break down into ten categories that segregate into 25 themes. The E pillar contains three categories: (1) Emission, (2) Innovation, and (3) Resource Use. The S pillar holds four categories: (1) Community, (2) Human Rights, (3) Product Responsibility, and (4) Workforce. The G pillar includes: (1) Corporate Social Responsibility (CSR) Strategy, (2) Management, and (3) Shareholders. Further down the hierarchy within Workforce, we find four themes that are of particular interest for our study, since they are related to working conditions that could potentially contribute to our understanding of the differential stock market performance of firms during the COVID-19 crisis. These four themes are (1) Diversity and Opportunity, (2) Training and Development, (3) Employment Quality, and (4) Health and Safety.

Refinitiv evaluates ESG in a bottom-up approach, which starts with the assessment of the firms' ESG category performance (see Appendix A). Note that Refinitiv does not provide the ESG performance within the above-mentioned ESG themes (e.g., Employment Quality). These theme classifications are only required for the construction of a category materiality matrix, which is then needed to aggregate category ratings into pillar ratings (as will be outlined in the next paragraph). To calculate ES (G) category ratings, Refinitiv ranks industry (country) peers by their overall performance across a category-specific set of ES (G) metrics and subsequently derives a percentile rating for each firm (see Appendix A). Thus, the ESG category ratings reflect a firm's relative ESG performance compared to its industry peers (for E and S ratings) and its country peers (for G ratings). We note that these "relative" ESG ratings may be perceived as unfitting in the sense that, for example, firms in an industry that is more "polluting" in an absolute sense (such as the airline industry) may have a higher E ratings than firms in a "sustainable" industry (such as a recycling firm). However, from the perspective of proper identification of the effect of ESG ratings on stock market performance, it is imperative to control for such industry (and country) effects, since otherwise it is impossible to ascertain that any detected relation between ESG ratings and stock returns is really due to the ESG performance itself, as opposed to some unobservable industry or country characteristic. We will return to this issue in more detail below.

To calculate ESG pillar ratings, Refinitiv first assigns *weights* to the underlying category ratings, which reflect the industry/country-specific materiality⁴ of each category. Pillar ratings are subsequently calculated as the *weighted* aggregate of the subsidiary ESG category ratings. We include all firms in the Refinitiv database for which we extract the ESG pillar ratings and category ratings. In line with Albuquerque et al. (2020ab), we drop the G pillar rating and calculate our primary ESG measure as the average of the E and S pillar ratings (*ES*). We examine other ESG measures, including the overall ESG rating, calculated as the average of the E, S, and G pillar ratings and more specific ESG sub-ratings, such as the category ratings for *Community*, *Human Rights*, and *Workforce* theme ratings within the latter category. Since Refinitiv does not provide ESG theme ratings, we manually construct *theme* ratings using Refinitiv's *category* rating approach.⁵ First, we obtain all thirty Workforce-related metrics (see Appendix A). Next, we allocate each metric to its related theme, which is indicated by the third and fourth string in the metrics' ID. For example, metric SOHSDP004 reports whether "the firm has an employee health and safety team" and therefore is assigned to the Health and Safety theme. Similarly, metric SOTDDP018 reports the "average hours of training per year per employee" and is therefore allocated to the Training and Development theme.

Next, we apply a Refinitiv materiality scheme (which outlines ESG materiality by TRBC industry group⁶) to classify each metric as material or immaterial to a particular industry. We rank TRBC industry peers on all thirty metrics separately and convert the rankings into percentile ratings. To calculate a theme rating, we select all *material* theme-specific percentile ratings, take their simple average⁷, and rank the industry peers accordingly. Again, we translate the rankings into percentile ratings, which now make up the desired theme ratings.

⁴ The Sustainable Accounting Standards Board defines material ESG factors as "reasonably likely to impact the financial condition or operating performance of a firm and therefore are most important to investors."

⁵ Refinitiv. April 2020. 'Environmental, Social and Governance (ESG) ratings from Refinitiv'. Downloaded from: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/esg-ratings-methodology.pdf

⁶ See https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/trbc-business-classification-methodology.pdf. We obtained the Refinitiv materiality scheme from Refinitiv via email; we are not aware of any public source for this scheme.

⁷ By taking the simple average (sum of percentile ratings / number of non-missing percentile ratings), we deviate from Refinitiv's approach, which prescribed taking the total sum of percentile ratings. We do this because we constructed two Workforce category ratings: (1) based on averaging percentile ratings and (2) based on adding up percentile ratings. We compared this to the Workforce rating that is available in Datastream. We found that the Workforce category rating based on the simple average best fitted the Workforce rating in Datastream.

In our main analyses, we assess the effect of ESG on stock returns during the COVID-19 stock market crisis (2020:Q1) in the cross-section. We use abnormal stock returns as our main measure of the stock market performance of firms. To calculate abnormal stock returns, we apply a domestic market model (Griffin, 2002) and regress monthly arithmetic stock returns on a constant and the domestic market factor, where the latter equals the monthly arithmetic return of the local benchmark. We turn to Thomson Reuters' Datastream and retrieve the monthly stock Return Index (RI) and Local Market Index (LI) from 2015 to 2020:Q1. RI, which we use to calculate raw stock returns, reflects the value of an equity investment, assuming that all dividends are re-invested at the closing price on the ex-dividend date. LI is the price index of the stock's local benchmark. Both variables are expressed in US\$. To estimate the equity betas, we adhere to common practice (e.g., Lins et al., 2017) and maintain a 60-month beta estimation period running up to the COVID-19 outbreak (i.e., from 2015 to the end of 2019). We include firms with a minimum of 36 monthly returns in our estimation window to preserve the quality of our beta estimates. Next, we calculate quarterly abnormal stock returns in 2020:Q1⁸ as the stock's arithmetic quarterly return in 2020:Q1 minus the estimated equity beta times the arithmetic return of the local benchmark in 2020:Q1. To reduce the impact of potential outliers, we winsorize the estimated betas as well as the 2020:Q1 abnormal returns at the 1st and 99th percentiles.

An important pitfall of the use of abnormal returns is the underlying assumption that a stock's market beta remained constant between the estimation window and the period of interest (in our case, 2020:Q1). As argued by Ramelli and Wagner (2020), this assumption might be audacious since, in response to the COVID-19 outbreak, investors largely changed their view on the susceptibility of firms to market risk. Using a sample of Russell 3,000 constituents, Ramelli and Wagner show a 17% correlation between beta estimates based on daily returns over 2020:Q1 and beta estimates based on daily returns over 2019, while they report an average 2019 inter-quarter beta correlation of 51%. This indicates that beta estimates based on pre-2020 stock returns are a rather crude indicator of the true exposure of firms to market risk from 2020 onwards. Hence, in a robustness check we repeat our analysis based on raw returns.

⁸ We calculate the 2020:Q1 returns from January 1, 2020 until April 1, 2020.

To control for the impact of the financial condition of firms on stock returns, we obtain accounting data for the year 2018 (expressed in US\$) from the Worldscope database via Thomson Reuters' Datastream. We follow Albuquerque et al. (2020ab) in our selection and calculation of *Tobin's Q*, firm size (natural log of market capitalization), *cash ratio*, *leverage*, and return on equity (*ROE*) for our first set of financial controls. Demers, Hendrikse, Joos, and Lev (2020) determine that the findings of Albuquerque et al. (2020ab) disappear when including a broader set of control variables. We thus expand our first set of control variables by including a limited number of additional controls from Demers et al. (2020). These include return on assets (*ROA*); a loss-indicating dummy and industry-adjusted *inventory turnover* as measures of accounting performance; research and development, and selling, general and administrative (*R&D & SGA*) expenses as a measure of internally developed intangibles; a book-to-market ratio (*BTM*); a dummy to indicate a *negative BTM*; stock price *momentum* and *idiosyncratic risk* to capture growth opportunities and risks; and finally, *market share* and *dividend payout*. All continuous control variables are winsorized at 1st and 99th percentile. Note that we do not include all other controls of Demers et al. (2020) because not all databases were available to us, but also because adding these further controls reduce the sample size for the U.S. by almost one third, and likely by even more for our global sample. After merging the datasets, we drop all countries that are represented by fewer than ten firms in our sample. We further exclude all firms for which the local benchmark is dead or inactive. Our final cross-sectional sample contains 6,824 firms from 45 countries. Appendix B provides a detailed description of all variables included in the cross-sectional regressions.

3. Methods

We analyze the impact of ESG on stock returns during the COVID-19 stock market crisis in 2020:Q1 using a cross-sectional ordinary least squares regression model, following previous literature (Lins et al., 2017; Albuquerque et al., 2020ab)). Our cross-sectional model can be expressed as follows:

$$Return_i = b_0 + b_1 ESG_i + b_2 Controls_i + \Lambda + \varepsilon_i, \quad (1)$$

where $Return_i$ is our main measure of return for firm i (the *abnormal* return of its stock over 2020:Q1 or AR_{Q1}); $Controls_i$ represent a set of variables that measure the financial condition of firm i ; ESG_i represents a measure of ESG performance for firm i . We follow the standards in the asset pricing

literature in lagging the ESG and control variables to make sure that they are observable before we measure the stock market performance in 2020:Q1. For stock market based variables such as *Size* (market cap), we use data from the end of 2019 since these data are immediately available to investors. For accounting variables such as *Leverage*, we use data from 2018 since the values of these variables over 2019 are not known to investors by 2020:Q1. For the ESG rating, we follow Albuquerque et al. (2020ab) and also use the values for 2018, since these ratings are published with a considerable delay and hence the 2019 rating was not known by 2020:Q1.

The variable Λ represents industry and country dummies (fixed effects). Including these fixed effects is important since, if any relation between ESG and (abnormal) stock returns is purely driven by industry or country effects, it is impossible to determine whether it is ESG itself that can explain why some firms perform better than others or the industry or country they belong to. We adopt 2-digit SIC code industry classifications in our industry-fixed effects, in line with related studies (Lins et al., 2017; Demers et al., 2020; Ding et al., 2020). To compare the effect of ESG across different regions worldwide, we split our sample in five different regions (Europe, Japan, Asia-Pacific, North America, and Emerging Countries) according to a geographic classification scheme provided on Ken French's website.⁹

We assess the robustness of our baseline results by several means. First, we re-estimate our baseline model using *log-based* abnormal returns, *log-based* stock momentum, and *log-based* idiosyncratic risk, thereby replacing their *arithmetic-based* variants. Second, we exchange AR_{Q1} with *raw returns* over 2020:Q1 (RR_{Q1}). Third, we acknowledge that AR_{Q1} captures investors' reaction to the spread of COVID-19 as well as the subsequent introduction of fiscal policies in many countries around the world through March 2020. To exclude the impact of U.S. policies on stock returns in particular, we investigate the impact of *ES* on the *raw return* over the so-called "fever period" (RR_{fever}); this showed the greatest stock price movements as a result of COVID-19, as defined in the U.S. study by Ramelli and Wagner (2020). This period stretches from February 24, which was the first trading day after Italy went into a lockdown, to at least March 6, when the first of a series of Coronavirus Emergency Aid Packages was introduced

⁹ Obtained from Ken French's data library at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

in the U.S. (Albuquerque et al., 2020ab)), but it possibly extends to March 23, when the Federal Reserve (FED) strongly intervened in the corporate bond markets (Ramelli and Wagner, 2020). We match Albuquerque et al. (2020ab) and set the fever period to the period from February 24 to March 18, when the second Coronavirus Emergency Aid Package was announced in the U.S. and the FED started to relieve the short-term debt markets via the Commercial Paper Funding Facility (Albuquerque et al., 2020ab)). As our fourth robustness test, we replace the 2-digit SIC industry classification by Thomson Reuters' TRBC industry classification in our industry-fixed effects.

4. Results

In this section, we present summary statistics for all variables included in our cross-sectional analysis (Section 4.1), our global regression results for the COVID-19 crisis in 2020:Q1 (Section 4.2), the results of our regressions estimated by region (Section 4.3), our robustness tests (Section 4.4), and more specific analyses of ESG sub-ratings instead of the broad *ES* rating (Section 4.5).

4.1. Cross-sectional summary statistics

Table 1 presents summary statistics for all variables in our cross-sectional analysis. The number of observations is capped at the number of observations for *ES*, in the sense that we only report summary statistics for the various variables included in our cross-sectional analysis for firms for which we also have an *ES* rating. The RR_{Q1} mean statistic in the top row shows that the average raw return across our global sample was -32% over 2020:Q1, and the mean AR_{Q1} was close to -5%, indicating that the firms in our sample performed poorly and even underperformed versus their local market benchmarks considering the firms' historical beta. The firms in our sample held on average 17% of their assets in cash, maintained an average leverage ratio of 25%, and 17% of the firms in our sample was not profitable (as defined by a negative *ROA*). Data on most variables are available for more than 6,000 firms globally.

Table 1: Summary statistics for the cross-sectional analysis over 2020:Q1

This table shows summary statistics for the full sample (N=6,824). The number of observations for all variables is capped at the number of observations for *ES*. The abnormal returns and all continuous controls are winsorized at the 1st and 99th percentile in the cross-section. Appendix B provides a detailed variable description. Market and financial data are from Thomson Reuters' Datastream/Worldscope, and ESG data are from Thomson Reuters' Refinitiv ESG Database.

	Count	Mean	SD	Min	p25	Median	p75	Max
<i>AR_{Q1}</i>	6,824	-5.065	20.539	-57.097	-17.775	-5.035	6.945	53.281
<i>AR_{fever}</i>	6,824	-3.683	19.821	-56.092	-15.739	-3.289	7.641	52.361
<i>RR_{Q1}</i>	6,824	-32.284	25.214	-100.000	-45.892	-32.154	-18.620	841.021
<i>RR_{fever}</i>	6,824	-35.528	20.186	-100.000	-48.111	-35.453	-22.409	400.000
<i>ESG</i>	6,824	0.420	0.206	0.001	0.253	0.393	0.578	0.943
<i>ES</i>	6,824	0.386	0.243	0.001	0.176	0.343	0.580	0.969
<i>E</i>	6,824	0.322	0.289	0.000	0.031	0.267	0.564	0.987
<i>S</i>	6,824	0.450	0.233	0.002	0.264	0.429	0.630	0.978
<i>G</i>	6,824	0.487	0.225	0.002	0.308	0.494	0.669	0.982
<i>Community</i>	6,027	0.505	0.288	0.002	0.262	0.488	0.735	0.999
<i>Human Rights</i>	6,026	0.515	0.271	0.031	0.272	0.449	0.754	0.998
<i>Workforce</i>	6,027	0.520	0.286	0.002	0.276	0.527	0.769	0.999
<i>Health and Safety</i>	6,398	0.527	0.275	0.024	0.253	0.531	0.771	0.999
<i>Training and Development</i>	6,398	0.525	0.274	0.052	0.276	0.505	0.766	0.999
<i>Employment Quality</i>	5,891	0.496	0.288	0.001	0.245	0.500	0.745	0.999
<i>Diversity and Inclusion</i>	6,455	0.529	0.275	0.031	0.320	0.527	0.770	0.999
<i>Custom Theme</i>	6,398	0.527	0.274	0.024	0.259	0.531	0.764	0.999
<i>Tobin's Q</i>	6,653	2.092	2.113	0.563	1.027	1.312	2.179	13.749
<i>Size</i>	6,657	7.944	1.596	3.637	6.858	7.951	8.995	11.862
<i>Leverage</i>	6,819	25.466	20.029	0.000	8.586	23.233	38.115	86.148
<i>Cash Ratio</i>	6,129	16.847	19.206	0.169	4.148	10.144	21.376	92.400
<i>ROE</i>	6,814	7.484	35.799	-201.085	4.002	10.020	16.716	148.886
<i>ROA</i>	6,819	2.900	11.503	-58.332	0.917	3.501	7.477	29.627
<i>Loss (dummy)</i>	6,824	0.166	0.372	0.000	0.000	0.000	0.000	1.000
<i>R&D & SGA</i>	6,690	19.081	27.404	0.000	1.382	8.996	25.304	156.823
<i>Inventory turnover</i>	6,824	58.913	93.385	0.000	0.000	31.040	79.298	609.833
<i>BTM</i>	6,653	0.710	0.736	-0.230	0.251	0.513	0.906	4.513
<i>Negative BTM (dummy)</i>	6,824	0.024	0.152	0.000	0.000	0.000	0.000	1.000
<i>Momentum</i>	6,824	22.430	40.021	-72.526	-0.642	20.038	41.011	179.874
<i>Idiosyncratic risk</i>	6,824	8.912	5.029	2.947	5.516	7.428	10.705	29.861
<i>Market share</i>	6,819	0.071	0.166	0.000	0.003	0.014	0.051	1.082
<i>Dividend payout</i>	6,808	41.407	81.724	-228.321	0.000	27.815	57.869	530.830

Figure 1: Performance of top and bottom global ES stocks during COVID-19

This figure shows top 20% vs. bottom 20% ES global portfolio cumulative market-weighted returns over 2020:Q1.

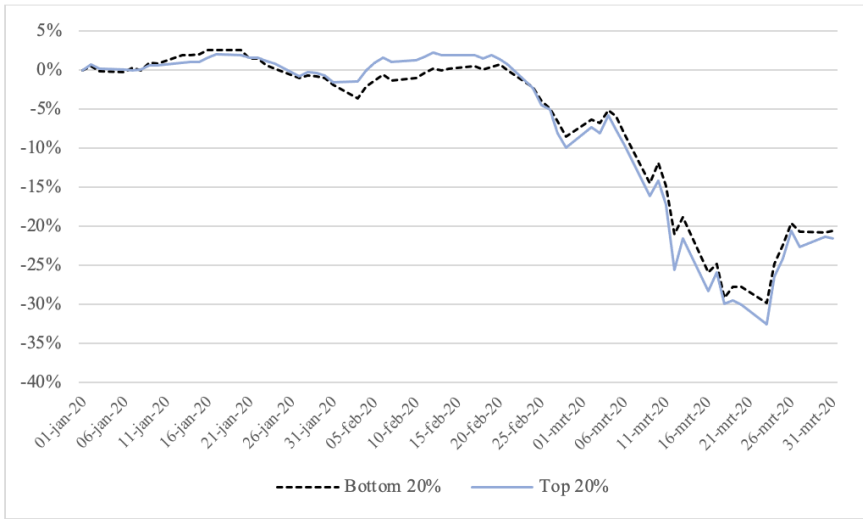
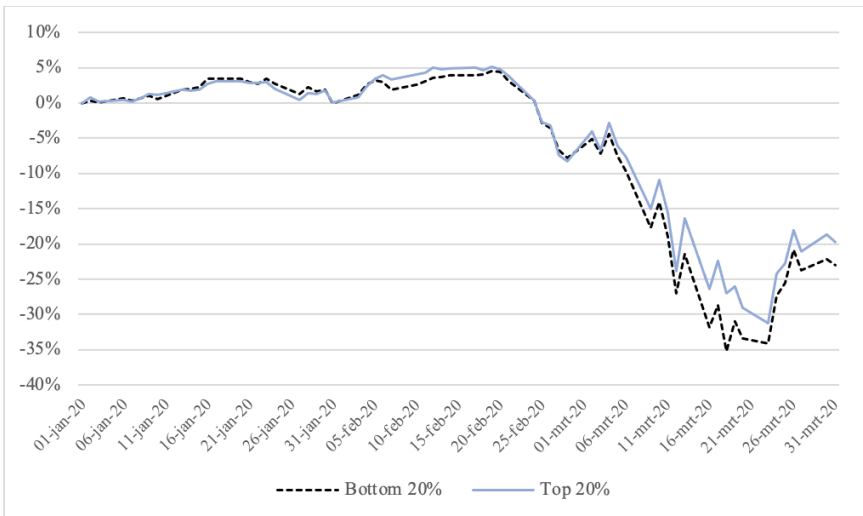


Figure 2: Performance of top and bottom North American ES stocks during COVID-19

This figure shows top 20% vs. bottom 20% ES North American portfolio cumulative market-weighted returns over 2020:Q1.



Covid Economics 83, 2 July 2021: 25-53

Table 2: Cross-sectional regressions of stock returns during COVID-19 on ES rating

This table presents regression results of abnormal returns over 2020:Q1 (AR_{Q1}) on the ESG performance of firms, measured as the average of Refinitiv's Environmental and Social pillar rating (ES) and various controls, across the global sample ($N=6,824$). We use 2-digit SIC codes industry classification in the industry-fixed effects. Market and accounting data are from Thomson Reuters' Datastream/Worldscope, and ESG data are from Thomson Reuters' Refinitiv ESG Database. AR_{Q1} and all continuous controls are winsorized at the 1st and 99th percentile. Appendix B provides a detailed variable description. All models include heteroscedasticity robust standard errors. Intercepts are suppressed to conserve space. Parentheses contain t -statistics. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

<i>Dependent variable:</i>	(1)	(2)	(3)	(4)	(5)	(6) – baseline
	AR_{Q1}	AR_{Q1}	AR_{Q1}	AR_{Q1}	AR_{Q1}	AR_{Q1}
<i>ES</i>	4.441*** (4.55)	4.099*** (4.35)	1.129 (1.08)	0.859 (0.86)	-0.859 (-0.65)	-0.775 (-0.59)
<i>Tobin's Q</i>					1.321*** (8.51)	0.841*** (4.82)
<i>Size</i>					1.541*** (6.69)	2.071*** (7.46)
<i>Leverage</i>					-0.0612*** (-3.94)	-0.0749*** (-4.70)
<i>Cash ratio</i>					0.110*** (6.18)	0.0809*** (4.27)
<i>ROE</i>					-0.0134 (-1.49)	-0.00499 (-0.47)
<i>ROA</i>						0.0371 (0.83)
<i>Loss (dummy)</i>						0.435 (0.38)
<i>R&D & SGA</i>						0.0314** (2.00)
<i>Inventory turnover</i>						-0.00250 (-0.89)
<i>BTM</i>						1.297** (2.34)
<i>Negative BTM (dummy)</i>						2.164 (0.97)
<i>Momentum</i>						0.0704*** (8.29)
<i>Idiosyncratic risk</i>						0.480*** (5.26)
<i>Market share</i>						1.673 (0.98)
<i>Dividend payout</i>						-0.000848 (-0.26)
N	6,824	6,824	6,824	6,824	5,972	5,965
Adj. R^2	0.003	0.123	0.040	0.163	0.225	0.247
Industry FE	No	Yes	No	Yes	Yes	Yes
Country FE	No	No	Yes	Yes	Yes	Yes

4.2 Cross-sectional regressions results on the performance of ESG stocks during COVID-19

Before turning to our cross-sectional regression results, we first present (in Figures 1 and 2 for our global and North American samples, respectively) a graph of the performance (cumulative raw stock returns over 2020:Q1) of the global top and bottom ES stocks, respectively, defined by stocks in the top and bottom 20% of their 2018 ES rating. These initial, crude graphical representations show little difference between the performance of the top and bottom 20% of global ES stocks (Figure 1). The top 20% ES stocks did seem to outperform the bottom ES stocks by a few percentage points over 2020:Q1 in North America (Figure 2), although unreported tests show that the difference in cumulative returns across the two portfolios was not statistically significant. We obtain similar graphs when we use the ESG instead of the ES rating.

Next, we discuss the results of our cross-sectional regressions of quarterly abnormal stock returns in 2020:Q1 (AR_{Q1}) on the stock-level ES rating and a varying set of control variables in Eq. (1). Table 2 shows estimation results of six different regression models for the global sample. Model (1) in Table 2 estimates the effect of ES on AR_{Q1} without any controls or fixed effects. We extend the first model and introduce industry-fixed effects in model (2) and substitute these for country-fixed effects in model (3). In model (4) we include both industry- and country-fixed effects. We add controls from Albuquerque et al. (2020ab) in model (5). In our baseline model (6), we furthermore include additional control variables from Demers et al. (2020). We use heteroscedasticity robust standard errors in all models.

Model (1) in Table 2 suggests that global stocks with higher ES ratings did achieve higher abnormal returns over 2020:Q1. The coefficient on our key ES variable is 4.441, which indicates that a one standard deviation (SD) gain in ES (0.243; from Table 1) was associated with a 1.08% higher AR_{Q1} (0.243×4.441) on average. This effect is statistically significant at the 1% level (t -stat = 4.55), is

economically meaningful (although also not huge in light of the large overall stock market decline in 2020:Q1), and persists after the introduction of industry-fixed effects in model (2). The subsequent inclusion of country-fixed effects in model (3), however, diminishes the economic magnitude of this effect by almost three quarters to 1.129 (or only a 0.27% change in abnormal returns associated with a one SD change in the *ES* rating) and renders it statistically insignificant (t -stat = 1.08). The inclusion of both country- and industry-fixed effects in model (4) weakens the economic impact and statistical significance of *ES* slightly further. As we gradually add control variables in model (5) and our “baseline” model (6), the coefficient on *ES* actually turns negative – although still not statistically significant.

Of course, average ESG ratings may well differ across countries, and country-fixed effects also absorb such average differences in ESG ratings across countries and are in that sense rather crude. That being said, it seems perilous to attribute any significant effect of the *ES* rating on a firm’s stock market performance in the absence of a country-fixed effect to a true effect stemming from the differential ESG performance of firms. After all, countries differ along many, sometimes unobservable dimensions that could influence their stock market performance during 2020:Q1, including susceptibility to COVID-19 as well as medical and economic policy responses to the crises. Since it is virtually impossible to control for all of these country-level characteristics, the inclusion of country-fixed effects is simply a conservative way to assess the relation between ESG and stock market performance, which our results in Table 2 suggest is not driven by firm-specific ESG performance. We thus infer from Table 2 that, from a global perspective, stocks of firms with higher *ES* ratings did on average outperform their peers during the COVID-19 pandemic-induced market turmoil, but mainly due to the fact that these firms tend to be located in countries whose stock market performed relatively well in 2020:Q1.¹⁰ All in all, we

¹⁰ These findings deviate from Ding, Levine, Lin, and Xie (2020), who examine raw stock returns during the corona crash and conclude that, across a global sample, stocks of higher ESG-rated firms achieved higher weekly returns. It is

conclude that the evidence in support of the hypothesis that firms with higher ESG ratings showed better stock market performance during the COVID-19 crisis for our global sample is weak.

Regarding the results on the control variables in Table 2, we observe that stocks of financially stronger firms (i.e., firms with lower leverage and more cash) showed higher abnormal returns, which aligns with previous studies in the sustainable finance literature on stock reactions to exogenous shocks (Lins et al., 2017; Albuquerque et al., 2020ab); Demers et al., 2020; Ding et al., 2020; Ramelli and Wagner, 2020). For example, according to our baseline model, a one SD decrease in leverage (20.03) and cash ratio (19.21) was associated with, respectively, a 1.50% and 1.55% higher AR_{QI} . Consistent with the U.S. results of Albuquerque et al. (2020ab), we find that firm *size* has a positive effect on AR_{QI} , indicating that large firms showed better stock returns during 2020:Q1. As a robustness check, we reassess our baseline findings when excluding the dominant U.S. stocks Facebook, Apple, Amazon, Netflix, and Google, as well as Microsoft and Tesla (FAANGMT), from our global sample and obtain similar results (ES coefficient = -0.690 with a t -stat of -0.52; not tabulated).

4.3 Regional differences in the performance of ESG stocks during COVID-19

In this subsection, we shed more light on the relation between ES and AR_{QI} in various regions to examine whether ESG stocks do serve as “rainy day” assets in specific regions around the world. We adopt the classification scheme (largely geographic) from Ken French’s website and estimate our baseline model for the following regions: Europe, Asia-Pacific, Japan, North America, and Emerging Countries.

not clear what drives the differences in results; candidate explanations include using abnormal instead of raw returns in our baseline regression analysis, differences in the data frequency, and differences in the control variables.

Table 3: Cross-sectional regressions of stock returns during COVID-19 on ES rating by region

This table presents baseline regression (see Table 2) results of abnormal returns over 2020:Q1 (AR_{Q1}) on ESG performance, measured as the average of Refinitiv's Environmental and Social pillar ratings (ES) and various controls across different regions. Geographical classification is derived from Ken French's website. We use 2-digit SIC codes industry classification in the industry-fixed effects. Market and financial data are from Thomson Reuters' Datastream/Worldscope, and ESG data are from Thomson Reuters' Refinitiv ESG Database. AR_{Q1} and all continuous controls are winsorized at the 1st and 99th percentile. Appendix B provides a detailed variable description. Appendix C presents a sample distribution by region. All models include heteroscedasticity robust standard errors. Intercepts are suppressed to conserve space. Parentheses contain t -statistics. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	Global	Europe	Japan	Asia-Pacific	North America	Emerging Countries
<i>Dependent variable:</i>	AR_{Q1}	AR_{Q1}	AR_{Q1}	AR_{Q1}	AR_{Q1}	AR_{Q1}
<i>ES</i>	-0.775 (-0.59)	-3.842 (-1.39)	2.542 (0.64)	-2.923 (-0.54)	4.198* (1.69)	0.0706 (0.03)
<i>Tobin's Q</i>	0.841*** (4.82)	0.504 (1.30)	1.070 (1.22)	0.487 (1.01)	1.057*** (3.87)	1.039*** (2.86)
<i>Size</i>	2.071*** (7.46)	2.577*** (4.46)	1.311 (1.20)	1.845** (2.03)	2.037*** (4.21)	0.857 (1.33)
<i>Leverage</i>	-0.0749*** (-4.70)	-0.0133 (-0.38)	-0.141** (-2.43)	-0.0358 (-0.62)	-0.117*** (-4.73)	-0.0102 (-0.29)
<i>Cash ratio</i>	0.0809*** (4.27)	0.0301 (0.77)	0.0488 (0.75)	0.176*** (3.00)	0.108*** (3.00)	0.0391 (1.02)
<i>ROE</i>	-0.00499 (-0.47)	-0.0465* (-1.71)	-0.0233 (-0.17)	-0.0621 (-1.28)	0.00755 (0.63)	-0.0283 (-0.69)
<i>ROA</i>	0.0371 (0.83)	0.109 (1.07)	0.111 (0.25)	0.0849 (0.53)	0.104* (1.71)	-0.0646 (-0.44)
<i>Loss (dummy)</i>	0.435 (0.38)	-4.300* (-1.70)	-3.523 (-0.63)	-1.921 (-0.50)	2.030 (1.24)	-0.678 (-0.23)
<i>R&D & SGA</i>	0.0314** (2.00)	0.0459 (1.46)	-0.141** (-2.04)	-0.0755 (-1.49)	0.0343 (1.46)	0.0507 (0.94)
<i>Inventory turnover</i>	-0.00250 (-0.89)	-0.00760 (-1.37)	0.0102 (0.81)	0.00485 (0.57)	0.00301 (0.57)	-0.0139** (-2.52)
<i>BTM</i>	1.297** (2.34)	0.263 (0.19)	-1.291 (-0.55)	-0.242 (-0.17)	2.224** (2.15)	0.462 (0.50)
<i>Negative BTM (dummy)</i>	2.164 (0.97)	-0.820 (-0.18)	0 (.)	5.347 (0.46)	1.726 (0.66)	33.63*** (4.29)
<i>Momentum</i>	0.0704*** (8.29)	0.0929*** (4.93)	0.127*** (4.44)	0.0802*** (3.37)	0.0688*** (4.73)	0.0469*** (2.68)
<i>Idiosyncratic risk</i>	0.480*** (5.26)	1.113*** (4.27)	0.352 (0.72)	0.341 (1.41)	0.461*** (3.41)	0.450** (3.41)
<i>Market share</i>	1.673 (0.98)	-6.697* (-1.94)	1.932 (0.36)	11.29 (1.51)	3.856 (1.26)	4.332 (1.13)
<i>Dividend payout</i>	-0.000848 (-0.26)	-0.0166*** (-2.63)	-0.0256 (-0.97)	0.00167 (0.16)	0.0127** (2.44)	-0.00564 (-0.97)
N	5,965	1,365	402	639	2,338	1,221
Adj. R^2	0.247	0.221	0.273	0.199	0.312	0.187
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 3 shows the results. For purposes of comparison, we include the baseline regression results for the global sample (from Table 2) in the first column. It is clear from Table 3 that the relation between 2020:Q1 stock returns and *ES* differs markedly across the different regions around the world. The coefficient on *ES* is negative in two of the five regions and positive in three regions. The only significant coefficient on *ES* is observed for North America, for which the coefficient estimate of 4.198 is significant at the 10% level (t -stat=1.69). A one SD (0.216 for North America, not tabulated) increase in *ES* is associated with a 0.9% ($0.216 * 4.198$) higher AR_{Q1} within the North American region; we interpret this as a non-negligible but also not major “rainy day” effect of North American stocks during the COVID-19 crisis.

This result is weaker compared to findings by Albuquerque et al. (2020ab), who report that a similar increase in *ES* was associated with a 1.8% higher AR_{Q1} among U.S. firms. To illustrate, consider two North American firms that differ in *ES*, at the 75th percentile (0.431) and the 25th percentile (0.146). Our baseline model predicts that the firm with an *ES* rating at the 75th percentile showed a 1.20% ($(0.431 - 0.146) * 4.198$) higher AR_{Q1} . To put this in perspective, our model estimates that the impact of *ES* on AR_{Q1} is approximately 2/5th the magnitude of *cash ratio* and *leverage* among North American firms, respectively, which are both dominant predictors of corporate resilience in times of crisis. A likely explanation for the weaker impact of *ES* in our analysis compared to Albuquerque et al. (2020ab) is the inclusion of additional control variables that are correlated with *ES*.

Overall, the results in Tables 2 and 3 indicate that firms with higher *ES* ratings around the world did not have greater abnormal stock returns during 2020:Q1, except for North American firms.

4.4 Robustness tests of the performance of ESG stocks during COVID-19

Table 4 presents a number of robustness tests of our baseline results of the performance of ESG stocks during COVID-19. For comparison purposes, we include the baseline regression results for the global sample (from Table 2) in the first column. In model (2) of Table 4, we replace the *arithmetic return-based* AR_{Q1} , *stock momentum* and *idiosyncratic risk* with their *log return-based* variants. In models (2) and (3), we replace the *abnormal* returns as dependent variable by *raw buy-and hold* returns.

Table 4: Robustness of regression results of stock returns during COVID-19 on ES rating

This table presents the results of our robustness tests. Column 1 contains the baseline model (see Table 2). Column 2 shows regression results after we replace *arithmetic-based returns* with *log-based returns* in the calculations of our independent and dependent variables. Columns 3 and 4 show regression results when we replace *abnormal returns* by *raw returns*, over 2020:Q1 and the “fever period” (Ramelli and Wagner, 2020) from February 24 to March 18 (Albuquerque et al., 2020ab)). In models 1 to 4, we maintain a 2-digit SIC codes industry classification in the industry-fixed effects, and in model 5 we apply a TRBC industry group classification. Market and financial data are from Thomson Reuters’ Datastream/Worldscope, and ESG data are from Thomson Reuters’ Refinitiv ESG Database. (*log*) AR_{Q1} and all continuous controls are winsorized at the 1st and 99th percentile. Appendix B provides a detailed variable description. All models include heteroscedasticity robust standard errors. Intercepts are suppressed to conserve space. Parentheses contain *t*-statistics. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

<i>Dependent variable:</i>	(1) AR_{Q1}	(2) $Log AR_{Q1}$	(3) RR_{Q1}	(4) RR_{fever}	(5) AR_{Q1}
<i>ES</i>	-0.775 (-0.59)	-2.548 (-1.21)	-4.112*** (-3.19)	-2.180* (-1.95)	-0.241 (-0.18)
<i>Tobin's Q</i>	0.841*** (4.82)	1.228*** (4.81)	1.373*** (3.66)	0.767*** (5.89)	0.712*** (3.90)
<i>Size</i>	2.071*** (7.46)	2.997*** (6.63)	1.616*** (4.48)	0.856*** (3.60)	2.035*** (7.44)
<i>Leverage</i>	-0.0749*** (-4.70)	-0.165*** (-6.35)	-0.0864*** (-4.91)	-0.0850*** (-6.36)	-0.0561*** (-3.43)
<i>Cash Ratio</i>	0.0809*** (4.27)	0.127*** (4.47)	0.0333 (1.59)	0.0389*** (2.67)	0.0510*** (2.62)
<i>ROE</i>	-0.00499 (-0.47)	-0.00656 (-0.38)	0.00225 (0.15)	0.00220 (0.24)	-0.0110 (-1.06)
<i>ROA</i>	0.0371 (0.83)	0.0745 (1.03)	0.0587 (0.82)	0.0664* (1.80)	0.108** (2.33)
<i>Loss</i>	0.435 (0.38)	0.125 (0.07)	1.836 (1.50)	0.810 (0.85)	1.027 (0.89)
<i>R&D & SGA</i>	0.0314** (2.00)	0.0685*** (2.78)	0.0787*** (2.83)	0.0169 (1.34)	0.0227 (1.39)
<i>Inventory turnover</i>	-0.00250 (-0.89)	-0.00100 (-0.22)	0.000376 (0.13)	-0.000575 (-0.24)	-0.00151 (-0.55)
<i>BTM</i>	1.297** (2.34)	0.308 (0.32)	-0.974 (-1.31)	0.0229 (0.05)	1.348** (2.40)
<i>Negative BTM (dummy)</i>	2.164 (0.97)	1.820 (0.50)	2.078 (0.79)	1.678 (0.94)	2.326 (1.05)
<i>Momentum</i>	0.0704*** (8.29)	0.135*** (7.96)	0.0389*** (2.91)	-0.0149** (-2.17)	0.0739*** (8.53)
<i>Idiosyncratic risk</i>	0.480*** (5.26)	0.303* (1.91)	-0.273** (-2.56)	-0.521*** (-7.75)	0.444*** (4.72)
<i>Market share</i>	1.673 (0.98)	5.045* (1.88)	-0.262 (-0.14)	-0.136 (-0.08)	-2.338 (-1.47)
<i>Dividend payout</i>	-0.000848 (-0.26)	0.00255 (0.44)	0.00842** (2.37)	0.00449 (1.58)	0.00234 (0.73)
N	5,965	5,957	5,965	5,965	5,965
Adj. R ²	0.247	0.305	0.281	0.376	0.284
Industry FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes

We measure raw returns over 2020:Q1 in model (2) of Table 4 and over the “fever period” in model (3), which we define as the period from February to March 18, as proposed by Ramelli and Wagner (2020) and applied by Albuquerque et al. (2020ab). Model (5) uses the TRBC industry group classification instead of the 2-digit SIC code classification to define industry-fixed effects. Overall, Table 4 shows that none of the four robustness analyses indicates that stocks with higher ES ratings in our global sample significantly outperformed stocks with lower ES ratings during the COVID-19 crisis. In fact, two of the four tests (those based on raw returns in 2020:Q1 and in the fever period) suggest that firms with higher ES ratings actually underperformed significantly, although again the economic magnitudes of these effects are limited.

4.5 Specific measures of ESG performance and abnormal stock returns during COVID-19

In this subsection, we examine whether more detailed ESG sub-ratings can explain abnormal stock returns during 2020:Q1. We are specifically interested in ESG category ratings within the S pillar (except for Product Responsibility) and ESG theme ratings under the Workforce category, as these ratings are arguably of specific interest for understanding how well different firms handled the consequences of COVID-19 and thus their stock price responses in 2020:Q1. Table 5 presents a correlation matrix of all ESG (sub-)ratings under consideration and shows that most ESG (sub-)ratings are highly intercorrelated, which raises concerns about multicollinearity in our regressions. To mitigate these concerns, we include one ESG (sub-)rating in our cross-sectional regressions at the time.

The results of these more specific analyses are presented in Table 6. This table presents the coefficient on the ESG (sub-)rating from our cross-sectional regressions with abnormal returns over 2020:Q1 (AR_{Q1}) as dependent variable, as well as its t -statistic, the adjusted R^2 , and the number of observations used in each regression. All regressions in Table 6 include the full set of control variables and industry- and country-fixed effects, as in our baseline model in Table 2. For each ESG (sub-)rating that we consider, we present results both for the global sample and for North America. Again, we include the baseline regression results of ES for the global sample (from Table 2) in the first row. In the second row, we include the ES results for North America (from Table 3).

Table 5: ESG correlation matrix

This table presents a correlation matrix of all Refinitiv ESG (sub-)ratings in this study (N=6,824), including the average of Environmental and Social pillar (*ES*), the average of Environmental, Social and Governance pillar (*ESG*), the Environmental pillar (*E*), the Social pillar (*S*), the Governance pillar (*G*), the Community category (*Com*), the Human Rights category (*Rights*), the Workforce category (*Work*), the Health and Safety theme (*HS*), the Training and Development theme (*TD*), the Employment Quality theme (*EQ*), the Diversity and Opportunity theme (*DO*), and a Custom theme (*CT*). Appendix B provides a detailed variable description. Market and financial data are from Thomson Reuters' Datastream/Worldscope and ESG data are from Thomson Reuters' Refinitiv ESG Database. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	<i>ES</i>	<i>ESG</i>	<i>E</i>	<i>S</i>	<i>G</i>	<i>Com</i>	<i>Rights</i>	<i>Work</i>	<i>HS</i>	<i>TD</i>	<i>DI</i>	<i>EQ</i>	<i>CT</i>
<i>ES</i>	1												
<i>ESG</i>	0.945***	1											
<i>E</i>	0.948***	0.893***	1										
<i>S</i>	0.916***	0.869***	0.740***	1									
<i>G</i>	0.423***	0.695***	0.394***	0.396***	1								
<i>Com</i>	0.513***	0.517***	0.364***	0.624***	0.306***	1							
<i>Rights</i>	0.753***	0.707***	0.630***	0.794***	0.305***	0.364***	1						
<i>Work</i>	0.689***	0.652***	0.605***	0.691***	0.294***	0.250***	0.536***	1					
<i>HS</i>	0.650***	0.620***	0.579***	0.641***	0.290***	0.257***	0.517***	0.719***	1				
<i>TD</i>	0.618***	0.597***	0.548***	0.614***	0.296***	0.231***	0.487***	0.713***	0.515***	1			
<i>DI</i>	0.645***	0.617***	0.564***	0.649***	0.294***	0.301***	0.498***	0.755***	0.477***	0.478***	1		
<i>EQ</i>	-0.0498***	-0.0653***	-0.0557***	-0.0350***	-0.0715***	-0.148***	-0.0310**	0.187***	-0.0422***	-0.00687	-0.0288**	1	
<i>CT</i>	0.688***	0.656***	0.615***	0.677***	0.306***	0.279***	0.529***	0.776***	0.831***	0.535***	0.694***	-0.0396***	1

In the remaining rows of Table 6, we first present the results when we use the overall ESG rating or the individual E, S, and G pillar ratings separately instead of the ES rating. We then add a first layer of detail to our analysis and estimate the effect of the ESG category ratings of firms on AR_{Q1} . We include category ratings for *Community*, *Human Rights*, and *Workforce*. Next, we introduce a further layer of detail as we narrow down the *Workforce* context and estimate the effects of ESG theme ratings, including *Diversity and Inclusion* (DI), *Health and Safety* (HS), *Training and Development* (TD), and *Employment Quality* (EQ). From the latter theme, we select six metrics that we perceive as particularly material in the context of the COVID-19 pandemic, which we label as “*Custom Theme*” (CT). Appendix B provides a detailed description of all these ESG (sub-)ratings.

Taken together, the results in Table 6 indicate that there is little evidence that the abnormal stock returns of around 6,000 firms around the world are systematically related to ESG (sub-)ratings. Of the 13 ESG (sub-)ratings included in the table, only one has a coefficient that is statistically significant, namely the G rating with a coefficient of -2.032 (t -stat = -1.83). This finding suggests that firms with better governance actually showed worse stock market performance during 2020:Q1. However, both the statistical and the economic significance of this effect are modest, and we need to take into account the possibility that this is a “false positive” (type I error), given the multiple tests presented in this table.

For North America, the picture is somewhat different. The positive and significant effect of ES on 2020:Q1 abnormal returns for North American firms documented in Table 3 appears to stem from the S pillar rating, which has a similar coefficient and t -statistic as *ES*, while the coefficients on neither E nor G are significant. Within the S pillar, the category rating for *Workforce* is the only one with a significant coefficient (4.404, with a t -statistic of 2.42), in line with the impression that the abnormal performance of firms during the COVID-19 crisis may be partly explained by how well firms treat their employees. At the same time, these results seem to suggest a different (or complementary) explanation for the outperformance of North American firms with higher ESG ratings than the customer and investor loyalty explanation proposed by Albuquerque et al. (2020ab).

Table 6: Cross-sectional regressions of stock returns during COVID-19 on ESG (sub-)ratings

This table presents baseline model regressions (see Table 2) of abnormal stock returns over 2020:Q1 (AR_{Q1}) on various ESG (sub-)ratings for the global sample (N=6,824) and the North American subsample (N=2,734). ESG (sub-)ratings include the average of the Environmental and Social pillar (ES), the average of the Environmental, Social pillar and Governance pillar (ESG), the Environmental pillar (E), the Social pillar (S), the Governance pillar (G), the Community category (Com), the Human Rights category ($Rights$), the Workforce category ($Work$), the Health and Safety theme (HS), the Training and Development theme (TD), the Employment Quality theme (EQ), the Diversity and Opportunity theme (DO), and a Custom theme (CT). Geographical classification is derived from Ken French's website. We use 2-digit SIC codes industry classification in the industry-fixed effects. Market and financial data are from Thomson Reuters' Datastream/Worldscope, and ESG data are from Thomson Reuters' Refinitiv ESG Database. AR_{Q1} and all continuous controls are winsorized at the 1st and 99th percentile. Appendix B provides a detailed variable description. All models include heteroscedasticity robust standard errors. Intercepts are suppressed to conserve space. Parentheses contain t -statistics. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	Sample	AR_{Q1}	t -stat	Adj. R^2	N	Controls	FE
ES	Global	-0.775	(-0.59)	0.247	5,965	Yes	Industry + Country
	North America	4.198*	(1.69)	0.312	2,338	Yes	Industry + Country
ESG	Global	-1.907	(-1.27)	0.247	5,965	Yes	Industry + Country
	North America	3.498	(1.21)	0.312	2,338	Yes	Industry + Country
E	Global	-1.358	(-1.21)	0.247	5,965	Yes	Industry + Country
	North America	2.845	(1.39)	0.312	2,338	Yes	Industry + Country
S	Global	0.301	(0.23)	0.247	5,965	Yes	Industry + Country
	North America	4.813*	(1.67)	0.312	2,338	Yes	Industry + Country
G	Global	-2.032*	(-1.83)	0.248	5,965	Yes	Industry + Country
	North America	0.207	(0.10)	0.311	2,338	Yes	Industry + Country
Com	Global	-0.832	(-0.80)	0.268	5,373	Yes	Industry + Country
	North America	-1.783	(-0.78)	0.337	2,103	Yes	Industry + Country
$Rights$	Global	-0.116	(-0.11)	0.268	5,373	Yes	Industry + Country
	North America	1.257	(0.68)	0.337	2,103	Yes	Industry + Country
$Work$	Global	0.802	(0.79)	0.268	5,373	Yes	Industry + Country
	North America	4.404**	(2.42)	0.339	2,103	Yes	Industry + Country
DI	Global	-0.473	(-0.46)	0.260	5,763	Yes	Industry + Country
	North America	1.039	(0.54)	0.326	2,235	Yes	Industry + Country
HS	Global	-0.0276	(-0.03)	0.263	5,713	Yes	Industry + Country
	North America	1.601	(0.91)	0.327	2,233	Yes	Industry + Country
TD	Global	-0.0844	(-0.09)	0.263	5,713	Yes	Industry + Country
	North America	1.137	(0.68)	0.327	2,233	Yes	Industry + Country
EQ	Global	1.355	(1.47)	0.261	5,231	Yes	Industry + Country
	North America	1.366	(0.93)	0.329	2,079	Yes	Industry + Country
CT	Global	0.508	(0.51)	0.263	5,713	Yes	Industry + Country
	North America	1.073	(0.60)	0.327	2,233	Yes	Industry + Country

None of the five theme ratings within the *Workforce* category in Table 6 (that is, DI , HS , TD , EQ , and CT) have a significant coefficient for North America, and we thus fail to uncover the specific aspect of the *Workforce* rating that explains cross-sectional differences in stock market performance during the COVID-19 crisis. All in all, the results in Table 6 support our earlier conclusion that firms with higher ESG ratings did not outperform their peers in 2020:Q1, with the exception of firms in North America.¹¹

¹¹ We obtain similar results when excluding the FAANGMT stocks from the global and North American samples in Table 6.

5. Conclusion

Are firms with superior ESG performance more resilient to external unexpected shocks? We empirically investigate this issue by testing whether firm-level ESG ratings are positively associated with (abnormal) stock returns during the COVID-19 crisis in the first quarter of 2020 in a global setting. We show that, when we control for country-wide effects, stronger ESG firms did not show better stock returns because of the ES (or ESG) rating, but because they are domiciled in countries whose stocks in general showed a relatively better (abnormal) stock return. Such country-wide effects cannot be reliably attributed to the sustainability profile of individual firms per se. Only for the North America region do we find some evidence of greater resilience in terms of stock market performance during crisis periods of firms with higher ESG ratings.

The reliability of the ESG ratings used in the empirical analysis is an important precondition for interpreting our results. Notoriously, the input for most ratings are based on self-reported and unaudited data. Also, the measurement and rating aggregation approaches differ across rating providers. As a consequence, ESG ratings, different from financial creditworthiness ratings, show a high degree of inconsistency (Berg et al., 2019; Busch et al., 2020). Besides the technical shortcomings of the ESG ratings currently available, these ratings may fail to fully capture intangible elements such as the social capital endowment, the corporate culture, and the perceived sustainability product differentiation, all of which are supposed to financially benefit the socially responsible firms and not the firms that invest less in sustainability. Future research should shed light on which ESG ratings are more reliable, and on whether our results depend on the ESG ratings used.

A further possible limitation of our study lies in the imperfect synchronicity of the shock across countries, as well as in the heterogeneous responses put in place by different governments around the world. The speed and quality of government responses (or the lack thereof) were markedly heterogeneous, and some specific national policy measures may have weighted in the market's perception of the ability of firms to weather the pandemic. Nonetheless, the face value of our findings supports the view that firms with higher ESG ratings were not more resilient to the COVID-19 shock than firms with lower ESG ratings. At best, the contribution of corporate sustainability to the financial resilience of firms is geography-dependent and limited to the North American market.

References

- Acharya, V. V. and Steffen, S., 2020. The risk of being a fallen angel and the corporate dash for cash in the midst of COVID. *Review of Corporate Finance Studies*, forthcoming.
- Albuquerque, R., Y. Koskinen, S. Yang, C. Zhang, 2020a. Love in the time of COVID-19: The resiliency of environmental and social stocks. *Covid Economics: Vetted and Real-Time Papers* 11, CEPR.
- Albuquerque, R., Y. Koskinen, S. Yang, and C. Zhang, 2020b. Resiliency of environmental and social stocks: An analysis of the exogenous COVID-19 market crash. *Review of Corporate Finance Studies*, 593-621.
- Albuquerque, R., Y. Koskinen, and C. Zhang, 2019. Corporate social responsibility and firm risk: Theory and empirical evidence. *Management Science* 65, 4451-4469.
- AXA Investment Management, 2020. Coronavirus: How ESG scores signalled resilience in the Q1 market downturn, April 7, 2020.
- Berg, F., Kolbel, J., and Rigobon, R., 2019. Aggregate confusion: the divergence of ESG ratings. MIT Sloan Research Paper No. 5822-19.
- Busch, T., Johnson, M., and Pioch, T., 2020. Corporate carbon performance data: Quo vadis? *Journal of Industrial Ecology*, 1-14.
- Cheema-Fox, A., LaPerla, B. R., Serafeim, G., and Wang, H., 2020. Corporate resilience and response during COVID-19. Working Paper, Harvard Business School.
- Demers, E., Hendrikse, J., Joos, P. and Lev, B., 2020. ESG didn't immunize stocks against the Covid-19 market crash, SSRN working paper.
- Dyck, A., Lins, K.V., Roth, L., and Wagner, H., 2019. Do institutional investors drive corporate social responsibility? International evidence. *Journal of Financial Economics* 131, 693-714.
- Fahlenbrach, R., Rageth, K., and Stulz, R., 2020. How valuable is financial flexibility when revenue stops? Evidence from the Covid-19 crisis. Ohio State University working paper.
- Ferriani, F. and Natoli, F., 2020. ESG risks in times of Covid-19, *Applied Economics Letters*, forthcoming.

- Gantchev, N., Giannetti, M., and Li, Q., 2019. Does money talk? Market discipline through selloffs and boycotts, CEPR Discussion Paper No. DP14098.
- Gibson, R., Glossner, S., Krueger, P., Matos, P., and Steffen, T., 2020. Responsible institutional investing around the world. Swiss Finance Institute Research Paper No. 20-13.
- Griffin, J.M., 2002. Are the Fama and French factors global or country specific? *Review of Financial Studies* 15, 783-803.
- Heinkel, R., Kraus, A., and Zechner, J., 2001. The effect of green investment on corporate behavior. *Journal of Financial and Quantitative Analysis* 36, 431-449.
- Lins, K., Servaes, H., and Tamayo, A., 2017. Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *Journal of Finance* 72, 1785-1824.
- Pagano, M., Wagner, C. and Zechner, J., 2020. Disaster resilience and asset prices. SSRN working paper.
- Ramelli, S. and Wagner, A., 2020. Feverish stock price reactions to Covid-19, SSRN working paper.
- Renneboog, L., ter Horst, J., and Zhang, C., 2011. Is ethical money financially smart? Nonfinancial attributes and money flows of socially responsible investment funds. *Journal of Financial Intermediation* 20, 562-588.

Appendix A: Refinitiv’s ESG scoring hierarchy

Figure A1 shows Refinitiv’s data structure to evaluate corporate ESG performance. 186 ESG Metrics, divided over ten categories, are used to construct an Environmental, Social, and Governance rating. For example, the Workforce rating (under the Social pillar) is calculated based on thirty different ESG metrics. Figure A2 shows the steps involved in the Refinitiv ESG rating methodology.

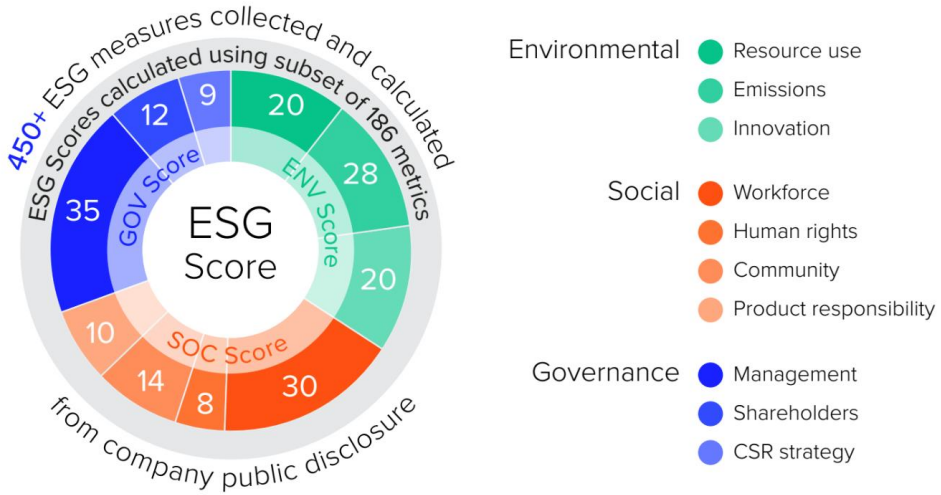


Figure A1. Source: Refinitiv. April 2020. “Environmental. Social and Governance (ESG) ratings from Refinitiv”. Downloaded from: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/esg-ratings-methodology.pdf

The Refinitiv ESG scoring methodology can be summarized and illustrated by means of a five-step process flow.

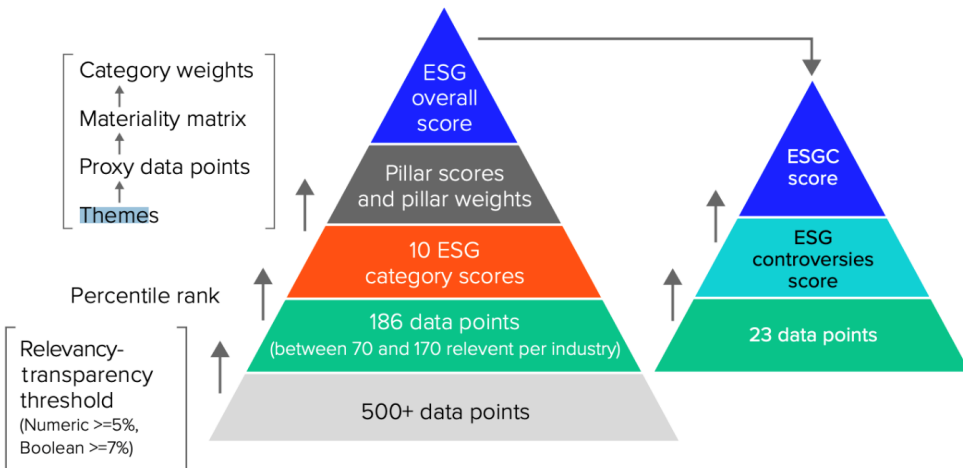


Figure A2. Source: Refinitiv. April 2020. “Environmental. Social and Governance (ESG) ratings from Refinitiv”. Downloaded from: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/esg-ratings-methodology.pdf

Appendix B: Variable descriptions

This table presents variable descriptions. Market and financial data are from Thomson Reuters' Datastream/Worldscope, and sustainability data are from Thomson Reuters' Refinitiv ESG Database. Datastream identifiers are shown in parentheses.

Variable	Definition	Source
<i>ES</i>	The simple average of Refinitiv's Environmental (ENSCORE) and Social (SOSCORE) Pillar Ratings divided by 100.	Refinitiv
<i>ESG</i>	The simple average of Refinitiv's Environmental (ENSCORE), Social (SOSCORE), and Governance (CGSCORE) Pillar Ratings divided by 100.	Refinitiv
<i>E</i>	Refinitiv's Environmental (ENSCORE) Pillar Ratings divided by 100.	Refinitiv
<i>S</i>	Refinitiv's Social (SOSCORE) Pillar Ratings divided by 100.	Refinitiv
<i>G</i>	Refinitiv's Governance (CGSCORE) Pillar Ratings divided by 100.	Refinitiv
<i>Community</i>	Refinitiv's rating for the Community category (TRESGSOCOS), which is subject to the Social Pillar. Measured for the year 2018.	Refinitiv
<i>Human Rights</i>	Refinitiv's rating for the Human Rights category (TRESGSOHRS), which is subject to the Social Pillar. Measured for the year 2018.	Refinitiv
<i>Workforce</i>	Refinitiv's rating for the Workforce category (TRESGSOWOS), which is subject to the Social Pillar. Measured for the year 2018.	Refinitiv
<i>DI</i>	Average relative performance rating across 10 Diversity and Inclusion-related ESG metrics under the Workforce category. Measured for the year 2018.	Refinitiv
<i>HS</i>	Average relative performance rating across 9 Health and Safety-related ESG metrics under the Workforce category. Measured for the year 2018.	Refinitiv
<i>EQ</i>	Average relative performance rating across 6 Employment Quality-related ESG metrics under the Workforce category. Measured for the year 2018.	Refinitiv
<i>TD</i>	Average relative performance rating across 5 Training and Development-related ESG metrics under the Workforce category. Measured for the year 2018.	Refinitiv
<i>CT (custom theme)</i>	Average relative performance rating across 6 Health and Safety-related ESG metrics under the Workforce category. Measured for the year 2018.	Refinitiv
<i>AR_{Q1}</i>	The difference between 2020:Q1 arithmetic stock return (derived from RI) and beta, times the 2020:Q1 return of the local benchmark (derived from LI). 2020:Q1 returns are estimated from January 1, 2020 to April 1, 2020. Beta is estimated by using 60-month estimation window between 2015 and 2019, where the market factor is arithmetic return of the local benchmark (derived from LI). All values are measured in US\$.	Datastream
<i>Tobin's Q</i>	Total asset value (WC02999) minus book value of equity (WC05476 * WC05301) + market value of equity (WC08001), all divided by total assets (WC02999). All values are measured in US\$.	Worldscope
<i>Size</i>	Natural log of market cap (WC08001) in million US\$.	Worldscope
<i>Cash ratio</i>	Cash and short-term investments (WC02001) over total assets (WC02999) and multiplied by 100. All values are measured in US\$.	Worldscope
<i>Leverage</i>	Total debt (WC03255) over total assets (WC02999) and multiplied by 100. All values are measured in US\$.	Worldscope
<i>ROE</i>	Net income (WC01706) over book value of equity (WC05476 * WC05301) and multiplied by 100. All values are measured in US\$.	Worldscope
<i>ROA</i>	Net income (WC01706) minus extraordinary items (WC04225) minus discontinued operations (WC4054), all divided by total assets (WC02999) and multiplied by 100. Set to zero if missing. All values are measured in US\$.	Worldscope

Variable	Definition	Source
<i>Loss</i>	A dummy set to one if ROA < 0.	Worldscope
<i>Inventory turnover</i>	Firm-specific inventory turnover (WC01051/WC02101) divided by the average industry inventory turnover and multiplied by 100. Set to zero if missing. Industry classification by 2-digit SIC codes. All values are measured in US\$.	Worldscope
<i>R&D & SGA</i>	R&D (WC01202) + 1/3*SGA (WC01101) using 5-year amortization and divided by total assets in year t (WC02999). Set to zero if missing. All values measured in US\$.	Worldscope
<i>BTM</i>	Book value of equity (WC05476 * WC05301) divided by market value of equity (WC08001). All values measured in US\$ for the year 2019.	Worldscope
<i>BTM negative</i>	A dummy set to one if BTM < 0.	Worldscope
<i>Momentum</i>	Buy-and-hold return (RI) of the year prior to the start of the return period. All values expressed in US\$. Used in cross-sectional analysis.	Worldscope
<i>Idiosyncratic risk</i>	The firm-specific root mean squared error of the domestic market model regression estimations.	Worldscope
<i>Market share</i>	Revenues (WC01001) divided by total industry revenues and multiplied by 100. Industry classification by 2-digit SIC codes. All values measured in US\$.	Worldscope
<i>Dividend payout</i>	Dividends (WC05376) divided by net income (WC01706) and multiplied by 100. All values measured in US\$.	Worldscope

The COVID-19 bailouts¹

Jean-Marie Meier² and Jake Smith³

Date submitted: 22 June 2021; Date accepted: 22 June 2021

We use hand-collected data to investigate the COVID-19 bailouts for all publicly listed US firms. The median tax rate is 4% for bailout firms and 16% for no-bailout firms. The bailouts are expensive when compared to past corporate income tax payments of the bailout firms. We compute the number of years a bailout recipient has to pay corporate income tax to generate as much tax revenue as it received in bailouts: 135.0 years for the Paycheck Protection Program and 267.9 years for the airline bailouts. We also document a dark side of the bailouts. For many firms, the bailouts appear to be a windfall. Numerous bailout recipients made risky financial decisions, so bailing them out might induce moral hazard. Moreover, lobbying expenditures positively predict the bailout likelihood and amount.

- 1 We thank David Schoenherr, Henri Servaes, Jan Starmans, Sorabh Tomar, and Steven Xiao for helpful comments and discussions. We thank Keerthi Bisaramanempalli Reddiwanda and Mohit Sanjay Sahasrabudhe for excellent research assistance.
- 2 Assistant Professor of Finance, University of Texas at Dallas, Jindal School of Management.
- 3 PhD Candidate, University of Texas at Dallas, Jindal School of Management.

Copyright: Jean-Marie Meier and Jake Smith

1 Introduction

COVID-19 hit the world with unprecedented force. The responses by the US government and the Federal Reserve have been similarly unprecedented. This paper investigates two bailout programs that have experienced significant uptake: the \$659 billion Paycheck Protection Program (PPP) for “small” businesses and the \$32 billion Payroll Support Program for the aviation industry (henceforth “airline bailouts”).

Critics of these interventions argue that bailouts socialize losses, while past profits have been paid out to shareholders.¹ Moreover, bailouts can create moral hazard and lead to excessive risk taking (Farhi and Tirole, 2012; Duchin and Sosyura, 2014). Poorly designed bailouts can also be expensive for taxpayers and generate windfalls for the private sector, while being ineffective in alleviating a crisis.² In contrast, bailout supporters argue that bailouts are necessary to keep workers employed and avoid the crisis from worsening.

To shed light on this debate, one would, in an ideal setting, conduct a firm-level analysis of the bailouts. This is not possible, however, since firm-level data is not available for privately held firms. Therefore, we focus on publicly listed US firms and combine this with detailed hand-collected bailout data from corporate filings with the Securities and Exchange Commission (SEC).

Our paper has two main contributions. First, the bailouts are expensive—both on a bailout funds-per-employee basis and when compared to past corporate income tax payments of the bailout firms. Second, we find a dark side of the bailouts, which appear to be a windfall for some firms and potentially induce moral hazard for other firms that made risky financial decisions.

We collect data on 755 bailouts worth \$17.9 billion. The mean and median airline bailout per employee are \$34.39 thousand and \$31.76 thousand, respectively. The corresponding

¹Amit Seru and Luigi Zingales (2020): “Save Capitalism From the Cares Act”, March 30, www.wsj.com/articles/save-capitalism-from-the-cares-act-11585608917 (accessed May 1, 2020).

²Amit Seru and Luigi Zingales (2020): “The Stimulus Package Is Too Expensive and Poorly Targeted”, March 29, <https://promarket.org/the-stimulus-package-is-too-expensive-and-poorly-targeted-the-waste-contained-in-the-cares-act> (accessed April 28, 2020)

values for PPP bailouts are \$21.70 thousand and \$16.16 thousand.

We compute the number of years a bailout recipient has to pay corporate income tax to generate as much tax revenue as it received in bailouts. The mean and median number of years for a PPP recipient are 135.0 and 22.1, respectively. The numbers for the airline bailouts are 267.9 and 138.3 years. These values are driven by low effective tax rates and the size of the bailouts. The median tax rate is 4% for bailout firms and 16% for no-bailout firms, while the current statutory corporate income tax rate is 21%. A small number of bailout firms are resident in tax havens such as Bermuda or Ireland. Given the substantial size of the bailouts and the significant US fiscal deficits in 2020 and 2021, the stylized facts on the aforementioned effective tax rates and “years to repay” might be relevant for the ongoing policy debate on US corporate taxation.

We now turn to the dark side of the bailouts. First, 66 firms paid out more in dividends and net repurchases from 2015-2019 than they received in bailouts, potentially inducing future moral hazard problems. Second, 437 firms had more cash and cash equivalents at the end of 2019 than they received in bailouts, suggesting that the bailouts might be a windfall for these firms. Third, a substantial fraction of the bailout firms can be considered start-up like firms, for which the bailouts are likely a windfall, as well. Fourth, many of the recipients of the bailouts are quite large, implying that some firms might have been able to raise additional financing absent a bailout (Hadlock and Pierce, 2010).

Next, we run cross-sectional regressions to ascertain the determinants of the incidence and magnitude of the bailouts. Greater assets, cash/assets and Tobin’s Q are associated with a lower bailout likelihood and amount. Surprisingly, firm age and sales both have a positive effect on the bailout probability and amount. A dummy for firms with a persistent negative EBITDA has a powerful effect on the bailout probability and amount, supporting our earlier point about these start-up like firms. Lobbying expenditures have a sizable effect on the bailout likelihood and amount. Firms that lobby might be experienced in navigating bureaucracy and red tape, and might therefore be in a better position to disentangle the

bailout rules.

To what extent can our results be extrapolated to privately held firms? The average privately held bailout recipient will most likely be more financially constrained than the average publicly listed one. Therefore, the likelihood that a bailout of a privately held firm involves any of the documented “dark sides” of bailouts will be lower than for a publicly listed firm. However, because the vast majority of PPP bailout funds went to privately held firms, it is reasonable to assume that the “dark side” of the bailouts for privately held firms will be quantitatively large due to the sheer size of the private sector and the bailouts. There will be many large privately held firms that will not be financially constrained and do not need a bailout. In addition, the bailouts do not condition on whether a firm was affected by or could have survived the COVID-19 crisis without a bailout.

One limitation of our paper is that we cannot establish causality, so the results should be interpreted as suggestive evidence.

Lastly, we discuss the policy implications. One implication is that the bailouts should have been conditioned on whether a firm has been affected by the COVID-19 crisis. The airline bailouts also appear overly generous on a bailout-per-employee basis and expensive when compared to recent corporate income tax payments by the airlines. Moreover, the large publicly listed airlines paid out more to their shareholders in the last couple of years than they received in bailouts, suggesting that these bailouts might be inducing moral hazard. Delta, for instance, had a pre-tax income in 2019 of \$5.7 billion, on which it received a tax refund of \$95.00 million. The bailout Delta received was \$3.8 billion, while its aggregate payouts to shareholders from 2015-2019 were \$13.6 billion. In addition, the four largest airlines on average went bankrupt 4.25 times since the 1980s. This raises the question of why bankruptcy (or fire sales) could not have been used instead of bailouts to restructure the airlines.

Due to the vast number of papers on COVID-19, our literature discussion focuses only on the most closely related papers. Our paper contributes to the literature on the COVID-19

bailouts. To the best of our knowledge, our paper is the first to document the dark sides of the COVID-19 bailouts and the high cost of the bailouts when compared to the corporate income tax payments of the bailout firms. Elenev et al. (2020) use a macroeconomic model to document that the bailouts prevented a much deeper crisis. Granja et al. (2020) study the congressional district-level distribution of PPP bailouts using confidential data from the SBA, and find that PPP funds initially flowed more to areas less affected by COVID-19. Using survey data on small businesses, Bartik et al. (2020) investigate the employment effects of the PPP bailouts and find a positive but insignificant effect. Chetty et al. (2020) find that the PPP increased employment at small businesses by 3%, implying a cost of \$290 thousand per job saved. Autor et al. (2020) estimate that the cost per job saved by PPP is \$224 thousand.

Moreover, our paper contributes to the literature on bailouts more broadly by documenting the dark sides of the COVID-19 bailouts. For a discussion of the 2008/9 bailouts, see Calomiris and Khan (2015) and Goolsbee and Krueger (2015). Bailouts can increase moral hazard in the future (Farhi and Tirole, 2012; Duchin and Sosyura, 2014), and can be distorted because of political connections (Faccio et al., 2006; Duchin and Sosyura, 2012). Moreover, receiving a bailout can subject firms to political influences (Chavaz and Rose, 2019). Meier and Servaes (2019) argue that fire sales, an alternative for firms that do not receive a bailout, are not as costly from a welfare perspective as previously argued by policymakers.

2 Institutional Details of the Bailouts

The US government and the Federal Reserve have approved five private-sector bailout programs worth \$1.8 trillion in total. See Table 1 for an overview of these programs. The Federal Reserve enacted additional emergency initiatives to fight the crisis that are not direct bailouts. Table A1 provides an overview of the different emergency programs by the Federal Reserve, their dollar volume, the eligible borrowers/beneficiaries, the collateral/assets, and a classification of whether a program involves a bailout. In this paper, we focus on the two

bailout programs that have been widely used—the \$32 billion airline bailouts and the \$659 billion “small” business bailouts through the Small Business Administration’s (SBA) PPP. For a legislative history of the federal government’s emergency measures, see Appendix A.1.

Table 1: Five Direct Private Sector Bailout Programs

Bailout Program	Size	Provider	Financing
Airline Bailouts: Payroll Support Program	32	Government	CARES Act
Industries Required for National Security	17	Government	CARES Act
Small-Business Bailouts: Pay-check Protection Program	659	Government	\$349 billion from CARES Act and \$310 billion from Paycheck Protection Program and Health Care Enhancement Act
Large-Business Bailouts: Main Street New Loan Facility, Main Street Priority Loan Facility, and Main Street Expanded Loan Facility	600	Fed	\$75 billion equity investment from the Treasury through the CARES Act and self-made leverage from the Federal Reserve
Mega-Firm Bailouts: Primary Market Corporate Credit Facility	500	Fed	\$50 billion equity investment from the Treasury through the CARES Act and self-made leverage from the Federal Reserve
Total	1,808		

This table provides an overview of five direct private sector bailout programs from the federal government (abbreviated government) and the Federal Reserve (abbreviated Fed). Size is the amount of the program in billions of dollars.

Airline Bailouts through the Payroll Support Program

The \$32 billion allocated for airline bailouts by the CARES Act includes \$25 billion for passenger airlines, \$4 billion for cargo airlines, and \$3 billion for airline contractors. Passenger airlines that receive more than \$100 million (\$50 million for cargo carriers) are required to issue a loan and warrants to the Treasury. Airline contractors that receive more than \$37.5 million must issue a loan only to the Treasury. 100% of the funds received must

be used for “employee wages, salaries, and benefits.”³ The face value of the loans is up to 30% of the total funds received for passenger airlines, and 49% for Atlas Air, the only cargo airline that received more than \$50 million in payroll support. The warrants are issued at-the-money, with a term of five years. The amount of the warrants is such that the strike price times the number of warrants is approximately equal to 10% of the face value of the loan (less than 3% of the total funds received for the passenger airlines). Not included in the numbers for the airline industry was the suspension of aviation excise taxes through January 1, 2021. The Treasury publishes a list of recipients of the airline bailouts on its homepage. *“Small” Business Bailouts through the Paycheck Protection Program*

The first tranche of the \$349 billion PPP bailouts from the CARES Act became available for payout on April 3, 2020 and was depleted within two weeks. The second \$310 billion tranche from the Paycheck Protection Program and Health Care Enhancement Act became available for payout on April 27, 2020. The application deadline for PPP bailouts was August 8, 2020. After the top off, every firm eligible for a PPP bailout was able to get a bailout if it applied. PPP funds come in the form of forgivable loans. The PPP loan amount is equal to 2.5 times the average monthly payroll costs pre-COVID-19 (capped at \$100,000 per employee), with a maximum PPP loan amount of \$10 million. Loan payments are not required for the first six months after issuance. In the Paycheck Protection Program Flexibility Act (PPPFA) (signed into law on June 5, 2020), the period during which no loan payments have to be made was extended by several months. Each PPP loan, if not converted into a grant, has an interest rate of 1%. Each PPP loan approved before the PPPFA (and not converted into a grant) has a maturity of two years. PPP bailouts approved after the PPPFA was passed (and not converted into a grant) have a maturity of five years. At least 75% (lowered by PPPFA to 60%) of the PPP loan amount must be used for payroll costs. Allowable non-payroll costs include payments of mortgage interest, other interest, rent, and utilities.

³US Department of the Treasury (2020): “Payroll Support Programs,” home.treasury.gov/policy-issues/cares/preserving-jobs-for-american-industry/payroll-support-program-payments (accessed September 12, 2020).

Importantly, the PPP loan can be fully forgiven if the recipient maintains employment and pay levels during an 8-week (24-week after PPPFA) period after the origination of the loan and subject to other conditions, such as the aforementioned 75% (60%) rule.

Eligibility for the Paycheck Protection Program

The eligibility rules for the PPP bailouts are opaque, complex, and contradictory. Information on eligibility for PPP funds is provided in a Frequently Asked Questions (FAQ) document on the Treasury's homepage. This frequently changing FAQ document is the main framework used to administer the \$659 billion PPP bailout program. In practice, the PPP rules do not condition eligibility on the financial health of a firm. Thus, the PPP is a "first-come, first-served" program with eligibility criteria primarily based on firm size that intends to enable firms to retain employment during the economic shock caused by COVID-19. The PPP has three main eligibility rules based on firm size and industry classification. A firm must meet at least one of the criteria to be eligible for a PPP bailout.⁴

First, as a rule of thumb, most firms with at most 500 employees are eligible. Second, there are exceptions for all firms whose NAICS code starts with 72, which includes hotels and restaurants. These firms can obtain bailouts even if they have more than 500 employees, as long as each location or legal entity that applies for the bailout has at most 500 employees. Third, there are additional opportunities for firms with more than 500 employees to obtain PPP bailouts through the special SBA industry size cutoffs, which includes firms with up to 1,500 employees.

3 Data

We collect data on PPP bailouts from the SEC and airline bailouts from the Treasury. We remove firms that received but subsequently repaid their entire bailout. See Appendix B.1 for how we treat the special case of one airline holding company. Since we are only interested in the grant portion of the bailouts, we interpret the grant for the airline bailouts

⁴See Appendix A.2 for additional details on the eligibility criteria.

as the difference between the total funds received and the face value of the loan. No such adjustment is needed for the PPP bailouts.

The sample period for our firm data is 2010-2019. Firm data is from the following sources: Compustat North America for accounting and stock price data, CRSP for stock price data, Compustat Snapshot for historical company names, and OpenSecrets.org for lobbying expenditure data (Center for Responsive Politics, 2018). The lobbying data is from 2010-2018, and is matched with Compustat using a fuzzy name match. All variables are defined in Table 2. For variables with a measure of income in the denominator, we require that income is positive. All variables are winsorized at the 1% level except for the bailout variables. Tax rate variables are further winsorized to ensure that they are between 0 and 100%. For variables using stock price data, we first define the variable using data from CRSP. If the variable is missing, then we replace it with data from Compustat Security. Lastly, we keep only the “primary” instance of dual-listed firms using the data from Meier and Smith (2020).

Table 2: Variable Definitions

SEC and Treasury	
Bailout Amount	Bailout Amount
Bailout Dummy	Equals 100 if the firm received a bailout and 0 otherwise
Compustat/CRSP	
Market Cap	Price-per-share times shares outstanding ($prccm \times cshoq$)
Book Assets	Total book assets (at)
Sales	Sales ($revt$)
Number of Employees	Number of employees (emp)
Firm Age	Firm age in years based on the IPO date, or, if missing, the date the firm first appeared in Compustat
$\mathbb{1}(\text{Total Debt} > 0)$	Dummy that equals 1 if total debt is positive ($dltt + dlc$)
Crisis Return	Stock return including dividends from 2-19-20 to 3-23-20 ($(prcd_t/prcd_{t-1}) \times (trfd_t/trfd_{t-1}) - 1$)
Payouts	Dividends plus buybacks from 2015-2019 ($\sum_t (dvc + prstkc - sstk)$)
Payout Ratio	Payouts divided by income before extraordinary items from 2015-2019 ($\sum_t (dvc + prstkc - sstk) / \sum_t ib$)
Tobin's Q	Market assets over book assets ($(lt + pstk - txditc + prccm \times cshoq) / at$)
Book Leverage	Total book liabilities divided by book assets (lt/at)
Market Leverage	Total book liabilities divided by market assets ($lt / (lt + pstk - txditc + prccm \times cshoq)$)
Sales Growth	Growth in sales from 2018 to 2019 ($revt_t / revt_{t-1} - 1$)
EBITDA/Assets	EBITDA divided by book assets ($ebitda/at$)
$\mathbb{1}(\text{3-Yr EBITDA} < 0)$	Dummy that equals 1 if the firm has a negative EBITDA for each of 2017, 2018, and 2019, and 0 otherwise ($ebitda$)
Capex/Assets	Capital expenditures divided by book assets ($capx/at$)
R&D/Assets	R&D expenditures divided by book assets (rnd/at)
Cash/Assets	Cash and cash equivalents divided by book assets (che/at)
Quick Ratio	Current assets minus inventory all divided by current liabilities ($(act - invt) / lct$)
ETR Pre-2018	Taxes paid from 2010-2017 divided by pre-tax income from 2010-2017 ($\sum_t txpd / \sum_t pi$)
ETR Post 2017	Taxes paid from 2018-2019 divided by pre-tax income from 2018-2019 ($txpd/pi$)
OpenSecrets.org	
Lobbying Amount	Total lobbying expenditures from 2010-2018

This table lists the variable definitions by source. Unless otherwise stated, variables are measured as of 2019. The variable names from the relevant databases are in parenthesis. All variables are winsorized at the 1% level except for the bailout variables. Tax rate variables are further winsorized to ensure that they are between 0 and 100%.

4 Results

4.1 Summary Statistics

Bailout Data

We have information on 755 publicly listed firms that received a total of \$17.9 billion in bailouts (Table 3, Panel A). This consists of 13 airline bailouts worth a total of \$16.5 billion and 742 PPP bailouts worth a total of \$1.4 billion. The mean and median airline bailout are \$1.3 billion and \$336.62 million, respectively. The mean and median PPP bailout are \$1.86 million and \$0.80 million, respectively.

The airline bailouts are, on a per-employee basis, much more generous than the PPP bailouts. The mean and median airline bailout per employee are \$34.39 thousand and \$31.76 thousand, respectively. The corresponding numbers for PPP bailouts are \$21.70 thousand and \$16.16 thousand.

Bailouts and Taxes Paid

To put the size of the bailouts in perspective, we compute the number of years a publicly listed bailout recipient has to pay corporate income tax to generate as much tax revenue as it received in bailouts (see Table 3, Panel B). The mean and median number of years for a PPP recipient are 135.0 and 22.1. The corresponding numbers for the airline bailouts are 267.9 and 138.3 years. The numbers are biased downward since companies that received a tax refund are excluded from the calculations. Because the airline bailouts are so large, both in absolute terms and relative to the size of the firm, we report the pre-tax income, taxes paid, effective tax rates, bailout amount, and years to repay for all publicly listed airlines (see Table 3, Panel C). For example, American Airlines has a pre-tax income of \$2.1 billion, \$13.00 million in taxes paid, \$4.1 billion in bailouts, and 315.4 years to repay.

The high number of years is driven by low effective tax rates and the enormous size of the bailouts. The median tax rate is 4% for bailout firms and 16% for no-bailout firms (see Table 4), while the current statutory corporate income tax rate is 21%. Firms with low long-term

Table 3: Summary Statistics - Bailout Data

Panel A: Bailout Summary Statistics

	Total	Mean	Median	N
PPP Loans	1,383.62	1.86	0.80	742
Per Employee		21.70	16.16	545
Airline Bailouts	16,489.47	1,268.42	336.62	13
Per Employee		34.39	31.76	13
All Bailouts	17,873.09	23.67	0.81	755
Per Employee		21.99	16.40	558

Panel B: Summary Years to Repay

	Mean	Median	N
PPP Loans	135.0	22.1	262
Airline Bailouts	267.9	138.3	8

Panel C: Airline Bailout Recipients

	Pre-Tax Inc.	Taxes	ETR	Payouts	Bailout	Years to Repay
Delta	5,674.5	-95.00	-1.67	13,569.0	3,835.4	End of Time
United	3,286.0	24.00	0.73	8,547.0	3,500.9	145.9
Southwest	3,060.5	553.00	18.07	9,784.0	2,311.4	4.2
American	2,070.0	13.00	0.63	12,963.0	4,100.2	315.4
Alaska	800.5	15.50	1.94	1,615.0	720.0	46.5
JetBlue	493.5	-20.50	-4.15	1,414.0	685.0	End of Time
SkyWest	406.3	2.58	0.63	256.3	336.6	130.7
Spirit	320.7	-33.82	-10.55	267.7	264.3	End of Time
Hawaiian	303.1	20.94	6.91	374.1	234.7	11.2
Allegiant	250.3	-21.88	-8.74	529.4	150.3	End of Time
Air Wisconsin	78.6	0.04	0.05	.	51.0	1,259.3
Mesa	39.6	0.40	1.02	.	92.5	230.0
Atlas	-81.7	-0.51	.	67.4	207.0	End of Time

Panel A summarizes the bailouts by type. Panel B summarizes the number of years it would take bailout firms to pay enough corporate income taxes to cover the bailout amount. As such, *Years to Repay* is calculated as *Bailout Amount* divided by *Taxes* (taxes paid, Compustat variable *txpd*). To be included in Panel B, *Taxes* must be positive. Panel C provides a breakdown of the relevant variables for the airline bailout recipients. *Pre-Tax Income* (Compustat variable *pi*) and *Taxes* represent averages from 2018-2019. Since taxes paid is missing for Delta in Compustat, we replace it with income tax expense (Compustat variable *txt*) from the income statement minus deferred taxes (Compustat variable *txdc*) from the statement of cash flows. *ETR* is the ratio of *Taxes* and *Pre-Tax Income*. *Payouts* is defined in Table 2. Dollar figures are in millions of USD except for the per-employee figures, which are in thousands. None of the variables in this table are winsorized.

effective tax rates are aggressive tax planners or tax avoiders (Dyreng et al., 2008). Given the substantial size of the bailouts and the significant US fiscal deficits in 2020 and 2021, the stylized facts on the aforementioned effective tax rates and “years to repay” might be relevant for the ongoing policy debate on US corporate taxation.

We use the global tax residence database from Meier and Smith (2020) to investigate the tax residence of bailout firms. 736 out of 755 bailout firms with tax residence data reside in the US, followed by 12 Canadian firms. There are a number of firms that reside in tax havens: 1 in Bermuda, 1 in the Cayman Islands, and 2 in Ireland.

Industry Distribution of Bailout Firms

The industry distribution of the bailout recipients differs substantially from that of the 2008/9 bailouts (see Table C1). Back then, with the exception of General Motors and Chrysler, the bailout recipients were banks and other financial institutions. This time, financial institutions are almost absent in the list of bailout recipients. Among the current bailout recipients, pharmaceutical products and medical equipment comprise about 25.8% of the bailout recipients, which is about twice their share among publicly listed firms. Computer software is also overrepresented (10.2% among the bailout recipients compared to 7.3% overall), when one would have expected that these firms are less affected by COVID-19.

Summary Statistics of Full Sample

Next, we compare bailout and non-bailout firms (Table 4). Since most of the bailout firms received PPP bailouts, the mean market capitalization, book assets, sales, and employees are all significantly smaller than those of non-bailout firms. In addition, bailout firms tend to have a higher ratio of R&D to book assets than non-bailout firms. This could be due to the fact that bailout firms are disproportionately from the computer software or pharmaceutical industries. Bailout firms also have lower EBITDA/assets than non-bailout firms. In contrast, bailout firms have larger cash/assets than non-bailout firms.

Table 4: Summary Statistics - Full Sample

	Bailout			No Bailout			Difference
	Mean	Median	N	Mean	Median	N	Mean
Bailout Amount	31	1	579	0	0	6,529	31***
Bailout/Emp	22	16	557	0	0	5,304	22***
Market Cap	286	29	564	6,432	539	6,182	-6,146***
Book Assets	537	30	579	11,696	827	6,529	-11,158***
Sales	389	19	578	3,732	294	6,488	-3,343***
Employees	1,090	73	559	11,068	1,039	5,501	-9,978***
Firm Age	17	13	577	19	14	6,521	-2**
Crisis Return	-35	-40	566	-38	-39	6,199	3***
Payouts	103	-3	398	1,043	11	4,856	-939***
Payout Ratio	68	21	111	52	51	3,175	16
Tobin's Q	615	155	564	557	135	6,159	58
Book Leverage	165	58	579	108	59	6,504	58***
Market Leverage	40	37	564	41	37	6,173	-1
Sales Growth	29	2	513	18	5	5,517	11***
EBITDA/Assets	-66	-13	573	-28	5	6,256	-37***
Capex/Assets	3	1	577	4	2	6,423	-1***
R&D/Assets	34	13	378	15	4	2,813	18***
Cash/Assets	25	14	579	21	8	6,513	4***
Quick Ratio	203	104	556	301	125	5,240	-97***
ETR Pre-2018	32	23	185	26	23	3,864	6***
ETR Post 2017	16	4	153	20	16	3,771	-4**
Lobbying Amount	271	0	579	1,146	0	6,529	-875***

This table provides summary statistics for firms based on whether they received a bailout. All variables are defined in Table 2. Accounting dollar figures and *Bailout Amount* are in millions. *Bailout/Emp* and *Lobbying Amount* are in thousands. Ratios are multiplied by 100 for presentation purposes. ***, ** and * denote 1%, 5% and 10% significance levels.

Summary Statistics of High Payout Firms

We also provide summary statistics on bailouts for subsets of firms. First, some bailout recipients made risky financial decisions. In particular, 66 firms paid out more in dividends and net repurchases from 2015-2019 than they received in bailouts (see Table 5). These high-payout firms are somewhat different than the typical bailout recipient; they are older, larger, profitable, and have low levels of R&D. For the high-payout group, the median aggregate

payouts from 2015-2019 is \$18 million and the median bailout is \$2 million, which implies that they could have easily produced the amount of the bailouts internally by withholding payouts. The mean and median payout ratio for the high payout firms are 142 and 81. Thus, instead of suffering for potentially reckless past behavior, these firms are helped at the taxpayer's expense, thereby potentially inducing future moral hazard problems.

Table 5: Summary Statistics - High Payout, High Cash, and Low EBITDA

	High Payout			High Cash			Low EBITDA		
	Mean	Median	N	Mean	Median	N	Mean	Median	N
Bailout Amount	242	2	66	21	1	437	1	1	279
Bailout/Emp	15	15	65	20	17	426	24	19	264
Market Cap	2,021	64	62	246	34	427	57	25	275
Book Assets	3,929	98	66	392	33	437	39	15	279
Sales	2,883	76	66	274	19	436	23	5	279
Employees	7,246	197	65	782	67	428	99	38	265
Firm Age	28	24	66	17	14	436	13	9	279
Crisis Return	-41	-41	62	-37	-41	428	-31	-39	277
Payouts	767	18	66	51	-4	301	-43	-21	175
Payout Ratio	142	81	45	80	25	83	290	111	3
Tobin's Q	143	119	62	451	155	427	974	248	275
Book Leverage	46	42	66	110	52	437	242	68	279
Market Leverage	42	40	62	36	33	427	36	31	275
Sales Growth	5	1	66	31	2	382	48	0	228
EBITDA/Assets	4	5	65	-54	-13	431	-120	-55	279
Capex/Assets	3	2	66	3	1	435	2	1	278
R&D/Assets	6	3	26	32	14	308	49	24	224
Cash/Assets	22	16	66	31	23	437	33	25	279
Quick Ratio	318	165	60	254	139	415	225	94	276
ETR Pre-2018	27	25	47	35	25	136	37	15	16
ETR Post 2017	20	12	41	18	5	111	2	0	8
Lobbying Amount	1,964	0	66	196	0	437	46	0	279

This table provides summary statistics for three groups of bailout firms. *High Payout* includes firms where $Payouts > Bailout\ Amount$, *High Cash* includes firms with 2019 cash and cash equivalents greater than the bailout amount, and *Low EBITDA* includes firms with a negative EBITDA in each year from 2017-2019. All variables are defined in Table 2. Accounting dollar figures and *Bailout Amount* are in millions. *Bailout/Emp* and *Lobbying Amount* are in thousands. Ratios are multiplied by 100 for presentation purposes.

Covid Economics 83, 2 July 2021: 54-87

Summary Statistics of High Cash Firms

Second, we identify 437 “high cash” firms that had more cash and cash equivalents at the end of 2019 than they received in bailouts (see Table 5). Since the majority of bailout firms are high-cash firms, their characteristics are similar to the overall sample of bailout firms; they are small, liquid, and unprofitable on average. For the high-cash group, the median cash holdings are \$23 million, while the median bailout is \$1 million, implying that the bailouts are small compared to the cash holdings of these firms. As a result, the bailouts are likely a windfall for the high-cash firms that may have been capable of producing these funds internally.

Summary Statistics of Firms with Persistent Negative Cash Flow

Third, we investigate firms with a persistent negative cash flow (Table 5), which we define as a negative EBITDA in each of 2017, 2018, and 2019. 279 such firms received a bailout, with the mean and median EBITDA/Assets being -120% and -55%. The mean and median Tobin’s Q of these firms are 9.74 and 2.48. The average book leverage of these firms is a very high 242%, while the average market leverage is only 36%. The mean and median R&D over assets of these bailout firms are 49% and 24%, respectively, while the mean and median cash over assets are 33% and 25%. 84 of the 279 “windfall” bailout firms with a negative EBITDA for three consecutive years are pharmaceutical and biotechnology firms, which are likely in the initial stages of developing new healthcare innovations—validating our proxy. In unreported results, we find that the second largest lender by loan value to publicly listed PPP recipients is Silicon Valley Bank, a bank based in Silicon Valley that focuses on start-ups. Silicon Valley Bank is the PPP lender to 32 publicly listed firms with non-missing EBITDA from 2017-2019, of which 26 are classified as having a persistent negative EBITDA. In addition, the firm age of negative EBITDA firms is clearly lower than for the average and median no-bailout firm in the sample. Overall, these numbers suggest that a large fraction of the firms that received bailouts appear to be start-up like firms that would have been unprofitable in 2020 absent COVID-19, so these bailouts might be a windfall for these firms.

Summary Statistics of Large Firms

Fourth, many of the recipients of the bailouts are quite large. Hadlock and Pierce (2010) document that firm size is a powerful proxy for financial constraints. This suggests that many of these firms might have been able to raise additional financing on the capital markets without a bailout, but at the cost of diluting or potentially even wiping out their existing shareholders and creditors. In particular, 104 firms with a market cap of at least \$100 million at the end of 2019 received a median bailout of \$4 million (see Table C2). The average Tobin's Q, market leverage, and cash/assets of these firms are 6.40, 31%, and 22%, respectively.

Summary Statistics by 500 Employee Cutoff

Fifth, we split the sample into those above and below the 500 employee cutoff in Table C3 since this cutoff was one focus of the policy debate on the PPP bailouts (see Section 2 for a discussion of eligibility rules). We only include bailouts from the PPP for this analysis. There are 52 bailout firms with more than 500 employees and 495 with at most 500 employees at the end of 2019. The average market capitalization, sales, and number of employees for a recipient of the PPP funds with more than 500 employees are \$99 million, \$273 million, and 1,683. Bailout firms with more than 500 employees have high market leverage (the median is 61%) and low liquidity (median cash/assets 6% and Quick Ratio 0.89). Bailing out these firms could imply that firms that deliberately took high risk before the crisis are now saved by the taxpayer.

4.2 Regression Analysis

Next, we employ a regression analysis of the bailout data. To make the sample more homogeneous, we restrict the sample to publicly listed firms with at most 500 employees that did not receive an airline bailout. We analyze the determinants of the bailout probability in models 1-3 using OLS, and the bailout magnitude in models 4-6 using Tobit. These are cross-sectional regressions at the firm level. Industry fixed effects using three digit SIC codes are included in all models.

Table 6: Bailout Determinants

	Dep Var: Bailout Dummy			Dep Var: ln(Bailout Amount)		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(Book Assets)	-4.34*** (0.49)	-9.09*** (0.90)	-9.20*** (0.92)	-0.13*** (0.02)	-0.33*** (0.04)	-0.34*** (0.04)
Firm Age	0.31*** (0.07)	0.27*** (0.08)	0.33*** (0.08)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)
1(Total Debt > 0)	7.27*** (2.26)	5.36 (3.31)	3.92 (3.43)	0.31*** (0.11)	0.16 (0.13)	0.11 (0.13)
Book Leverage	-0.49 (0.36)	0.32 (0.64)	0.24 (0.52)	-0.03** (0.01)	0.03* (0.02)	0.04** (0.02)
Cash/Assets	-23.21*** (3.45)	-18.02*** (4.78)	-18.74*** (4.88)	-0.97*** (0.13)	-0.60*** (0.17)	-0.60*** (0.17)
Capex/Assets	-14.05 (14.29)	-18.52 (16.91)	-16.35 (17.58)	-0.31 (0.71)	-0.19 (0.84)	-0.04 (0.86)
EBITDA/Assets	1.69* (0.92)	0.35 (1.67)		0.04 (0.03)	-0.05 (0.05)	
1(3-Yr EBITDA < 0)			8.18*** (2.97)			0.27*** (0.10)
Tobin's Q		-0.40*** (0.05)	-0.43*** (0.05)		-0.04*** (0.01)	-0.04*** (0.01)
ln(Sales)		3.05*** (0.69)	3.86*** (0.74)		0.16*** (0.03)	0.18*** (0.03)
Crisis Return		-11.14** (4.34)	-9.46** (4.40)		-0.59*** (0.17)	-0.55*** (0.17)
1(Lobbying Amount > 0)	6.92*** (2.65)	6.98** (2.92)	6.55** (2.97)	0.34*** (0.10)	0.29*** (0.11)	0.29** (0.11)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Estimation Method	OLS	OLS	OLS	Tobit	Tobit	Tobit
Bailout Prob.	18.73	21.07	21.04			
N	2,493	1,922	1,868	2,544	1,972	1,916
Adjusted R ²	0.212	0.250	0.259			
Pseudo R ²				0.230	0.262	0.268

This table analyzes the determinants of the incidence (using OLS) and magnitude (using Tobit) of the bailouts using cross-sectional regressions at the firm-level. In the first three specifications, the dependent variable is *Bailout Dummy* (100 or 0). In models 4-6, the dependent variable is *ln(Bailout Amount)*. The sample includes firms with at most 500 employees that did not receive an airline bailout. All variables are defined in Table 2. *Industry FE* refers to three digit SIC code fixed effects. Heteroskedasticity robust standard errors are reported in parentheses. ***, ** and * denote 1%, 5% and 10% significance levels.

Firms with lower levels of assets and higher sales are more likely to be bailed out and tend to receive a greater bailout amount. Firm age has a positive effect on the bailout likelihood and amount, a surprising result given that firm age is regarded as a proxy for financial constraints (Hadlock and Pierce, 2010). Reassuringly, crisis return⁵ has a negative and significant effect on the likelihood of receiving a bailout, implying that firms more affected by the crisis are more likely to receive a bailout. In addition, cash over assets is negatively associated with the bailout likelihood and amount. The negative and significant estimate for Tobin's Q suggests that firms with a higher Tobin's Q might be better able to raise financing without government support. Past lobbying has a positive and significant effect on the incidence and magnitude of the bailout. Firms that lobby might be experienced in navigating bureaucracy and red tape, and might therefore be in a better position to disentangle the frequently changing, opaque, and contradictory bailout rules.

The dummy for persistent negative EBITDA is positive and significant. For instance, in column 3 of Table 6, the results indicate that having a negative EBITDA in 2017, 2018, and 2019 increases the probability of receiving a bailout by 8.18 percentage points. Since the unconditional bailout likelihood is 21.04%, this is a quantitatively large effect. This supports our prior interpretation that a large fraction of the firms that received bailouts appear to be start-up firms that would have been unprofitable in 2020 absent COVID-19, so these bailouts might be a windfall for these firms.

4.3 Discussion of Results

Interpretation of Results

One limitation of our paper is that we cannot establish causality, so the results should be interpreted as suggestive evidence.

Generalizability of Results to Privately Held Bailout Recipients

In an ideal setting, one would conduct a firm-level analysis of the bailouts. This is not

⁵The crisis return is the return of a stock, including dividends, from February 19, 2020, when the S&P 500 hit an all-time high, to March 23, 2020, when the Federal Reserve announced emergency measures.

possible, however, since firm-level data is not available for privately held firms. Therefore, we focus on publicly listed US firms and combine this with detailed hand-collected bailout data from corporate filings with the Securities and Exchange Commission SEC. This raises the question: to what extent can our results be extrapolated to privately held firms? The average privately held bailout recipient will most likely be more financially constrained than the average publicly listed one. Therefore, the likelihood that a bailout of a privately held firm involves any of the documented “dark sides” of bailouts will be lower than for a publicly listed firm. However, because the vast majority of PPP bailout funds went to privately held firms, it is reasonable to assume that the “dark side” of the bailouts for privately held firms will be quantitatively large due to the sheer size of the private sector and the bailouts. There will be many large privately held firms that will not be financially constrained and do not need a bailout. Bailouts are also not conditioned on whether a firm was affected by the COVID-19 crisis, implying that bailouts have been given to firms that were not affected by the crisis or may even have benefited from it (e.g., software). In addition, the bailouts do not condition on whether a firm can survive the crisis without a bailout. Moreover, the risk for abuse or outright fraud with bailout funds for privately held firms is larger than for publicly listed firms due to the lower transparency and accountability to outside stakeholders.⁶

5 Policy Implications

One implication from the preceding subsection is that the payout of bailouts should have been conditioned on whether a firm has been affected by the COVID-19 crisis, by, for instance, comparing 2020 to 2019 revenue. This design flaw is one of the drivers of the “dark side” of the bailouts such as the windfalls that we have documented.

Since the main aim of the PPP is to protect employment, it is unclear why policymakers use firms as an intermediary in achieving this goal. A better approach could be one similar

⁶New York Times (2020): “Spotting \$62 Million in Alleged P.P.P. Fraud Was the Easy Part”, August 28, <https://www.nytimes.com/2020/08/28/business/ppp-small-business-fraud-coronavirus.html?action=click&module=Top%20Stories&pgtype=Homepage> (accessed September 13, 2020).

to the German “Kurzarbeitergeld,” which is a wage subsidy scheme that allows companies to lower their operating costs by immediately reducing their payroll while maintaining employment (Möller, 2010). Importantly, this approach avoids using firms as intermediaries, and instead directly pays the employees.

The goal of the airline bailouts was to allow the industry to maintain employment at existing salary levels. This was generous, since the bailouts per employee for the airlines are more than 50% larger than for the PPP bailouts (Table 3), and since airline employees are better paid (median hourly wage of \$30.04) than the majority of the labor force (median hourly wage of \$19.14).⁷ Moreover, airlines paid out more to shareholders in the last five years than they received in bailouts, suggesting that these bailouts might be inducing moral hazard by rewarding aggressive financial strategies. The airline bailouts are also expensive when compared to recent corporate income tax payments by the airlines. Delta, for instance, had a pre-tax income in 2019 of \$5.7 billion, on which it received a tax refund of \$95.00 million. The bailout Delta received was \$3.8 billion, while its aggregate payouts to shareholders from 2015-2019 were \$13.6 billion (see Panel C of Table 3 for data on all publicly listed US airlines). Thus, without the bailouts, many airline employees might have lost their jobs. In addition, few industries have undergone as many bankruptcies as the airlines—Delta, American, Southwest, and United (or their predecessors) went bankrupt 4.25 times on average since since the 1980s (see Table C4). In line with Morrison and Saavedra (2020), we argue that chapter 11 would have been an effective tool instead of bailouts to restructure the airlines.

Morrison and Saavedra (2020) argue that policymakers have minimized the role of bankruptcy law in mitigating the financial fallout from COVID-19. They suggest that Chapter 11 bankruptcy is an effective tool for dealing with the financial distress of large corporations during the COVID-19 crisis that should be used more often. Therefore, it seems plausible

⁷Bureau of Labour Statistics (2019): “National Occupational Employment and Wage Estimates United States”, https://www.bls.gov/oes/current/oes_nat.htm#00-0000; https://www.bls.gov/oes/current/naics3_481000.htm#00-0000 (accessed September 13, 2020).

that, at least at the margin, corporate bankruptcies would have been an effective alternative to bailouts for airlines and other large firms, such as the bailout recipients with more than \$100 million in market capitalization (see Table C2).

One difference between the current crisis and that of 2008/9 is that there seems to be a lack of fire sales of struggling companies or investments into such companies at fire-sale prices. Warren Buffett's Berkshire Hathaway, for instance, invested \$5 billion in Goldman Sachs in September 2008 and \$3 billion in General Electric in October 2008, while, Warren Buffett's firm has not undertaken any major investments during the COVID-19 crisis.⁸ A key reason for the lack of fire sale acquisitions or investments at fire-sale prices is most likely that the bailouts have been so large in size that there are few profitable investment opportunities for private investors to purchase assets cheaply.⁹ As a consequence, existing shareholders and creditors have been bailed out by taxpayers. Due to the severity of the current crisis, some bailouts likely would have been necessary to prevent the free-fall of the economy, but considering the evidence in Meier and Servaes (2019), private investors could have stepped in and taken care of some struggling companies in fire sales, with limited welfare implications for the rest of the economy.

6 Conclusion

We use hand-collected data to investigate the COVID-19 bailouts for all publicly listed US firms. We document a dark side of the bailouts. For many firms, the bailouts appear to be a windfall. Numerous bailout recipients made risky financial decisions, so bailing them out might induce moral hazard. The bailouts are expensive when compared to past corporate income tax payments of the bailout firms. We compute the number of years a bailout recipient has to pay corporate income tax to generate as much tax revenue as it received in

⁸Financial Times (2020): "Famed Investor Tells Virtual Annual Meeting Berkshire Hathaway Can Find Nothing to Buy", May 2, <https://www.ft.com/content/4b707086-4b48-48ab-9369-a3fcf5dd1af3> (accessed September 13, 2020).

⁹Wall Street Journal (2020): "How Fed Intervention Saved Carnival", April 26, <https://www.wsj.com/articles/how-fed-intervention-saved-carnival-11587920400> (accessed April 28, 2020).

bailouts: 135.0 years for the Paycheck Protection Program and 267.9 years for the airline bailouts.

References

- Autor, D., D. Cho, L. D. Crane, M. Goldar, B. Lutz, J. Montes, W. B. Peterman, D. Ratner, D. Villar, and A. Yildirmaz (2020). An Evaluation of the Paycheck Protection Program Using Administrative Payroll Microdata. Working Paper. Massachusetts Institute of Technology, Board of Governors of the Federal Reserve System, and ADP.
- Bartik, A. W., Z. B. Cullen, E. L. Glaeser, M. Luca, C. T. Stanton, and A. Sunderam (2020). The Targeting and Impact of Paycheck Protection Program Loans to Small Businesses. Working Paper. University of Illinois at Urbana-Champaign and Harvard University.
- Calomiris, C. W. and U. Khan (2015). An Assessment of TARP Assistance to Financial Institutions. *Journal of Economic Perspectives* 29(2), 53–80.
- Center for Responsive Politics (2018). Bulk Data Download. <https://www.opensecrets.org/bulk-data/downloads> (accessed April 7, 2020).
- Chavaz, M. and A. K. Rose (2019). Political Borders and Bank Lending in Post-crisis America. *Review of Finance* 23(5), 935–959.
- Chetty, R., J. N. Friedman, N. Hendren, M. Stepner, and the Opportunity Insights Team (2020). The Economic Impacts of COVID-19: Evidence from a New Public Database Built from Private Sector Data. Working Paper. Harvard University and Brown University.
- Duchin, R. and D. Sosyura (2012). The Politics of Government Investment. *Journal of Financial Economics* 106(1), 24–48.
- Duchin, R. and D. Sosyura (2014). Safer Ratios, Riskier Portfolios: Banks' Response to Government Aid. *Journal of Financial Economics* 113(1), 1–28.
- Dyreng, S. D., M. Hanlon, and E. L. Maydew (2008). Long-Run Corporate Tax Avoidance. *The Accounting Review* 83(1), 61–82.
- Elenev, V., T. Landvoigt, and S. Van Nieuwerburgh (2020). Can the Covid Bailouts Save the Economy? Working Paper. Johns Hopkins University, University of Pennsylvania, NBER, CEPR, and Columbia University.
- Faccio, M., R. W. Masulis, and J. J. McConnell (2006). Political Connections and Corporate Bailouts. *Journal of Finance* 61(6), 2597–2635.
- Farhi, E. and J. Tirole (2012). Collective Moral Hazard, Maturity Mismatch, and Systemic Bailouts. *American Economic Review* 102(1), 60–93.
- Goolsbee, A. D. and A. B. Krueger (2015). A Retrospective Look at Rescuing and Restructuring General Motors and Chrysler. *Journal of Economic Perspectives* 29(2), 3–24.
- Granja, J., C. Makridis, C. Yannelis, and E. Zwick (2020). Did the Paycheck Protection Program Hit the Target? Working Paper. University of Chicago, Massachusetts Institute of Technology, and NBER.
- Hadlock, C. J. and J. R. Pierce (2010). New Evidence on Measuring Financial Constraints: Moving Beyond the KZ Index. *Review of Financial Studies* 23(5), 1909–1940.
- Meier, J.-M. and H. Servaes (2019). The Bright Side of Fire Sales. *Review of Financial Studies* 32(11), 4228–4270.
- Meier, J.-M. and J. Smith (2020). Tax Avoidance through Cross-Border Mergers and Ac-

- quisitions. Working Paper. University of Texas at Dallas.
- Möller, J. (2010). The German Labor Market Response in the World Recession—De-Mystifying a Miracle. *Zeitschrift für Arbeitsmarktforschung* 42(4), 325–336.
- Morrison, E. R. and A. C. Saavedra (2020). Bankruptcy’s Role in the COVID-19 Crisis. *Working paper. Columbia University.*

For Online Publication

Appendix to

“The COVID-19 Bailouts”

A Appendix: Further Institutional Details

A.1 *Legislative History of Bailout Programs*

On March 25, 2020, the Senate passed the “Coronavirus Aid, Relief, and Economic Security Act” or the “CARES Act.” The House agreed to the Senate amendment on March 27, 2020 and it was signed into law by the President on March 27, 2020. The overall volume of the CARES Act is approximately \$2.1 trillion. The second wave of bailouts for the private sector from the federal government are included in the \$484 billion “Paycheck Protection Program and Health Care Enhancement Act”. It passed the Senate on April 21, 2020 and the House of Representatives on April 23, 2020, before it was signed into law by the President on April 24, 2020. The Paycheck Protection Program Flexibility Act (PPFPA) was passed by the House of Representatives on May 27, 2020 and the Senate approved it by unanimous consent on June 3, 2020. The president signed it into law on June 5, 2020. The PPFPA loosened many of the rules with regards to the PPP bailouts. On June 30, 2020, the PPP application deadline was extended from June 30, 2020 to August 8, 2020.

A.2 *Eligibility for the Paycheck Protection Program*

A firm must meet at least one of the criteria to be eligible for a PPP bailout. First, as a rule of thumb, most firms with at most 500 employees are eligible for PPP funds.

Second, there are exceptions for all firms whose North American Industry Classification System (NAICS) code starts with 72, which includes hotels and restaurants. The FAQ discusses eligibility criteria for firms with separate locations and separate legal entities (even if these separate legal entities are affiliated with the same parent entity, including 100% ownership). If each location of a business with a NAICS code starting with 72 has at most 500 employees, such a business is also eligible for PPP funds. In addition, a NAICS 72-code-business is eligible for PPP funds if each separate legal entity (even if affiliated through 100% ownership) has at most 500 employees. NAICS 72-code-businesses are also eligible for PPP funds even if they have more than 500 employees in any particular location, as long

as these employees are employed by separate legal entities (even if affiliated through 100% ownership links) that each have at most 500 employees across all the locations in which a particular legal entity operates.

Third, the SBA has size cut-offs to determine whether a firm is eligible for SBA funding. The size cut-offs differ across NAICS codes and are either provided in the dollar amount of “annual receipts,” the number of employees, or, in the case of financial institutions, in millions of assets. The largest values for “annual receipts,” employees, and assets are \$41.5 million (63 NAICS codes), 1,500 employees (44 NAICS codes) and \$600 million of assets. We use the size cut-offs that the SBA provides on www.ecfr.gov/cgi-bin/text-idx?SID=b919ec8f32159d9edaaa36a7eaf6b695&mc=true&node=pt13.1.121&rgn (accessed June 4, 2020).

A.3 COVID-19 Emergency Programs by the Federal Reserve

Table A1: COVID-19 Emergency Programs by the Federal Reserve

Program	Size	Eligible Borrowers/Beneficiaries	Collateral/Assets	Bailout
Repurchase Operations.	1,500	24 broker dealers in US government debt.	treasuries, agencies	
Swap Lines.		Bank of England, Bank of Japan, European Central Bank, Bank of Canada, Swiss National Bank.	foreign currency	
Commercial Paper Funding Facility.		US issuers of commercial paper rated at least A-1/P-1/F-1 by major nationally recognized statistical rating organization.	commercial paper	
Primary Dealer Credit Facility.		24 broker dealers in US government securities.	treasuries, agencies, corporate bonds, equities	
MMF Liquidity Facility.		Depositories, bank holdings companies, US branches and agencies of foreign banks lending to prime money market mutual funds.	treasuries, agencies, commercial paper	
Swap Lines Extension.		Central banks of Australia, Brazil, Denmark, South Korea, Mexico, Norway, New Zealand, Singapore, and Sweden.	foreign currency	
Term Asset-Backed Securities Loan Facility.	100	Companies with eligible collateral and account relationships with one of 24 primary broker dealers.	asset-backed securities	
Primary Market Corporate Credit Facility.	500	Investment grade US companies headquartered in US with material US operations.	corporate bonds, business loans	Yes
Secondary Market Corporate Credit Facility.	250	Investment grade US companies headquartered in the US with material US operations.	corporate bonds, ETFs	Indirect bailout
Foreign and International Monetary Authorities.		Foreign central banks and monetary authorities with accounts at the New York Fed.	treasuries	

Table A1: COVID-19 Emergency Programs by the Federal Reserve (Continued)

Program	Size	Eligible Borrowers/Beneficiaries	Collateral/Assets	Bailout
Municipal Liquidity Facility.	500	States (+ DC), counties with 500,000+ residents, cities with 250,000+ residents; direct borrowing from Fed.	muni bonds	Yes: States, counties, cities
Main Street New Loan Facility, Main Street Priority Loan Facility, Main Street Expanded Loan Facility.	600	Businesses with 15,000 employees or up to \$5B sales; Fed will buy 95% of loans from lenders who retain 5%.	business loans	Yes

This table lists major emergency programs by the Federal Reserve in response to the economic crisis caused by COVID-19. Size is the amount of the program in billions dollars. Bailout indicates a private sector bailout. The table is based on our modification, update, and extension of an article on the Columbia Law School's Blog on Corporations and the Capital Markets by Lev Menand (2020): "Fed to the Rescue: Unprecedented Scope, Stretched Authority," April 27, <https://clsbluesky.law.columbia.edu/2020/04/27/fed-to-the-rescue-unprecedented-scope-stretched-authority/> (accessed May 5, 2020).

B Appendix: Additional Data Adjustments

B.1 *Harbor Diversified, Inc*

Harbor Diversified, Inc, a holding company whose main operating subsidiary is Air Wisconsin, received both a PPP bailout of \$10 million and an airline bailout of \$41 million. This is the only company that we have found that has received bailouts from both programs. Since Harbor Diversified is an airline holding company, and since the airline bailout is larger than the PPP bailout, we simplify the analysis by considering the bailouts as a single \$51 million airline bailout. When listed among the other airlines in Table 3, Panel C, we list the name of the operating subsidiary, Air Wisconsin, for the sake of simplicity and clarity. In addition, since 2018-2019 data is not available for this firm in Compustat, we manually fill in pre-tax income, taxes paid, and the number of employees using their 2019 10-K.

C Appendix: Additional Tables

Table C1: Industry Breakdown

	Overall	Bailout	No Bailout	Ratio
Pharmaceutical Products	11.5	18.0	10.7	1.6
Banking	9.6	0.9	10.5	0.1
Computer Software	7.3	10.2	7.1	1.4
Trading	6.8	3.5	7.1	0.5
Industrial Mining	5.3	1.7	5.7	0.3
Petroleum and Natural Gas	5.0	4.3	5.0	0.9
Precious Metals	3.8	0.3	4.1	0.1
Utilities	3.5	0.3	3.9	0.1
Electronic Equipment	3.4	6.4	3.1	1.9
Business Services	3.2	5.0	3.0	1.6
Retail	2.9	1.2	3.1	0.4
Medical Equipment	2.8	7.8	2.2	2.8
Transportation	2.6	3.1	2.6	1.2
Insurance	2.2	0.3	2.4	0.2
Wholesale	2.2	2.9	2.1	1.3
Communication	2.1	1.6	2.2	0.7
Machinery	2.1	3.3	2.0	1.6
Chemicals	1.4	2.6	1.3	1.8
Construction Materials	1.4	1.2	1.4	0.9
Real Estate	1.2	0.9	1.2	0.7
Restaurants, Hotels, Motels	1.2	2.1	1.0	1.8
Measuring and Control Equip.	1.1	2.6	0.9	2.3
Automobiles and Trucks	1.2	2.1	1.1	1.8
Food Products	1.1	0.7	1.1	0.6
Entertainment	1.0	1.6	1.0	1.5
Healthcare	1.0	1.2	1.0	1.2
Electrical Equipment	0.9	2.8	0.8	3.0
Personal Services	0.9	0.2	1.0	0.2
Construction	0.8	0.7	0.8	0.9
Computer Hardware	0.8	1.4	0.7	1.8

This table provides a percentage breakdown of the Fama French 49 industries for all firms, those that received a bailout, and those that did not receive a bailout. In addition, the *Ratio* column divides the bailout percentage by the overall percentage. All firms with non-missing book assets in Compustat are included. Industry classification is as of 2019. Industries are sorted in descending order of their overall share among Compustat firms.

Table C1: Industry Breakdown (Continued)

	Overall	Bailout	No Bailout	Ratio
Consumer Goods	0.8	1.4	0.8	1.7
Steel Works Etc	0.7	0.7	0.8	0.9
Apparel	0.6	0.2	0.6	0.3
Business Supplies	0.6	0.0	0.6	0.0
Recreation	0.5	1.2	0.4	2.6
Printing and Publishing	0.4	0.0	0.4	0.0
Rubber and Plastic Products	0.4	0.5	0.4	1.3
Aircraft	0.3	0.3	0.3	1.0
Beer & Liquor	0.3	0.3	0.3	1.2
Candy & Soda	0.3	0.7	0.2	2.6
Coal	0.3	0.7	0.2	2.8
Agriculture	0.2	0.3	0.2	1.6
Shipbuilding, Railroad Equip.	0.2	0.3	0.2	2.1
Shipping Containers	0.2	0.0	0.2	0.0
Fabricated Products	0.2	0.0	0.2	0.0
Textiles	0.1	0.3	0.1	3.1
Defense	0.1	0.0	0.1	0.0
Tobacco Products	0.1	0.0	0.1	0.0
Almost Nothing	3.8	2.1	4.0	0.6

Table C2: Summary Statistics - Market Capitalization

	Mean	Median	N
Bailout Amount	162	4	104
Bailout/Emp	22	19	102
Market Cap	1,423	217	104
Book Assets	2,653	171	104
Sales	1,933	82	104
Employees	4,924	271	102
Firm Age	18	14	103
Crisis Return	-46	-50	104
Payouts	664	-12	70
Payout Ratio	2	26	28
Tobin's Q	640	189	104
Book Leverage	85	53	104
Market Leverage	31	26	104
Sales Growth	53	6	92
EBITDA/Assets	-32	-4	102
Capex/Assets	4	2	104
R&D/Assets	23	13	63
Cash/Assets	22	12	104
Quick Ratio	230	104	100
ETR Pre-2018	29	17	44
ETR Post 2017	13	2	35
Lobbying Amount	1,392	0	104

This table provides summary statistics for firms with a market capitalization of at least \$100 million. All variables are defined in Table 2. Accounting dollar figures and *Bailout Amount* are in millions. *Bailout/Emp* and *Lobbying Amount* are in thousands. Ratios are multiplied by 100 for presentation purposes.

Table C3: Summary Statistics - Number of Employees

	Employees > 500			Employees ≤ 500		
	Mean	Median	N	Mean	Median	N
Bailout Amount	7	7	52	2	1	495
Bailout/Emp	7	7	52	23	17	493
Market Cap	99	62	52	64	27	482
Book Assets	238	193	52	86	24	495
Sales	273	198	52	41	14	494
Employees	1,683	894	52	105	57	495
Firm Age	23	24	52	16	13	493
Crisis Return	-43	-46	52	-34	-39	483
Payouts	-10	0	42	-23	-4	335
Payout Ratio	9	12	17	86	14	82
Tobin's Q	121	105	52	642	170	482
Book Leverage	68	59	52	168	55	495
Market Leverage	63	61	52	37	33	482
Sales Growth	13	1	51	31	2	433
EBITDA/Assets	2	5	52	-70	-17	489
Capex/Assets	4	3	52	3	1	493
R&D/Assets	3	0	23	35	14	344
Cash/Assets	8	6	52	27	18	495
Quick Ratio	97	89	52	225	116	474
ETR Pre-2018	41	36	28	33	23	144
ETR Post 2017	33	18	21	14	3	119
Lobbying Amount	23	0	52	41	0	495

This table provides summary statistics for two groups of bailout firms based on the number of employees. All variables are defined in Table 2. Accounting dollar figures and *Bailout Amount* are in millions. *Bailout/Emp* and *Lobbying Amount* are in thousands. Ratios are multiplied by 100 for presentation purposes.

Table C4: Bankruptcies by Publicly Listed US Airlines Since 1980

Airline	Predecessor	Date	Remarks
Delta	Pan Am	1/8/1991	Delta acquired Pan Am in 1991
Delta	Northwest	9/14/2005	Delta acquired Northwest in 2008
Delta	Comair	9/14/2005	Delta acquired Comair in 1999
Delta	Pinnacle	4/2/2012	Emerged from bankruptcy as a subsidiary of Delta
United		12/9/2002	
United	Pan Am	1/8/1991	Pan Am sold its Latin American and Caribbean Routes to United airlines in 1991
United	Continental	9/23/1983	United acquired Continental in 2010
United	Continental	12/3/1990	United acquired Continental in 2010
American	Eastern	3/9/1989	American acquired Eastern in 1989
American	TWA	1/31/1992	American acquired TWA in 2001
American	TWA	6/30/1995	American acquired TWA in 2001
American	TWA	1/10/2001	American acquired TWA in 2001
American	US Airways	8/11/2002	US Airways and American merged in 2013
American	US Airways	9/12/2004	US Airways and American merged in 2013
American		11/29/2011	
Southwest	ATA	10/26/2004	Southwest acquired ATA in 2008
Southwest	ATA	4/2/2008	Southwest acquired ATA in 2008
Hawaiian		9/1/1993	
Hawaiian		3/1/2003	
Allegiant		12/14/2000	
Mesa		1/5/2010	
Atlas		1/30/2004	
Atlas	Southern	9/28/2012	Atlas acquired Southern in 2016

This table provides an overview of all bankruptcies since 1980 by all US airlines that were publicly listed as of the end of 2019. A predecessor is listed if it went bankrupt, otherwise, it was the airline in column 1 that went bankrupt.

Everyone is a stock trader now: Retail investors and Covid-19¹

Katrin Tinn²

Date submitted: 22 June 2021; Date accepted: 22 June 2021

There was a surge in the participation of retail investors in the stock market during the year 2020, with a large set of new investors starting to trade stocks on fintech platforms for the first time. This development could seem surprising, as the Covid-19 pandemic likely increased uncertainty and entailed negative wealth effects. In most canonical models of stock trading, at least one of these effects would imply a reduced demand for risky assets (such as stocks). This paper develops a model which incorporates both effects and maintains the assumption of weak form efficient markets. It shows that the observed surge of demand is best explained by there being investors whose trading is based on common sentiment analysis rather than fundamental analysis. Reduced opportunity costs of participation can help further. We provide arguments that both trends have increased over the past year. The paper also contributes to the REE literature by considering wealth effects and sentiment effects jointly in a stylized setting that has an analytical solution. It derives new predictions on the relationship between stock market participation and asset prices.

- 1 I thank Chaire Fintech AMF -- Finance Montreal who kindly provided a research grant for this project within their call the call for projects on "Fintechs and COVID-19" in Canada, and I thank Ananya Nair, BCom student at McGill, who provided excellent research assistantship for this paper. Her analysis of the market and user trends have been essential for this project.
- 2 Assistant Professor of Finance, Desautels Faculty of Management at McGill University and CEPR.

Copyright: Katrin Tinn

1 Introduction

Covid-19 was identified by the World Health Organization as a global pandemic risk in early 2020, and by March 2020, infection and death statistics had got out of hand worldwide, so that most countries decided to implement measures of social distancing, encouraging working and studying from home. It was then feared that many sectors (retail, restaurants and hospitality services, etc) would experience a contraction. As immediate reaction to the fear of a global supply chain shock, many stock prices and indexes fell rapidly around the same period of time; for example, NASDAQ Composite fell from about 9700 in mid February 2020 to about 6900 around late March 2020 (about 40% loss), and similar patterns were echoed in other stock indexes: e.g., Dow Jones retail Index (USA), FTSE (UK), DAX (Germany), S&P/TSX (Canada) all experienced sudden losses of value in the magnitude of 40-50% during the same interval.

At the time, it was not obvious whether this stock market movement would persist or not. On the one hand investors might have believed that the fall in stock prices was too large, and bound to be temporary - because firm values largely reflect cash flows that will also materialize over several periods of time after the Covid-19 crisis is over. On the other hand, there was an increased risk that many firms would not survive and go out of business. Both of these effects increase uncertainty about asset returns, and may have also lead to an unusually large wedge between investors' beliefs and stock prices.

Relatedly, the Covid-19 situation posed an interesting question regarding stock market participation, especially when we consider the participation of retail investors. One hypothesis was that the participation in stock market would fall because individuals faced negative wealth and liquidity shocks. Another hypothesis was that some previously not participating investors would join the stock market, because of the extra time they had to overcome the fixed cost of participation, and/or because they might expect high returns, be it for rational or behavioral reasons. The patterns of new participants are arguably most visible when we focus on FinTech platforms that bring greater convenience in establishing an investment account for online trading. Well known examples of such platforms are Robinhood in the USA and Wealthsimple in Canada.

In this paper, we will provide a brief overview of platforms that offer online stock trading in Canada, and highlight stylized facts about the dynamics of participation in these platforms as well as in their USA equivalents during the period of interest. There appears to be evidence that these platforms have onboarded large number of new clients during the period. As it would be a stretch to argue that investors have become wealthier or less risk averse, it calls for a theoretical explanation based on subjective expectations of high returns, together with a

possible greater willingness to pay for the non-monetary fixed costs needed for participation. Among the stylized facts presented, we highlight the indications of prices being driven by retail investors demanding assets for non-fundamental reasons. The most famous incidence in this spirit is the GameStop Corp share price that experienced an unprecedented price increase and volatility in early 2021, which was arguably heavily affected by discussions in social media and by Elon Musk's tweet of Jan 26, 2021, cryptically stating "Gamestonk!!". We provide other examples in this spirit.

To interpret these stylized facts in the context of the aforementioned hypotheses, we then develop an illustrative model of retail investors' stock market participation incentives and their price impact. For the sake of clarity, there are just two assets: a risky asset (e.g., a stock or a portfolio of stocks), and a risk-free asset (e.g., savings in a bank or a portfolio of risk-free bonds). In the model, there are three types of investors. First, there are sophisticated informed investors who trade based on their superior knowledge about the fundamental value of the risky asset (e.g., professional investment and hedge funds). Second, there are non-sophisticated retail investors, who first choose whether to participate in the stock market, and then trade based on their common subjective beliefs about the risky asset value. Retail investors are heterogeneous in wealth and their risk tolerance is endogenous - a higher wealth makes investors more risk tolerant/less risk averse. This implies that there is a threshold level of wealth below which investors choose not to participate. Third, there are (risk-neutral) competitive market makers who do not know the fundamental value and the beliefs of retail investors, but learn noisy information about the fundamental value from prices and order flows. Market makers implement the market efficiency condition by setting the asset price to be equal to the expected value of the fundamental based on all public information.

The model shows that in the context of the Covid-19 shock, where wealth distribution is unlikely to have had a positive mean shift, the wedge between the beliefs of retail investors about risky asset values and the price on the one hand, and the non-monetary aspects of participation costs on the other hand, may be the more plausible explanations for the surge in the demand of retail investors. The analysis of the implied equilibrium price patterns highlights further predictions. For example, the increased demand by retail investors, trading on the basis of non-fundamentals signals, monotonically reduces the importance of fundamental information as a driver of equilibrium asset prices. At the same time, it increases the weight on prior beliefs about fundamental values and generates non-monotonic price effects regarding the role of retail investors' common signals. At the limit and in our baseline model, retail investors' signals are uncorrelated with the fundamental, and the model highlights intuitive parameters under which price distortion is maximized. As the assumption that retail

investors' demand is always driven by non-informative signals may seem a bit too radical, we also provide an extension where some retail investors have fundamental information. As one would expect, all the main effects remain, but are smaller in their magnitude.

From a theoretical modelling perspective, our paper builds on methods from the large literature of REE models (see e.g., seminal papers by Grossman and Stiglitz (1980, 1988), Brown and Jennings (1989), Vives (1995), Kyle (1985) and Glosten and Milgrom (1985), and many models that build on these settings, see Brunnermeier (2001) and Vives (2008) for surveys). Our formal solution method is most similar to the one in Vives (1995), where the asset price is set to implement the (weak form) market efficiency condition (as in Kyle (1985)), while considering traders that operate in a competitive market and are risk averse. Many classical settings in the REE literature abstract from considerations of wealth effects for the benefit of analytical solutions (by adopting the CARA assumption). At the same time, there is overwhelming evidence that wealthier investors are less risk averse and more likely to participate in the stock market. To capture such a realistic wealth effect, we incorporate the modelling method in Peress (2004).¹

Our assumption that retail investors' trading incentives are primarily driven by considerations outside the superior knowledge of the fundamental value, is consistent with many models and empirical evidence in the literature focused on behavioral finance (see e.g. Barber and Odean (2000, 2001), Kahneman and Tversky (1980, 2013)). These arguments also point towards retail investors tending to have losses on average when competing against informed traders and rational market makers. There is also empirical evidence that the possibility to trade on online platforms encourages retail traders to take excessive risks and buy "lottery-like" stocks² (see e.g., Kumar (2009), Kalda (2021)) and to participate more (Choi et. al. (2002), Bogan (2008)). Our baseline modelling approach regarding retail investors' information is closest to Mendel and Shleifer (2012), where retail investors are "noise traders" with correlated beliefs.

The existing empirical and theoretical literature on Covid-19's impact on stock trading is still at its early stages. There are some notable empirical findings, which tend to be consistent with our findings. For example, van der Beck and Jaunin (2021) confirm the insight that retail investors' demand soared during the Covid-19 pandemic. By focusing on

¹In such a setting, each individual investor's utility function is still CARA. However, different investors have different risk tolerance parameters, which depend on their wealth. While there are many other utility functions (the family of homothetic preferences, including CRRA) that produce similar aggregate wealth effects in models with representative agents, this modelling choice enables an informative analytical solution. For the application of such utility functions in the context of micro founded trading models, see e.g. Kyle and Xiong (2001).

²Stocks that have low expected value compared to the price, high variance and positive skewness, based on public data.

the Robinhood Market Inc. platform, they find large price impacts driven by retail investors. The paper by Glossner et. al. (2021) documents that institutional investors experienced client outflows, and platforms like Robinhood's clients acted as liquidity providers. While there are other papers confirming such key empirical patterns, the set of theoretical models and interpretations is more scarce.

Finally, our paper relates to the broader macroeconomic literature on stock market participation. For example, Mankiw and Zeldes (1991) find that the percentage of U.S households with direct holdings of at least \$1000 in the stock market is only 23.2%, Bertaut and Starr-McCluer (2000) find that in the US less than 50% of households own some form of stock. Based on Giannetti and Koskinen (2010) Canadian investors' participation rate in the domestic stock markets is 25%, and foreign equity held by domestic investors is 30.2%. While the focus of this paper is a partial equilibrium one, it is plausible that over the long term there could be some positive effects of engaging new investors, who may become more sophisticated over time, which may bring some positive macroeconomic effects.

2 Online trading platforms and stylized facts

2.1 Players in the Canadian Market

All the big five banks that dominate the banking industry in Canada, BMO, Scotiabank, CIBC, RBC and TD, offer self-directed trading platforms. These banks offer both mobile and desktop trading while requiring varying levels of account minimum and charging varying levels of commission per equity trade. Table B.1 in the Appendix details how these platforms differ in terms of charges, app store ratings among other things. Typically, the pricing structure offered by these banks is dependent on the frequency of trades, where a lower flat fee is charged for clients with a higher number of trades per quarter.

In addition to these traditional banks, there exist brokerage firms, like Interactive Brokers (IB), that operate in the Canadian market. IB's pricing structure is different from those offered by traditional brick and mortar banks as it offers a fixed and tiered pricing structure. The tiered pricing structure sets up varying fee levels dependent on the monthly volume of shares; it further imposes a minimum and a maximum commission per order.³ Whereas the fixed pricing structure imposes a flat fee along with a maximum and a minimum cap on commission per order. IB proposes more sophisticated trading platforms and tools that are targeted towards professional traders, while being largely accessible to retail investors at the same time.

³See IB - Pricing structure at <https://www.interactivebrokers.ca/en/index.php?f=45251&p=stocks2>

In addition to Interactive Brokers, there exist firms that specifically provide online brokerage services, such as Questrade, Qtrade Investor, Virtual Brokers, but who have managed to set themselves apart from traditional financial institutions. For instance, commission free trading was first offered in Canada in 2009 by Virtual Brokers.⁴

Finally, Wealhsimple Trade is the online trading app established by Wealhsimple Inc. an investment management service that was founded in 2014. Wealhsimple trade is a mobile-only platform whose specificity is to offer unlimited commission free trades along with no account minima.

These mobile-only platforms typically tend to have more easy-to-use interfaces and garner higher ratings and positive reviews on the Apple app store. Some of these platforms, such as Wealhsimple and Questrade, even offer to their clients services that go beyond what is referred to as “self-directed investing”, where clients can buy and sell various investments (stocks, ETFS, etc), and the option to invest in managed portfolios.

It should be noted, however, that professional traders do not seem to favor these platforms, possibly due to delays in price quotes, or changes in exchange rates that can impact their potential gains.⁵

2.2 Observed surge in online stock trading by retail investors

In 2020, many trading platforms experienced a sharp increase in the number of trading accounts. There is strong evidence supporting increased user registration and activity across trading platforms operating in the United States and Canada. For example, Etrade Financial Corporation, a popular American electronic trading platform, saw a record 260,493 accounts open in March, 2020, which is a sharp increase from the numbers observed in the past.⁶ In Canada, Wealhsimple Trade saw an increase of its user base by 80%, reaching 380,000, between July and December 2020.⁷ Interactive Brokers reported \$523 million in revenues Q2 2020, to be contrasted with \$488 million in Q2 2019, and which was attributed to a strong growth in commission revenue (+ 55% from Q2 2019 to Q2 2020).⁸

⁴Goldman, Andrew 2021, “Best Trading Platforms in Canada - 2021 Guide” Wealhsimple, <https://www.wealhsimple.com/en-ca/learn/best-trading-platforms-canada>

⁵See e.g., reviews at the App Store, <https://apps.apple.com/>

⁶Wursthorn, Michaelm Mischa Frankl-Duval and Gregory Zuckerman 2020, “Everyone’s a Day Trader Now”, The Wall Street Journal July 25, 2020. <https://www.wsj.com/articles/everyones-a-day-trader-now-11595649609>

⁷Balakrishnan, Anita 2020, “As thousands of Canadians flock to day trading during lockdown, experts urge caution: A growing number of investors are trying their hand at day trading”, December 17, 2020, IE|Investment Executive, <https://www.investmentexecutive.com/news/research-and-markets/active-trading-more-popular-among-canadians-in-2020/>

⁸Businesswire 2020 “Interactive Brokers Group Announces 2Q2020 Results” July 21, 2020 <https://www.businesswire.com/news/home/20200721005909/en/Interactive-Brokers-Group-Announces->

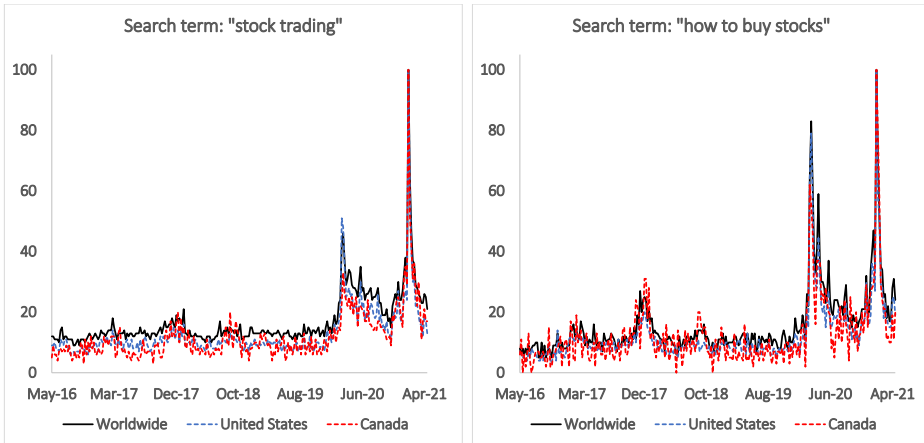


Figure 1: **Google trends.** Both panels plot the number of searches each week relative to the number of searches in the week with the highest number of searches (normalized to 100). Source: <https://trends.google.com/>

Figure 1 provides a different indication of increased interest in stock trading since Covid-19 was identified as a pandemic. It documents Google search trends data over the past 5 years. The left panel plots the search interest in stock trading in general, and the right panel provides an arguably more accurate proxy of the interest from unexperienced investors. Both figures clearly indicate a peak of interest after the mid-March of 2020, and another one in 2021 (possibly associated with the widespread coverage on the unusual price patterns of GameStop Corp stock). On average, interest in stocks trading seems to be noticeably higher during the year from the spring 2020 to spring 2021 compared to the previous years. The patterns in Canada, United States and Worldwide are similar.

2.3 Possible drivers for the observed surge in online trading

2.3.1 A lowering of the barriers to entry

Platforms such as Wealthsimple Trade in Canada and Robinhood in the US have minimized several barriers to entry to the trading market. The easy-to-use interface, combined with commission-free trading, no account minima, and with the minimum requirement of having a bank account, in practice enables any 18 and above adult to engage with this platform and become an online investor. More than half of the customers of Robinhood, a popular

2Q2020-Results

commission-free trading and investing app based in the US, opened a brokerage account for the first time ever during the period of Covid-19 pandemic.⁹

2.3.2 Stimulus packages and changes in the consumption/investment pattern

Some commentators have considered that one of the implications of the lockdown was that there were fewer opportunities to buy goods, while there was more time to manage one's finances online. It was also noted that, during the pandemic, many students found themselves eligible for CERB payments in Canada, seeing an increase of income of nearly 2k each month for 7 months, and that some of these students used the entirety of the CERB payment (14k) to invest in a small number of individual stocks.

Similar trends were present in the United States. According to a CNBC article, Americans who received a stimulus check during the pandemic increased their spending by 81% and some of the spending was directed towards buying stocks.^{10,11} Across various income brackets, securities trading seems to have been one of the expenditures that saw an increase after the disbursement of the stimulus check (As seen in Figure 2 in that article).

The overall macro impact is however difficult to establish, as in many cases, the stimulus only partially offset a loss of income during the lockdown and the contraction of economic activity. Given the overall slowdown of the economy, it is more likely that the overall effects of Covid-19 on the wealth distribution were not favourable. At the same time this evidence points towards reduced opportunity costs of participation in the stock market and we will incorporate this aspect to our model in Section 3.

2.3.3 Prevalence of sentiment analysis over fundamental analysis

Stock prices in the Covid period have shown significant movement in response to positive or negative signals on social media platforms; for example, the stock prices of several companies like Etsy Inc, Signal Advance Inc and GameStop Corp. have responded strongly to Elon Musk's tweets. Notable examples of these tweets did not contain more than a hint to the company's name. Nevertheless, as shown in the article in Bloomberg BusinessWeek (2021) these tweets led to an immediate upward jump in the mentioned firm's share price.¹² Figure

⁹Rooney, Kate 2020, "Fintech app Robinhood is driving a retail trading renaissance during the stock market's wild ride", CNBC, June 17 2020. <https://www.cnbc.com/2020/06/17/robinhood-drives-retail-trading-renaissance-during-markets-wild-ride.html>

¹⁰Fitzgerald, Maggie 2020, "Many Americans used part of their coronavirus stimulus check to trade stocks", CNBC May 21 2020. <https://www.cnbc.com/2020/05/21/many-americans-used-part-of-their-coronavirus-stimulus-check-to-trade-stocks>

¹¹See also Wursthorn et. at 2020 article in the Wall Street Journal, cited above.

¹²Gambrell, Dorothy 2020 "A Brief History of Elon Musk's Recent Market-Moving Tweets", Bloomberg BusinessWeek, February 11, 2021. <https://www.bloomberg.com/news/articles/2021-02-11/how-elon-musk->

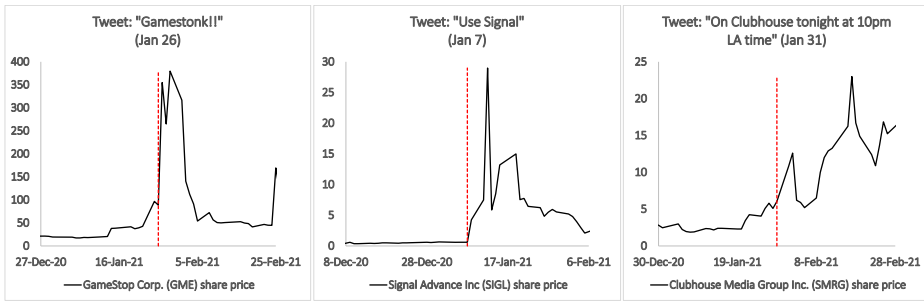


Figure 2: Share price impact of Elon Musk’s tweets in 2021. These figures are based on daily opening price data. The dates of the tweets are marked with red dotted lines.

2 highlights the timing (dotted red line) and content of these tweets along with the mentioned firms’ share price dynamics (solid black line) 30 days before and after Elon Musk posted the tweets. It shows that the share price impact of these tweets was rather persistent and lasted at least several days or weeks.

This seems to illustrate an increasing reliance on sentiment analysis vs fundamental analysis as a driver of investors’ demand and prices. This seems to illustrate an increasing reliance on sentiment analysis vs fundamental analysis as a driver of investors’ demand and prices.

2.3.4 New communities/fora and the democratization of investment advice

As more novice traders enter the market, they rely on community support to move along the learning curve. There has been a lot of mobilization and social buzz on different platforms like reddit and discord. This allows would-be investors to take confidence and to learn more easily about what used to constitute complex technical issues reserved for the specialists.

2.3.5 New forms of coordinated action via social media

The retail trending frenzy that was seen in early 2021 has had considerable spillovers in other regions of the world. FreeTrade, a prominent British trading platform, reported an increase of nearly 160,000 users in January 2021 which was accredited to the increased interest in the “reddit-rally stocks”. Their daily onboard rate for new users during this “retail trading boom” nearly tripled.¹³

s-tweets-moved-gamestop-gme-bitcoin-dogecoin-and-other-stocks

¹³White, Lawrence and Anna Irrera, 2021, "Reddit-fuelled retail trading frenzy spreads to Europe", Reuters, February 9, 2021. <https://www.reuters.com/article/us-retail-trading-europe-idUSKBN2A9284>

The GameStop Corp short squeeze that was fuelled by a social media mobilization put the spotlight on stock trading and stock trading platforms. As novice traders began sharing on social media the “gains” they had made during this short squeeze on different social media platforms, and as more media coverage on this issue was brought to attention, more retail investors were tempted into trying stock trading for the first time, on platforms like Robinhood, to buy stocks like AMC (AMC Entertainment Holdings Inc) and GME (GameStop Corp).

2.3.6 The millennial investor (passion for technology firms, gaming, cryptocurrencies)

With the increased ease of navigation and accessibility, trading apps such as Robinhood and Wealthsimple have been successful in attracting and targeting young individuals/ millennials into the trading scene (Robinhood has reported its median age of customers as being 31).¹⁴ The sharp increase in the number of novice traders across the board can be attributed to the entry of young people into the online trading landscape. The spike in users was particularly sharp during the first quarter of 2020 where the stock market experienced a downturn and a subsequent recovery; according to some equity strategists, these accounts reflect the entry of new investors who perceived this downturn to be a “generational- buying moment” while having limited knowledge of the equity space.¹⁵ The interface built by Robinhood and to some extent Wealthsimple Trade is often compared to that of a social media app or a mobile game. Wealthsimple Trade offers its customers 25\$ when a friend signs up and trades at least 100\$. Robinhood, in particular, has been remarked to have a gamified easy-to-use interface that further attracts novice investors.

As previously remarked, platforms like Robinhood have led to the gamification of investing and trading to a considerable extent.¹⁶ This gamification can explain the investing activity typically found on this platform. According to the New York Times, it was found that in the first quarter of 2020 Robinhood users traded 40 times as many shares as Charles Schwab (a more traditional brokerage) customers. Gamification comes with the risk of large losses as it may encourage retail investors to pursue trading strategies that are systematically disconnected from fundamental drivers of share prices (e.g., too frequent stock trading

¹⁴See Rooney 2020 article in CNBC, cited above.

¹⁵Fitzgerald, Maggie 2020, “Young investors pile into stocks, seeing ‘generational-buying moment’ instead of risk”, CNBC, May 12 2020. <https://www.cnbc.com/2020/05/12/young-investors-pile-into-stocks-seeing-generational-buying-moment-instead-of-risk.html>

¹⁶Massa, Annie and Edward Robinson, 2020, “Robinhood’s Role in the ‘Gamification’ of Investing, Bloomberg Wealth, December 19 2020. <https://www.bloomberg.com/news/articles/2020-12-19/robinhood-s-role-in-the-gamification-of-investing-quicktake>

is rarely a profitable strategy for retail investors), and instances of large individual losses due to gamification have indeed been documented.¹⁷

3 The model

3.1 The setting

There are two assets: a risk-free asset, which offers certain return normalized to one, and a risky asset with an uncertain fundamental value

$$v = \theta + u, \tag{1}$$

where θ is the explainable part of the fundamental value, and u is the unexplainable part of the fundamental value. Both θ and u are Normally distributed and independent: $\theta \sim \mathcal{N}\left(\theta_0, \frac{1}{\beta_\theta}\right)$, $u \sim \mathcal{N}\left(0, \frac{1}{\beta_u}\right)$. These distributions are common knowledge.

Three types of investors trade these assets: 1) a mass I of sophisticated informed investors, 2) a mass N of retail investors and 3) competitive market makers. At date 0 retail investors choose whether to participate, at date 1 all investors trade, and at date 2 the fundamental value v is realized and investors consume.

Sophisticated informed investors (e.g., institutional investors) are a homogenous set of investors who each observe the value of θ , and the same CARA utility with risk tolerance parameter τ (inverse of risk aversion), and always participate. Each of them chooses demand schedule d , to maximize his/her expected utility function

$$\mathbb{E} \left[- \exp \left(- \frac{d(v - p) + w_I}{\tau} \right) \mid \theta \right], \tag{2}$$

where w_I is the informed investor’s wealth.¹⁸ When choosing their optimal demand schedule, investors know the equilibrium asset price p .¹⁹ The aggregate demand of informed investors is $D = Id$.

Retail investors are heterogenous in their wealth, and the wealth distribution is $G(w_i)$ with support $[w_m, \theta)$, where $w_m > 0$, and w_i is independent of θ and u . At date 0 each

¹⁷Popper, Nathaniel 2020, "Robinhood Has Lured Young Traders, Sometimes With Devastating Results", The New York Times, July 8, 2020. <https://www.nytimes.com/2020/07/08/technology/robinhood-risky-trading.html>

¹⁸The assumption that w_I is the same across all informed investors is inconsequential because their demand does not depend on w_I . This is a well known feature of CARA utility and their demand will be specified shortly.

¹⁹This can be interpreted as the investor observing the market price p and submitting a market order in a competitive market, or him submitting a set of limit orders for each price.

retail investor i chooses whether to participate in the stock market or not, and his/her choice to participate is denoted with $\mathbf{1}_i \in \{0, 1\}$. If a retail investor chooses to participate, he/she must pay a fixed participation cost F in monetary equivalent units. This cost reflects monetary costs such as the fees of establishing an account, but also non-monetary costs like spending time to find out more about trading and the risky asset. As in Peress (2004), each consumer's risk tolerance parameter $\gamma(w_i)$ is increasing in the consumer's wealth, $\gamma'(w_i) > 0$. As in Mendel and Shleifer (2012) retail investors have biased beliefs about the explainable component of the fundamental value, and do not learn fundamental information from prices. Namely, they consider that the fundamental value of the risky asset is

$$S + u,$$

where S is Normally distributed, $S \sim \mathcal{N}\left(S_0, \frac{1}{\beta_S}\right)$ and u is the unexplainable component as in (1), and S is independent of u , θ and w_i . Retail investors observe the value S if they choose to participate in the stock market, and participating retail investors observe the same value of S . The value S can be viewed as the "sentiment", and may for example be driven by non-fundamentals based coordination in social media platforms.²⁰

At date 1, each retail investor i who has decided to participate in the stock market observes the sentiment value S , the risky asset price p , and chooses his/her demand schedule h_i to maximize

$$U_i(\mathbf{1}_i = 1) \equiv \mathbb{E} \left[-\exp \left(-\frac{h_i(S + u - p) + w_i - F}{\gamma(w_i)} \right) \mid S, w_i \right]. \quad (3)$$

The utility of a retail investor who does not participate in the stock market is certain and given by

$$U_i(\mathbf{1}_i = 0) = -\exp \left(-\frac{w_i}{\gamma(w_i)} \right).$$

Anticipating her/his expected utility of at date 1, each investor i chooses to participate, if and only if,

$$\mathbb{E}[U_i(\mathbf{1}_i = 1)] \geq U_i(\mathbf{1}_i = 0).$$

When retail investors decide whether to participate, they know the price but do not know the realization of S . The aggregate demand of retail investors is $H = N \int_i \mathbf{1}_i h_i di$.

²⁰An example of this would be the coordinated retail trading of GameStop shares. This approach also provides one possible way to endogenize noise traders' demand. A model where retail investors observe slightly different signals would deliver qualitatively similar results and would needlessly complicate the model. In Section 4 we will also analyze an extension where some retail investors trade based on fundamental information.

Finally, there are competitive market makers who implement the market efficiency condition, as they set the price equal the expected fundamental value of the risky asset based on all public information. Namely, market makers do not observe the random variables θ , S , and u , but observe the total order flow $D + H$, which they use to update their beliefs on the fundamental value of the asset, i.e.,

$$p = \mathbb{E} [v|D + H].$$

Market makers also know the structure of the model and distributions from which random variables are drawn.²¹

3.2 Retail investors' participation

Before analyzing the full problem, let us focus on the drivers of retail investors' participation. For that, we need to first derive the retail investor's optimal demand and utility, should he/she participate. To shorten the notation, denote $c_i \equiv h_i (S + u - p) + w_i - F$. Using the observation that the conditional distribution $c_i|S \sim \mathcal{N} (\mathbb{E} [c_i|S], Var [c_i|S])$, and the functional form of Normal distribution, we can simplify (3) to obtain

$$U_i (\mathbf{1}_i = 1) = - \exp \left(- \frac{1}{\gamma (w_i)} \left(\mathbb{E} [c_i|S] - \frac{1}{2\gamma (w_i)} Var [c_i|S] \right) \right), \tag{4}$$

where $\mathbb{E} [c_i|S] = h_i (S - p) + w_i - F$ and $Var [c_i|S] = h_i^2 \frac{1}{\beta_u}$. From there, it follows that the retail investor's demand has a familiar form

$$h_i = \gamma (w_i) \beta_u (S - p). \tag{5}$$

The optimal demand is higher if the difference between the investors' subjective beliefs about the fundamental value are higher and when there is less uncertainty about the unexplainable components of the fundamental value (higher β_u). Furthermore, the demand is higher when investors are more risk tolerant, and as we assume that risk tolerance is increasing with wealth, wealthier investors' demand for risky asset is higher.

We can then derive the participation condition

Lemma 1 *There exists a threshold level of wealth \bar{w} such that all retail investors with $w_i \geq \bar{w}$ participate in trading the risky asset and all retail investors with $w_i < \bar{w}$ do not. The threshold*

²¹The solution method used in this paper is based on Vives (1995), which considers competitive traders. A similar market efficiency condition also features in models with traders who have market power such as Kyle (1985) and Holden and Subrahmanyam (1992). Many other papers build on these seminal settings.

$$\bar{w} = \gamma^{-1} \left(\frac{F}{\frac{(S_0-p)^2}{2(\sigma_S^2+\sigma_u^2)} + \frac{1}{2} \ln \left(1 + \frac{\sigma_S^2}{\sigma_u^2} \right)} \right), \tag{6}$$

where γ^{-1} is the inverse function of γ , $\sigma_S^2 \equiv \frac{1}{\beta_S}$ and $\sigma_u^2 \equiv \frac{1}{\beta_u}$ are variances of the sentiment and unexplainable part of the fundamental.

Proof. See Appendix A.1. ■

Lemma 1 features the realistic property that retail investors only participate in trading risky assets if they are sufficiently wealthy. Furthermore, the condition (6) is helpful for interpreting the Covid-19 situation. First of all, it is natural to expect that the variance, σ_u^2 , increased when the Covid-19 pandemic emerged: there was greater uncertainty about the value of firms and the economy in general. This effect alone would suggest that the threshold level of wealth needed for participating in the stock markets would have increased and led to less participation. This however is inconsistent with the stylized facts discussed in Section 2, which indicate that more investors participated, and arguably new less wealthy investors started to trade stock. Based on (6) there are two realistic effects that could explain the surge of new investors, by reducing the wealth threshold of participation. First, \bar{w} is lower when the cost of participation, F , is lower. It is plausible that some non-monetary costs, like the opportunity cost of time, could have become smaller. This effect also suggests why new investors would have incentives to use trading apps and platforms that do not charge fixed fees. Second, a larger perceived wedge between the prices and the expected sentiment, $(S_0 - p)^2$ is another realistic driver of greater participation.²² Finally, a higher perceived precision of the sentiment, β_S , (lower perceived variance) has an ambiguous effect on participation incentives. If the wedge $(S_0 - p)^2$ was small, a higher perceived precision (lower variance) would reduce participation. However, when the wedge is large enough, a higher perceived precision of S would further increase participation.

Given this participation threshold, we can derive the aggregate demand by retail investors as

$$H = N \int_i \mathbf{1}_i h_i di. = N \int_{\bar{w}}^{\infty} \gamma(\bar{w}) \beta_u (S - p) dG(w_i) = N \beta_u (S - p) \gamma(\bar{w}) (1 - G(\bar{w})). \tag{7}$$

²²Both of these effects seem crucial for explaining the stylized facts discussed. It is worth noting that if retail investors were to trade primarily based on fundamental information, we would have a similar participation condition with S_0 replaced by θ_0 , and σ_S^2 replaced by σ_θ^2 . However, it would be much more challenging to explain why unbiased prior beliefs about the fundamental value would be far from the prices. This together with empirical evidence of behavioural motives of online trading seems more consistent with a sentiment based explanation. See also Section 4.

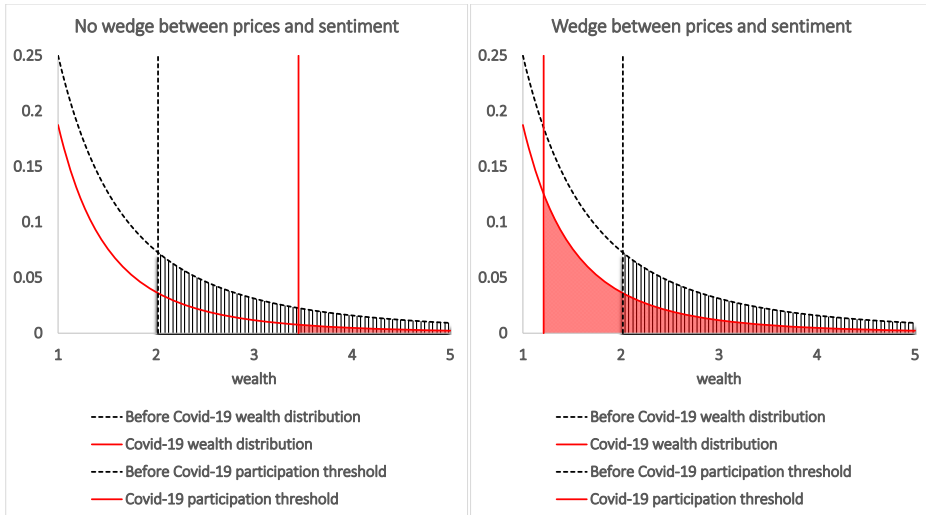


Figure 3: **Covid-19 impact and stock market participation.** These figures consider Pareto wealth distribution (originally applied and often used to describe the distribution of wealth in a society). The before Covid-19 distribution assumes scale parameter $w_m = 1$ and shape parameter $\alpha = 2$. Other parameters in the before Covid-19 economy are: $\gamma_c = 1$, $\sigma_u^2 = 1$, $\sigma_S^2 = 1$, $F = 0.7$, $S_0 = p$. The Covid scenarios assume that: 1) wealth distribution becomes less favourable in the sense of first order stochastic dominance, captured by $\alpha = 3$; 2) Uncertainty increases, $\sigma_u^2 = 1$. The wedge $|S_0 - p| = 0$ on the left panel and $|S_0 - p| = 1.5$ on the right panel.

To shorten the notation, we will use $\bar{\gamma}(\bar{w}) \equiv \gamma(\bar{w})(1 - G(\bar{w}))$ to denote the average risk aversion of participating retail investors, which in turn is determined by the threshold \bar{w} , and the wealth distribution.

The analysis so far has kept the wealth distribution unchanged. As discussed in the introduction, it is also plausible that the Covid-19 shock had direct negative wealth effects. To capture this, consider that the pre-Covid-19 wealth distribution first order stochastically dominates the during-Covid-19 wealth distribution, i.e., $G_{before}(w_i) \leq G_{during}(w_i)$ for all w_i , with strict inequality for some w_i . This effect alone would have lowered the retail investors' demand, ceteris paribus, as it would imply that $1 - G_{during}(\bar{w}) \leq 1 - G_{before}(\bar{w})$, and for many distributions and values of \bar{w} this inequality would be strict.

To summarize, a less favorable wealth distribution and a higher uncertainty ($\sigma_u^2 = 1/\beta_u$) would both reduce retail investors' participation. These effects are offset if the retail investors perceive a large wedge between the asset prices and their (possibly biased) expectations and the effective participation costs are lower. These considerations are necessary to explain the

surge of new investors. Figure 3 illustrates these effects, by considering that risk tolerance is linearly increasing in wealth $\gamma(w_i) = \gamma_c w_i$, and by considering Pareto wealth distribution with scale parameter w_m and shape parameter $\alpha > 1$. On both figures, the black dotted lines give an illustrative example of pre-Covid-19 wealth distribution and threshold (6). The shaded area with black vertical stripes represents the mass of stock market participants pre-Covid-19. The solid red lines cover the Covid-19 period, and on both panels of Figure 3, we assume that uncertainty increases (i.e., higher $\sigma_u^2 = 1/\beta_u$), and the pandemic leads to a wealth distribution that is stochastically dominated by the pre-crisis wealth distribution (higher parameter α). On the left panel, we consider that there is no wedge, i.e., $S_0 = p$, and on the right panel we consider that retail investors prior beliefs indicate a large enough positive wedge, $S_0 > p$. The red area represents the mass of stock market participants under these Covid-19 scenarios. Whether stock market participation increases or decreases due to Covid-19 is determined by comparing the red and the striped area on each graph. We can see that a realistic explanation of increased stock market participation seems to require a large enough difference $|S_0 - p|$ to offset the negative wealth and uncertainty effects that would alone reduce participation.

3.3 Equilibrium asset prices

The aggregate demand by retail investors was specified in the previous Section, and following a similar derivation, the individual and aggregate demand by sophisticated informed investors are respectively

$$d = \tau\beta_u (\theta - p) \tag{8}$$

and $D = I\tau\beta_u (\theta - p)$. From this and (7) the total order flow/limit order schedule by these two investor groups is

$$H + D = N\beta_u (S - p) \bar{\gamma}(\bar{w}) + I\tau\beta_u (\theta - p). \tag{9}$$

Competitive market makers observe the total demand, and use it to Bayesian update their beliefs about θ and the fundamental value v . Market makers do not consider S to be payoff relevant and it is a source uncertainty in the order flow.²³ Following a similar approach to Vives (1995), we can now solve for the equilibrium price.

²³Noise in the order flow is necessary to avoid the Grossman and Stiglitz (1980) paradox. If all retail investors would be informed about θ instead of following the signal S and/or the market makers would observe S , it would be necessary to introduce other sources of noise. Such modifications would not alter the key messages.

Lemma 2 *The equilibrium price is*

$$p = (1 - \kappa_\theta) \theta_0 + \kappa_\theta \theta + \kappa_S S, \tag{10}$$

where

$$\begin{aligned} \kappa_\theta &\equiv \frac{(I\tau)^2 \beta_S}{(N\bar{\gamma}(\bar{w}))^2 \beta_\theta + (I\tau)^2 \beta_S} \\ \kappa_S &\equiv \frac{N\bar{\gamma}(\bar{w}) (I\tau) \beta_S}{(N\bar{\gamma}(\bar{w}))^2 \beta_\theta + (I\tau)^2 \beta_S}. \end{aligned} \tag{11}$$

Proof. See Appendix A.2. ■

Equilibrium price is a linear function of the prior mean of the fundamental, θ_0 , the explainable component of the fundamental, θ , and retail investors' sentiment, S . We can see that the equilibrium price does not directly depend on the degree of uncertainty, $\sigma_u^2 = 1/\beta_u$, and it only depends on this parameter via its impact on retail investors' participation decisions and the average risk tolerance of $\bar{\gamma}(\bar{w})$ participating retail investors. This is because market makers are risk neutral and retail and institutional investors have the same assessment of unexplainable part of the fundamental. This also implies that asset price volatility is affected by the degree of uncertainty $\sigma_u^2 = 1/\beta_u$ solely via retail investors' participation decisions.

We also see that an increased retail investor participation, which translates into a higher $\bar{\gamma}(\bar{w})$ reduces the relevance of information about the fundamental as the weight $\frac{\partial \kappa_\theta}{\partial \bar{\gamma}(\bar{w})} < 0$. At the same time it increases the impact of prior expectations θ_0 . Interestingly, the impact of retail investor sentiment is non-monotonic in retail investor participation.

Corollary 3 *The impact of sentiment, S , on asset prices is maximized at an intermediate level of retail investor participation. Namely, weight κ_S is at its highest when $\bar{\gamma}(\bar{w}) = \gamma^* = \tau \frac{I}{N} \left(\frac{\beta_S}{\beta_\theta} \right)^{\frac{1}{2}}$*

Proof. This follows from $\gamma^* = \arg \max_{\bar{\gamma}(\bar{w})} \kappa_S$, where κ_S is defined in (11). The first order condition implies γ^* , and the second order condition confirms it as maximum. ■

When there are few retail investors, their total demand, which depends on S , is small and does not have that much impact on prices. At the same time, when there are many retail investors, their total order flow becomes very noisy and less informative about the fundamental value. Namely, the variance of the signal that the market makers obtain from the order flow is $Var(z|\theta) = \frac{(N\bar{\gamma}(\bar{w}))^2}{(I\tau)^2} \frac{1}{\beta_S}$ increasing in $\bar{\gamma}(\bar{w})$ (see Appendix A.2). Consequently, the market makers set the price close to the prior mean θ_0 and both the sentiment and

the informed traders' information has less impact. The value of $\bar{\gamma}(\bar{w}) = \gamma^*$ that maximizes the weight on the sentiment is proportional to the informed investors' risk tolerance and higher if retail investors believe that their signal is more precise, β_S . A higher precision of fundamental information, β_θ , reduces γ^* .

As argued in the previous section, large enough shocks to the wealth distribution, uncertainty in the economy, and the wedge between prices and sentiment, can all lead to fast changes in the retail investors' participation in trading, which this section shows can lead to fast price changes and changes in the sensitivity of prices to different sources of uncertainty.

As more standard results from a comparative statics standpoint, if the prior mean θ_0 is more precise (higher β_θ) then the prices are closer to the prior mean and information shocks, θ and S , have a lesser impact. At the same time, when the retail investors consider their signal S to be more precise, the weight of the prior mean is relatively lower. It is also intuitive that when informed investors are more risk tolerant (higher τ), they trade more in the equilibrium and their informative signal θ has a greater impact on equilibrium prices.

Finally, we can derive the unconditional variance of the risk asset price, given information that is available at date 0.²⁴ From (10) and (11) we obtain that

$$\text{Var}(p) = \frac{\kappa_\theta^2}{\beta_\theta} + \frac{\kappa_S^2}{\beta_S} = \frac{1}{\left(\frac{N\bar{\gamma}(\bar{w})}{I\tau}\right)^2 \frac{\beta_\theta}{\beta_S} + 1} \cdot \frac{1}{\beta_\theta}$$

We can see that even though a greater retail investor participation has complex effect on different drivers of asset prices, it always reduces the unconditional volatility of prices by making the market more liquid. The effect of other parameters is also intuitive: a higher precision of informed signals, β_θ , reduces the variance of the price and if retail investors perceive the sentiment signal, β_S , to be more precise, the prices are more volatile.

4 Extension: presence of informed retail investors.

The assumption in the baseline model is that all retail investors are trading based on a sentiment signal, and do not have fundamental information. Consequently, these traders make losses in expectations. While the idea that retail investors tend to make losses on average is consistent with a number of empirical findings discussed in the introduction, this assumption may sound strong. Certainly some retail investors could have fundamental information and profit from trading on it.

To allow this possibility, we extend the setting as follows. A fraction $\lambda \in [0, 1)$ of retail

²⁴Note that $\bar{\gamma}(\bar{w})$ is known at this point as it depends on publicly observable variables.

investors are informed and observe θ . The probability of being informed is independent of the investor's wealth. If an informed retail investor participates, the optimal demand schedule h_i^I must maximize

$$U_i^I (\mathbf{1}_i^I = 1) \equiv \mathbb{E} \left[- \exp \left(- \frac{h_i^I (\theta - p) + w_i - F}{\gamma (w_i)} \right) \mid \theta, w_i \right].$$

The utility from non-participation remains the same, and an informed retail investor i participates, if and only if, $\mathbb{E} [U_i^I (\mathbf{1}_i^I = 1)] \geq U_i (\mathbf{1}_i = 0)$. The remaining fraction $(1 - \lambda)$ of retail investors solve the same problem as in in the main setting. The wealth threshold for participation may now differ among retail investors, and we denote the demand of informed retail investors and the demand of sentiment following retail investors with H^I and H^S , respectively. The total demand of retail investors is $H = H^I + H^S$. The informed institutional investors' problem remains unchanged and the market makers take the new structure into account.

The solution method of this problem follows the same steps as in Section 3. Naturally, the only difference in the informed retail trader's problem compared to earlier is that his/her decision is based on the distribution of θ , rather than S . There are now two, potentially different wealth levels, at which different types of retail investors participate. These are

$$\begin{aligned} \bar{w}^I &= \gamma^{-1} \left(\frac{F}{\frac{(\theta_0 - p)^2}{2(\sigma_\theta^2 + \sigma_u^2)} + \frac{1}{2} \ln \left(1 + \frac{\sigma_\theta^2}{\sigma_u^2} \right)} \right) \\ \bar{w} &= \gamma^{-1} \left(\frac{F}{\frac{(S_0 - p)^2}{2(\sigma_S^2 + \sigma_u^2)} + \frac{1}{2} \ln \left(1 + \frac{\sigma_S^2}{\sigma_u^2} \right)} \right), \end{aligned}$$

where informed retail investors participate as long as their wealth is at least \bar{w}^I , and the rest participates as long as their wealth is at least \bar{w} , as before. In stable times, these thresholds may be similar or may be ranked either way. However, as it is arguably plausible that the wedge between the price and the expected sentiment, $|S_0 - p|$, during Covid-19 crisis was higher than the wedge between the price and prior beliefs, $|\theta_0 - p|$, it seems more likely that \bar{w} was lower than \bar{w}^I , and new investors who started to participate were more often retail investors who rely on sentiment rather than on unbiased fundamental information. Furthermore, it is hard to find compelling justification why a wedge $|\theta_0 - p|$ would remain persistently high, as in many rational settings, prior beliefs often reflect information revealed by past prices. As shown in Section 3.2, such a wedge would need to be there to justify the surge in new retail investors during Covid-19.

The derivation of aggregate demand and equilibrium prices can nevertheless maintain flexibility regarding the comparison of \bar{w}^I and \bar{w} . Provided these threshold, we derive that $H^I = N\lambda\beta_u(\theta - p)\bar{\gamma}(\bar{w}^I)$ and $H^S = N(1 - \lambda)\beta_u(S - p)\bar{\gamma}(\bar{w})$, where $\bar{\gamma}(\bar{w}) = \gamma(\bar{w})(1 - G(\bar{w}))$ and $\bar{\gamma}(\bar{w}^S) = \gamma(\bar{w}^S)(1 - G(\bar{w}^S))$.

We can then derive the equilibrium price²⁵ in this setting as

$$p = (1 - \kappa_\theta)\theta_0 + \kappa_\theta\theta + \kappa_S S,$$

where

$$\begin{aligned}\kappa_\theta &= \frac{(N\lambda\bar{\gamma}(\bar{w}^I) + I\tau)^2 \beta_S}{(N\bar{\gamma}(\bar{w}))^2 \beta_\theta + (N\lambda\bar{\gamma}(\bar{w}^I) + I\tau)^2 \beta_S} \\ \kappa_S &= N(1 - \lambda) \frac{\bar{\gamma}(\bar{w})(N\lambda\bar{\gamma}(\bar{w}^I) + I\tau) \beta_S}{(N\bar{\gamma}(\bar{w}))^2 \beta_\theta + (N\lambda\bar{\gamma}(\bar{w}^I) + I\tau)^2 \beta_S}.\end{aligned}$$

Comparing this with (10) and (11) we can see that all the main results hold and the weights on components that determine the equilibrium price have slightly changed. Namely, when a larger proportion of retail investors are informed, the equilibrium price has a higher weight on the signal about the fundamental, θ . This is a natural outcome of there being more informed traders and thus the variance of the signal that the market makers obtain from the order flow is now $Var(z|\theta) = \frac{(N\bar{\gamma}(\bar{w}))^2}{(N\lambda\bar{\gamma}(\bar{w}^I) + I\tau)^2 \beta_S} \frac{1}{\beta_S}$, which is decreasing in λ (see Appendix A.3).

Overall all highlighted effects of the main model remain qualitatively the same, while the impact of sentiment on the overall demand by retail investors and the effects of sentiment and the prior on prices are smaller in magnitude.

5 Final remarks

The paper documented stylized patterns of stock trading by retail investors during the year 2020. It also provided an REE model to highlight the drivers of these patterns and derived new results on the relationship between stock market participation and sentiment analysis. The increased participation by retail investors during the time of negative shocks to wealth distributions seems to be best explained by the increased role of common (possibly fundamentals unrelated) signals that drive the new investors' beliefs. While the observation that a common belief shock (i.e., a common sentiment) may be an important driver of asset prices and lead to a wedge between the price and fundamental value of an asset, has been noted

²⁵See details in Appendix A.3.

before²⁶, the Covid-19 period together with easier access to trading platforms via Fintechs provided an interesting case for better understanding the sentiment effects on stock market participation. Furthermore, it could be argued that the combined effect of greater uncertainty due to Covid-19, lower participation costs, and increased coordination possibilities via social media, makes the role of sentiment starker and more visible.

The surge of these new retail investors in the market during the Covid-19 period raises further interesting questions for the persistence of their future participation and for the persistence of the coordination of sentiment as a driver of asset prices in the future periods to come.

²⁶See e.g., Allen, Morris and Shin 2006.

A Proofs

A.1 Proof of Lemma 1

Using $\mathbb{E}[c_i|S] = h_i(S - p) + w_i - F$, $Var[c_i|S] = h_i^2 \frac{1}{\beta_u}$, and the optimal demand (5) in (4), we obtain

$$U_i(\mathbf{1}_i = 1) = -\exp\left(-\frac{1}{\gamma(w_i)}(w_i - F)\right) \exp\left(-\frac{\beta_u(S - p)^2}{2}\right)$$

When the retail investors decide whether to participate, they know their wealth and price, but do not know S . It follows that

$$\mathbb{E}[U_i(\mathbf{1}_i = 1)] = -\exp\left(-\frac{1}{\gamma(w_i)}(w_i - F)\right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi/\beta_S}} \exp\left(-\frac{\beta_u(S - p)^2}{2}\right) \exp\left(-\frac{\beta_S(S - S_0)^2}{2}\right) dS$$

After simplifying, we obtain

$$\mathbb{E}[U_i(\mathbf{1}_i = 1)] = -\exp\left(-\frac{1}{\gamma(w_i)}(w_i - F) - \frac{\beta_u\beta_S(S_0 - p)^2}{2(\beta_u + \beta_S)}\right) \left(\frac{\beta_S}{(\beta_u + \beta_S)}\right)^{\frac{1}{2}}$$

It then follows that $\mathbb{E}[U_i(\mathbf{1}_i = 1)] \geq U_i(\mathbf{1}_i = 0)$ if and only if

$$\gamma(w_i) \geq \frac{F}{\frac{(S_0 - p)^2}{2\left(\frac{1}{\beta_S} + \frac{1}{\beta_u}\right)} + \frac{1}{2} \ln\left(1 + \frac{\beta_u}{\beta_S}\right)}$$

As $\gamma(w_i)$ is monotonically increasing in w_i , we obtain that investor i participates, if and only if,

$$w_i \geq \bar{w} \equiv \gamma^{-1}\left(\frac{F}{\frac{(S_0 - p)^2}{2\left(\frac{1}{\beta_S} + \frac{1}{\beta_u}\right)} + \frac{1}{2} \ln\left(1 + \frac{\beta_u}{\beta_S}\right)}\right).$$

A.2 Proof of Lemma 2

From (9) the total limit order schedule is

$$L(p) = H + D = N\beta_u\bar{\gamma}(\bar{w})S + I\tau\beta_u\theta - \beta_u(N\bar{\gamma}(\bar{w}) + I\tau)p. \tag{12}$$

The signal that the market makers observe from this is

$$z \equiv \frac{L(p) + \beta_u(N\bar{\gamma}(\bar{w}) + I\tau)p}{I\tau\beta_u} = \theta + S\frac{N\bar{\gamma}(\bar{w})}{I\tau},$$

and it follows that $z|\theta \sim \mathcal{N}\left(\theta, \frac{(N\bar{\gamma}(\bar{w}))^2}{(I\tau)^2} \frac{1}{\beta_S}\right)$. By Bayes's Rule and properties of the normal distribution, it then follows that the conditional distribution

$$v|z = \theta + u|z \sim \mathcal{N}\left(\frac{\beta_\theta \theta_0 + \beta_S \frac{(I\tau)^2}{(N\bar{\gamma}(\bar{w}))^2} z}{\beta_\theta + \beta_S \frac{(I\tau)^2}{(N\bar{\gamma}(\bar{w}))^2}}, \frac{1}{\beta_\theta + \beta_S \frac{(I\tau)^2}{(N\bar{\gamma}(\bar{w}))^2}} + \frac{1}{\beta_u}\right)$$

Setting the price equal to the fundamental value based of all public information

$$p = E[v|H + D] = E[v|z] = \frac{\beta_\theta \theta_0 + \beta_S \frac{(I\tau)^2}{(N\bar{\gamma}(\bar{w}))^2} \left(\frac{L(p) + \beta_u(N\bar{\gamma}(\bar{w}) + I\tau)p}{I\tau\beta_u}\right)}{\beta_\theta + \beta_S \frac{(I\tau)^2}{(N\bar{\gamma}(\bar{w}))^2}}$$

and rearranging

$$L(p) = p \left(\frac{(N\bar{\gamma}(\bar{w}))^2}{I\tau} \frac{\beta_\theta}{\beta_S} - N\bar{\gamma}(\bar{w}) \right) \beta_u - \frac{(N\bar{\gamma}(\bar{w}))^2}{I\tau} \frac{\beta_\theta}{\beta_S} \beta_u \theta_0$$

Equating this with (12), we obtain the equilibrium price as

$$p = \frac{(N\bar{\gamma}(\bar{w}))^2 \beta_\theta}{(N\bar{\gamma}(\bar{w}))^2 \beta_\theta + (I\tau)^2 \beta_S} \beta_\theta \theta_0 + \frac{(I\tau)^2 \beta_S}{(N\bar{\gamma}(\bar{w}))^2 \beta_\theta + (I\tau)^2 \beta_S} \theta + \frac{N\bar{\gamma}(\bar{w}) I\tau \beta_S}{(N\bar{\gamma}(\bar{w}))^2 \beta_\theta + (I\tau)^2 \beta_S} S$$

Defining the coefficients κ_θ and κ_S , then proves Lemma 2.

A.3 Derivation of equilibrium price in a setting where some retail investors are informed

The total limit order schedule is now $H^I = N\lambda\beta_u(\theta - p)\bar{\gamma}(\bar{w}^I)$ and $H^S = N(1 - \lambda)\beta_u(S - p)\bar{\gamma}(\bar{w})$

$$L(p) = H + D = H^I + H^S + D = N\lambda\beta_u(\theta - p)\bar{\gamma}(\bar{w}^I) + N(1 - \lambda)\beta_u(S - p)\bar{\gamma}(\bar{w}) + I\tau\beta_u(\theta - p)$$

The signal that the market makers observe from this is

$$z \equiv \frac{L(p) + \beta_u(N\lambda\bar{\gamma}(\bar{w}^I) + N(1 - \lambda)\bar{\gamma}(\bar{w}) + I\tau)p}{\beta_u(N\lambda\bar{\gamma}(\bar{w}^I) + I\tau)} = \theta + S \frac{N\bar{\gamma}(\bar{w})}{\beta_u(N\lambda\bar{\gamma}(\bar{w}^I) + I\tau)},$$

and $z|\theta \sim \mathcal{N}\left(\theta, \frac{(N\bar{\gamma}(\bar{w}))^2}{(N\lambda\bar{\gamma}(\bar{w}^I)+I\tau)^2} \frac{1}{\beta_S}\right)$. The market makers set the price

$$p = E[v|z] = \frac{\beta_\theta \theta_0 + \beta_S \frac{(N\lambda\bar{\gamma}(\bar{w}^I)+I\tau)^2}{(N\bar{\gamma}(\bar{w}))^2} \left(\frac{L(p)+\beta_u(N\lambda\bar{\gamma}(\bar{w}^I)+N(1-\lambda)\bar{\gamma}(\bar{w})+I\tau)p}{\beta_u(N\lambda\bar{\gamma}(\bar{w}^I)+I\tau)} \right)}{\beta_\theta + \beta_S \frac{(N\lambda\bar{\gamma}(\bar{w}^I)+I\tau)^2}{(N\bar{\gamma}(\bar{w}))^2}}$$

Following the same steps as in the Proof of 2, we obtain that the equilibrium price is

$$p = \frac{\beta_\theta (N\bar{\gamma}(\bar{w}))^2}{\beta_\theta (N\bar{\gamma}(\bar{w}))^2 + \beta_S (N\lambda\bar{\gamma}(\bar{w}^I) + I\tau)^2} \theta_0 + \frac{\beta_S (N\lambda\bar{\gamma}(\bar{w}^I) + I\tau)^2}{\beta_\theta (N\bar{\gamma}(\bar{w}))^2 + \beta_S (N\lambda\bar{\gamma}(\bar{w}^I) + I\tau)^2} \theta + N(1-\lambda) \frac{\bar{\gamma}(\bar{w}) \beta_S (N\lambda\bar{\gamma}(\bar{w}^I) + I\tau)}{\beta_\theta (N\bar{\gamma}(\bar{w}))^2 + \beta_S (N\lambda\bar{\gamma}(\bar{w}^I) + I\tau)^2} S$$

B Canadian platforms for online stock trading

Table B.1 Traditional Banks Offering Online Trading Services

Name	Year established/ notable acquisitions	Account Minimum	Commission per equity trade	Apple App Store rating
BMO InvestorLine	1988	\$15,000	\$9.95	1.6/5
CIBC Investor's Edge	Key acquisitions in the brokerage business: Wood, Gundy & Co (1988); Merrill Lynch & Company's retail brokerage business (2001)	\$10,000	\$4.95 -\$6.95	3.3/5
RBC Direct Investing	Brokerage division of Royal Bank of Canada (1860)	\$15,000	\$6.95-\$9.95	4.8/5
Scotia iTRADE	Acquired TradeFreedom Securities and ETrade (2007)	\$10,000	\$4.99	1.2/5
TD Direct Investing	2012 TD Green Line Investor Services (1984)	\$15,000	\$7-\$9.99	4.5/5

Table B.2 Specialized Online Trading Platforms

Name	Year established	Account Minimum	Commission per equity trade	Apple App Store rating
Interactive Brokers	1978	US \$10,000 (or non-USD equivalent) - first 8 months	Tiered vs Fixed	2.6/5
Questrade	1999	\$1000	1¢ per share (\$6.95 max)	1.3/5
Qtrade Investor	2000	-	\$6.95 for equity trades	2.2/5
Wealthsimple Trade	2014	0	0	4.7/5
Virtual Brokers	2009	\$5000	0	2.5/5

Covid Economics 83, 2 July 2021: 88-115

References

- Allen, Franklin, Stephen Morris, and Hyun Song Shin, 2006. "Beauty contests and iterated expectations in asset markets." *The Review of Financial Studies* 19.3: 719-752.
- Barber, Brad M., and Terrance Odean., 2000 "Trading is hazardous to your wealth: The common stock investment performance of individual investors." *The Journal of Finance* 55, no. 2: 773-806.
- Barber, Brad M., and Terrance Odean, 2001 "Boys will be boys: Gender, overconfidence, and common stock investment." *The Quarterly Journal of Economics* 116, no. 1: 261-292.
- Bertaut, C.C. and Starr-McCluer, M., 2000. "Household portfolios in the United States" (No. 2000-26). Board of Governors of the Federal Reserve System (US).
- Bogan, V., 2008. "Stock market participation and the internet." *Journal of Financial and Quantitative Analysis*, pp.191-211.
- Brown, David P., and Robert H. Jennings, 2001, "On Technical Analysis." *Review of Financial Studies* 2, 527-551.
- Brunnermeier, Markus K., 2001, *Asset Pricing Under Asymmetric Information—Bubbles, Crashes, Technical Analysis and Herding* Oxford, England: Oxford University Press.
- Choi, J.J., Laibson, D. and Metrick, A., 2002. "How does the internet affect trading? Evidence from investor behavior in 401 (k) plans." *Journal of Financial economics*, 64(3), pp.397-421.
- Giannetti, M. and Koskinen, Y., 2010. "Investor protection, equity returns, and financial globalization." *Journal of Financial and Quantitative Analysis*, pp.135-168.
- Glossner, S., Matos, P., Ramelli, S. and Wagner, A.F., 2021. "Do institutional investors stabilize equity markets in crisis periods? Evidence from COVID-19." Available at SSRN
- Glosten, Lawrence R., and Paul R. Milgrom, 1985, "Bid, Ask and Transaction Prices in a Specialist Market With Heterogeneously Informed Traders." *Journal of Financial Economics* 14, 71-100.
- Grossman, Sanford J., 1988, "An Analysis of the Implications for Stock and Futures Price Volatility of Program Trading and Dynamic Hedging Strategies." *Journal of Business* 61, 275-298.
- Grossman, Sanford J., and Joseph E. Stiglitz, 1980, "On the Impossibility of Informationally Efficient Markets." *American Economic Review* 70, 393-408.
- Holden, Craig W., and Avanidhar Subrahmanyam, 1992, "Long-Lived Private Information and Imperfect Competition." *Journal of Finance* 47, 247-270.
- Kahneman, D. and Tversky, A., 1980. "Prospect theory." *Econometrica*, 12.
- Kahneman, D. and Tversky, A., 2013. "Prospect theory: An analysis of decision under risk.

- In Handbook of the fundamentals of financial decision making: Part I" (pp. 99-127)
- Kalda, A., Loos, B., Previtero, A. and Hackethal, A., 2021. "Smart (Phone) Investing? A within Investor-time Analysis of New Technologies and Trading Behavior." (No. w28363). National Bureau of Economic Research
- Kumar, A., 2009. "Who gambles in the stock market?" *The Journal of Finance*, 64(4), pp.1889-1933.
- Kyle, Albert S., 1985, "Continuous Auctions and Insider Trading." *Econometrica* 53, 1315-1336.
- Kyle, Albert S., and Wei Xiong. "Contagion as a wealth effect." *The Journal of Finance* 56, no. 4 (2001): 1401-1440.
- Mankiw, N.G. and Zeldes, S.P., 1991. "The consumption of stockholders and nonstockholders." *Journal of Financial Economics*, 29(1), pp.97-112.
- Mendel, B. and Shleifer, A., 2012. "Chasing noise." *Journal of Financial Economics*, 104(2), pp.303-320.
- Peress, J., 2004. "Wealth, information acquisition, and portfolio choice." *The Review of Financial Studies*, 17(3), pp.879-914.
- Van der Beck, P. and Jaunin, C., 2021. "The Equity Market Implications of the Retail Investment Boom." Available at SSRN 3776421.
- Vives, Xavier, 2008, *Information and Learning in Markets*, Princeton, NJ: Princeton University Press.
- Vives, Xavier, 1995."Short-term investment and the informational efficiency of the market." *The Review of Financial Studies* 8, no. 1: 125-160.

Consumption spending and inequality during the Covid-19 pandemic¹

Christopher D. Cotton,² Vaishali Garga³ and Justin Rohan⁴

Date submitted: 23 June 2021; Date accepted: 29 June 2021

We use a novel empirical approach to decompose the impact of different economic, demographic, and Covid factors (such as lockdowns, fear of the virus, and vaccination rate) on consumption spending and spending inequality on a week-by-week basis throughout the Covid-19 pandemic. This allows us to study how different demographic and economic groups were differentially affected by the pandemic while crucially controlling for other factors. We find that Hispanic and college-educated populations show particularly large and persistent falls in relative consumption. Spending inequality is persistently driven by political affiliation, age, education, and Covid factors. At a more disaggregated level of spending, political affiliation and Covid factors have a much stronger and more persistent impact on spending that is social-distancing-sensitive (SDS) such as travel and restaurant dining than non-SDS spending.

- 1 We thank Daniel Cooper and Giovanni Olivei for comments. The views expressed here are our own, and do not represent the views of the Federal Reserve System.
- 2 Economist, Research Department Federal Reserve Bank of Boston.
- 3 Economist, Research Department, Federal Reserve Bank of Boston.
- 4 Senior Research Assistant, Research Department, Federal Reserve Bank of Boston.

Copyright: Christopher D. Cotton, Vaishali Garga and Justin Rohan

1 Introduction

The Covid-19 pandemic has dramatically affected everyday life since March 2020. It had enormous economic and public health effects as countries recorded unprecedented rates of unemployment, sharp contractions in output, and massive death counts. Only recently have the economic losses begun to be recovered. Recent studies have highlighted the concentration of these negative effects among certain disadvantaged groups as well as unequal recoveries among these groups. For example, [Alon et al. \(2020\)](#) emphasize the larger decrease in labor force participation rates among women compared to men during the pandemic; [Abedi et al. \(2020\)](#) show that counties with higher poverty rates suffered worse health outcomes from the virus; [Couch, Fairlie and Xu \(2020\)](#) document a disproportionately higher rise in unemployment among hispanic Americans compared to white Americans.

An important aspect of the economic effect of Covid-19 is on consumption spending. [Blundell and Preston \(1998\)](#) and [Krueger and Perri \(2006\)](#) show that the distribution of consumption expenditures is a key measure of the inequality in household well-being, surpassing income-based inequality measures. This motivates our two key questions. One, which factors have driven consumption spending during the pandemic and were there inequities in the impact of Covid-19 on personal spending along demographic and economic dimensions? Two, what was the importance of these factors in explaining consumption inequality? Throughout the pandemic, the situation was constantly and swiftly changing, and any comprehensive answer to these questions needs to take this into account.

In this paper, we use high-frequency credit/debit card spending data to determine which factors drove the differences in consumption spending throughout the pandemic and quantify the relative contribution of these factors in explaining consumption inequality on a week-by-week basis. We are particularly interested in analyzing the impact of demographic, economic, and Covid factors on consumption spending. Our demographic factors include age, political affiliation, race, ethnicity, gender; economic factors include education, income, occupation; and Covid factors include fear of the virus (proxied using Covid-positive case count), restrictions such as lockdowns/containment measures, vaccination rates. Crucially, we find the impact of each of these factors holding other factors constant. This allows us to approximately find the degree to which each factor causes consumption spending and inequality rather than the degree to which each factor is associated with consumption spending and inequality.

Our analysis employs a novel empirical approach to high-frequency spending data. First, we conduct regressions of consumption spending on each of our factors. We allow for a differential week-by-week impact of each factor on consumption spending. This is necessary to identify the effects of demographic and economic variables that are fixed across weeks. Second, we use a Fields' decomposition to assess which factors are most important in explaining differences in consumption spending. We conduct this analysis for each week of the pandemic and are therefore able to assess how the importance of each factor in explaining consumption inequality has changed through every week of the pandemic.

We have three main findings. First, there were disparities in the impact of Covid-19 on consumption spending along several demographic and economic dimensions. These disparities were starker at the start of the pandemic than later on. Controlling for other factors, total spending by

Republican voters rose relative to Democratic voters at the start of the pandemic while spending by older people fell relative to younger people at the start of the pandemic. There were also inequalities among racial and ethnic groups. Spending by black people dropped at the start of the pandemic relative to white people before recovering. Spending by Asian-Americans and Hispanics fell at the start of the pandemic relative to white people and non-Hispanics, respectively, but still has not fully recovered. Higher-income, higher-educated people spent less on consumption than lower-income, less-educated people throughout the pandemic, and the relative spending of the higher-educated group has still not recovered. Therefore, overall we find evidence that the pandemic differentially affected some groups over others but the situation seems to have improved over time.

Second, demographic, economic, and Covid factors were all important in explaining consumption inequality throughout the pandemic, albeit to varying degrees in different weeks. Political affiliation was a particularly important factor in driving consumption inequality throughout the pandemic. Age played a moderate role in explaining consumption inequality throughout the pandemic. Race and ethnicity played a key role in driving consumption dynamics in the early part of the pandemic but play a much smaller role as the pandemic proceeded. Median income and education were particularly important in explaining changes in consumption later on in the pandemic and once vaccinations began. Covid case counts played an important role in explaining consumption patterns earlier in the pandemic. Restrictions have explained consumption differences, particularly in weeks with stricter lockdowns. Covid vaccination was an important factor in explaining differential consumption differences in early-2021.

Third, trends in aggregate spending mask heterogeneity in the disaggregated spending data in terms of inequalities in the impact of Covid-19 as well as the relative importance of factors in explaining consumption inequality. We decompose total spending into social-distancing-sensitive (SDS) spending and non-SDS spending using credit card spending categories. By SDS spending, we mean spending on categories such as restaurant dining and travel that were very sensitive to social distancing. Unsurprisingly, we see that there were much larger and more persistent falls in SDS spending than non-SDS spending. While political affiliation- and median income-based inequalities appear to have recovered in aggregate spending, they persist for SDS spending even as of May 2021. Most factors more strongly explain SDS rather than non-SDS spending, which makes sense since there is more variation in SDS spending to explain. We observe that political affiliation and Covid-specific factors appear to particularly strongly affect SDS spending relative to non-SDS spending. Republican voters show higher SDS spending than Democratic voters but smaller differences in non-SDS spending throughout the pandemic. Similarly, Covid factors substantially lower SDS spending but have only limited effects on non-SDS spending throughout the pandemic.

Related Literature Our paper contributes to four main literatures. First, we contribute to the literature studying the heterogeneity in the economic impact of Covid-19. [Fan, Orhum and Turjeman \(2020\)](#) document heterogeneity in risk tolerance as measured by mobility trends among groups. They find Democratic voters, women, and high-income individuals were more likely to limit their social interactions during the pandemic. [Baker et al. \(2020\)](#) show that there is

heterogeneity in spending across demographics such as age and family structure. They conclude that at the start of the pandemic, these factors explained a larger degree of the heterogeneity in the ratio of spending to income. In contrast, [Chetty et al. \(2020\)](#) and [Cox et al. \(2020\)](#) consider a longer time frame and find that while spending decreased across all households, high-income households made bigger cuts in spending and it took a longer time for their spending to recover to its pre-pandemic level. The novel contribution of our paper is that we control for a large set of demographic, economic, and Covid factors at the same time and allow these factors to have a differential effect over time. This allows us to estimate an approximately causal impact of each factor on consumption spending throughout the pandemic on a week-by-week basis.

Second, we contribute to the literature studying inequality in household consumption during recessions. As highlighted in [Meyer and Sullivan \(2013\)](#), it is important to consider inequality in consumption as well as income inequality during a recession as these two variables do not always behave in the same way. Moreover, [Blundell and Preston \(1998\)](#) and [Krueger and Perri \(2006\)](#) show that the distribution of consumption expenditures, not of income, gives greater insight into the distribution of household well-being. As previously discussed, the Covid-19 pandemic has had a large and heterogeneous economic impact leading to consumption inequality between groups. Relative to these papers, we use high-frequency spending data to provide a real-time dynamic analysis of inequality of consumption spending as driven by demographic, economic, and Covid factors.

Third, we contribute to the growing literature studying the impact of Covid-19 and associated lockdown measures on household consumption. The first weeks of the pandemic saw increases in consumption as people stockpiled goods. This was quickly followed by decreases in consumption as the disease spread and lockdown measures increased ([Cox et al., 2020](#)) ([Baker et al., 2020](#)). Our novel empirical approach allows us to isolate which demographic, economic, and Covid factors drove these changes in consumption on a week-by-week basis. We also decompose spending into SDS and non-SDS consumption to reveal the trends in disaggregated spending that are sometimes masked by aggregate spending. ([Baker et al., 2020](#)), ([Chetty et al., 2020](#)), and ([Alexander and Karger, 2020](#)) also analyze different categories of spending and find that the increase in spending at the onset of the pandemic was due to an increase in household and grocery spending, and once the virus began to spread more broadly, spending decreased among “non-essential” goods and services (such as restaurants, retail, travel).

Fourth, we contribute to the debate on whether the decrease in consumption spending during the pandemic was due to the virus itself or the policies (such as lockdowns) that were put in place to counteract the virus. [Coibion, Gorodnichenko and Weber \(2020\)](#) find that lockdown measures account for majority of the decrease in spending that occurred during the pandemic. However, [Goolsbee and Syverson \(2020\)](#) argue that while stay-at-home orders played a significant role in shifting spending from “non-essential” to “essential” goods and services, they do not account for decreases in aggregate economic activity – this was due to the fear of the pandemic itself. [Chetty et al. \(2020\)](#) comes to a similar conclusion, finding that state reopening only had a small impact on returning spending to pre-pandemic levels. [Chen, Wenlan and Qiang \(2021\)](#) find that in a sample of Chinese cities during the initial lockdown period, day-to-day changes in case counts negatively impacted spending, even after controlling for lockdown measures. Our decomposition allows us

to find the impact of lockdowns controlling for other factors that explain spending. Our results appear to suggest that lockdowns explain only a small component of consumption spending. We next look at SDS spending, which is more likely to be affected by lockdowns and restrictions. We find that lockdowns explain a larger but still small proportion of SDS spending. Therefore, our results appear to support [Chetty et al. \(2020\)](#) and [Chen, Wenlan and Qiang \(2021\)](#).

The rest of the paper proceeds as follows. Section 2 discusses the data used in the analysis. Section 3 discusses our empirical approach. Section 4 discusses the impact of demographic, economic, and Covid factors on consumption spending through various phases of the pandemic. Section 5 quantifies the relative contribution of these factors in explaining consumption inequality during the various phases. Section 6 concludes.

2 Data

In analyzing the role of demographic and economic factors on consumption spending during the pandemic, we use data from two main sources: high-frequency credit card spending data from Affinity Data Solutions (Affinity) and survey results from the American Community Survey (ACS). We supplement this with three sources of Covid-19 data: data on vaccines from the Centers for Disease Control and Prevention (CDC), case counts from the New York Times (NYT), and a measure of restrictions from IHS Markit/Macroeconomic Advisors.

Affinity Affinity has relationships with many banks and financial institutions for marketing purposes, and as part of this arrangement Affinity gets access to detailed information on credit/debit card transactions from these institutions. The Affinity data measures activity on a sample of 40 million active credit/debit cards, capturing 6-9% of all card transactions in the U.S.¹ These data are aggregated and available at various levels of geography and spending category disaggregation. In particular, we have weekly data on spending at the zip code level by Merchant Classification Category (MCC) – the worldwide classification system used by Visa and other credit card providers that assigns a transaction category to each merchant. The sample is fairly representative if slightly weighted towards the relatively affluent.² However, our approach looks at the consumption differences across groups rather than in aggregate, so the only impact of imperfect representation is lower significance of results for groups with less data, and low significance does not appear to be a problem for our analysis.

Since we only observe high-frequency card spending, we miss cash transactions, and other transactions not paid for using credit/debt cards. This includes payments for health procedures that go directly through the insurance company, mortgage/rent payments, cash payments to contractors,

¹According to our back of the envelope calculations based on [Kumar and O'Brien \(2019\)](#) this represents 3.06-4.59% of total transactions in the U.S.

²In terms of geography, the sample is fairly representative, with the Middle Atlantic states being slightly under-represented. When comparing the share of spending by region in Affinity versus in the 2017 Economic Census, Middle Atlantic States make up 14.58% of spending in the Census compared to 8.29% in Affinity. Other states seem to be represented approximately correctly. In terms of age, conditional on knowing age, and adjusting for number of unique cards for a given age group in a spending category at a point in time, roughly 17% of the sample is 18-34 years, 40% is 35-54 years, and 42% is 52+. In terms of income, 5% of the sample earns \$0-35K, 50% earns \$35-85K, and 45% earns \$85K+ annually. Thus, the sample is comprised of the relatively affluent banked population. We thank Daniel Cooper for providing these details on the sample representativeness of Affinity.

or auto purchases through loans. One could think of the dataset as providing spending on retail (except autos), some service expenditures, and some durable expenditures, such as refrigerators and recreation equipment.

ACS While the Affinity data provides weekly zip code level card spending, there is no demographic or economic information in this dataset. For this reason, we combine the data with demographic and economic information from the 2014-2018 ACS 5-year sample. The lowest level of geography which is easily compatible with the Affinity data and for which demographic and economic information is available in the ACS is a ZCTA.³ Thus, we convert the zip code level spending data from Affinity to the ZCTA level using a crosswalk provided by the U.S. Census Bureau.

We obtain the ZCTA level demographic and economic data on age, gender, race, ethnicity, median income, and educational attainment from the ACS sample. For employment, the most precise measure of occupation available at the ZCTA level are SOC major groups.⁴ The major groups are defined by a 2-digit code and sort an occupation into one of 22 categories. These categories include occupations such as “personal care and service”, “legal”, and “protective services”. We supplement this with the percent of the ZCTA population living in rural/urban regions from the 2010 Census, and county level results from the 2016 presidential election from the MIT Election Lab.

Covid data We obtain the ZCTA level Covid information from a variety of sources. We obtain the percent of the ZCTA’s county population (over age 18) that is fully vaccinated from the CDC (using state-level data when county-level data is not available); the weekly average Covid-positive case count (per 100,000) of the ZCTA’s county from the NYT github, and the containment index of the ZCTA’s state from IHS Markit/Macroeconomic Advisors.

3 Empirical Approach

Our empirical strategy in decomposing the role of different factors exploits the variation in the pre-pandemic demographic and economic composition of ZCTAs to estimate the causal effect of each factor on consumption spending during each week of the pandemic. We capture the time-varying impact of time-invariant variables by allowing for the impact of these variables to vary on a week-by-week basis. For consistency, we also allow for the impact of the Covid factors (which are not time-invariant) to vary across weeks.

³Zip code is a trademark of the U.S. Postal Service, whereas ZCTA is the Census Bureau’s measure of a zip code. The main difference between them is that while zip codes can correspond to stand-alone buildings and often change over time, ZCTAs are slightly less disaggregated but tend to be stable over time. There are 42,000 zip codes versus 32,000 ZCTAs in the U.S.

⁴The ACS classifies occupations using 2018 Census Occupation Codes (OCC) which are derived from the Office of Management and Budget’s Standard Occupation Classification (SOC) codes. SOC codes classify occupations on 4 different levels with “Major Group” being the broadest, and “Detailed Occupation” being the narrowest.

3.1 Empirical Specification

In simple terms, our empirical approach is to examine how the consumption spending of a ZCTA with a higher share of a given factor compares to a similar ZCTA with a lower share of that factor during each week of the pandemic. Equation (1) summarizes our empirical approach. Observations are compared by ZCTA, which is represented by z , and by week, which is represented by t .⁵ The dependent variable is the five-week moving average of the log of spending by consumers residing in a ZCTA z in week t . We use the five-week moving average to reduce the noise in our high-frequency data. We take the log of spending to look at the percentage change in spending across ZCTAs so that spending differences in larger ZCTAs do not dwarf spending differences in smaller ZCTAs. We also include ZCTA fixed effects (α_z) and weekly time fixed effects (γ_t) in all our regressions. To avoid multicollinearity, we omit the week of Jan 27–Feb 2, 2020, which then becomes the baseline week relative to which all the coefficients are indexed.

$$\begin{aligned} \log_spending_{z,t} = & \alpha_z + \sum_{t=1}^{T-1} \gamma_t \times week_dummy_t + \sum_{j=1}^{J-1} \sum_{t=1}^{T-1} \beta_{j,t} \times fixedfactor_z^j \times week_dummy_t \\ & + \sum_{j=1}^{J-1} \sum_{t=t_j^0}^T \chi_{j,t} \times covidfactor_{z,t}^j \times week_dummy_t + u_{z,t} \end{aligned} \quad (1)$$

$fixedfactor_z$ refers to the ZCTA level demographic and economic controls that are time-invariant. The demographic controls include: the share of the ZCTA population over age 65, and between ages 18 to 29; the share of the ZCTA that identifies as male; the share of the ZCTA that identifies as black, asian-american, multiple races, and other non-white; the share of the ZCTA that identifies as hispanic; the share of the ZCTA that is in a rural area; the share of the ZCTA's county that voted for the Republican Party in the 2016 Presidential Election, and the share that voted for a third party. The economic controls include: the median income of the ZCTA, the share of the ZCTA population with a college degree; and the share of the ZCTA employed in the different 2-digit SOC groups. We include these fixed factors interacted with the week dummies to capture the fact that the consumption spending of various demographic and economic groups may differ relatively during each week of the pandemic. For example, this would capture the fact that older people may choose to stay at home relatively more and spend relatively less than the younger people in response to the health risks posed by the pandemic, and the degree to which they do so may vary throughout the pandemic.

For all demographic and economic factors that we consider (except median income), we need to exclude a dummy variable baseline category to avoid multicollinearity. The coefficients will be relative to this baseline category. For example, we include controls for the share of the population that is aged under-30 and over-65, but we have to exclude a control for the share of the population that is 30 to 64 years old. Therefore, the coefficient for the spending of the over-65 years old population reflects the impact of having more over-65s and a corresponding fall in the share of the

⁵To prevent the results from being driven by large changes in the number of cards within a ZCTA (due to, for example, consumers moving), we exclude ZCTAs for which the absolute log difference between average number of cards used in January 2020 and January 2021 is greater than 0.5.

30-64 years old population. This is true for all demographic and economic factors (except median income). The baseline categories are: age group ‘share of the ZCTA population aged 30-64 years’; gender group ‘the share of the ZCTA that identifies as female’; race group ‘the share of the ZCTA that identifies as white’; rural group ‘the share of the ZCTA that is in an urban area’; political group ‘the share of the ZCTA’s county that voted for the Democratic Party in the 2016 Presidential Election’; education group ‘the share of the ZCTA population without a college degree’; and occupation group ‘Protective Services’ that includes occupations such as correctional officers, firefighters, and police.⁶ We normalize each fixed factor by its standard deviation before interacting it with the week dummies. Therefore, the $\beta_{j,t}$ coefficient on $(fixedfactor_z^j \times week_dummy_t)$ represents the percentage increase in the consumption of a ZCTA with a 1 standard deviation (s.d.) higher share of fixed-factor j and a corresponding lower share of the omitted group of the fixed factor in week t relative to the baseline week (Jan 27–Feb 2, 2020).

We also control for a number of ZCTA-relevant Covid factors $covidfactor_{z,t}$ interacted with the week dummies. These factors include: the share of the ZCTA’s county that is fully vaccinated, the weekly Covid-positive case count (over 100,000) of the ZCTA’s county (or the ZCTA’s state when the county-level data is unavailable); and the Covid-related restrictions captured by the containment index of the ZCTA’s state.⁷ Unlike the fixed factors, these Covid factors vary over time. However, we also interact them with week dummies to capture the fact that they may affect spending differently during each week of the pandemic and to maintain a consistent approach. For example, Covid-positive case count that proxy for the fear of the virus may have had a larger negative impact on spending at the start of the pandemic when there was less of an understanding of the virus transmission and greater fear of hospitals reaching capacity. Since $covidfactor_{z,t}$ is not absorbed by the ZCTA fixed effect, we do not need to omit a baseline week dummy-Covid factor interaction unlike for the fixed factors. As we did for the fixed factors, we normalize each Covid factor by its standard deviation (by week) across ZCTAs before interacting it with the week dummies. Therefore, the $\chi_{j,t}$ coefficient on $(covidfactor_{z,t}^j \times week_dummy_t)$ represents the percentage increase in the consumption of a ZCTA with a 1 s.d. higher share of Covid factor j in week t relative to the baseline week (Jan 27–Feb 2, 2020).

3.2 Aggregate versus Disaggregated Spending

We first use our empirical approach to identify the approximately causal impact of each of the demographic, economic, and Covid factors on aggregate consumption spending during each week of the pandemic. This involves estimating equation 1 with the log of total spending as the dependent variable.

To better understand the relationship between these factors and consumption spending during various phases of the pandemic, we then examine how spending has varied by disaggregated spending categories. This involves estimating equation 1 with the log of a disaggregated spending measure as the dependent variable.

⁶The reason we use Protective Services as our omitted category is because it has a fairly average education and income level compared to other occupations.

⁷We could instead include county or state fixed effects in the regression but those are highly correlated with our demographic and economic controls, so the impact of these demographic and economic controls could be misattributed to county or state fixed effects.

We consider two separate disaggregated categories of spending: social-distancing-sensitive (SDS) and non-social-distancing-sensitive (non-SDS). SDS spending is comprised of “non-essential” goods and services that require close-proximity between people and is likely most impacted by lockdown measures. This includes spending on leisure, transportation, food-out, and beauty/massage. Non-SDS spending includes spending on grocery stores, utilities, retail, and all other spending. These two categories provide a framework for categorizing all the available consumption spending data.

Affinity allows us to perform this disaggregation by providing a Merchant Classification Code (MCC) associated with each transaction. For each spending category we select a basket of MCCs which are representative of spending in the category. We further classify these spending categories into SDS or non-SDS spending. To see which MCCs are sorted into which spending categories, please see Appendix B.⁸

4 Results

In this section, we discuss the results of the analysis of the impact of demographic, economic, and Covid factors on spending on a week-by-week basis. We look at both aggregate spending and disaggregated spending. The results of the total, SDS, and non-SDS spending regressions are presented in columns 1, 2 and 3, respectively, of Figures 1, 2, and 3. We consider demographic, economic, and then Covid factors in the following three subsections.

4.1 Impact of Demographic Factors

The impact of demographic factors on consumption spending during each week of the pandemic is shown in Figure 1.⁹ Each column of this figure plots the coefficients estimated from a different regression as described in Section 3. Each row plots the coefficients corresponding to different factors. The y-axis shows the impact of one standard deviation (SD) increase in the factor on spending (total spending in column 1, SDS spending column 2, non-SDS spending in column 3) while the x-axis shows the date/week. The impact coefficients are re-scaled so that the baseline week (Jan 27–Feb 2, 2020) is at 100. It is worth noting that the impact of a factor on total spending does not equal the sum of the impact on SDS and non-SDS spending since we are conducting regressions with log dependent variables, which will emphasize different ZCTAs in each case.

Age. Panel (a) shows the impact on consumption spending of a one standard deviation higher share of older population relative to middle-aged population in each week relative to the baseline week. We see that relative to the middle-aged population, older people spent less during the pandemic. We might have expected this difference to recover as vaccination rates increase and the risks of catching the virus, which tends to have worse health outcomes for older people, diminish. A deeper look into disaggregated spending reveals that while SDS spending does seem to be rising relatively for older people in line with this explanation, non-SDS spending has actually fallen

⁸We thank Daniel Cooper for his help with organizing these spending categories.

⁹The graphs for gender, rural, multiple/other non-white race, and third-party political affiliation are relegated to Appendix Figure 5.

relatively in recent months. This could be because large stimulus payments have prompted credit-constrained households to spend more on non-SDS items and older people are less likely to be credit constrained.

Panel (b) shows the corresponding relative impact on consumption spending of a one standard deviation higher share of younger population relative to the middle-aged. Younger people spend much more on SDS spending throughout the pandemic in line with their perceived lower cautiousness given their lower risk of significant negative health outcomes from Covid-19. They spend similar or slightly less on non-SDS spending, and overall the impact on total spending is therefore similar to the middle-aged.

Political Affiliation. Panel (c) shows the impact on consumption spending of a one standard deviation higher share of Republican voters relative to Democratic voters in each week relative to the baseline week. Republican voters spent relatively more earlier in the pandemic but their aggregate spending is statistically indistinguishable from spending by Democratic voters since July 2020. The aggregate spending masks heterogeneity in trends between SDS and non-SDS spending. Republican voters had relatively higher SDS spending at the onset of the pandemic and it remains elevated even now. On the other hand, they had relatively lower non-SDS spending at the onset of the pandemic but this has more or less converged with the Democratic voters' non-SDS spending now. This result is consistent with [Barrios and Hochberg \(2020\)](#) which finds that Republican counties were less likely to follow social distancing measures and seek out information about the virus. It suggests that there is room for Democratic voters to spend more on SDS categories going forward.

Race. Panel (d) shows the impact on consumption spending of a one standard deviation higher share of Asian-American population relative to White population in each week relative to the baseline week while panel (e) shows the corresponding relative spending impact of a one standard deviation higher share of black population. The Asian-American population spent relatively less throughout the pandemic. Their aggregate spending trend mirrors their non-SDS spending trend. Their SDS spending was relatively lower until Oct 2020 but has since converged to the level the SDS spending by Whites. The black population spent less relative to whites initially in the pandemic but their spending recovered from the summer onwards. This is true for both SDS and non-SDS categories. This suggests that racial inequity in spending is still a concern, albeit less so than at the onset of the pandemic.

Ethnicity. Panel (f) shows the impact on consumption spending of a one standard deviation higher share of Hispanic population relative to non-Hispanic population in each week relative to the baseline week. Hispanic populations spent relatively less throughout the pandemic. The difference was very pronounced early in the pandemic especially for SDS spending. This suggests that ethnic inequity in spending is still a concern although less severe than at the start of the pandemic.

Overall, we find that there were significant differences in spending across demographic groups at the start of the pandemic but these differences have diminished as the pandemic has continued. The populations that still show relatively lower total spending are Asian-American, Hispanic,

and older people. The relative spending of the black population has recovered to pre-pandemic levels, and the spending gap between the Hispanic and non-Hispanic population has also lessened. Therefore, racial and ethnic inequities are still issues but less so than at the onset of the pandemic. We also find that aggregated spending sometimes masks the trends in disaggregated spending. This is particularly true for political affiliation, where Republicans spend statistically similarly to Democrats in the aggregate, but this is driven by opposing trends in SDS and non-SDS spending – they spend persistently more on SDS and persistently less on non-SDS consumption.

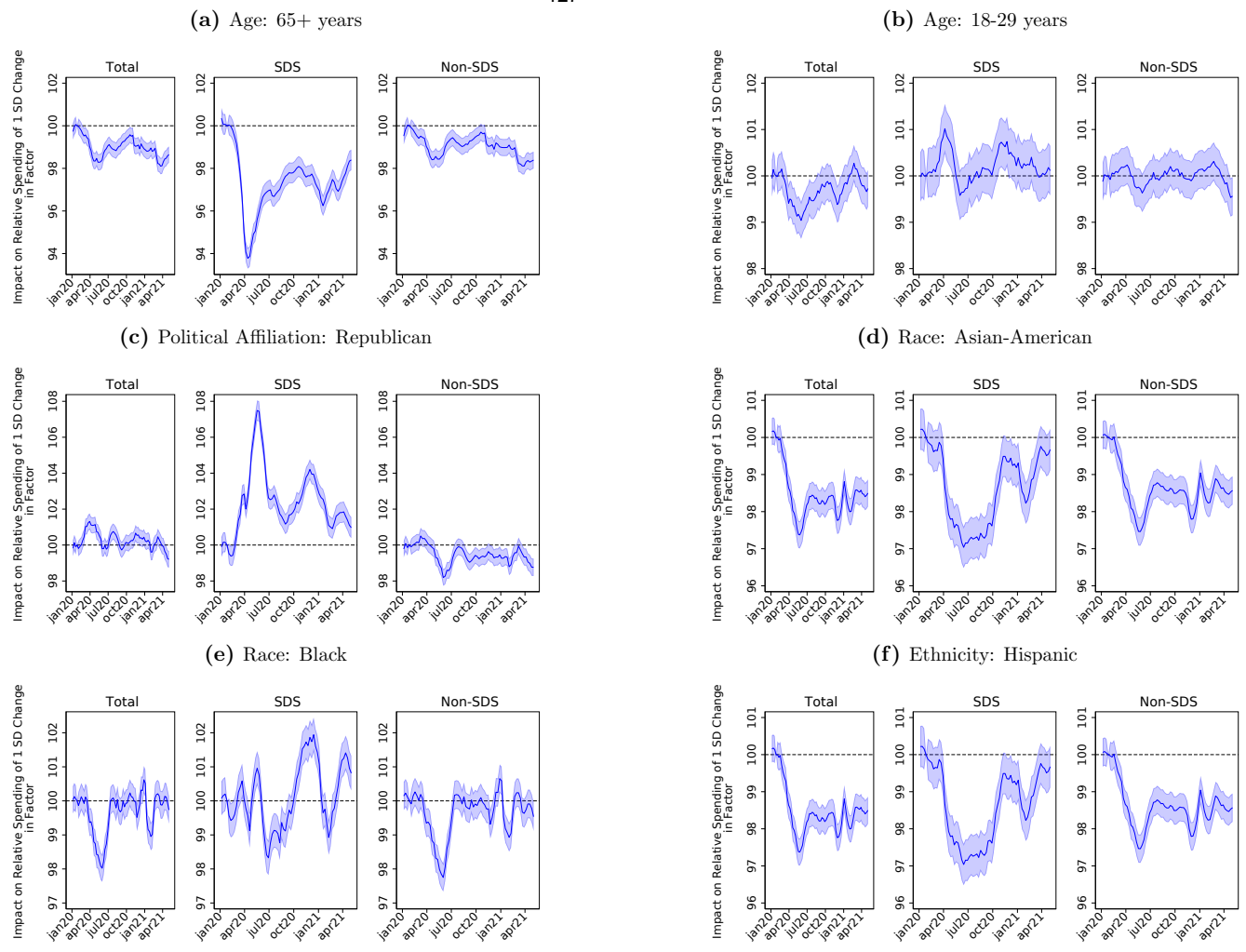


Figure 1. Demographic Factors: Aggregate versus SDS versus Non-SDS Spending By Factor.

Source: Coefficients from authors' calculations.

4.2 Impact of Economic Factors

Next, we discuss the impact of economic factors on consumption spending during each week of the pandemic. This is shown in Figure 2. The columns of this figure plot the estimated coefficients from the same three regressions that produced the impact estimates of the demographic factors – column 1 for total spending, column 2 for SDS spending, and column 3 for non-SDS spending. As in Figure 1, each column corresponds to a level of spending, and each row corresponds to an economic factor. The y-axis shows the impact of one standard deviation increase in the ‘row’ factor on the relative ‘column’ spending while the x-axis shows the date/week. The impact coefficients are re-scaled so that the baseline week is at 100.

Education. Panel (a) shows the impact on consumption spending of a one standard deviation higher share of college-educated population relative to non-college-educated population in each week relative to the baseline week. College-educated people spent relatively less throughout the pandemic. The effect is substantial relative to other factors. In terms of aggregate spending, the impact is very persistent and comparable to the start of the pandemic. However, most of the spending differences are due to SDS spending – non-SDS spending of the groups with different education levels converged pretty quickly after the pandemic began.

Income. Panel (b) shows the impact on consumption spending of a one standard deviation higher median income in each week relative to the baseline week. Higher income groups had relatively lower aggregate spending at the onset of the pandemic that mostly recovered by July 2020. There was then a seasonal spike around Christmas time followed by another dip in relative spending. The bulk of the decline in relative spending as well as its persistence is due to SDS rather than non-SDS spending. The relative decline in SDS spending is also very persistent but has very recently recovered. Interestingly, the lowest points in relative spending of the higher income groups coincide with the weeks of the three economic impact stimulus payments made by the government – April 2020, Jan 2021, and March 2021. This suggests that the relative spending decline represents the spending increase of the lower income groups (with higher marginal propensity to consume out of stimulus checks) rather than a spending decrease of the higher income groups in absolute terms.

Occupation. Panels (c)-(f) show the impact on consumption spending of a one standard deviation higher share of a the stated occupation relative to protective services in each week relative to the baseline week.¹⁰ We see that at the start of the pandemic, occupations with higher human capital requirements such as business & finance or management occupations had relatively larger declines in spending compared to occupations with lower human capital requirements such as food preparation & services or sales & related occupations. More recently, the relative spending differences are muted and the occupations have converged to some degree in terms of both SDS as well as non-SDS spending.

Overall, we find that the main economic factor that caused lower relative spending during the

¹⁰There are 21 different plots corresponding to the different 2-digit SOC occupations. For space considerations, we present four here. The corresponding plots for the remaining occupations are relegated to Appendix Figure 6.

pandemic was college education. College-educated groups continue to spend relatively less even now as the pandemic has begun to recede. This is especially true for SDS spending. This suggests that going forward, as the economy hopefully returns to normal, we should expect the higher educated groups to drive the recovery in spending, and most of this will be in the SDS categories of spending.

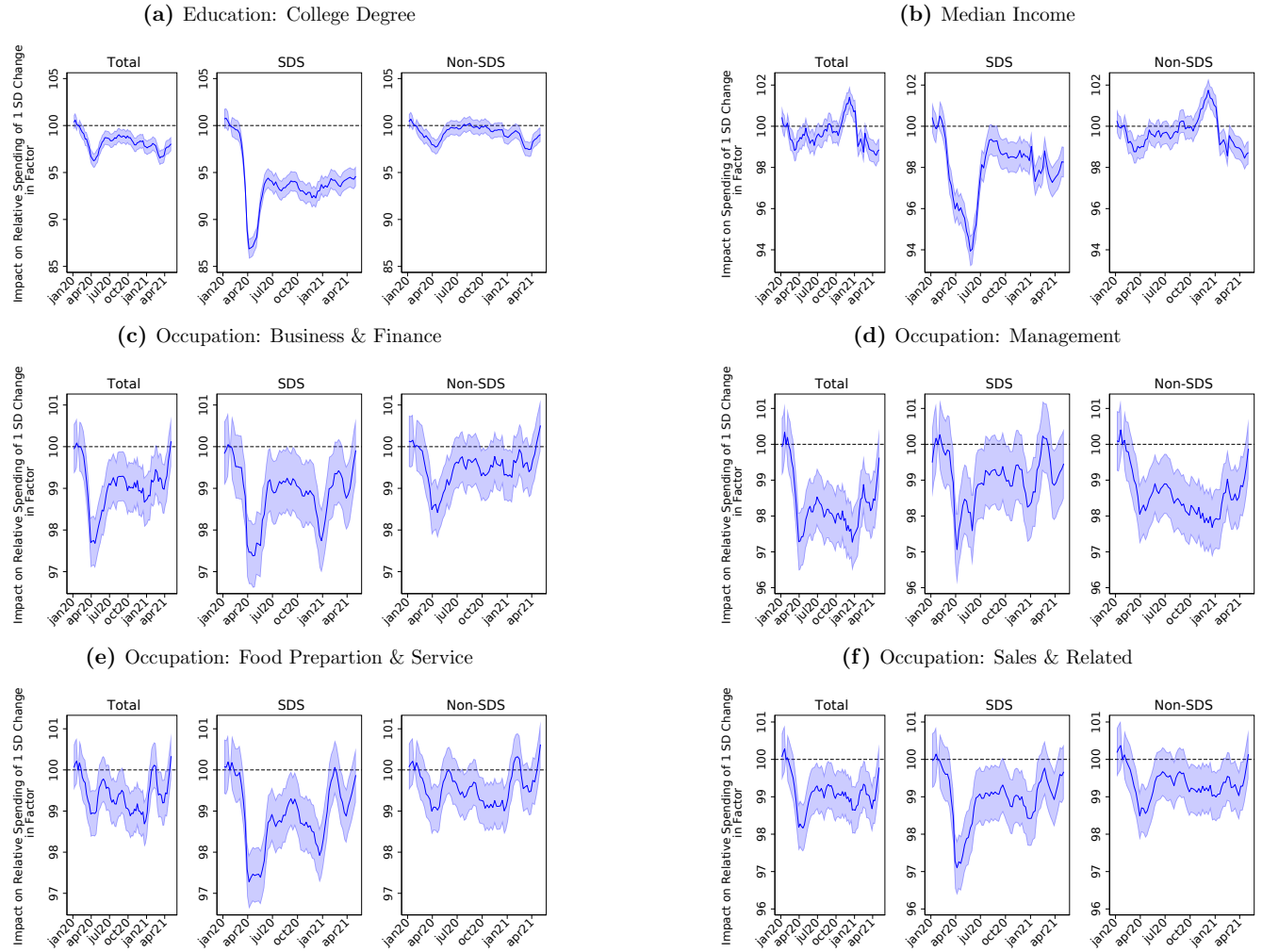


Figure 2. Economic Factors: Aggregate versus SDS versus Non-SDS Spending.

Source: Coefficients from authors' calculations.

4.3 Impact of Covid Factors

Finally, we discuss the impact of Covid factors on consumption spending during each week of the pandemic. This is shown in Figure 3. As in the preceding two subsections, the y-axis shows the impact of one standard deviation increase in the ‘row’ factor on the ‘column’ spending while the x-axis shows the date/week. The impact coefficients are re-scaled so that the baseline week is at 100.

Covid case count. Panel (a) shows the impact on consumption spending of a one standard deviation higher Covid case count (of the ZCTA’s county) in each week. Higher case count caused only slightly lower consumption spending earlier in the pandemic when case counts were very inaccurate. However, it had a larger negative impact in July 2020 and late-2020 as cases spiked and case counts were more accurate. The impact is mostly driven by SDS spending, perhaps unsurprisingly as non-SDS consumption is not as likely to risk virus spread.

Vaccination rate. Panel (b) shows the impact on consumption spending of a one standard deviation higher vaccination rate (of the 18+ population and of the ZCTA’s county where vaccination means a fully completed regiment of vaccines) in each week. The estimated impact is at 100 (implying no relative difference from the baseline week) until December 2020 when the vaccines first began to be administered. Starting in 2021, fully vaccinated groups had higher spending, both in SDS and non-SDS categories. The effect for SDS categories is more persistent presumably because it took some time for people to get used to returning to pre-pandemic activities such as travel and indoor recreation. More recently, as the vaccination roll-out has continued and most adults who want a vaccine have been able to get one, its effect on spending has become insignificant.

Containment index. Panel (c) shows the impact on consumption spending of a one standard deviation higher containment index (of the ZCTA’s state) in each week. Higher containment index, which implies stricter lockdown measures, had a consistently negative impact on consumption spending throughout the pandemic, most of which was driven by SDS spending. The effect is largest from the start of the pandemic through July 2020 and then again in late-2020 to early-2021 as lockdown measures were reinforced in some states.

Overall, we find that Covid factors had a large and significant impact on spending driven by SDS consumption. As the pandemic recedes further and case counts fall, lockdowns are removed, and more people become fully vaccinated, we should expect to see a further recovery in SDS spending.

5 Relative Contribution of Factors to Spending Inequality

In this section we investigate the relative importance of each factor in driving spending inequality through each week of the pandemic. To do this, we follow a modified approach to the decomposition in Fields (2003) that studied which factors drive income inequality in the United States.

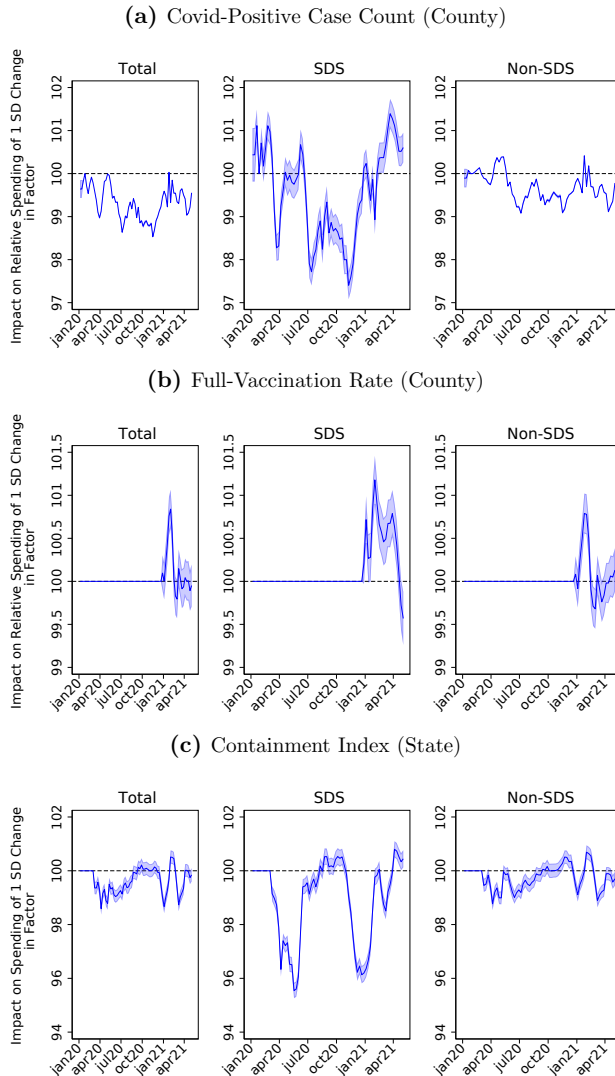


Figure 3. Covid Factors: Aggregate versus SDS versus Non-SDS Spending.

Source: Coefficients from authors' calculations.

5.1 Fields' Decomposition Application

The basic idea behind a Fields' decomposition is that if we believe that two factors explain consumption then we can decompose how much of the R^2 is explained by each of the two factors. To see this mathematically, we can consider a regression of y_i on $x_{1,i}, x_{2,i}$, which is shown in equation (2). Equation (3) holds by the definition of covariance. We can then separate this out into equation (4). Dividing by the variance of y_i allows us to get the Fields' decomposition in

equation (5).

$$y_i = \hat{\alpha} + \hat{\beta}_1 x_{1,i} + \hat{\beta}_2 x_{2,i} + \hat{u}_i \tag{2}$$

$$Var(y_i) = Cov(\hat{\alpha} + \hat{\beta}_1 x_{1,i} + \hat{\beta}_2 x_{2,i} + \hat{u}_i, y_i) \tag{3}$$

$$Var(y_i) = \hat{\beta}_1 Cov(x_{1,i}, y_i) + \hat{\beta}_2 Cov(x_{2,i}, y_i) + Cov(\hat{u}_i, y_i) \tag{4}$$

$$1 = \underbrace{\hat{\beta}_1 \frac{Cov(x_{1,i}, y_i)}{Var(y_i)}}_{R^2 \text{ due to } x_{1,i}} + \underbrace{\hat{\beta}_2 \frac{Cov(x_{2,i}, y_i)}{Var(y_i)}}_{R^2 \text{ due to } x_{2,i}} + \underbrace{\frac{Cov(\hat{u}_i, y_i)}{Var(y_i)}}_{\text{Unexplained Sum of Squares}} \tag{5}$$

We want to find how important different factors were in explaining consumption through each week of the pandemic, so we conduct these regressions on a weekly basis. We demean the variables to remove the fixed effects, which would otherwise be the dominant driver of spending between ZCTAs of different sizes. To find the Fields’ decomposition for a given week t , we therefore run the regression in equation (6). We again use the five-week moving average for spending.

$$\begin{aligned} & \log_spending_{z,t} - \overline{\log_spending_{z,t}} \\ &= \sum_{j=1}^{J-1} \beta_{j,t} \times (\overline{fixedfactor_z^j \times week_dummy_t} - fixedfactor_z^j \times week_dummy_t) \\ &+ \sum_{j=1}^{J-1} \chi_{j,t} \times (\overline{covidfactor_{z,t}^j \times week_dummy_t} - covidfactor_{z,t}^j \times week_dummy_t) + u_{z,t} \tag{6} \end{aligned}$$

The Fields’ decomposition for total spending is plotted in figure 4a. The x-axis shows the date while the y-axis shows the share of R^2 due to each factor. The variables are ordered with the demographic variables at the top of the graph followed by the economic factors in the middle and then the Covid factors at the bottom of the graph. Sex is the first factor at the top of the graph with vaccinations the last factor to be shown at the bottom of the graph. It is possible for a factor to negatively explain inequality. As can be seen in equation (5), this happens if a variable is positively correlated with spending ($cov(x_{1,i}, y_i) > 0$) but the coefficient for the variable is negative ($\hat{\beta}_1 < 0$).¹¹ These negative explanatory factors are pretty small but are why there are the small movements below zero in some weeks, which also explains why the positive factors sometimes sum to more than 100%. The Fields’ decomposition for SDS and non-SDS spending is plotted in figure 4b and figure 4c, respectively.

5.2 Fields’ Decomposition Results

Looking first at demographic variables, we observe that the share of Republican voters is an important factor in driving consumption inequality throughout the pandemic. This is particularly true in the later lockdown periods around January 2021. This fits with the common observation that there were significant differences in the degree to which Democratic and Republican voters engaged in social distancing, and also with the fact that these differences seemed to grow as the

¹¹This may seem counterintuitive but can occur when $\hat{\beta}_1 < 0$ but variable $x_{1,i}$ is strongly positively correlated with another variable for which the coefficient is strongly positive.

pandemic continued (Clinton et al., 2020). Living in an urban area was also an important factor at certain points during the pandemic, particularly again in the later lockdown periods. Age plays a moderate role in explaining consumption throughout the pandemic. This is perhaps surprising, since older people were more at risk of the virus. However, we control for some of this risk using our case count measure, and it does correspond with anecdotes that suggest older people were not social distancing more despite these higher risks.¹² Gender seems to have very little explanatory power, perhaps unsurprisingly given the similar gender composition of most ZCTAs.

Race and ethnicity played a key role in driving consumption inequality in the early part of the pandemic. We see that in May and June 2020, race and ethnicity variables explained around 50% of consumption inequality. However, as the pandemic has continued, these differences have diminished and race and ethnicity variables play a much smaller role in explaining consumption inequality. This corresponds with our finding that the degree of racial and ethnic differences has fallen as the pandemic has continued.

Economic factors played an important role in explaining consumption inequality during the pandemic. Median income and education were particularly important in explaining changes in consumption later on in the pandemic and once vaccinations began. The occupations that people work in has consistently been an important factor throughout the pandemic, but this factor has not varied in its importance over time.

Unsurprisingly, factors associated with Covid-19 play an important role in explaining consumption inequality. The number of Covid-positive cases played an important role as cases spiked in July 2020 and then again in late-2020. During this period, there were notable differences in cases of Covid-19 across different areas of the United States. Interestingly, case counts had only a very small impact at the start of the pandemic but this is presumably because with low testing it was very difficult to get an accurate picture of the number of cases at the start of the pandemic. Lockdowns have played a more consistent role during the pandemic with increases in importance around high-case periods at the start of the pandemic, in July 2020 and in early-2021. Vaccinations seemed to have had a significant impact at the start of the vaccination roll-out. However, as the roll-out has continued, the impact of the vaccinations has diminished.

We also compare the differential impact of the factors on SDS and non-SDS spending inequality in figure 4b and figure 4c, respectively. Politician affiliation and Covid factors seem more important in driving SDS than non-SDS spending differences. This makes sense given that differences in the degree to which Republicans and Democrats social distanced primarily affected social-distancing-sensitive spending (SDS). And the degree of cases/restrictions/vaccines are also likely to drive how comfortable people feel spending on social-distancing-sensitive (SDS) categories more than on non-SDS categories. On the flip side, race/ethnicity and economic variables seem to play a larger role in driving non-SDS spending. Comparing spending to the pre-pandemic period, we see that this may partly just reflect that non-SDS spending was less affected by the pandemic and continued to be driven by more standard factors.

¹²<https://abcnews.go.com/Health/frustration-millennials-older-people-taking-coronavirus-precautions/story?id=69618912>

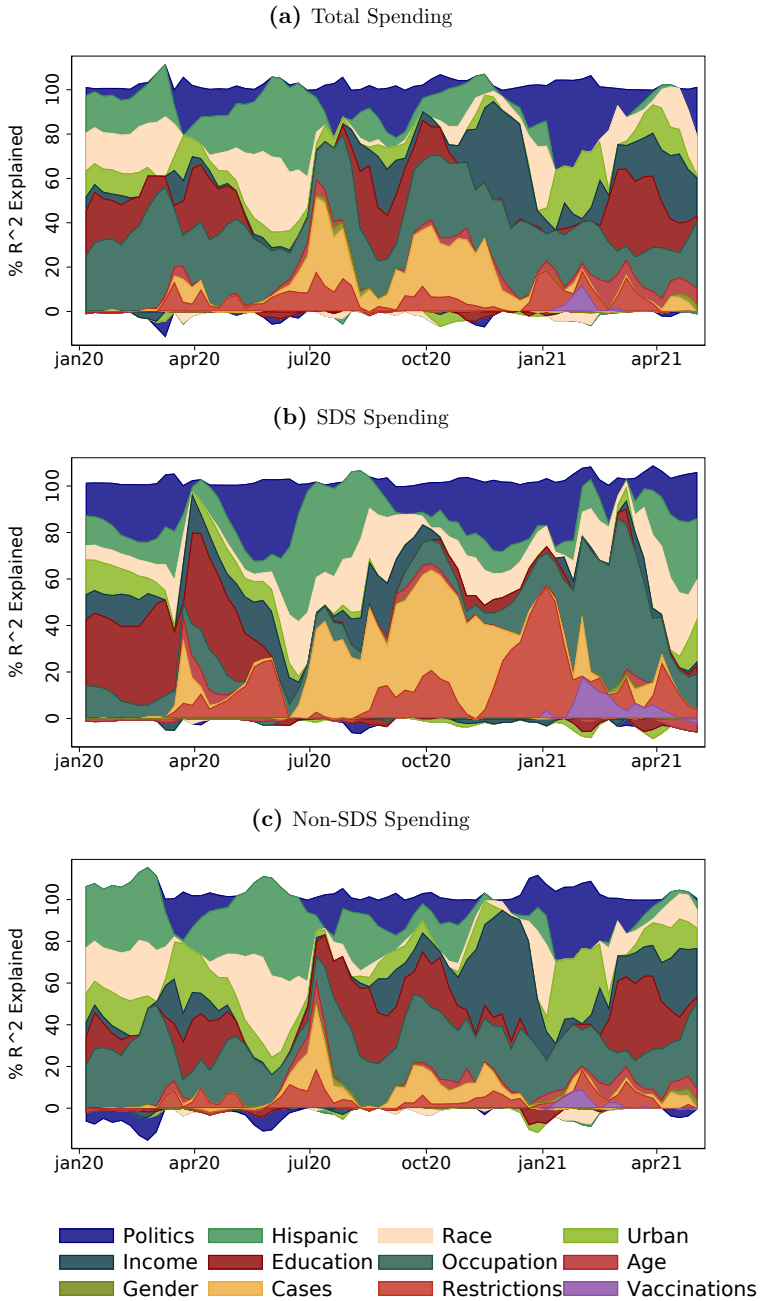


Figure 4. Fields' Decompositions.

Source: Coefficients from authors' calculations.

6 Conclusion

We have used a novel empirical approach to study which factors have caused differences in high-frequency household spending and inequality throughout the pandemic. First, we interacted time-invariant demographic and economic factors with weekly dummy variables to study the differential impact of these factors over time. Second, we used Fields' decompositions to decompose the importance of each of these factors in driving consumption inequality. We find that demographic, economic, and Covid factors are all important in explaining household spending differences, albeit in varying degrees throughout the pandemic. Within each group, having high numbers of Hispanic populations, high numbers of college-educated people and high Covid case numbers are especially likely to lower consumption.

We show that there are significant differences along racial and ethnic dimensions even after controlling for other factors. This raises important questions about why such differences exist that we hope future research will tackle.

References

- Abedi, Vida, Oluwaseyi Olulana, Venkatesh Avula, Durgesh Chaudhary, Ayesha Khan, Shima Shahjouei, Jiang Li, and Ramin Zand.** 2020. "Racial, Economic, and Health Inequality and COVID-19 Infection in the United States." *Journal of Racial and Ethnic Health Disparities*, 732–742.
- Alexander, Diane, and Ezra Karger.** 2020. "Do stay-at-home orders cause people to stay at home? Effects of stay-at-home orders on consumer behavior." Federal Reserve Bank of Chicago Working Paper 2020-12.
- Alon, Titan, Matthias Doepke, Jane Olmstead-Rumsey, and Michèle Tertilt.** 2020. "The Impact of COVID-19 on Gender Equality." National Bureau of Economic Research Working Paper 26947.
- Baker, Scott R, R.A. Farrokhnia, Steffen Meyer, Michaela Pagel, and Constantine Yannelis.** 2020. "How Does Household Spending Respond to an Epidemic? Consumption During the 2020 COVID-19 Pandemic." National Bureau of Economic Research Working Paper 26949.
- Barrios, John M, and Yael Hochberg.** 2020. "Risk Perception Through the Lens of Politics in the Time of the COVID-19 Pandemic." National Bureau of Economic Research Working Paper 27008.
- Blundell, Richard, and Ian Preston.** 1998. "Consumption Inequality and Income Uncertainty." *Quarterly Journal of Economics*, 113(2): 603–640.
- Chen, Haiqiang, Qian Wenlan, and Wen Qiang.** 2021. "The Impact of the COVID-19 Pandemic on Consumption: Learning from High-Frequency Transaction Data." *AEA Papers and Proceedings*, 111.
- Chetty, Raj, John N. Friedman, Nathan Hendren, Michael Stepner, and the Opportunity Insights Team.** 2020. "The Economic Impacts of COVID-19: Evidence from a New Public Database Built from Private Sector Data." Harvard University Working Paper.
- Clinton, Joshua, Jon Cohen, John S. Lapinski, and Marc Trussler.** 2020. "Partisan Pandemic: How Partisanship and Public Health Concerns Affect Individuals' Social Distancing During COVID-19." SSRN Working Paper No. 3633934.
- Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber.** 2020. "The Cost of the Covid-19 Crisis: Lockdowns, Macroeconomic Expectations, and Consumer Spending." National Bureau of Economic Research Working Paper 27141.
- Couch, Kenneth A., Robert W. Fairlie, and Huanan Xu.** 2020. "Early evidence of the impacts of COVID-19 on minority unemployment." *Journal of Public Economics*, 192: 104287.
- Cox, Natalie, Peter Ganong, Pascal Noel, Joseph Vavra, Arelen Wong, Diana Farrell, Fiona Greig, and Erica Deadman.** 2020. "Initial Impacts of the Pandemic on Consumer Behavior: Evidence from Linked Income, Spending, and Savings Data." *Brookings Papers on Economic Activity*.
- Fan, Ying, A. Yesim Orhun, and Dana Turjeman.** 2020. "Heterogeneous Actions, Beliefs, Constraints and Risk Tolerance During the COVID-19 Pandemic." National Bureau of Economic Research Working Paper 27211.
- Fields, Gary.** 2003. "Accounting for Income Inequality and Its Change: A New Method, with Application to the Distribution of Earnings in the United States, Research in Labor Economics." *Research in Labor Economics*.

- Goolsbee, Austan, and Chad Syverson.** 2020. "Fear, Lockdown, and Diversion: Comparing Drivers of Pandemic Economic Decline 2020." National Bureau of Economic Research Working Paper 27432.
- Hacioglu Hoke, Sinem, Diego R Kanzig, and Paolo Surico.** 2020. "Consumption in the time of Covid-19: Evidence from UK transaction data." C.E.P.R. Discussion Papers CEPR Discussion Papers 14733.
- Krueger, Dirk, and Francisco Perri.** 2006. "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory." *Review of Economic Studies*, 73: 163–193.
- Kumar, Raynil, and Shaun O'Brien.** 2019. "2019 Findings from the Diary of Consumer Payment Choice." Federal Reserve Bank of San Francisco FedNotes.
- Meyer, Bruce D., and James X. Sullivan.** 2013. "Consumption and Income Inequality and the Great Recession." *The American Economic Review*, 103(3): 178–183.

Appendices

A Additional Results

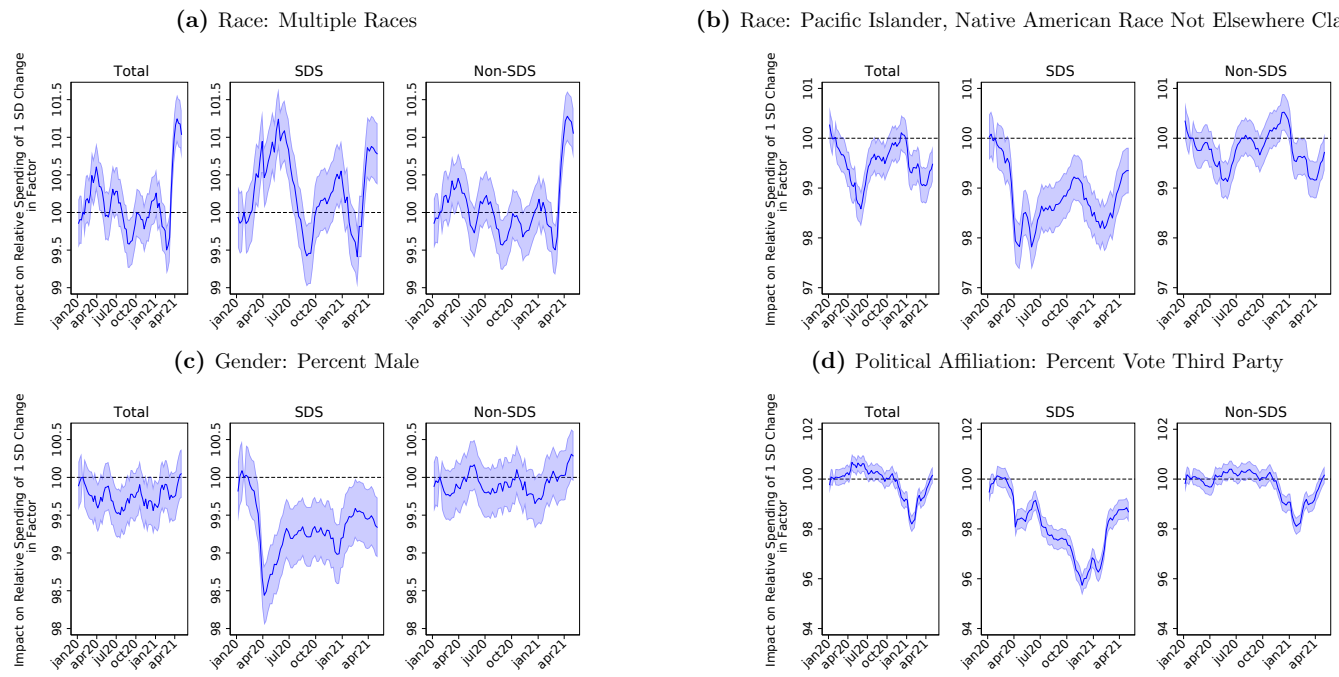


Figure 5. Additional Demographic Factors

Source: Coefficients from authors' calculations.

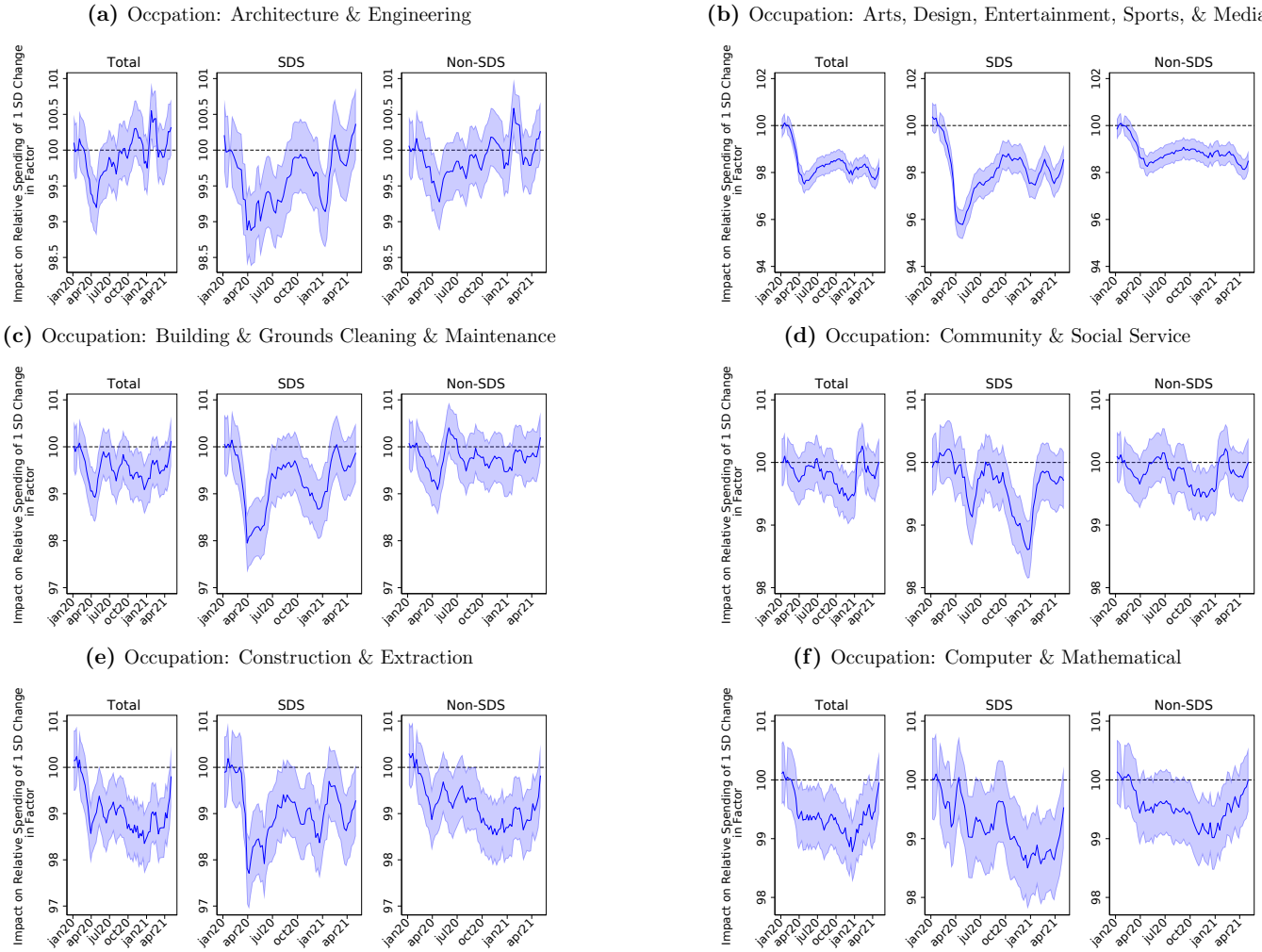


Figure 6. Additional Occupation Plots.

Source: Coefficients from authors' calculations.

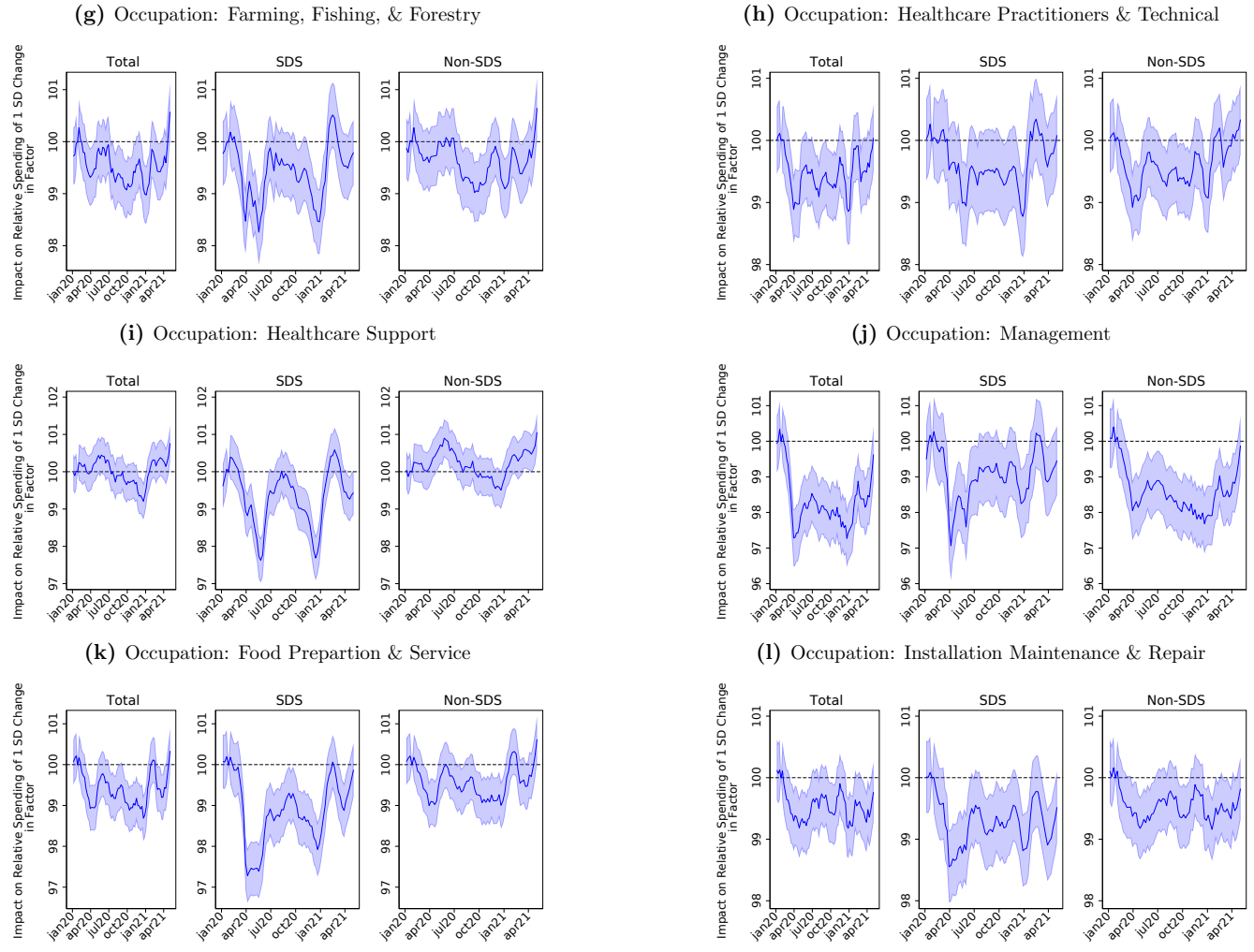


Figure 6. Additional Occupation Plots.

Source: Coefficients from authors' calculations.

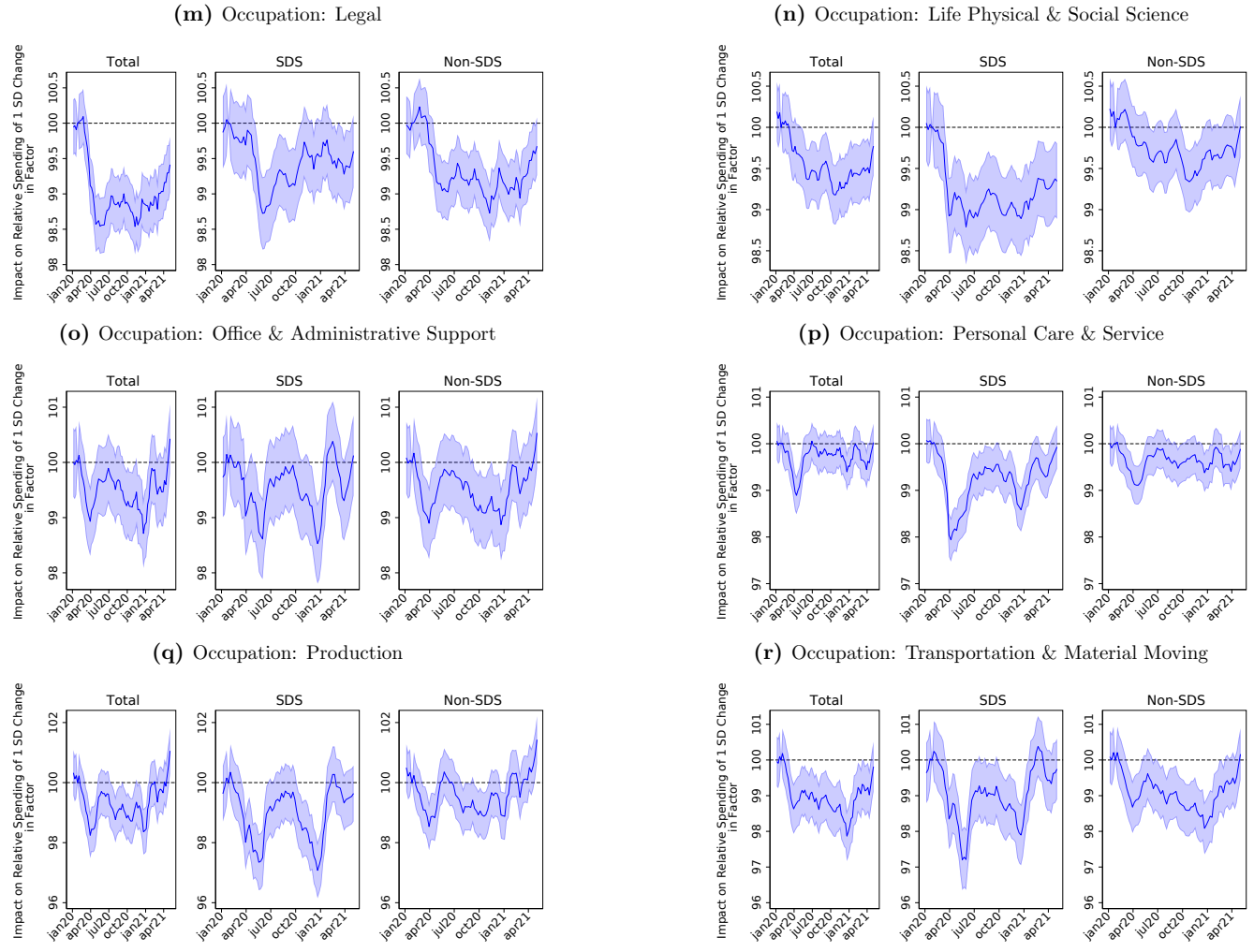


Figure 6. Additional Occupation Plots.

Source: Coefficients from authors' calculations.

B Spending Category Definitions

For each spending category we include the following Merchant Classification Codes:

Leisure: Airlines:

- All MCC Codes between 3000 and 3299, plus MCC 4511

Leisure: Car Rentals:

- All MCC Codes between 3351 and 3411, plus MCC 7512-7513

Leisure: Hotels:

- All MCC Codes between 3501 and 3900, plus MCC 7011

Leisure: Recreation/Entertainment:

- 4457: Boat Leases and Boat Rentals
- 4468: Marinas, Marine Service/Supplies
- 7032: Recreational and Sporting Camps
- 7829: Motion Picture and Video Tape Production and Distributio
- 7832: Motion Picture Theaters
- 7841: Video Entertainment Rental Stores
- 7911: Dance Halls, Schools, and Studios
- 7922: Theatrical Producers (except Motion Pictures), Ticket Agencies
- 7929: Bands, Orchestras, and Miscellaneous Entertainers—not elsewhere classified
- 7932: Pool and Billiard Establishments
- 7933: Bowling Alleys
- 7941: Athletic Fields, Commercial Sports, Professional Sports Clubs, Sports Promoters
- 7991: Tourist Attractions and Exhibits
- 7992: Golf Courses, Public
- 7993: Video Amusement Game Supplies
- 7994: Video Game Arcades/Establishments
- 7995: Gambling Transactions
- 7996: Amusement Parks, Carnivals, Circuses, Fortune Tellers
- 7997: Clubs—Country Clubs, Membership (Athletic, Recreation, Sports), Private Golf Courses
- 7998: Aquariums, Dolphinariums, Zoos, and Seaquariums
- 7999: Recreation Services—not elsewhere classified
- 7800: Government Owned Lottery (U.S. Region Only)
- 7801: Internet Gambling (U.S. Region Only)
- 7801: Internet Gambling (U.S. Region Only)

Leisure: Travel:

- 7519: Motor Home and Recreational Vehicle Rental
- 7012: Timeshares
- 7033: Campgrounds and Trailer Parks
- 4411: Cruise Lines
- 4582: Airports, Airport Terminals, Flying Fields
- 4722: Travel Agencies and Tour Operators
- 4723: Package Tour Operators – Germany Only

Leisure: Entertainment Away from Home:

- 7832: Motion Picture Theaters
- 7911: Dance Halls, Studios and Schools
- 7994: Video Game Arcades/Establishments
- 7996: Amusement Parks, Circuses, Carnivals, and Fortune Tellers
- 7998: Aquariums, Seaquariums, Dolphinariums, and Zoos
- 7922: Ticket Agencies and Theatrical Producers (Except Motion Pictures)
- 7929: Bands, Orchestras, and Miscellaneous Entertainers (Not Elsewhere Classified)
- 7932: Billiard and Pool Establishments
- 7933: Bowling Alleys
- 7941: Commercial Sports, Professional Sports Clubs, Athletic Fields, and Sports Promoters
- 7991: Tourist Attractions and Exhibits
- 7992: Public Golf Courses

Transportation:

- 4111: Transportation—Suburban and Local Commuter Passenger, including Ferries
- 4112: Passenger Railways
- 4119: Ambulance Services
- 4121: Limousines and Taxicabs
- 4131: Bus Lines
- 7523: Automobile Parking Lots and Garages
- 4784: Bridge and Road Fees, Tolls
- 4789: Transportation Services—not elsewhere classified

Food-Out:

- 5811: Caterers
- 5812: Eating Places and Restaurants
- 5813: Drinking Places (Alcoholic Beverages) – Bars, Taverns, Nightclubs, Cocktail Lounges, and Discotheques

- 5814: Fast Food Restaurants

Utilities/Telecom:

- 4812: Telecommunication Equipment and Telephone Sales
- 4814: Telecommunication Services, including Local and Long Distance Calls, Credit Card Calls, Calls Through Use of Magnetic Stripe-Reading Telephones, and Fax Services
- 4816: Computer Network/Information Services
- 4821: Telegraph Services
- 4829: Money Transfer
- 4899: Cable, Satellite and Other Pay Television/Radio/Streaming Services
- 4900: Utilities – Electric, Gas, Water, and Sanitary
- 5983: Fuel Dealers – Fuel Oil, Wood, Coal, and Liquefied Petroleum

Beauty/Massage:

- 7297: Massage Parlors
- 7298: Health and Beauty Spas
- 7230: Beauty and Barber Shops

Grocery Stores:

- 5411: Grocery Stores and Supermarkets

Retail Stores:

- 5300: Wholesale Clubs
- 5309: Duty Free Stores
- 5310: Discount Stores
- 5311: Department Stores
- 5331: Variety Stores
- 5399: Miscellaneous General Merchandise
- 5422: Freezer and Locker Meat Provisioners
- 5441: Candy, Nut, and Confectionery Stores
- 5451: Dairy Products Stores
- 5462: Bakeries
- 5499: Miscellaneous Food Stores – Convenience Stores and Specialty Markets
- 5732: Electronics Stores
- 5733: Music Stores – Musical Instruments, Pianos, and Sheet Music
- 5734: Computer Software Stores
- 5735: Record Stores
- 5912: Drug Stores and Pharmacies

- 5921: Package Stores – Beer, Wine, and Liquor
- 5931: Used Merchandise and Secondhand Stores
- 5932: Antique Shops – Sales, Repairs, and Restoration Services
- 5933: Pawn Shops
- 5935: Wrecking and Salvage Yards
- 5937: Antique Reproductions
- 5992: Florists
- 5993: Cigar Stores and Stands
- 5994: News Dealers and Newsstands
- 5995: Pet Shops, Pet Foods and Supplies Stores
- 5997: Electric Razor Stores – Sales and Service
- 5998: Tent and Awning Shops
- 5999: Miscellaneous and Specialty Retail Shops
- 5942: Book Stores
- 5943: Stationery Stores, Office and School Supply Stores
- 5944: Jewelry Stores, Watches, Clocks, and Silverware Stores
- 5945: Hobby, Toy, and Game Shops
- 5946: Camera and Photographic Supply Stores
- 5947: Gift, Card, Novelty and Souvenir Shops
- 5948: Luggage and Leather Goods Stores
- 5949: Sewing, Needlework, Fabric and Piece Goods Stores
- 5950: Glassware/Crystal Stores
- 5970: Artist's Supply and Craft Shops
- 5971: Art Dealers and Galleries
- 5972: Stamp and Coin Stores
- 5973: Religious Goods Stores
- 5975: Hearing Aids – Sales, Service, and Supply
- 5976: Orthopedic Goods – Prosthetic Devices
- 5977: Cosmetic Stores
- 5978: Typewriter Stores – Sales, Rentals, and Service
- 7296: Clothing Rental – Costumes, Uniforms, Formal Wear
- 7622: Electronics Repair Shops
- 7623: Air Conditioning and Refrigeration Repair Shops
- 7629: Electrical and Small Appliance Repair Shop

- 7631: Watch, Clock and Jewelry Repair
- 7641: Furniture – Reupholstery, Repair, and Refinishing
- 7692: Welding Services
- 7699: Miscellaneous Repair Shops and Related Services

Clothing Retail:

- 5611: Men's and Boys' Clothing and Accessories Stores
- 5621: Women's Ready-To-Wear Stores ..
- 5631: Women's Accessory and Specialty Shops
- 5641: Children's and Infants' Wear Stores
- 5651: Family Clothing Stores
- 5655: Sports and Riding Apparel Stores
- 5661: Shoe Stores
- 5681: Furriers and Fur Shops
- 5691: Men's and Women's Clothing Stores
- 5697: Tailors, Seamstresses, Mending, and Alterations
- 5698: Wig and Toupee Stores
- 5699: Miscellaneous Apparel and Accessory Shops

Contractors:

- 1520: General Contractors – Residential and Commercial
- 1711: Heating, Plumbing, and Air Conditioning Contractors
- 1731: Electrical Contractors
- 1740: Masonry, Stonework, Tile Setting, Plastering and Insulation Contractors
- 1750: Carpentry Contractors
- 1761: Roofing, Siding, and Sheet Metal Work Contractors
- 1771: Concrete Work Contractors
- 1799: Special Trade Contractors (Not Elsewhere Classified)

Delivery Services:

- 4214: Motor Freight Carriers, Trucking—Local/Long Distance, Moving and Storage Companies, Local Delivery
- 4215: Courier Services—Air and Ground, Freight Forwarders
- 4225: Public Warehousing—Farm Products, Refrigerated Goods, Household Goods Storage

Building/Hardware Stores:

- 5200: Home Supply Warehouse Stores
- 5211: Lumber and Building Materials Stores

- 5231: Glass, Paint, and Wallpaper Stores
- 5251: Hardware Stores
- 5261: Nurseries and Lawn and Garden Supply Stores
- 5039: Construction Materials (Not Elsewhere Classified)

Motor Vehicle and Parts Dealers:

- 5511: Car and Truck Dealers (New and Used) Sales, Service, Repairs, Parts, and Leasing
- 5521: Car and Truck Dealers (Used Only) Sales, Service, Repairs, Parts, and Leasing
- 5532: Automotive Tire Stores
- 5533: Automotive Parts and Accessories Stores
- 5541: Service Stations (With or without Ancillary Services)
- 5542: Automated Fuel Dispensers
- 7531: Automotive Body Repair Shops
- 7534: Tire Retreading and Repair Shops
- 7535: Automotive Paint Shops
- 7538: Automotive Service Shops (Non-Dealer)
- 7542: Car Washes
- 7549: Towing Services
- 5013: Motor Vehicle Supplies and New Parts
- 5599: Miscellaneous Automotive, Aircraft, and Farm Equipment Dealers (Not Elsewhere Classified)

Recreation Equipment:

- 5551: Boat Dealers
- 5552: Electric Vehicle Charging
- 5561: Camper, Recreational and Utility Trailer Dealers
- 5571: Motorcycle Shops and Dealers
- 5592: Motor Homes Dealers
- 5598: Snowmobile Dealers
- 5940: Bicycle Shops – Sales and Service
- 5941: Sporting Goods Stores
- 5271: Mobile Home Dealers
- 5996: Swimming Pools – Sales and Service

Furniture/Appliances:

- 5712: Furniture, Home Furnishings, and Equipment Stores, Except Appliances
- 5713: Floor Covering Stores
- 5714: Drapery, Window Covering, and Upholstery Stores

- 5718: Fireplace, Fireplace Screens and Accessories Stores
- 5719: Miscellaneous Home Furnishing Specialty Stores
- 5722: Household Appliance Stores

Direct Marketing:

- 5960: Direct Marketing – Insurance Services
- 5962: Direct Marketing – Travel-Related Arrangement Services
- 5963: Door-To-Door Sales
- 5964: Direct Marketing – Catalog Merchant
- 5965: Direct Marketing – Combination Catalog and Retail Merchant
- 5966: Direct Marketing – Outbound Telemarketing Merchant
- 5967: Direct Marketing – Inbound Teleservices Merchant
- 5968: Direct Marketing – Continuity/Subscription Merchant
- 5969: Direct Marketing – Other Direct Marketers (Not Elsewhere Classified)

Insurance Premiums:

- 6300: Insurance Sales, Underwriting, and Premiums
- 6012: Financial Institutions – Merchandise, Services, and Debt Repayment
- 6051: Non-Financial Institutions – Foreign Currency, Non-Fiat Currency (for example: Cryptocurrency), Money Orders (Not Money Transfer), Account Funding (not Stored Value Load), Travelers Cheques, and Debt Repayment
- 6211: Security Brokers/Dealers

Personal Care Services:

- 7210: Laundry, Cleaning, and Garment Services
- 7211: Laundries – Family and Commercial
- 7216: Dry Cleaners
- 7299: Miscellaneous Personal Services (Not Elsewhere Classified)
- 7251: Shoe Repair Shops, Shoe Shine Parlors, and Hat Cleaning Shops

Household Services:

- 7276: Tax Preparation Services
- 7277: Counseling Services – Debt, Marriage, and Personal
- 7278: Buying and Shopping Services and Clubs
- 7311: Advertising Services
- 7321: Consumer Credit Reporting Agencies
- 7333: Commercial Photography, Art, and Graphics
- 7338: Quick Copy, Reproduction, and Blueprinting Services
- 7339: Stenographic and Secretarial Support
- 7342: Exterminating and Disinfecting Services

- 7349: Cleaning, Maintenance, and Janitorial Services
- 7361: Employment Agencies and Temporary Help Services
- 7372: Computer Programming, Data Processing, and Integrated Systems Design Services
- 7375: Information Retrieval Services
- 7379: Computer Maintenance, Repair and Services (Not Elsewhere Classified)
- 7392: Management, Consulting, and Public Relations Services
- 7393: Detective Agencies, Protective Services, and Security Services, including Armored Cars, and Guard Dogs
- 7394: Equipment, Tool, Furniture, and Appliance Rental and Leasing
- 7395: Photofinishing Laboratories and Photo Developing
- 7399: Business Services (Not Elsewhere Classified)
- 0780: Landscaping and Horticultural Services
- 7217: Carpet and Upholstery Cleaning
- 7261: Funeral Services and Crematories
- 742: Veterinary Services
- 0763: Agricultural Co-operatives
- 2741: Miscellaneous Publishing and Printing
- 2791: Typesetting, Plate Making and Related Services
- 2842: Specialty Cleaning, Polishing and Sanitation Preparations
- 7221: Photographic Studios

Healthcare:

- 8011: Doctors and Physicians (Not Elsewhere Classified)
- 8021: Dentists and Orthodontists
- 8031: Osteopaths
- 8041: Chiropractors
- 8042: Optometrists and Ophthalmologists
- 8043: Opticians, Optical Goods, and Eyeglasses
- 8049: Podiatrists and Chiropodists
- 8050: Nursing and Personal Care Facilities
- 8062: Hospitals
- 8071: Medical and Dental Laboratories
- 8099: Medical Services and Health Practitioners (Not Elsewhere Classified)

Professional Services:

- 8911: Architectural, Engineering, and Surveying Services
- 8931: Accounting, Auditing, and Bookkeeping Services

- 8999: Professional Services (Not Elsewhere Classified)
- 8111: Legal Services and Attorneys

Education Services:

- 8211: Elementary and Secondary Schools
- 8220: Colleges, Universities, Professional Schools, and Junior Colleges
- 8241: Correspondence Schools
- 8244: Business and Secretarial Schools
- 8249: Vocational and Trade Schools
- 8299: Schools and Educational Services (Not Elsewhere Classified)
- 8351: Child Care Services

Other Services and Fees:

- 9211: Court Costs, Including Alimony and Child Support
- 9222: Fines
- 9223: Bail and Bond Payments
- 8398: Charitable Social Service Organizations
- 8641: Civic, Social, and Fraternal Associations
- 8651: Political Organizations
- 8661: Religious Organizations
- 8675: Automobile Associations
- 8699: Membership Organizations (Not Elsewhere Classified)
- 8734: Testing Laboratories (Non-Medical Testing)
- 9399: Government Services (Not Elsewhere Classified)
- 9402: Postal Services – Government Only
- 9950: Intra-Company Purchases
- 9702: Emergency Services (GCAS) (Visa use only)

Computer/Software:

- 5045: Computers and Computer Peripheral Equipment and Software

Business Supplies:

- 5021: Office and Commercial Furniture....
- 5039: Construction Materials (Not Elsewhere Classified)
- 5044: Photographic, Photocopy, Microfilm Equipment and Supplies
- 5046: Commercial Equipment (Not Elsewhere Classified)
- 5047: Medical, Dental, Ophthalmic and Hospital Equipment and Supplies
- 5051: Metal Service Centers and Offices

- 5065: Electrical Parts and Equipment
- 5072: Hardware, Equipment and Supplies
- 5074: Plumbing and Heating Equipment and Supplies
- 5085: Industrial Supplies (Not Elsewhere Classified)
- 5094: Precious Stones and Metals, Watches and Jewelry
- 5099: Durable Goods (Not Elsewhere Classified)
- 5111: Stationery, Office Supplies, Printing and Writing Paper
- 5122: Drugs, Drug Proprietaries, and Druggist Sundries
- 5131: Piece Goods, Notions, and Other Dry Good
- 5137: Men's, Women's, and Children's Uniforms and Commercial Clothing
- 5139: Commercial Footwear
- 5169: Chemicals and Allied Products (Not Elsewhere Classified)
- 5172: Petroleum and Petroleum Product
- 5192: Books, Periodicals and Newspapers
- 5193: Florists Supplies, Nursery Stock and Flowers
- 5198: Paints, Varnishes and Supplies.....
- 5199: Nondurable Goods (Not Elsewhere Classified)

Digital Goods:

- 5815: Digital Goods Media – Books, Movies, Music
- 5816: Digital Goods – Games
- 5817: Digital Goods – Applications (Excludes Games)
- 5818: Digital Goods – Large Digital Goods Merchant

Real Estate Payments:

- 6513: Real Estate Agents and Managers

The tortoise and the hare: The race between vaccine rollout and new COVID variants¹

David Turner,² Balázs Égert,³ Yvan Guillemette⁴ and
Jarmila Botev⁵

Date submitted: 17 June 2021; Date accepted: 23 June 2021

New variants of the virus are spreading, which, together with seasonal effects, are estimated to be able to raise effective reproduction numbers by up to 90%. Meanwhile, many countries are rolling out vaccination programmes, but at varying speeds. Hence the race is on to beat the variants with the vaccines. Vaccination is very powerful at reducing virus transmission: fully vaccinating 20% of the population is estimated to have the same effect as closing down public transport and all-but-essential workplaces; fully vaccinating 50% of the population would have a larger effect than simultaneously applying all forms of containment policies in their most extreme form (closure of workplaces, public transport and schools, restrictions on travel and gatherings and stay-at-home requirements). For a typical OECD country, relaxing existing containment policies would be expected to raise GDP by about 4-5%. Quick vaccination would thus help limit the extent to which containment policies need to be escalated in future epidemic waves, providing huge welfare benefits both in terms of fewer infections and stronger economic activity.

1 The authors would like to thank Laurence Boone, Luiz de Mello, Alain de Serres and colleagues from the OECD Economics Department for useful comments and suggestions. The views and opinions expressed in the paper do not reflect the official views of the OECD.

2 OECD Economics Department.

3 OECD Economics Department and EconomiX at the University of Paris X-Nanterre.

4 OECD Economics Department.

5 OECD Economics Department.

Copyright: David Turner, Balázs Égert, Yvan Guillemette and Jarmila Botev

1. Introduction

The ever-growing body of the literature has provided useful insights with regard to the effects of individual lockdown and other public health policies on the reproduction rate and economic activity. There is now much research to suggest that containment policies are associated with a reduction in infections, in COVID-19-related deaths and in reproduction numbers (Conyon et al., 2020; Gros et al., 2020) but there is disagreement as to which interventions are most effective, with different studies emphasising different policies. Containment policies are also found to depress economic activity, with workplace closure often found to have the strongest link with GDP, whereas stay-at-home requirements and school closures having weaker links (Chen et al., 2020; Demircuc-Kunt et al., 2020; Kok, 2020). The literature on the effectiveness of testing and tracing emphasises that a speedy identification of contact persons and the implementation of isolation orders are crucial (Hellewell et al. 2020). Empirical evidence also suggests that mask wearing might prevent the spread of the virus, though the magnitude of the effects is still not settled (Hatzius et al., 2020, Welsch, 2020).

The contribution of this paper to the literature is manifold. First, using OECD country experiences from the beginning of 2020 to mid-May 2021 it evaluates policy impacts for a set of policies and not for specific policies in isolation, as it done in much of the literature. Second, the paper studies the importance of vaccination. Many countries around the world are now faced with a race to rollout vaccination programmes against potential upsurges in infections caused by more transmissible COVID-19 variants. While mass vaccination should eventually win the race and bring the pandemic under control,¹ the rapid spread of new variants has the potential to cause new epidemic waves in the meantime, requiring further containment policies, which will slow or delay any economic recovery. In these circumstances, the appropriate policy responses, and their dependence on the speed of vaccine rollout. Third, the paper assesses the relative importance of differential variants of the virus as well as seasonal effects. Finally, the paper looks at the economic costs using the OECD Weekly GDP Tracker.

The empirical analysis consists of two equations estimated at high frequency: the first explains the daily evolution of the effective reproduction number, R (representing the spread of the virus), and the second explains a proxy measure of weekly GDP. The estimation results should not be interpreted as causal policy effects, as many of the explanatory policy variables might not be independent from the reproduction number. The estimations make extensive use of a set of COVID-19 policy trackers maintained by the Oxford Blavatnik School of Government (Hale et al., 2020), as well as data on the scale of vaccinations and the prevalence of different variants of the virus. t

These equations are then used to explore alternative scenarios that differ according to the speed of vaccine rollout and the infectiousness of new variants of the virus. A feature of all of the scenarios is that the number of infections is on a downward trajectory (i.e. the reproduction number remains below 1). This is because differing OECD country experiences since the start of the pandemic underline the advantage of suppressing the virus rather than pursuing a strategy of stop-and-go mitigation (Aghion et al, 2021). On that basis, the different scenarios consider whether new lockdown measures may be required to offset the effect of more transmissible variants of the virus and the extent to which this can be avoided by a faster vaccination rollout.

The remainder of the paper is organised as follows. Section 2 provides stylised facts. Section 3 describes the data and the estimation method for the reproduction number and the proxy measure of GDP, as well as the policies and other explanatory variables that influence them in the estimation framework. In section

¹ This assumes that a new variant does not prove to be resistant to existing vaccines, although were this to be the case, the hope is that existing vaccines could be modified relatively quickly.

4, estimation results for the drivers of the reproduction rate are presented for OECD countries, because sensitivity analysis suggests that responses from a wider set of countries are very different.² Section 5 reports the determinants of weekly GDP. In section 6, the estimation results are used to construct a number of stylised scenarios incorporating the prevalence of more transmissible variants, to consider the required setting of containment policies and how this varies with the proportion of the population that are vaccinated. Section 7 summarises the results.

2. Stylised facts

The median R estimate for OECD countries fell rapidly from around 2¼ in early March 2020 to around 0.75 the next month and has since fluctuated in the 0.75 to 1.50 range (Figure 1, panel A). Reasons behind these fluctuations include seasonal effects, the probably premature lifting of restrictions in the summer of 2020 and the delayed and muted policy reaction to the surge of the virus in the Autumn of 2020. Another reason is the ebb and flow of mutations in the virus. For instance, the emergence of the so-called UK or Alpha variant (B.1.1.7) in the fall of 2020, first in the United Kingdom and then in many other countries, was a major factor behind the increase in R in the first quarter of 2021 (Figure 1, panel B).

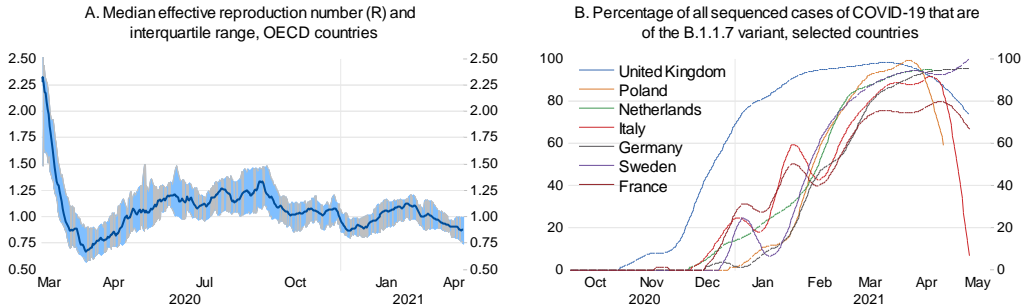
The OECD Weekly Tracker of GDP growth, developed by OECD (2020b) and Woloszko (2020a and 2020b)³ shows a large drop and quick recovery in 2020. The median weekly GDP tracker for OECD countries also shows that output has been about 5% below the no-pandemic counterfactual in late 2020 and much of 2021 (Figure 2).

The question that needs to be raised is how the typical policy response with regard to containment measures has evolved since the start of the pandemic. A cursory examination would suggest that among OECD countries, more recent responses have been characterised by less resort to the most stringent implementation of school closures, workplace closures, international travel restrictions and stay-at-home requirements, but more stringent restrictions on gatherings. Different waves of the virus have, however, hit countries at different times, so in order to define a typical policy response, a focus here is on all OECD European countries on the assumption that the timing may be more similar. Typically, policy response during the first wave contained hard lockdown measures such as stay-at-home requirements and workplace closures, whereas policies have been shifted towards less binding measures including mandatory mask wearing and to the scaling up test, trace and isolate policies and the roll-out of vaccination programmes.

² For example, many of the main containment measures (such as workplace closures and stay-at-home requirements) have much weaker effects on the reproduction number in non-OECD countries than OECD countries, perhaps reflecting weaker social security nets and/or more limited possibilities to work remotely.

³ applies a machine learning model to a panel of Google Trends data, aggregating together information about search behaviour related to consumption, labour markets, housing, trade, industrial activity and economic uncertainty to provide estimates of the growth rate of weekly GDP compared to a year earlier.

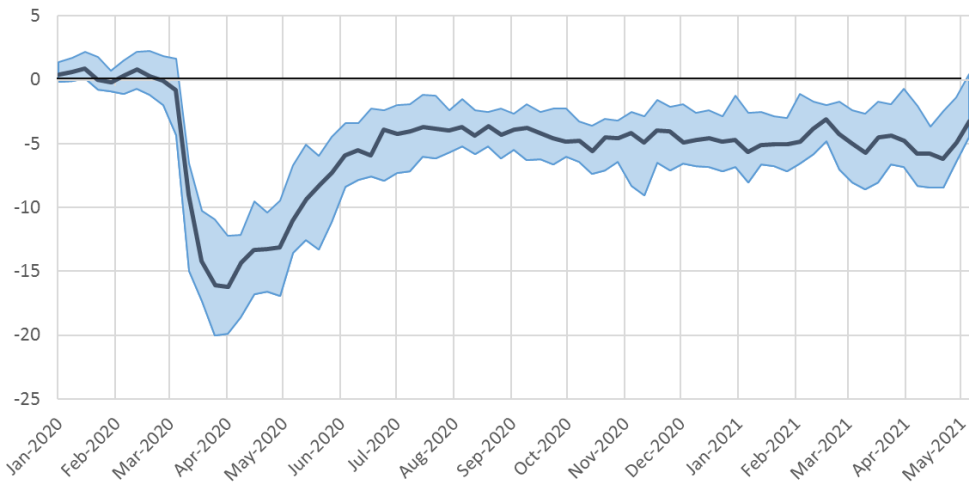
Figure 1. Reproduction number and emergence of the B.1.1.7 variant



Source: Arroyo Marioli, F. et al. (2020), "Tracking R of COVID-19: A New Real-Time Estimation Using the Kalman Filter"; Hodcroft, E.B. (2021), "CoVariants: SARS-CoV-2 Mutations and Variants of Interest" and authors' calculations.

Figure 2. A weekly measure of OECD GDP

Median and interquartile range, % difference from a no-pandemic counterfactual



Note: The chart shows the difference between the proxy measure of weekly GDP and a no-pandemic counterfactual projection provided by OECD Economic Outlook No. 106 (November 2019).

Source: OECD Weekly Tracker site: <http://www.oecd.org/economy/weekly-tracker-of-gdp-growth/>

3. Data and estimation issues

The empirical analysis relies on a set of variables representing containment measures maintained by the Oxford Blavatnik School of Government (Hale et al., 2020), which in their original form are scored according to the degree of stringency or comprehensiveness with which they are applied. Eight categories of containment measures are distinguished, being variously scored from 0 to 2, 0 to 3, or 0 to 4 (Table 1). The scoring of measures refers to their design, not necessarily how they have been applied. This

Covid Economics 83, 2 July 2021: 154-182

represents a potential weakness as the variables do not capture how enforcement and abidance by regulations has varied across countries.

For the purposes of estimation, the cardinal value of these scores are ignored (as there is no reason, for example, to expect a policy with a stringency value of 3 to have treble the effect of a policy with a value of 1) and instead the same policy at different levels of stringency are included as distinct dummy variables (taking the value of zero or one). Subsequently, if the estimation does not deliver the expected ordinal ranking in coefficients (so that a more stringent application of a policy has a greater effect), the same coefficient may be imposed across different levels of stringency by combining policy variables.⁴ For school closures, an indicator produced by UNESCO and measuring whether schools and universities are partially or completely closed due to COVID-19 is used in the empirical analysis instead of the one by the Oxford Blavatnik School of Government.

Table 1. Definition of different stringency levels for containment policies

Containment measure	Scoring of degree of stringency
School closures	1: Partial closure due to COVID 2: Complete closure due to COVID
Workplace closures	1: Recommend closing (or work from home) 2: Require closing (or work from home) for some sectors or categories of workers 3: Require closing (or work from home) all-but-essential workplaces (e.g. grocery stores, doctors)
Cancel public events	1: Recommend cancelling 2: Require cancelling
Restrictions on gatherings	1: Restrictions on very large gatherings (above 1000 people) 2: Restrictions on gatherings between 101-1000 people 3: Restrictions on gatherings between 11-100 people 4: Restrictions on gatherings of 10 people or less
Close public transport	1: Recommend closing (or significantly reduce volume/route/means of transport available) 2: Require closing (or prohibit most citizens from using it)
Stay at home requirements	1: Recommend not leaving house 2: Require not leaving house with exceptions for daily exercise, grocery shopping, and 'essential' trips 3: Require not leaving house with minimal exceptions (e.g. only once a week, or one person at a time)
Restrictions on internal movement	1: Recommend not to travel between regions/cities 2: Internal movement restrictions in place
International travel controls	1: Screening 2: Quarantine arrivals from high-risk regions 3: Ban on arrivals from some regions 4: Ban on all regions or total border closure

Note: Not shown in the table, but "No measures" or "No restrictions" are always scored 0.

Source: Oxford COVID-19 Government Response Tracker described in Hale et al. (2020) and UNESCO's [School closures caused by Coronavirus \(Covid-19\) \(unesco.org\)](https://unesco.org).

The Oxford COVID-19 Government Response Tracker includes the severity of restrictions on visitors to LTCFs, with the most stringent classification reserved for when all non-essential visitors are prohibited (Table 2). In this paper's empirical framework, facial coverings are investigated using the Oxford COVID-19 Government Response Tracker indicators, which denote the strength of government mandates in this regard. The empirical work will investigate the effect of the protection of the elderly.

To capture the effect of test and trace policies, the policy indicators from the Blavatnik School of Government at the University of Oxford (Hale et al., 2020) are used (Table 2). Again, for the purpose of

⁴ Multicollinearity, implying problems in being able to separately identify coefficients on different explanatory variables, is a concern given that containment measures were often introduced simultaneously or very close together. Earlier work, estimating similar equations over a shorter sample to mid-2020 (Égert et al. 2020), suggested there was collinearity between some of the main containment policies associated with lockdown, however this is much less apparent over a longer sample period.

the estimation work, the cardinal values of the trackers are ignored and instead different dummy variables are used to represent test-and-trace variables at different degrees of comprehensiveness. Many aspects of testing and tracing may be easier at low infection levels and this can be readily tested in the empirical framework. However, an important limitation of these indicators is that they do not capture timing or speed, which can be key to a successful strategy: test results must be returned with minimum delay and then contacts need to be traced quickly. A further important limitation of these indicators is there is no information on the extent to which test and trace polices are followed up with measures to isolate those infected or potentially infected.

Table 2. Public health policies against COVID-19

Protection of elderly people
1: Recommendations regarding isolation, hygiene and visitor restrictions in Long-term Care Facilities (LTCFs) and/or elderly to stay at home.
2: Narrow restrictions for isolation, hygiene and visitor restrictions in LTCFs and/or restrictions protecting elderly people at home.
3: Extensive restrictions for isolation, hygiene in LTCFs, all non-essential external visitors prohibited, and/or elderly required to stay at home with minimal exceptions and no visitors.
Testing policy¹
1: Only those who both (a) have symptoms AND (b) meet specific criteria (e.g. key workers, admitted to hospital, came into contact with a known case, returned from overseas)
2: Testing of anyone showing COVID-19 symptoms
3: Open public testing (e.g. "drive through" testing available to asymptomatic people).
Contact tracing
1: Limited contact tracing - not done for all cases
2: Comprehensive contact tracing - done for all identified cases.
Facial coverings
1: Recommend
2: Required in some specified shared/public spaces or some situations when social distancing not possible
3: Required in all specified shared/public spaces or all situations when social distancing not possible
4: Required outside the home at all times

Note: Not shown in the table, but "No measures" or "No restrictions" are scored 0.

(1) Testing variable relates to policies testing for infection (PCR test), not to policies testing for immunity (antibody tests).

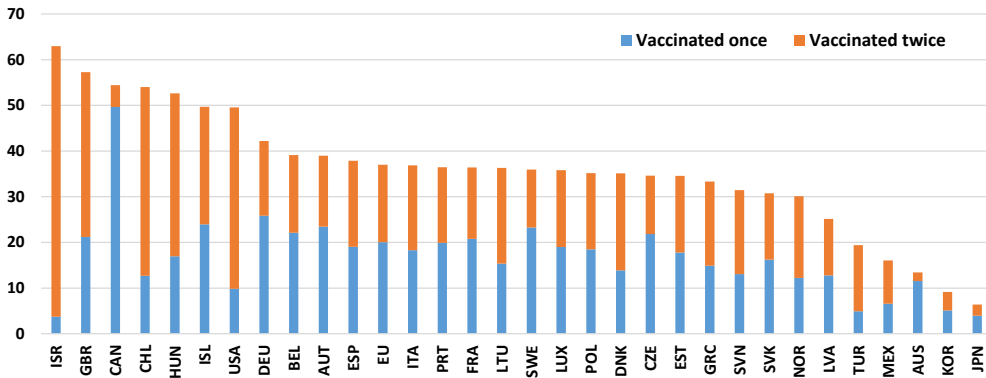
Source: Oxford COVID-19 Government Response Tracker described in Hale et al. (2020).

Data on vaccination rates are taken from [Our World in Data](#). The estimation results do not differentiate between the different vaccines, although there is some evidence to suggest that some vaccines -- such as the SINOVAQ vaccine which, among OECD countries, has been approved in Chile, Turkey, Mexico and Colombia -- has lower efficacy than other vaccines commonly used in OECD countries, especially after only one shot (Figure 3).⁵

⁵ The speed with which vaccination programmes are being rolled out varies greatly across OECD countries. By end-May, there were five countries where more than 50% of the population had received at least one vaccine done, but also five countries where this share had not reached 20%.

Figure 3. Vaccination rates vary greatly across OECD countries

Per cent of population vaccinated as of end-May 2021



Source: Official data collated by [Our World in Data](#). Some OECD countries are missing because the data do not allow splitting vaccination status between one and two doses.

The empirical work considers mutants of the virus that have had widespread prevalence in the estimation sample as well as the recent Indian variant. Data on the prevalence of different variants are from the GISAID database via the website [Covariants.org](#). The series measure the proportions of all virus sequences (not cases) submitted to GISAID classified as a specific variant. The names for the variants refer to the [Pango Lineage nomenclature](#). B.1.1.7 is commonly referred to as the Alpha or UK variant. Other variants considered (B.1.258, B.1.160 and B.1.177) have all been highly prevalent in at least a few OECD countries. The more recent Delta or Indian variant (B.1.617) is also considered

4. Explaining the reproduction number

Regression analysis is used to relate country-level daily reproduction numbers to several potential explanatory factors, including containment policies, public-health policies, seasonal conditions, the prevalence of variants, vaccination rates as well as proxies for spontaneous behavioural changes and natural immunity. In addition, the estimated equation includes a full set of country fixed effects to account for fixed country characteristics (such as population density, general social habits, etc.) that could affect virus transmission. This means that identification relies on variation of reproduction numbers and other variables in the time dimension.

Unlike many other studies that focus on a single explanatory factor, the approach used here has the advantage of considering all of the above factors at once within the same framework and using a large amount of daily data from several countries. Nevertheless, the approach is not able to identify causal policy effects, as many of the explanatory variables are not independent from the reproduction number. For instance, containment and other policies are taken in response to the evolution of the pandemic. Identifying pure causal effects would require exogenous sources of policy variation (e.g. natural experiments) or the use of instrumental variables, if good instruments could be found.

An important feature of the estimated equation explaining R is that the preferred functional form for the dependent variable is logarithmic; a formal test decisively rejects a linear form in favour of a logarithmic

one.⁶ This implies that any policy intervention will have a larger effect when R is initially high than when it is low.

4.1. Containment policies

4.1.1. Literature overview

Existing research suggests that containment policies are associated with a reduction in infections, in COVID-19-related deaths and in reproduction numbers (Conyon et al., 2020; Gros et al., 2020). However, there is disagreement as to which interventions are most effective. Some research suggests that stay-at-home requirements are most effective in curbing the propagation of the virus (Esra et al., 2020; Gapen et al., 2020; Spiegel and Tookes, 2020; Chernozhukov et al., 2020). A recent survey concludes that a majority of papers find significant effects of school closure on the spread of the virus (Walsh et al., 2020). School closure is found to be particularly effective by Liu et al. (2020) and Olczak et al. (2020). Chernozhukov et al. (2020) argue the effect is uncertain, whilst Spiegel and Tookes (2020) find no effect. Some studies find that limiting potential 'super-spreader' events -- such as NBA and NHL games in the United States, football matches in England or gatherings more generally -- are a very effective way of reducing the spread of the virus (Ahammer et al., 2020; Chen et al., 2020; Hunter et al., 2020; Olczak et al., 2020; Weber, 2020; Li et al., 2020). Other studies emphasise the importance of travel restrictions, particularly on international flights and at the early stages of the pandemic (Hubert, 2020; Keita, 2020, Leffler et al., 2020).

4.1.2. Estimation results

In the estimation, the coefficients on six containment policies -- workplace and school closures, restrictions on gatherings, stay-at-home requirements, international travel controls and closures of the public transport system -- are found to have a statistically significant effect in reducing R (

⁶ Testing the appropriate functional form of the dependent variable is not as straightforward as testing for the functional form of explanatory variables because the competing models cannot be nested within a general model. The test is therefore conducted by first transforming the dependent variable by dividing by its geometric mean to make the two competing models (log and linear) comparable. A formal test of model equivalence can be performed with the BoxCox statistic by comparing the relative goodness-of-fit of the two models. The test decisively favours the logarithmic form for R over different country samples. Conversely, when a similar test is carried out for the weekly GDP equation, the linear form is decisively preferred to the logarithmic form.

Table 3, Figure 4).

- The largest effects come from the most stringent application of workplace closures, which reduce logged R by 0.27 (which from an initial R number of 1.5, implies an absolute fall in R of 0.35). Resort to such stringent policies have, however, been very infrequent. Also rare is the accompaniment of workplace closures with a complete closure of the public transport system, which in combination reduces logged R by 0.37.
- Tight restrictions on gatherings and stay-at-home requirements have been widely applied across most OECD countries and are estimated to reduce logged R by up to 0.15 and 0.13, respectively.
- Complete closure of schools, which OECD countries have been more reluctant to impose since their widespread imposition in dealing with the first wave of the virus in the first half of 2020, reduce logged R by 0.13.
- International travel restrictions are estimated to reduce logged R by up to 0.12, although this may underestimate their importance in preventing the spread of new variants of the virus and in sustaining a situation in which the national rate of infections is much lower than in other countries.

The estimated equation suggests that tightening the stringency of policies increases their effectiveness, although not always in a linear fashion (Figure 4). The finding of stronger effects from more stringent application of containment policies contradicts the findings of Bendavid et al. (2021), who conclude that more restrictive containment policies do not have a significantly greater impact on the growth of infections.

Table 3. The drivers of the effective reproduction number, OECD countries

Sample period: 15 January 2020 to 16 May 2021

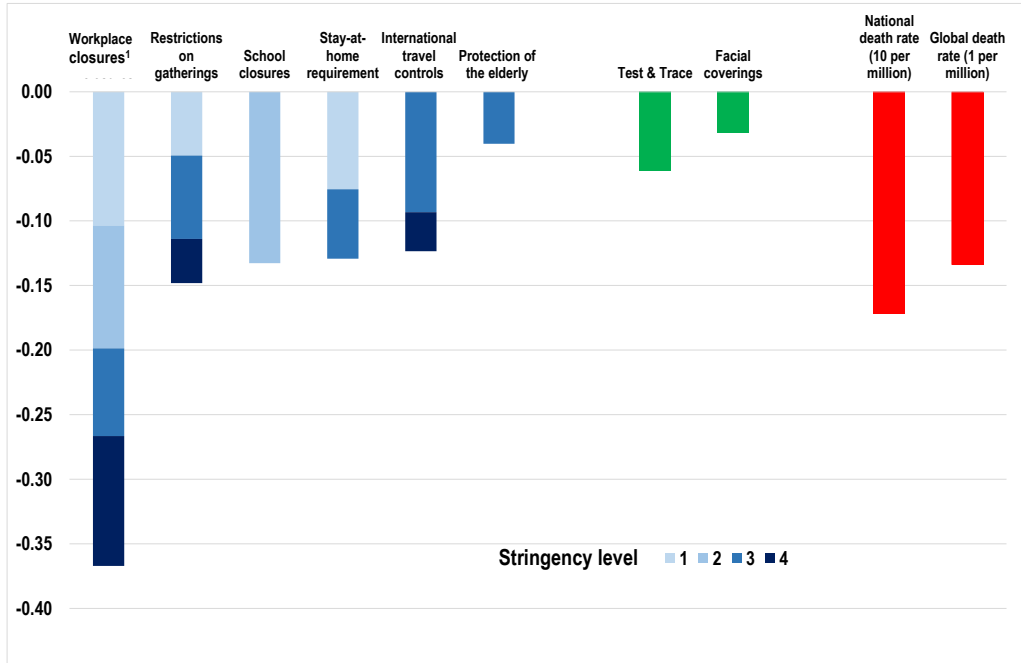
Dependent variable: ln(R)	
Constant	0.6999**
<i>Containment policies</i>	
Stay-at-home requirement (=1 & 2)	-0.0755**
Stay-at-home requirement (=3)	-0.1293**
Workplace closures (=1)	-0.1036**
Workplace closures (=2)	-0.1987**
Workplace closures (=3)	-0.2666**
Closure of public transport(=2) & workplaces (=3)	-0.1005**
School closures (=2)	-0.1327**
Restrictions on gatherings (=1 & 2)	-0.0494**
Restrictions on gatherings (=3)	-0.1139**
Restrictions on gatherings (=4)	-0.1481**
International travel controls (=3)	-0.0932**
International travel controls (=4)	-0.1235**
<i>Test and Trace policies</i>	
All Test & Trace combinations	0.0217
All Test & Trace combinations when deaths < 10 per million	-0.0612**
<i>Other non-containment policies</i>	
Facial coverings (=3 & 4)	-0.0157**
Protection of the elderly (=3)	-0.0403**
<i>Seasonal effect</i>	
Hours of daylight	-0.0203**
<i>Death rates (per million population)</i>	
Daily national death rate	-0.0172**
Daily global death rate	-0.1338**
Total national death rate	-0.0001**
<i>Vaccination rates (% of population)</i>	
Vaccinated once(-28)	-0.0054**
Vaccinated twice(-28)	-0.0181**
<i>Prevalence of virus variants (% of cases sequenced)</i>	
B.1.177	0.0014**
B.1.160	0.0009**
B.1.258	0.0043**
B.1.1.7 ('UK variant')	0.0035**
B.1.617 ('Indian variant')	0.0006
Adjusted R-squared	0.476
Daily observations	14167
Countries covered	35

Note: The policy variables are based on the variables described in Table 1 and Table 2, but re-normalised to be (0, 1) dummy variables as described in the main text. The notation in brackets "(=n)" after a containment policy variable denotes that the dummy variable is assigned a 1 if the original score for that policy was equal to n, whereas the notation "(>=n)" denotes that the dummy variable is assigned a 1 if the original score for that policy was greater than or equal to n. **** denotes statistical significance at the 5% level, based on heteroscedasticity-robust standard errors. The equation includes country fixed effects, which are not reported here. The dependent variable (R) has a lead of 12 days relative to the explanatory variables, to account for the delay between infection and case detection.

Source: Authors' calculations.

Figure 4. Estimated effects of policies and natural caution on logged R

A. Effect on logged R



Note: The chart shows the estimated effect on logged R of different containment policies at varying degrees of stringency, the effect of policies relating to test-and-trace and facial coverings as well as the effect of increased natural caution, proxied by the daily death rate. If a particular degree of policy stringency is not shown in the bar chart, then it either means there is no significant effect on GDP (this being the case when bar segments at a higher level of stringency are shown), or that it has the same effect as the policy at the previous level of stringency (when bar segments at a lower level of stringency are shown). The effect of the test-and-trace and facial coverings are the maximum effect, which in the former case depend on a low level of infections.

1) Workplace closures have a maximum stringency level of 3 according to the Oxford taxonomy, but are shown here as having a maximum stringency level of 4 when combined with a complete closure of public transport. Full closure of public transport is not shown separately because it has only rarely been applied and then in combination with full workplace closure and this is also reflected in the estimation.

Source: Authors' calculations.

4.2. Public health policies

4.2.1. Literature overview

The effectiveness of testing and tracing crucially depends on correctly and speedily identifying contact persons and on compliance with isolation orders (Hellewell et al. 2020). Empirical evidence, exploiting an administrative mistake in contact tracing in England, indicates potentially massive effects from timely contact tracing (Fetzer and Graeber, 2020). Testing and tracing appears to work well in countries with good social security and in particular high coverage of paid sick leave in case isolation is needed. This effect has been shown for earlier influenza episodes as well (Pichler et al., 2020). It has been argued that

Covid Economics 83, 2 July 2021: 154-182

repeated mass testing coupled with the isolation of the infected could eradicate the virus without reliance on contact tracing (Taipale, Romer and Linnarsson, 2020).⁷

A recent review of about 50 studies on the effectiveness of facial coverings in the community concluded that they have a “*small to moderate effect*” in preventing the spread of COVID-19 (ECDC, 2021). Measuring the health effect of facial coverings in multi-country empirical studies is difficult, because they depend on the take-up rate and it is difficult to account for the greater tendency to wear face masks, regardless of policy pronouncements, in some (mostly Asian) countries prior to the current virus outbreak. Nevertheless, some cross-country studies have identified important negative effects on mortality (Esra et al., 2020, Leffler et al., 2020, Hatzius et al., 2020) and the reproduction number (Égert et al., 2020). Beneficial effects of mask wearing on mortality are also reported at the regional level in the United States (Hatzius et al., 2020, Welsch, 2020; Spiegel and Tookes, 2020) and Germany (Mitze et al., 2020).

The elderly population is especially vulnerable to COVID-19 with much higher mortality rates than other demographic groups. A particular concern is that at the early stages of the pandemic, mortality rates were alarmingly high in long-term care facilities (LTCFs) in some OECD countries (ECDC, 2020b; Gandal et al. 2020). Evidence from influenza outbreaks as well as the current pandemic suggests that improving hygiene, including regular hand sanitising and disinfection at the establishment level (Koshkouei et al., 2020), as well as limiting the migration of staff across different care homes, help to reduce infection rates substantially (Koshkouei et al., 2020; Chen et al., 2020), although the latter may be difficult given widespread staff shortages. Recent work at the country level also underlines the importance of protecting the elderly (Égert et al. 2020).

4.2.2. Estimation results

The implementation of test and trace policies is found to reduce R, but the effects are relatively small. It is also difficult to distinguish the size of effects according to the comprehensiveness of the policy, so in the estimated specification the same coefficient applies to all combinations of test and trace stringencies. This result may be attributable to differences in how quickly and successfully test and trace policies are applied, as explained previously. All test and trace policies are found to be much more effective when the infection rate is low (which in this estimation is taken to be less than ten new daily cases per million population, a rate which many countries had well exceeded already in March and April 2020). This is a rather unsurprising finding, given the difficulties of tracking down all contact persons in a timely manner if the system is overwhelmed with new cases.

Recommendations to make facial coverings mandatory in public are found to reduce the spread of the virus, but the estimated effects are small, lowering logged R by less than 0.02. This result -- possibly explained by inadequate compliance, improper mask wearing and the fact that many infections take place in family settings where masks are impractical -- is consistent with the recent review of about 50 studies, referred to above, that concluded the use of facial coverings in the community have a “*small to moderate effect*” in preventing the spread of COVID-19 (ECDC, 2021).

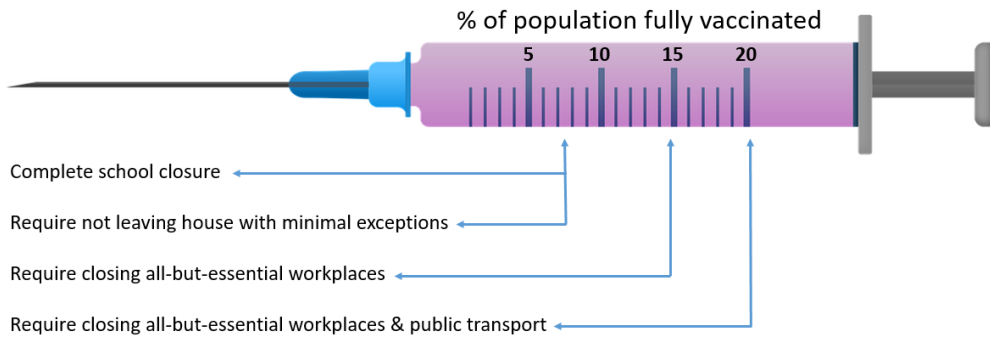
The most stringent interventions to protect the elderly, involving prohibiting all non-essential visitors from LTCFs, are found to reduce the spread of the virus, but the effects are relatively small, reducing logged R by about 0.04. However, this understates the importance of such measures given that the elderly are much more vulnerable to serious illness once infected.

⁷ China has demonstrated the potential for mass testing in the city of Wuhan in May 2020, with up to 1.5 million tests being processed in a single day (New York Times, 2020). Mass antigen testing has been considered a success in Slovakia, where the reproduction number declined by 0.3. Nevertheless, such a strategy can work in the longer term only if mass testing is repeated at regular intervals (Kahanec et al., 2021).

Vaccination is found to reduce the reproduction number substantially. The percentage of the population having received either one or two doses of vaccine enters the equation with a lag of four weeks, to account for the gradual build-up of immunity after vaccination. The estimated coefficients imply that each percentage point of the total population that is fully vaccinated lowers logged R by 0.018. Hence, fully vaccinating only 7% of the population reduces logged R by the same amount as closing down schools and fully vaccinating 20% of the population is as powerful as closing all-but-essential workplaces and public transport (Figure 5). Once 60% of the population has received two doses, as in Israel at the end of May 2021, then logged R is reduced by about 1.1. The results imply that two doses are much more effective at reducing transmission than a single dose, but the second-dose effect may also include some delayed immunity from the first dose.

Figure 5. The powerful effect of vaccination

Equivalency between estimated effects of selected containment policies and percentage of population fully vaccinated



Note: The figure is based on the coefficient estimates reported in Table 3.

4.3. Non-policy effects: Awareness of the virus and transition towards herd immunity

Voluntary social distancing and increased personal hygiene measures, such as increased handwashing, are important in containing the spread of the virus (Caselli et al., 2020). In some cases, voluntary behavioural changes preceded the implementation of containment measures (Audirac et al., 2020) and in some cases their effects might even have been superior to those of official containment policies (Berlemann and Haustein, 2020; Goolsbee and Syverson, 2020; Gupta et al., 2021).

In order to capture voluntary behavioural responses, the regression analysis includes different measures of the death rate from the virus (national and global daily deaths per million population) as proxies for general awareness of the pandemic prompting more cautious behaviour. The importance of these variables is that they substitute for changes in behaviour that are likely to be engendered regardless of government-mandated restrictions.

Total national deaths attributed to the virus, expressed as a share of the population, are also separately included as a proxy for the share of the population that has previously been infected, with the expectation of a negative coefficient: as the share of the population that has been infected rises (and presumably becomes immune), the speed with which the virus spreads should be reduced.

The death rate variables are statistically significant with the expected negative sign and their magnitudes imply they can play an important role in the evolution of R.

Covid Economics 83, 2 July 2021: 154-182

- The national daily death rate varies substantially, both across countries and over time, but for some OECD countries it was running at well over 10 per million going into the first lockdown in March 2020 and has exceeded this threshold in some countries in recent months. At such levels, the estimated coefficient implies precautionary behaviour to such an extent that that the reduction in R is larger than most of the containment policies implemented in isolation (Figure 4).
- The global daily death rate has exceeded 1 per million since November 2020, and the estimated coefficient implies this would further supplement this precautionary effect by reducing logged R by about 0.13.
- The coefficient on the total national death rate (i.e. based on cumulative deaths) is surprisingly small, implying that even in those OECD countries which have experienced the highest death rates of over 1 000 per million population, logged R would only be reduced by about 0.1.

4.4. Non-policy effects: Seasonal conditions

The global nature of the pandemic indicates that the virus can spread all over the world irrespective of seasonal conditions. Review papers trying to distil the several hundreds of studies analysing the potential seasonality of Coronavirus transmission struggle to come to a clear conclusion.⁸ Some evidence suggests that increased UV light exposure bears a stronger relation to transmission than temperature or humidity (Merow and Urban, 2020).⁹ For the purposes of this paper, seasonal effects (whether due to weather, behavioural changes or other factors) are proxied by a measure of daylight hours. For the median OECD country, the number of daylight hours varies by around seven hours between winter and summer and by up to 13 hours in some Nordic countries such as Finland, Norway or Sweden. The estimated coefficient implies that this would produce a seasonal variation in logged R of between 0.15 and 0.25. Climatic variables, including temperature, were found to generate similar seasonal variation, but multicollinearity between the different measures precludes the inclusion of more than one such variable.

4.5. Prevalence of virus variants

Relative to the original strain of the virus, to which the estimated constant assigns a basic reproduction number (R0) of about 2.0, four variants are found to increase the effective reproduction number substantially.

- The [B.1.177](#) variant initially expanded in Spain and spread widely across Europe via holiday travel (Hodcroft et al., 2021). In the summer of 2020, it became the most prevalent variant in western Europe.
- The [B.1.160](#) variant became the second most dominant variant in western Europe over the northern-hemisphere summer of 2020, after variant B.1.177.
- The [B.1.258](#) variant first circulated widely in Ireland in August 2020 and later on in central Europe, particularly the Czech Republic and Slovenia.
- The [B.1.1.7](#) variant, commonly referred to as the Alpha or UK variant because it appears to have arisen and/or initially expanded in the South East of England, became the dominant variant in many European countries during the first quarter of 2021. It has since been found throughout the world, notably the United States where it was the most prevalent variant according to the latest data.

⁸ Papers summarised in Mecenas et al. (2020) connect warm and wet weather with reduced virus transmission, the papers reviewed in Briz-Redon and Serrano-Aroca (2020), Shakil et al. (2020) and Paraskevis et al. (2021) provide more mixed evidence.

⁹ For European countries, Walrand (2021) identifies UV light as the only climatic factor explaining COVID-19 outbreaks. UV light exposure is strongly correlated with daylight hours, which determines time spent indoors and the emergence of vitamin D deficiency, which is thought to amplify the severity of COVID-19 (Weir et al., 2020).

Each of these variants is estimated to raise the logged reproduction number by between 0.01 and 0.04 for each percentage point of all cases sequenced attributable to it. If prevalence of the B.1.1.7 variant became 100%, for instance, then the effective reproduction number would be 2.9 instead of 2.0, a result consistent with Davies et al. (2021), who find that the Alpha variant is 43% to 90% more transmissible than the original strain of the virus. The [B.1.617](#) (aka Indian) variant is included in the estimated equation but its impact on the reproduction number is not statistically significant, most likely because prevalence was still very low in most OECD countries even toward the end of the estimation period.

5. Explaining weekly GDP over the pandemic period

5.1. Literature overview

Containment policies are also found to depress economic activity, for example König and Winkler (2020) and Gros et al. (2021) find a strong link between the overall stringency of containment policies and quarterly GDP growth rates. In terms of specific policies, workplace closure is often found to have the strongest link with GDP, with stay-at-home requirements and school closures having weaker links (Chen et al., 2020; Demirguc-Kunt et al., 2020; Kok, 2020). Other containment policies are less often singled out, although it is clear that international travel restrictions have more severely impacted those countries that are more reliant on travel and tourism. The GDP costs of containment measures must be understood in a short-run/immediate sense, because from a longer-run perspective, there may be no cost to containment policies if they prevent an epidemic from spreading, which would lead to larger GDP costs later on via either spontaneous changes in behaviour or the need for harsher containment measures. A recent study that attempts to take epidemiological dynamics into account indeed finds little output costs to additional containment policies that reduce deaths (Arias et al., 2021).

5.2. Estimation results

Estimation results suggest that six of the eight categories of containment policies have a negative short-run effect on GDP and a fiscal stimulus variable is found to have a significant positive effect (Table 4). Sensitivity testing suggests that there has been some decline in the (negative) GDP effects of containment policies over time, distinguishing between the first half of 2020 and the period since then.¹⁰ These differences are most significant for restrictions on gatherings and on internal movement (see the second column of Table 4).

Focussing on the effect over the last year of the most stringent application of each of the containment policies suggests a clear ranking of policies in terms of their adverse effect on GDP, with workplace closures having the largest effect on GDP (particularly in the extreme case of being accompanied by the closure of public transport), followed by stay-at-home requirements, with school closures, restrictions on gatherings and internal movement all having a lesser, but significant, cost to GDP (Figure 6).

¹⁰ The sensitivity of the estimation results to different sample periods is examined by maintaining a preferred specification estimated over the full sample period as a reference point, but then re-estimating it adding a time dummy (taking the value of 1 after the start of July 2020) interacted with each of the containment policy variables. Most of the interaction terms are found to be positive, although only those that are also statistically significant are reported in the second column of Table 4. A revised (and reduced) policy effect can then be computed over the second half of the sample by summing the coefficients estimated over the full sample with the corresponding coefficient interacted with the period dummy variable. When a similar procedure was followed for the R equation, it was difficult to discern any meaningful systematic pattern in the interaction coefficients, which were more often small and statistically insignificant. So, for this reason, it was decided to retain the R equation estimated over the full sample, not least to minimise problems of multicollinearity.

Table 4. The determinants of weekly GDP

Sample period: Mid-January 2020 to mid-May 2021

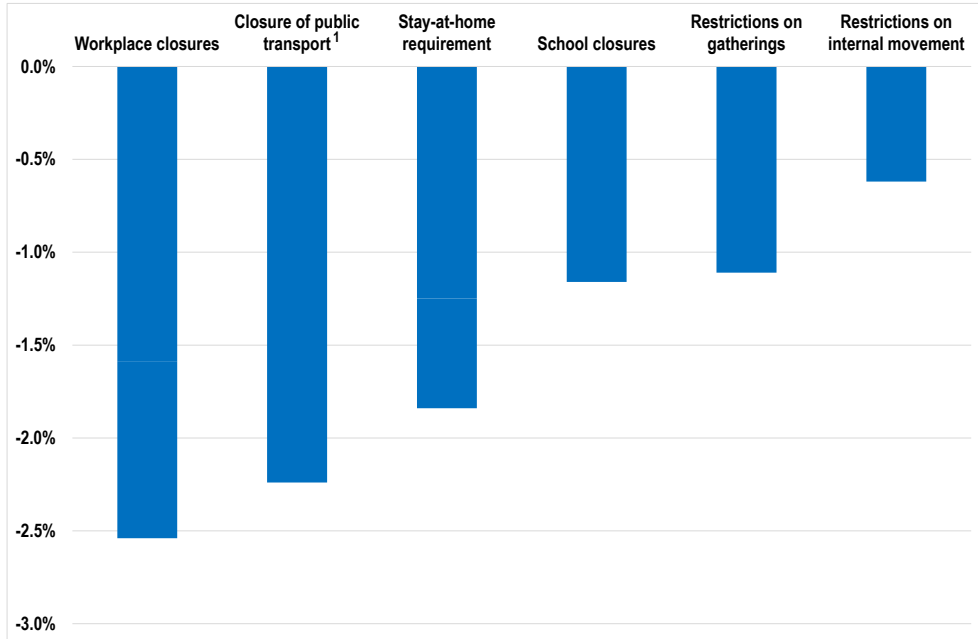
	Over entire period	Interacted with dummy from July
Dependent variable: GDP (% difference from counterfactual)		
Constant	-0.0109**	
<i>Containment policies</i>		
Stay-at-home requirement (=1)	-0.0125**	
Stay-at-home requirement (>=2)	-0.0184**	
Workplace closures (=1&2)	-0.0159**	
Workplace closures (=3)	-0.0254**	
School closures (=2)	-0.0116**	
Restrictions on gatherings (>=3)	-0.0257**	0.0146**
Restrictions on internal movement (=1)	-0.0231**	0.0169**
Restrictions on internal movement (=2)	-0.0425**	0.0381**
Closure of public transport (=2) * Workplace closure (=3)	-0.0224**	
<i>Fiscal policy controls</i>		
Structural budget balance (% of GDP)	-0.0015**	
<i>Death rates (per million population)</i>		
Daily national death rate	-0.0075**	0.0079**
Adjusted R-squared	0.656	
Daily observations	2122	
Countries covered	33	

Note: The containment policy variables are based on those summarised in Table 1, but re-normalised to be (0, 1) dummy variables as described in the main text. The notation in brackets "(=n)" after a containment policy variable denotes that the dummy variable is assigned a 1 if the original score for that policy was equal to n, whereas the notation ">=n" denotes that the dummy variable is assigned a 1 if the original score for that policy was greater than or equal to n. **** denotes statistical significance at the 5% level. The structural budget balance variable is taken from the *OECD Economic Outlook* No. 109 database at a quarterly frequency and is converted to a weekly frequency. The coefficients reported in the second column are those that are significant when interacted with a dummy taking the value of 1 from the beginning of July 2020; the effect of these policies from July is then estimated to be the sum of the coefficients in the two columns.

Source: Authors' calculations.

Figure 6. The effect of containment policies on GDP (per cent)

Maximum effect on GDP in period after July 2020



Note: The chart shows the maximum effect on GDP of different containment policies estimated for the period after July 2020.

1) The effect of full closure of public transport is estimated when implemented with full workplace closure.

Source: Authors' calculations based on estimation results in Table 4.

Modelling the effect of international travel restrictions on GDP is more difficult because it depends not only on restrictions imposed by the country in question, but also on the restrictions imposed on travellers by other countries. While international travel restrictions are not included in the equation reported in Table 4, alternative equations suggest the effect on GDP of international travel restrictions is proportional to the importance of tourism in the country.

Fiscal measures are also found to play a role in supporting GDP through the inclusion of the structural budget balance (expressed in per cent of GDP).¹¹ The estimated coefficient is highly statistically significant and implies a multiplier of about 0.2, which is consistent with much of the fiscal support having taken the form of income support and subsidies.

6. Scenario analysis

The estimations described earlier in the paper are used to construct scenarios that illustrate the benefits of a rapid vaccination rollout, particularly in the face of more contagious strains of the virus. The

¹¹ The structural budget balance is a measure of the cyclically-adjusted primary fiscal balance (expressed as a percentage of GDP) and is taken from the most recent quarterly *OECD Economic Outlook* database and is converted to a weekly frequency with interpolation.

assumptions are highly stylised, but nevertheless highlight some of the policy implications of different speeds of vaccination rollout.

6.1. Baseline scenario prior to vaccination rollouts and new variants

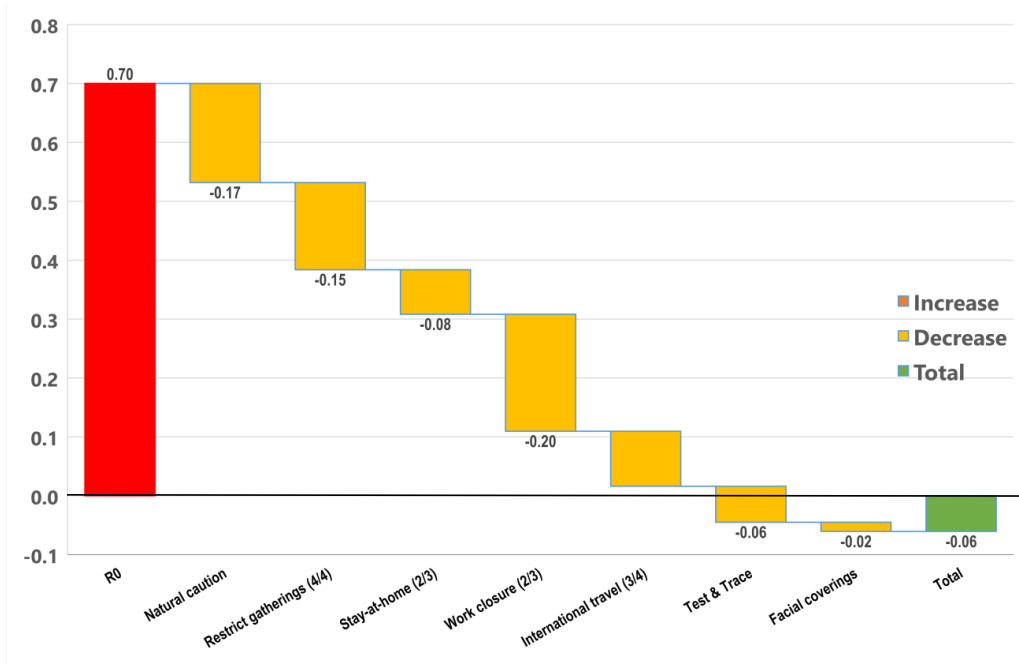
An initial baseline scenario is constructed to be representative of the position of a typical OECD country during the first quarter of 2021, before either vaccines or new variants were having any significant effect on the spread of the virus, on the following assumptions:

- Containment measures are applied at the median level of stringency observed across all OECD countries during the first quarter of 2021. This implies maximum stringency as regards restrictions on public events, gatherings and internal movement. It also implies some workplace closures as well as stay-at-home requirements, partial school closures and restrictions on international travel, although all these latter measures remain one stringency notch below the maximum specified by the Oxford taxonomy. The overall cost of these measures to the economy is estimated to be about 5% of GDP, which corresponds roughly with the hit to GDP in the median OECD country since the start of 2021.
- All public health measures, including test and trace policies and mandating facial coverings, which might help reduce the spread of the virus, have already been implemented and cannot be made more effective. However, protection for the elderly remains below the maximum level of stringency, so although there are some restrictions on visitors to long-term care facilities, these fall short of prohibiting all non-essential visitors.
- A 'fear factor', proxied by the global and national daily death rates (here assumed to be running at 1 per million and 2 per million, respectively), induces more cautious behaviour which further reduces the reproduction number.

Using the estimation results described in the previous section, the reproduction number implied by the combination of these assumptions is only just below 1.0 (equivalent to $\log R$ being below 0) as summarised in Figure 7. The situation represented by the scenario is, however, precarious, with many factors, including seasonal influences, having the potential to push the reproduction number above 1 and so lead to a surge in infections.

Figure 7. An initial baseline scenario before vaccinations and more transmissible variants

Effect on the logged reproduction number R



Note: The waterfall chart illustrates the component effects on logged R. The sum of the initial R0 (red bar) and the negative effects of natural caution and policies (yellow bars) give a total estimate for the logged reproduction number (green bar). The (small) negative number for the total logged R implies that R is (just) less than 1. The numbers in brackets after each containment policy (n/N) indicates that corresponding containment policy is assumed to be at stringency level n, compared to a possible maximum stringency level of N. The natural caution effect assumes a global daily death rate of 1 per million and a national daily death rate of 2 per million.

Source: Authors' calculations.

6.2. Scenario with more contagious new variant and moderate vaccination rollout

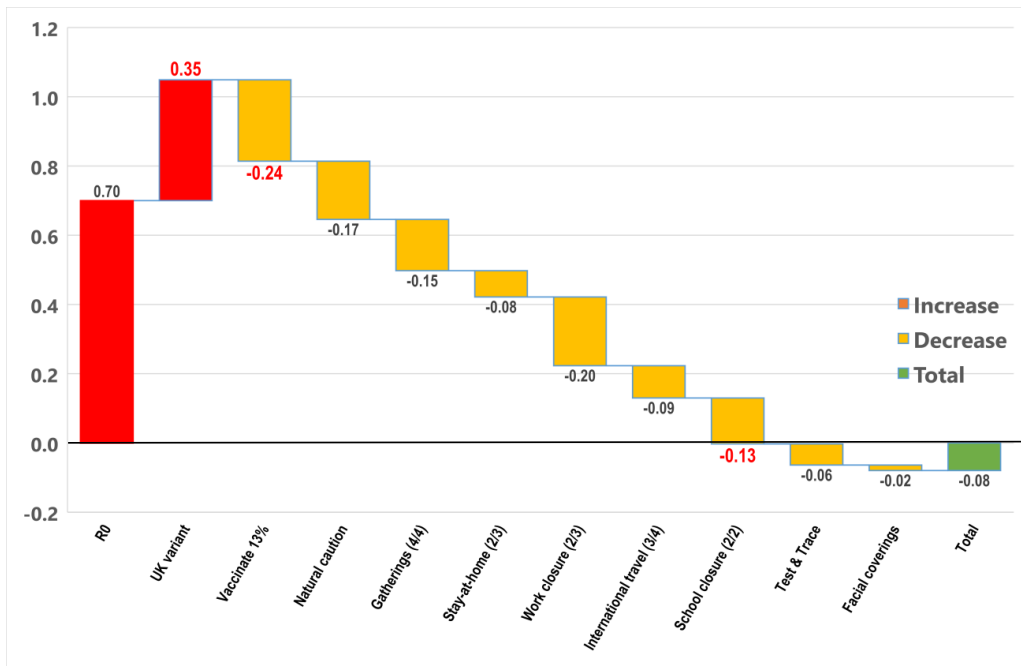
The scenarios that follow build on the initial baseline scenario by considering how the situation might develop in response both to more contagious variants and the rollout of vaccination programmes. For this purpose, all scenarios assume that the so-called Alpha variant becomes predominant, leading to an increase in transmissibility of the virus by 35%. If only 13% of the population have been fully vaccinated, which corresponds to the OECD median in mid-May, then vaccinations would be insufficient to keep the reproduction number below 1. Policy-makers would then face difficult choices about which containment policies to tighten further, given the raft of containment policies already implemented in the baseline scenario. In the scenario summarised in Figure 8, it is assumed that all schools are closed full-time, which would be just sufficient to keep the reproduction number below 1.

Of course, there is considerable uncertainty around the parameter estimates used to construct the scenarios as well as the assumptions underlying them. On the positive side, seasonal factors for many European countries coming into summer could serve to lower the reproduction number, by approximately 15% for a central European country relative to winter. Alternatively, the test, trace and isolation policy now operating in a particular country could be more effective than captured by the average estimate for all

countries over the whole pandemic period. On the other hand, a slower vaccination rollout (some OECD countries had only vaccinated around 5% of the population by mid-May) or even more transmissible variants would only exacerbate the dilemma about which containment policies to tighten; given lockdown fatigue, policy-makers might hesitate to implement the most stringent forms of stay-at-home requirements and the closure of all-but-essential workplaces that might then be required to keep the reproduction number below 1.

Figure 8. Alpha variant with limited vaccine rollout

Effect on the logged reproduction number R



Note: See the note to Figure 7. Changes relative to the baseline scenario are highlighted by the red bolded numbers.
 Source: Authors' calculations.

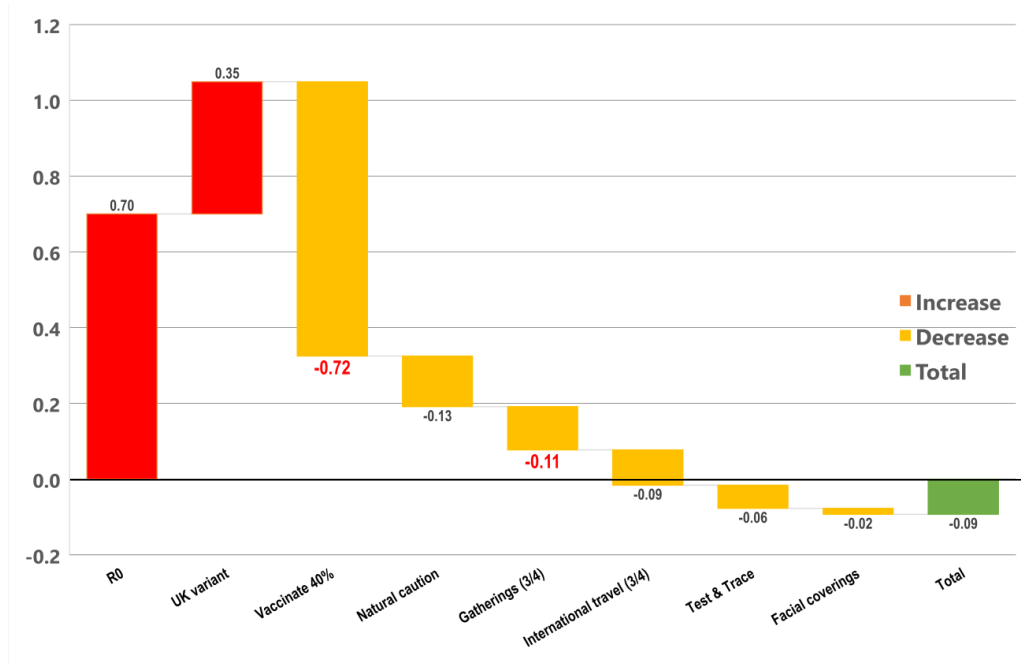
6.3. Scenario with more contagious new variant and faster vaccination rollout

A faster vaccine rollout would not only avoid the need to tighten containment policies, but also enable containment policies in the baseline scenario to be progressively relaxed. For example, if 40% of the population were fully vaccinated (which is close to the shares in the United States and United Kingdom at end-May), then there might be no need for any stay-at-home requirements or closure of workplaces and restrictions on gatherings could start to be relaxed, which would raise GDP by 4% relative to the baseline scenario (Figure 9).

Covid Economics 83, 2 July 2021: 154-182

Figure 9. Alpha variant with quick vaccine rollout

Effect on the logged reproduction number R



Note: See the note to Figure 7. Changes relative to the previous scenario illustrated in Figure 8 are highlighted by the red bolded numbers, in addition there is no need for stay-at-home requirements and closure of workplaces, which are consequently excluded from the chart
Source: Authors' calculations.

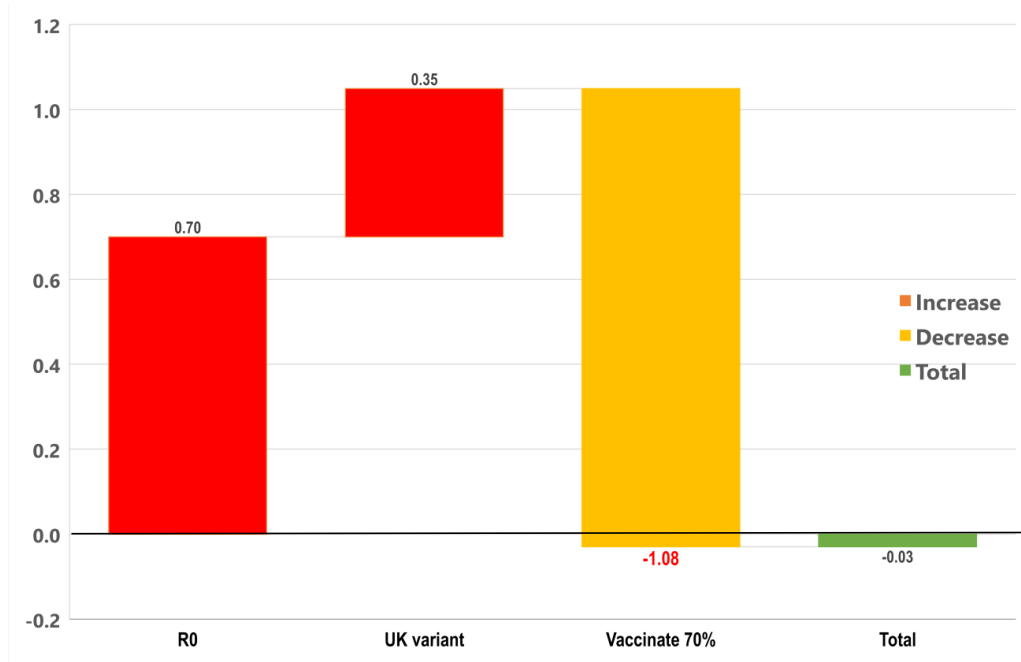
6.4. Scenario with comprehensive vaccination rollout

The estimates also suggest that, in the absence of any supporting measures, roughly 70% of the population would need to be vaccinated to ensure the reproduction number remains below 1 (Figure 3). While much of the focus has recently been on the speed with which vaccines have been rolled out, achieving such a high overall level of vaccination is likely to be a challenge in many countries, given that some vulnerable groups (such as young children) will not be vaccinated and given the observed slowdown in take-up when vaccination rates reach high levels. Moreover, if other more transmissible variants of the virus become predominant, then the overall required level of vaccination would rise further. This suggests that in some countries, it may be necessary to continue with some public health measures and limited containment measures over the longer term even after large shares of the population have been fully vaccinated.

Finally, it is important that policy-makers are not complacent over the summer as seasonal factors may temporarily contribute to suppressing the reproduction rate and so reduce the urgency to rollout vaccinations. For example, for a typical Central European country, seasonal factors could lower the reproduction rate by about 15%, which would mean that it could be kept below 1 with about 10 percentage points less of the population vaccinated. However, without further progress in vaccination, there could be a subsequent surge in the virus as seasonal factors reverse.

Figure 3. Comprehensive vaccination rollout

Effect on the logged reproduction number R



Note: See the note to Figure 7. Changes relative to the previous scenario are highlighted by the red bolded numbers.
Source: Authors' calculations.

7. Conclusions

This paper analysed empirically the impact of policies on the reproduction rate and the OECD Weekly GDP Tracker using high-frequency data for a set of OECD countries, based on which scenario analysis is carried out. The main findings of the paper can be summarised as follows.

First, some new variants of the virus are estimated to be able to boost the effective reproduction number by up to 50%. Seasonal effects are also found to increase the effective reproduction number in fall/winter, in some countries by up to 25% relative to summer. The rapidity of these adverse shocks represent a major challenge to policy-makers because they can coincide and take full effect over a matter of a few months. The two effects together can potentially boost reproduction numbers by up to 90%.

Second, vaccination is found to powerfully reduce the spread of the virus. Using the estimated equation, the effects can be stated in intervention-equivalent terms. Fully vaccinating...

- 7% of the population is equivalent to either complete school closure, requiring people not to leave the house with minimal exceptions, or banning all public gatherings;
- 15% of the population is equivalent to closing down all-but-essential workplaces;
- 20% of the population is equivalent to closing down all-but-essential workplaces as well as public transport.

- 50% of the population is equivalent to simultaneously applying all of the above restrictions as well as closing all international borders.

Third, scenarios based on the estimation results suggest a rapid rollout of vaccinations is needed to compensate for the pressure from more infectious variants, and so avoid an escalation of lockdown measures as well as the risk of the perpetuation of a cycle of stop-and-go mitigation policies. For a typical OECD country, fully vaccinating more than 40% of the population allows most containment policies to be either eliminated or substantially eased, typically raising GDP by 4-5%.

Finally, for those countries now going into summer, it is also important that policy-makers are not lulled into a false sense of security by the temporary decline in reproduction numbers due to seasonal factors, as in the summer of 2020. Failure to vaccinate a sufficient share of the population could then lead to a resurgence of the virus in the winter as seasonal factors reverse.

References

- Ahammer, A., M. Halla and M. Lackner (2020), "Mass Gathering Contributed to Early COVID-19 Spread: Evidence from US Sports", *CEPR Covid Economics*, No. 30.
- Aghion, P., P. Artus, M. Orlu-Barton and B. Pradelski (2021), "Aiming for zero Covid-19 to ensure economic growth", VoxEU.org, 31 March.
- Arias, J.E., J. Fernández-Villaverde, J.R. Ramírez and M. Shin (2021), "[Bayesian Estimation of Epidemiological Models: Methods, Causality, and Policy Trade-Offs](#)", *NBER Working Papers*, No. 28617, National Bureau of Economic Research, Cambridge, MA.
- Arroyo-Marioli, F., F. Bullano, S. Kučinskas, C. Rondón-Moreno (2020), "Tracking R of COVID-19: A New Real-Time Estimation Using the Kalman Filter", medRxiv 2020.04.19.20071886; <https://doi.org/10.1101/2020.04.19.20071886>
- Audirac, M.T., L.A. Meyers, S. Fox and C. Zigler (2020), "[How Timing of Stay-Home Orders and Mobility Reductions Impacted First-Wave Covid-19 Deaths in US Countries](#)", medRxiv.
- Berg, J., F. Bonnet and S. Soares (2020), "[Working from home: Estimating the worldwide potential](#)", VoxEU.
- Bendavid, E., C. Oh, J. Bhattacharya and J. Ioannidis (2021), "[Assessing mandatory stay-at-home and business closure effects on the spread of COVID-19](#)", *European Journal of Clinical Investigation*.
- Berlemann, M. and E. Haustein (2020), "Right and yet wrong: A spatio-temporal evaluation of Germany's Covid-19 containment policy", *CEPR Covid Economics*, No. 36.
- Boll, C. and T. Nikolka (2020), "Rightly blamed the 'bad guy'? Grandparental child care and Covid-19", *CEPR Covid Economics* No. 62.
- Blanco F., D. Emrullahu and R. Soto (2020), "Do Coronavirus Containment Measures Work? Worldwide Evidence", *World Bank Group Policy Research Paper* No. 9490.
- Briz-Redón, A. and A. Serrano-Aroca (2020), "[The effect of climate on the spread of the Covid-19 pandemic: A review of findings, and statistical and modelling techniques](#)", *Progress in Physical Geography*, 44(5), 591-604.
- Carraro, A., L. Ferrone and M. Squarcina (2020), "Are COVID-19 Containment Measures Equally Effective in Different World Regions?" *DISEI Working Papers* 11/2020.
- Caselli, F., F. Grigoli and D. Sandri (2021), "Protecting lives and livelihoods with early and tight lockdowns", *CEPR Covid Economics*, No. 66.

- Chernozhukov, V., H. Kasahara and P. Schrimpf (2020), "Casual impact of masks, policies, behaviour on early Covid-19 pandemic in the US", *CEPR Covid Economics*, No. 35.
- Canyon, M.J., L. He and S. Thomsen (2020), "Lockdowns and Covid-19 Deaths in Scandinavia", *CEPR Covid Economics*, No. 26, 5 June.
- Davies, N. et al. (2021). "Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England", *Science*, 372, April.
- Deb, P., D. Furceri, J. D. Ostry and N. Tawk (2020), "The effect of containment measures on the COVID-19 Pandemic", *CEPR Covid Economics*, No. 19.
- Demirgüç-Kunt, A. M. Lokshin and I. Torre (2020), "The Sooner, the Better: The Early Economic Impact of Non-Pharmaceutical Interventions during the COVID-19 Pandemic." *World Bank Policy Research Working Paper* No. 9257.
- Dylan H.M et al. (2020), "[The effect of temperature and humidity on the stability of SARS-CoV-2 and other enveloped viruses](https://doi.org/10.1101/2020.10.16.341883)", bioRxiv <https://doi.org/10.1101/2020.10.16.341883>.
- ECDC (2021), "[Using face masks in the community: Effectiveness in reducing transmission of COVID-19](https://www.ecdc.europa.eu/en/using-face-masks-in-the-community-effectiveness-in-reducing-transmission-of-covid-19)", 15 February.
- Égert, B., Guillemette, Y., Murtin, F. and D. Turner (2020), "Walking the tightrope: avoiding a lockdown while containing the virus", OECD Economics Department Working Papers 1633.
- Esra, R. et al. (2020), "[Evaluating the impact of non-pharmaceutical interventions for SARS-CoV-2 on a global scale](https://doi.org/10.1101/2020.07.30.20164939)", medRxiv <https://doi.org/10.1101/2020.07.30.20164939>.
- Fetzer, T. and T. Graeber (2020), "[Does Contact Tracing Work? Quasi Experimental Evidence from an Excel Error in England](https://www.cepr.org/workingpapers/2020/07/does-contact-tracing-work-quasi-experimental-evidence-from-an-excel-error-in-england)".
- Gapen, M., J. Millar, B. Uruci and P. Sriram (2020), "Assessing the effectiveness of alternative measures to slow the spread of Covid-19 in the United States", *CEPR Covid Economics*, No. 40.
- Goolsbee, A. and C. Syverson (2020), "[Fear, Lockdown, and Diversion: Comparing Drivers of Pandemic Economic Decline 2020](https://www.bfi.uchicago.edu/wp-content/uploads/2020/08/BFI-2020-80-Fear-Lockdown-and-Diversion-Comparing-Drivers-of-Pandemic-Economic-Decline-2020.pdf)", Becker-Friedman Institute for Economics at the University of Chicago *BF Working Paper*, No. 2020-80.
- Gottlieb, C et al. (2020), "Lockdown accounting", *CEPR Covid Economics*, No. 31.
- Gros D., A. Ounnas and T. Yeung (2021), "A New Covid Policy Stringency Index for Europe", *CEPR Covid Economics*, No. 66.
- Gupta, S., K.I. Simon and C Wing (2020), "Mandated and Voluntary Social Distancing during the Covid-19 Epidemic: A review", *NBER Working Paper* 28139.
- Hale, T. et al. (2020), "[Oxford COVID-19 Government Response Tracker](https://www.bsg.ox.ac.uk/oxford-covid-19-government-response-tracker)", Blavatnik School of Government, Oxford University.
- Hatzius, J., D. Struyven and I. Rosenberg (2020), "Face Masks and GDP", Goldman Sachs Global Economics Analyst, 29 June.
- Haug, N. et al. (2020), "Ranking the effectiveness of worldwide COVID-19 government interventions", *Nature Human Behaviour*, Vol. 4, pp. 1303-1312.
- Hellewell, J. et al. (2020), "[Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts](https://www.thelancet.com/journal/S2468-2667(20)30229-4)", *The Lancet Global Health*, Vol. 8(4), pp. e488-e496.
- Hodcroft, E.B. et al. (2021), "[Emergence and spread of a SARS-CoV-2 variant through Europe in the summer of 2020](https://www.medrxiv.org/content/10.1101/2021.05.11.21254441v1)", medRxiv.

- Hubert, O. (2020), "Spatial Distancing: Air Traffic, -19 Propagation and the Cost of Efficiency of Air Travel Restrictions", *CEPR Covid Economics*, Issue No. 24.
- Hunter, P.R., F.J. Colón-González, J. Brainard and S. Rushton (2020), "Impact of non-pharmaceutical interventions against COVID-19 in Europe: a quasi-experimental study", mimeo.
- Islam, N., S.J. Sharp, G. Chowell, S. Shabnam, I. Kawachi, B. Lacey, J.M. Massaro, R.B. D'Agostino and M. White (2020), "Physical distancing interventions and incidence of Coronavirus disease 2019: natural experiment in 149 countries." *BMJ* 370: m2743. 10.1136/bmj.m2743.
- Kahanec, M., L. Laffers and B. Schmidpeter (2021), "[The impact of antigen testing for Covid-19 on the prevalence of the disease](#)", *CELSI Discussion Paper* No. 59.
- Keita, S. (2020), "Air passenger mobility, travel restrictions and the transmission of the Covid-19 pandemic between countries", *CEPR Covid Economics*, No. 9.
- Kim, P. S., S.W. Read and A.S. Fauci (2020). "[Therapy for early COVID-19: A critical need](#)", *JAMA* 324 (21), 2149-2150.
- Koh, W.C., L. Naing and J. Wong (2020), "[Estimating the impact of physical distancing measures in containing COVID-19: An empirical analysis](#)", *International Journal of Infectious Diseases*, Vol.100, pp. 42–49.
- Kok, J.L.C. (2020), "Short-term trade-off between stringency and economic growth", *CEPR Covid Economics*, No. 60.
- König M. and A. Winkler (2020), "Monitoring in real time: cross-country evidence on the covid-19 impact on GDP growth in the first half of 2020", *CEPR Covid Economics*, No. 57.
- Leffler, G.T., E. Ing, J.D. Lykins, M.C. Hogan, C.A. McKeown and A. Grzybowski (2020), "[Association of country-wide coronavirus mortality with demographics, testing, lockdowns, and public wearing of masks](#)", medRxiv.
- Li Y., et al. (2020), "[The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number \(R\) of SARS-CoV-2: a modelling study across 131 countries](#)", *The Lancet Infectious Diseases*, Vol. 21 (2), pp. 193-202.
- Liu, Y., C. Morgenstern, J. Kelly and R. Lowe (2020), "[The impact of non-pharmaceutical interventions on SARS-CoV-2 transmission across 130 countries and territories](#)", medRxiv.
- Mecenas, P., R. Travassos da Ros Moreiro Bastos, A. C. R. Vallinoto, D. Normando (2020), "[Effects of temperature and humidity on the spread of Covid-19: A systematic review](#)", *PLOS ONE*.
- Merow C. and M.C. Urbam (2020), "[Seasonality and uncertainty in global Covid-19 growth rates](#)", *PNAS* 117(44), 27456-27464,
- Mills, M. (2020), "[Face masks and coverings for the general public: Behavioural knowledge, effectiveness of cloth coverings and public messaging](#)", June 26, University of Oxford.
- Mitze, T., R. Kosfeld, J. Rode and K. Wälde (2020), "Face Masks considerably reduce Covid-19 Cases in Germany: A Synthetic Control Method Approach", *CEPR Covid Economics*, No. 27.
- Morikawa, M. (2020), "Productivity of Working from Home during the Covid-19 Pandemic: Evidence from an Employee Survey", *CEPR Covid Economics*, No. 49.
- New York Times (2020), "[Here's How Wuhan Tested 6.5 Million for Coronavirus in Days](#)", updated June.
- OECD (2020a), "[Testing for COVID-19: A way to lift confinement restrictions](#)" (accessed on 26 June).
- OECD (2020b), "[The OECD Weekly Tracker of activity based on Google Trends](#)", Issue Note 1 in *OECD Economic Outlook*, Volume 2020 Issue 2, OECD Publishing, Paris.

- Olczak, J., J. Reade and M. Yeo (2020), "Mass outdoor events and the spread of an airborne virus: English football and Covid-19", *CEPR Covid Economics*, No. 47.
- Paraskevis E. et al.(2021), " A review of the impact of weather and climate variables to COVID-19: In the absence of public health measures high temperatures cannot probably mitigate outbreaks", *Science of The Total Environment*, Vol. 768.
- Pichler, S., K. Wen and N. R. Ziebarth (2020), "Positive Health Externalities of Mandating Paid Sick Leave." *IZA Discussion Papers* 13530.
- Piovani, D. et al. (2021), "[Effect of early application of social distancing interventions on COVID-19 mortality over the first pandemic wave: An analysis of longitudinal data from 37 countries](#)", *Journal of Infection*, Vol. 82, pp. 133-142.
- Shakil, M.H., Z.H. Munim, M. Tasnia and S. Sarowar (2020), "[Covid-19 and the environment: A critical review and research agenda](#)", *Sci Total Environ*, 745.
- Spiegel, M. and H. Tookes (2020), "Business restrictions and Covid fatalities", *CEPR Covid Economics*, No. 56.
- Taipale, J., Romer P. and S. Linnarsson (2020), "[Population-scale testing can suppress the spread of COVID-19](#)", medRxiv.
- Thunström, L., M. Ashworth, D. Finnoff and S.C. Newbold (2020), "Hesitancy towards a Covid-19 vaccine and prospects for herd immunity", *CEPR Covid Economics*, No. 35.
- Troeger, T. (2020), "Optimal testing and social distancing of individuals with private health signals", *CEPR Covid Economics*, No. 54.
- Walrand, S. (2021), "[Autumn Covid-19 surge dates in Europe correlated to latitudes not to temperature-humidity, pointing to vitamin D as contributing factor](#)", *Scientific Reports* 11 (1981).
- Walsh, S. et al. (2020), "[Do school closures reduce community transmission of Covid-19? A systematic review of observational studies](#)", medRxiv.
- Weber, E. (2020), "Which measures flattened the curve in Germany?", *CEPR Covid Economics*, No. 24.
- Weir, E.K., T. Thenappan, M. Bhargawa and Y. Chen (2020), "[Does vitamin D deficiency increase the severity of Covid-19?](#)", Vol. 20/4, *Clinical Medicine*.
- Welsch, D. M. (2020), "Do masks reduce Covid-19 deaths? A country-level analysis using IV", *CEPR Covid Economics*, No. 57.
- Woloszko, N. (2020a), "Tracking activity in real time with Google Trends", OECD Economics Department Working Papers, No. 1634, OECD Publishing, Paris, <https://doi.org/10.1787/6b9c7518-en>.
- Woloszko, N. (2020b), "Tracking GDP using Google Trends and machine learning: A new OECD model", VoxEU, December 2020.

Annex A. The importance of timeliness in introducing containment measures

The timing of the introduction of containment policies is very important for the control of the virus. It takes time for measures to become effective (Li et al., 2020) due to the disease's incubation period and the fact that some of the infected are asymptomatic. Moreover, the non-linear transmission of the virus implies that curbing it earlier on may lead to better outcomes than intervening after the spread has reached high growth rates.

To test whether an early timing of containment measures helped limit the spread of the epidemic, the total number of COVID-19-related deaths per million of population as of 15 July 2020 is regressed on the number of days between the local start of the epidemic (the day the country reached the first 100 cases per million) and the implementation of a typical or a maximum stringency of a package of containment measures for a cross-section of up to 148 countries. The stringency of the policy package is calculated as the sum of the scores of the following Oxford policy variables (thresholds for a typical policy in parentheses): stay-at-home requirements (1), workplace closures (2), school closures (2), restriction on public gatherings (3), international travel restrictions (1), testing (3) and tracing (2).¹² The median country reached typical policy levels 25 days and maximum policies four days before the start of the epidemic, reflecting the anticipation effect based on other countries' experience. In addition, the analysis controls for the age and density of the population, which may be associated with COVID-19 mortality and its speed of transmission.

The results suggest that delayed policy interventions lead to a significantly higher number of deaths relative to population by mid-2020 (Table A.1). The median country reached the maximum level of policies four days before the local start of the epidemic, which saved five lives per million, relative to the death toll of a country fully intervening only on the first day of the local epidemic outbreak. The fastest-acting (top 25%) countries saved about 80 lives per million more than the slowest-acting (bottom 25%) countries (Figure A.1)¹³. Countries which are relatively younger (top 25%) and less urbanized (bottom 25%), lost 30 and 22 lives per million less, respectively, than those with average population structure (Figure A.1).

These results are in line with other empirical studies which found that countries that implemented containment measures earlier, had a lower number of infections and/or deaths after a certain period than countries, that acted later (Deb et al., 2020; Caselli et al., 2021). While some authors find an effect of timing for a wide range of policies, including restrictions on public events, public transportation or mobility (Blanco

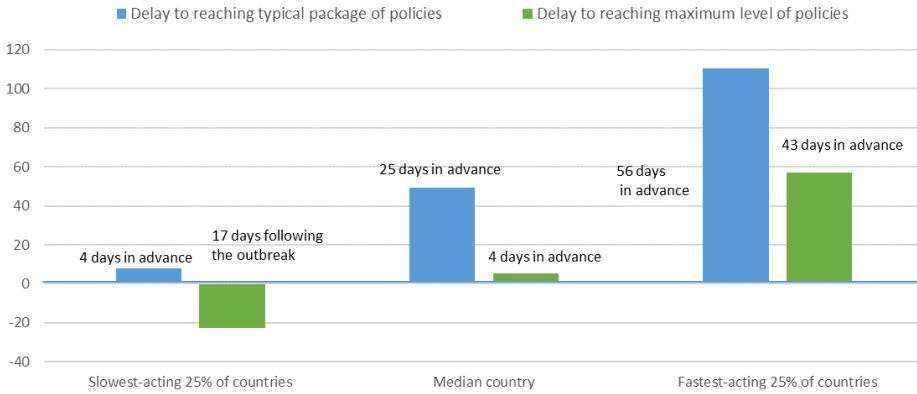
¹² This means that a package of the same overall strength could have consisted of various combinations of different instruments and their stringencies.

¹³ The results are robust to a number of robustness checks: 1) restricting the sample to countries with a minimum number of deaths (50) by mid-July 2020, 2) defining the local start of the epidemic as the date of the first death per million (rather than 100 new cases), 3) defining the end of the first wave as two months from the local start (instead of the fixed end-date of 15 July 2020). 4), controlling for the level (defined as either maximum or range) and/or duration of policies. Qualitatively similar results hold for a group of 28 European countries, except the share of older people does not play a role for mortality there.

et al.,2020), others find timing mostly matters for specific lockdown-type measures (Koh et al., 2020), or school closures and bans on mass-gatherings (Piovani et al., 2021; Haug et al.,2020).

Figure A.1. Lives saved (or lost [-]) by mid-July depending on population structure

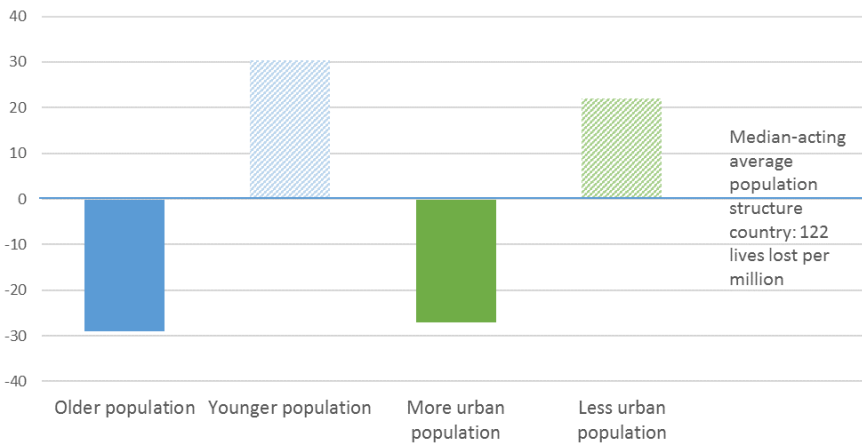
per million, relative to a country acting on the day when 100 cases were reached



Note: The day of the outbreak is the day when the country reach 100 cumulative cases per million of inhabitants.
Source: Authors' calculations.

Figure A.2. Lives saved (or lost [-]) by mid-July depending on population structure

per million, relative to median-acting country in terms of policy swiftness with average age and urban structure



Note: Older population means a median-acting country, with relatively older (top 25%) population and average urbanicity. Younger population means a median-acting country, with relatively younger (top 25%) population and average urbanicity. More urban population means a median-acting country, with relatively high (top 25%) urbanicity and average age structure. Less urban population means a median-acting country, with relatively low (bottom 25%) urbanicity and average age structure.
Source: Authors' calculations.

Covid Economics 83, 2 July 2021: 154-182

The effect of policy timing can be investigated relative to the local start of the epidemic as defined by some initial low number of cases or deaths (e.g. 1 or 100), as is commonly done in the literature. It could also be potentially assessed relative to the peak of new cases or deaths. While both points are not exogenous, the local start mainly due to anticipation effects and learning from other countries' experience, the peak is even less of an independent event, as policies directly affect its timing, severity, as well as the profile of the epidemic around the peak, so the former approach is used. The results are robust to choosing an alternative definition of the start of the local epidemic (one death per million) and even with this definition, the majority of countries implemented containment policies in advance of this event. Nevertheless, countries that acted faster than others had fewer deaths by mid-2020 than those acting relatively later.

Table A.1. Mortality and timing of policies regression results

Dependent variable: total deaths per million population

	(1)	(2)	(3)
Constant	157.212**	113.965**	-56.348
Policy response: time from first 100 cases per million to reach maximum policy		1.345**	0.804**
Policy response: time from first 100 cases per million to reach typical policy	1.973***		
Share of 65+			3.645#
Urban share			1.392*
Share of 65+ X policy (max) response			0.112*
Urban share X policy (max) response			0.022*
Number of countries	129	148	143
Adjusted R squared	0.165	0.124	0.257

Notes: Results from an OLS cross-section regression. * and ** denote statistical significance at the 5 and 1% levels, respectively. # indicates that the coefficient estimate is statistically significant at the 10% level. Variables used in the interaction terms are demeaned.

Source: Authors' calculations.

COVID-19 reporting and willingness to pay for leisure activities¹

Sonja Warkulat,² Sebastian Krull,³ Regina Ortmann,⁴
Nina Klocke⁵ and Matthias Pelster⁶

Date submitted: 15 June 2021; Date accepted: 22 June 2021

The containment of COVID-19 critically hinges on individuals' behavior. We investigate how individuals react to variations in COVID-19 reporting. Using a survey, we elicit individuals' perceived infection risk given various COVID-19 metrics (e.g., confirmed cases, reproduction rate, or case-fatality ratio). We proxy individuals' risk perception with their willingness to pay for the participation in everyday life and amusements events. We find that participants react to different COVID-19 metrics with varying sensitivity. We observe a saturation of sensitivity for several measures at critical limits used in the political discussion, making our results highly relevant for policy makers in their efforts to direct individuals to adhere to hygienic etiquette and social distancing guidelines.

- 1 Regina Ortmann gratefully acknowledges financial support by the German Research Foundation (DFG) - Collaborative Research Center (SFB/TRR) Project-ID 403041268 - TRR 266 Accounting for Transparency. Any errors, misrepresentations, and omissions are our own.
- 2 Paderborn University, Center for Risk Management.
- 3 Paderborn University, Center for Risk Management
- 4 Paderborn University, Center for Tax and Accounting Research.
- 5 Paderborn University, Center for Risk Management.
- 6 Paderborn University, Center for Risk Management.

Copyright: Sonja Warkulat, Sebastian Krull, Regina Ortmann, Nina Klocke and
Matthias Pelster

1 Introduction

Limiting close face-to-face contacts is the best way to reduce the spread of COVID-19. The compliance with social distancing regulations depends on people's willingness to restrict their daily lives to slow the spread of the pandemic. However, low levels of perceived infection risk may limit compliance with social distancing regulations. To assess the infection risk, people regularly browse news outlets for updates on the current state of the COVID-19 pandemic (Ries et al., 2020). With cases on the rise, a variety of COVID-19 metrics have emerged and are reported daily by news channels. Some of these metrics are commonly used as a basis for pandemic response plans of governments. Because people may perceive infection risk contingent on the used COVID-19 metric, people's sensitivities to individual metrics have to be taken into account when pandemic communication strategies are discussed. Thus, it is important to understand people's reactions to the reporting of different COVID-19 metrics and find out which metrics people react to the most and adjust their behavior to the rules accordingly.

In this paper, we study in how far various COVID-19 metrics potentially allow to nudge people to decrease their infection risk. We draw from the literature arguing that certain statistical information and verbal messages may alter individuals' risk perception and nudge them towards infection prevention behavior (Kim et al., 2020; Sasaki et al., 2020). Considering these findings and general insights on the so-called "framing"-effect (Tversky and Kahneman, 1985; Levin et al., 1998), we hypothesize that different COVID-19 metrics may result in systematically different perceptions of the infection risk. Thus, we study how sensitive individuals react i) to different commonly reported COVID-19 metrics, and ii) to different levels of these metrics. Given that individuals were to perceive infection risk differently for different metrics and behave according to their perception, a thoughtful choice of the metric may nudge consumers towards an objectively desired (health) decision.

We measure individuals' willingness to pay (WTP), for example, for a movie ticket, given different COVID-19 metrics and a range of infection rates to quantify their perceived

infection risk exposure.¹ WTP is an established measure in this context (see, e.g., Perry-Duxbury et al., 2019). With increasing case numbers individuals should either ask for a discount on movie-tickets, being exposed to infection risk, or be willing to pay a premium on movie tickets to be seated in a theatre with installed air filtration. In line with our expectations, we find that people are, on average, less willing to pay to participate in everyday life and amusement events as the pandemic spread increases. However, participants' reactions vary across COVID-19 metrics. Whereas participants react most sensitively to the incidence rate and to confirmed cases, they react most insensitively to the case-fatality ratio and to the "traffic light model", a three-tiered measure indicating the overall severeness of the current state of the pandemic. In addition, the sensitivity in participants' reactions varies for different levels in metrics. In particular, individuals seem to become insensitive to a change in metrics on intensive care cases, case-fatality ratios, and confirmed cases once a certain threshold is reached.

The results of our study are highly relevant for policy makers in their effort to achieve desired adherence of guidelines, in particular in light of recent evidence that suggests a "fatalism effect"; that is, in light of the observation that individuals become less willing to adhere to social distancing guidelines when they believe COVID-19 to be more infectious (Akesson et al., 2020). Most importantly, our results allow to draw conclusions on COVID-19 metrics and the levels for which people are or become insensitive. In such a state of insensitivity, pandemic self-regulation of the population may not work anymore and stronger governmental intervention may become necessary. Since we show that people respond differently to different COVID-19 metrics, we also provide examples of metrics that can be used to nudge people towards a desired infection prevention behavior.

¹In additional specifications, we use individuals' willingness to accept (WTA) a risk exposure, for example, by a discounted movie ticket. We find quantitatively similar results and therefore refer to both measures as WTP as a generic term hereafter.

2 Theoretical background and literature

“Nudging” utilizes changes in a decision’s environment to improve an individual’s behavior towards an objectively desired (health) decision (for an overview see Li and Chapman, 2013). In order to nudge individuals towards more optimal health decisions, descriptions of normative equivalent decision situations are commonly “framed” in different ways to evoke systematically different choices.

In the context of the still ongoing COVID-19 pandemic, health officials and media outlets make use of a large array of metrics to communicate the current state of the pandemic and to justify policy implications. Perhaps even unintentionally from communicators, individuals may perceive metrics differently and adjust their behavior accordingly. In a nudging-context, text-based frames for health decisions are at least to some extent understood (Sasaki et al., 2020). However, we know very little about the impact of number-based frames in a health context. Different COVID-metrics highlight different aspects of the state of the pandemic and, consequently, offer different frames. For example, the R-value emphasizes the infectiousness of the disease whereas the confirmed deaths stress the most severe consequence the pandemic might have for an individual.

In this paper, we shed light on the question to which degree various COVID-19 metrics allow to nudge individuals towards desirable health decisions. It is especially important to study the reactions to such metrics as individuals seem to have difficulties digesting numeric information. That is, for instance, humans tend to understand numerical information expressed as frequencies (e.g., 2 out of 10 chances for infection) more easily than the same message expressed as probability (e.g., 20% probability for infection) (see, e.g., Gigerenzer and Hoffrage, 1995) or react more sensitive to a relative risk increase compared to an absolute risk increase (Gigerenzer et al., 2007; Visschers et al., 2009). In the context of a pandemic, the threatening emotional component may even add to the already problematic processing of numerical information. Individuals seem to overestimate risks that are unfamiliar, outside of their control, inspire feelings of dread, and receive extensive media coverage (Slovic, 2000). In line with this notion, Akesson et al. (2020)

find that individuals dramatically overestimate the infectiousness of COVID-19 based on the R-value.

Our paper contributes to a fast growing literature on the framing of interventions to reduce the spread of COVID-19. Sasaki et al. (2020) investigate text-based nudging in a survey experiment and analyze how different verbal messages impact participants' intention to adopt infection prevention behavior. They find that only gain-framed, altruistically formulated messages promote the desired infection prevention behavior. Investigating private and public benefit frames, Blackman and Hoffmann (2021) assess the effectiveness of informational nudges about COVID-19, recent compliance with non-pharmaceutical interventions recommendations, and intended future compliance. Although these informational nudges boosted concern, they had limited effects on either recent or intended future compliance.

Only a few studies cover numeric nudging in a health context (see Shen and Hsee, 2017). Welkenhuysen et al. (2001) investigate characteristics of risk communication in the context of a decision regarding prenatal diagnosis for cystic fibrosis. In their design, they manipulate the message in two dimensions and study verbal versus numerical information on risk and a negative versus a positive framing of the problem and find the effect of framing to be significant only for verbal information. Kim et al. (2020) examine how additional relative statistical information in public service advertisements, for example on the flu, can alter the perception of a threat and the stockpiling behavior in times of crises. Their results provide empirical evidence that the addition of easy-to-compare statistics can lead to a reduction of the perceived threat of COVID-19 and the stockpiling intention. The authors conclude that the compilation of statistics on COVID-19 can influence individuals by increasing or decreasing their perceived level of threat by adding or omitting additional statistics depending on the desired outcomes. In contrast to Kim et al. (2020), we explicitly focus on different COVID-19 metrics and study how individuals react to different metrics.

Taking both the main findings of these COVID-19 studies and the general insights on the so-called "framing"-effect (Levin et al., 1998) into account, it is unclear if individuals are

able to digest numeric information and how they react to different metrics. Given that for the first time in decades, a pandemic impacts peoples' behaviors and the novelty of the COVID-19-metrics, we do not have any *ex ante* expectations to which metrics people may react particularly sensitive or insensitive. In other words, we do not know *ex ante*, how consumers may react to various metrics.

3 Methodology and data

3.1 Data collection

We surveyed a random sample of German residents via *clickworker* to elicit consumer's WTP to participate in leisure time activities such as going to the movies, visiting a soccer match, attending a concert, or going on vacation, with or without being exposed to COVID-19 infection risk. We label the different activities "scenarios" in the following. *Clickworker* has an extensive database of consumers. In particular, *clickworker* is one of the largest providers of online panels in Germany and has a database of more than 2.8 million potential respondents. *Clickworker* monitors its panel for duplicate, fraudulent, and suspicious records and thereby ensures high data quality. Importantly, *clickworker* provides respondents a strong assurance of anonymity, which may improve the response rate and quality of the data collected (Durant et al., 2002; Pearlin, 1961; Podsakoff et al., 2003) and is compliant with general data protection regulations. We posted the link to our survey on the *clickworker marketplace* at the beginning of December 2020 and thereby made the survey available to all German consumers who were registered with the platform at that time. We did not further specify a target group, because the COVID-19 pandemic affects all consumers. The survey was implemented using oTree (Chen et al., 2016).

Overall, 803 individuals started the survey. To ensure that responses do not suffer from careless responding of participants, we included two attention checks (Oppenheimer et al., 2009). 403 participants completed the survey and passed the attention checks. We

exclude scenarios in which participants made inconsistent choices, yielding a total of 9653 participant-scenario observations. On average, participants needed 23 minutes to complete the questionnaire and received monetary compensation.

The unit of analysis is consumers' willingness to pay for participation in everyday life and amusement events. Table 1 summarizes the personal characteristics of the respondents in our sample. The largest group of participants is between 25 and 44 years old (29%), and the majority of the sample identifies as male (61%). About 21% work in jobs that tend to have a high daily number of close contacts (education, health- and residential care, retail, and gastronomy) and more than one third (36%) of either respondents or members of the respondents' household belong to COVID-19 at-risk groups.

Table 1: Personal characteristics of the respondents

	Percent
Age	
18 to 24	15.38
25 to 34	29.28
35 to 44	27.54
45 to 54	13.40
55 to 64	12.66
65+	1.74
Gender	
Female	38.71
Male	61.29
Employment	
Close Contact Job	20.84
Other	79.16
Risk group	
Yes	36.23
No	63.77

3.2 Measures

We are interested in participants' WTP to take part in daily life during a pandemic while being exposed to COVID-19 infection risk or not. To this end, participants were

shown one randomly drawn COVID-19 metric in five different scenarios, and had to indicate their WTP to take part in the event for various randomly drawn realizations of the respective metric. Randomization of COVID-19 metrics and scenarios mitigates the concern that participants' choices were driven by scenario-specific considerations.² Metrics were selected based on German media coverage. The survey included definitions of relevant government mandated COVID-19 transmission, protective measures, and COVID-19 metrics to make sure that all participants had the same understanding of the metrics. COVID-19 metrics domains were based on the observed case numbers in Germany during the pandemic prior to the study. To take into account that a single metric may have to be considered in context to provide a thorough overview of infection incidence, we also showed participants a combination of metrics in five additional scenarios. In this case, we only altered realizations of one random metric to be able to attribute the change in WTP to the varying metric. As our conclusions remain valid for both individual and aggregate measures, we do not report them separately.

COVID-19 metrics. We include the following metrics: Case-fatality ratio, confirmed cases, confirmed deaths, hospital systems load, incidence rate, intensive care cases, probable cases, reproduction rate, traffic light model, and unrecorded cases. Detailed definitions of the metrics can be found in Table A.1 in the Appendix. These definitions were available to our participants at all times during the survey.

Willingness to pay (WTP). We measure WTP as the monetary amount people are willing to pay to participate in leisure activities during the pandemic. Given that going to the movies has a lower price tag than going on vacation, we use different price levels in the survey. For our analyses, we adapt the WTP to a scale from zero to ten. The average (median) WTP is 4.93 (5), with a standard deviation of 3.63 (see Table 2), indicating an important variation across our sample.

Controls. We control for individuals' overall *risk propensity* using the one-item scale by Dohmen et al. (2011), for their *social responsibility* using a short version of the scale by Berkowitz and Daniels (1964) (Cronbachs alpha = 0.83), and for their perceived usefulness

²We additionally include scenario-fixed effects to further alleviate the concern.

of the five most prominent government-mandated COVID-19 transmission and *protective measures*, the “AHA+L+A” guiding principles (distance, hygiene, masks, ventilation, and contact tracing app), on a 5-item Likert scale ranging from 0 to 4 that we aggregate using the average (Cronbachs alpha = 0.75). Especially among younger people, *social responsibility* has been reported to be a strong motivator for social distancing (Oosterhoff et al., 2020). Also, individual risk preferences are linked to health behaviors (Himmler et al., 2020).

Table 2 provides summary statistics of the variables used in our study. On average, participants in our study agree with the protective measures (mean = 3.18; median = 3.4; SD = 0.68), and are fairly socially responsible (mean = 3.12; median = 3.13; SD = 0.52). The average risk propensity of participants is 5.12 (median 5; SD 2.31).

Table 2: Summary statistics

	N	Mean	SD	Min	25	Median	75	Maximum
Willingness to pay	9653	4.9322	3.6331	0.000	1.0000	5.0000	8.000	10.0000
Protective measures	9653	3.1804	0.6776	0.200	2.8000	3.4000	3.600	4.0000
Social responsibility	9653	3.1199	0.5234	1.375	2.8750	3.1250	3.500	4.0000
Risk propensity	9653	5.1248	2.3071	1.000	3.0000	5.0000	7.000	10.0000

Variable definitions can be found in Table A.1 in the Appendix. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

We additionally control for individuals’ age, gender, close contact jobs, and at-risk groups using our demographic variables (see Table 1).

3.3 Model estimation

We estimate the following main model using a standard ordinary least squares (OLS) model with robust standard errors (MacKinnon and White, 1985):

$$WTP_i = \alpha + \beta \text{ case numbers}_i + \sum_{j=1}^J \gamma_j \text{ controls}_{ij} + \epsilon_i \tag{1}$$

Control variables include demographic controls of the consumer (gender, age, risk propensity, etc.). We also control for scenario-fixed effects to address the possibility that one scenario may be more or less appealing than the others.

4 Results

4.1 Pearson's correlation matrix

We begin our analysis by observing the bivariate correlations between our variables of interest. Table 3 reports the Pearson correlations. We observe a strong negative correlation between WTP and case numbers with a correlation coefficient of -0.3031 . We also observe a significantly negative correlation between WTP and all COVID-19 metrics. The estimates range between -0.2376 for the case-fatality ratio and -0.4585 for the traffic light model. Thus, bivariate correlations provide initial support for our hypotheses that (i) WTP is a decreasing function of the spread of COVID-19, and (ii) that individuals react with differing sensitivity to various COVID-19 metrics.

We also observe positive correlations between WTP and consumers' risk preferences (0.054), and a negative correlation between WTP and social responsibility (-0.034), risk group (-0.047), and protective measures (-0.126). The negative correlation between WTP and social responsibility is consistent with prior evidence reported by Oosterhoff et al. (2020). We also find significant and negative correlations between WTP and a dummy variable that indicates participants who identify as female (-0.030) and age (-0.023). Interestingly, we do not observe a significant correlation between WTP and close contact job. We also do not observe any significant correlations between case numbers or any of our COVID-19 metrics and our control variables, indicating that the relationship between the WTP and the COVID-metrics is not affected by our control variables.

Table 3: Pearson’s correlation table

	Willingness to pay	Case numbers	Reproduction rate	Incidence rate	Intensive care cases	Confirmed deaths	Confirmed cases	Case-Fatality ratio	Unrecorded cases	Probable cases	Hospital systems load	Traffic light model	Female	Age	Protective measures	Social responsibility	Close contact jobs	Risk group	
Case numbers	-0.3031***																		
Reproduction rate	-0.2448***																		
Incidence rate	-0.3385***																		
Intensive care cases	-0.2603***																		
Confirmed deaths	-0.2492***																		
Confirmed cases	-0.3199***																		
Case-fatality ratio	-0.2376***																		
Unrecorded cases	-0.3852***																		
Probable cases	-0.4291***																		
Hospital systems load	-0.3053***																		
Traffic light model	-0.4585***																		
Female	-0.0302**	-0.0039	0.0014	-0.0076	0.0031	0.0069	0.0043	-0.0054	-0.0083	0.0087	0.0249	0.0000							
Age	-0.0231*	-0.0005	0.0139	0.0041	0.0079	0.0148	-0.0052	-0.0036	0.0032	-0.0014	0.0010	0.0000	-0.0515***						
Protective measures	-0.1255***	-0.0162	0.0060	0.0023	-0.0135	-0.0006	-0.0048	0.0035	-0.0036	0.0090	0.0038	-0.0000	0.1184***	0.0668***					
Social responsibility	-0.0340***	-0.0061	-0.0064	-0.0010	-0.0025	-0.0086	-0.0010	0.0068	-0.0173	0.0136	0.0086	0.0000	0.0685***	0.2014***	0.2602***				
Close contact job	0.0156	0.0004	0.0062	0.0063	-0.0032	0.0046	0.0037	-0.0043	-0.0087	0.0125	0.0127	0.0000	0.1296***	0.0644***	-0.0301**	-0.0244*			
Risk group	-0.0469***	0.0038	-0.0044	0.0056	-0.0091	0.0027	0.0028	0.0110	0.0082	-0.0053	0.0053	0.0000	-0.0037	0.1183***	0.1072***	-0.0116	-0.0551***		
Risk propensity	0.0539***	0.0027	-0.0074	0.0028	0.0119	0.0099	0.0016	0.0080	0.0111	0.0061	-0.0204	0.0000	-0.1821***	-0.0829***	-0.1645***	-0.0659***	-0.0482***	-0.0639***	

Variable definitions can be found in Table A.1 in the Appendix. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

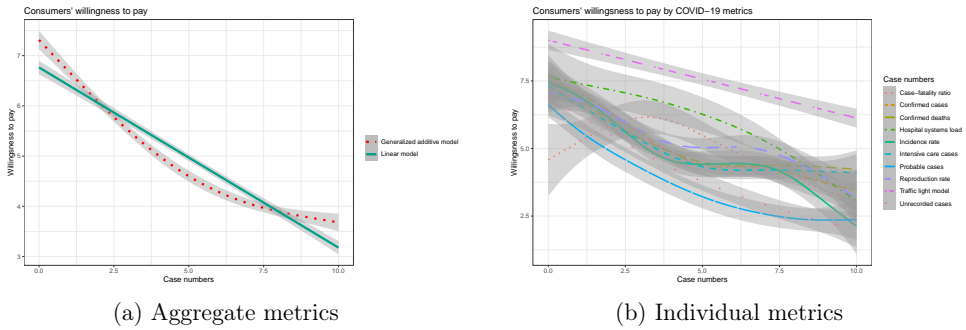
4.2 Hypothesis testing

We first test the hypothesis that WTP is a decreasing function of the spread of COVID-19. To this end, we jointly analyze all metrics. Given different levels of metrics, we adapt all metrics to a scale from zero to ten, thereby ensuring comparability. To conduct formal tests of the overall negative relationship, we estimate Equation (1). The negative coefficient of -0.41 (t -statistic of -36.48) in Table 4 indicates a negative correlation between case numbers and WTP. This is consistent with our expectations, as an increase in case numbers should be associated with a decrease in WTP or, in other words, an increased perceived infection risk.

Turning briefly to the control variables, we observe a significant positive coefficient for our measure of risk propensity (0.07, t -statistic of 4.64), showing that these individuals are more likely to take the risk of getting infected and are generally willing to pay more to take part in everyday-life. Negative and significant coefficients are reported for females (-0.12, t -statistic of -1.85), individuals in the age group 45+ (-0.22, t -statistic of -2.99), perceived usefulness of protective measures (-0.64, t -statistic of -11.83) and at-risk groups (-0.13, t -statistic of -1.92), indicating more risk aversion with regard to COVID-19 for these individuals. Overall, control variables are in line with expectations, and, where applicable, consistent with previous evidence in the literature finding, for example, that women are more likely than men to reduce their social interactions to decelerate the spread of the pandemic and therefore are less willing to take the risks of getting infected (Fan et al., 2020).

We visualize the relationship between consumers' WTP and the spread of COVID-19 in Figure 1. Panel A shows the linear fit with a negative slope and a generalized additive fit suggesting a non-linear relationship for the aggregated metrics. The non-linear fit provides first evidence in support of a decreasing marginal sensitivity. Panel B shows generalized additive models for individual metrics and highlights (i) that different metrics yield different WTP and (ii) marginal sensitivities vary across metrics, but are mostly non-linear.

Figure 1: COVID-19 metrics and consumers' willingness to pay



Covid Economics 83, 2 July 2021: 183-205

To formally test whether different COVID-19 metrics are perceived differently, we provide regression results on individual metrics in Table 5. The results support the graphical analysis. Case number estimates are negative and statistically significant, but with different effect sizes. With a coefficient of -0.59 (t -statistic of -16.48) we observe the largest effect size for the incidence rate. The smallest effect size belongs to the traffic light model with a coefficient of -0.29 (t -statistic of -10.26). Differences between coefficients are, with one exception (confirmed deaths & reproduction rate), statistically significant according to Chow's test.

Finally, we shed light on potential breakpoints in individuals' sensitivities to case numbers. We fit piece-wise linear relationships and estimate breakpoints according to Muggeo (2003). Essentially, we identify four different shapes and display these in Figure 2. The WTP for reproduction rate and incidence rate show shapes that align with natural and statutory thresholds. In particular, a reproduction rate of less than one implies a reduction in infection rates, providing a rationale for the breakpoints observed in Panel A. Similarly, at the time of the survey, Germany's most important policy-threshold was 50 infections per 100,000 inhabitants, which explains the breakpoint at 50 infections in Panel B. Confirmed cases level off at just about 10,000 confirmed COVID-19 infections per day, as an increase above 10,000 only has minimal impact on individuals' WTP (Slopes: -0.3631 vs -0.0727 (per 1000 confirmed cases); t -statistics of -6.45 and -1.65). Similar findings hold for confirmed deaths (breakpoint: 193) and intensive care cases

Table 4: COVID-19 reporting and willingness to pay

<i>Dependent variable:</i>	
Willingness to pay	
Constant	10.9577*** (35.2372)
Case numbers	-0.4093*** (-36.4831)
Female	-0.1247* (-1.8459)
Age: 45+	-0.2182*** (-2.9918)
Protective measures	-0.6411*** (-11.8301)
Social responsibility	0.0031 (0.0477)
Close contact jobs	0.1102 (1.3584)
Risk group	-0.1276* (-1.9200)
Risk propensity	0.0673*** (4.6383)
Scenario-Dummies	Yes
Observations	9,653
Adjusted R ²	0.2777
Residual Std. Error	3.0877 (df = 9630)
F Statistic	169.6898*** (df = 22; 9630)

Regression coefficients are presented with *t*-statistics in parentheses. Standard errors are robust. *p<0.1; **p<0.05; ***p<0.01. Variable definitions can be found in Table A.1 in the Appendix.

Table 5: Case numbers and consumers' willingness to pay

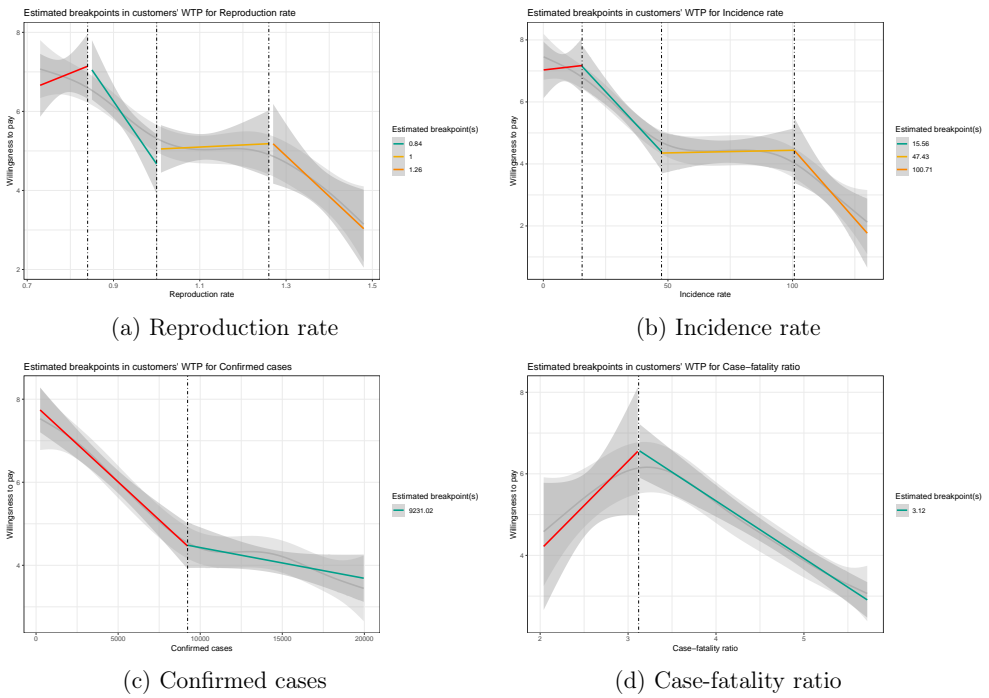
<i>Dependent variable: Willingness to pay</i>					
COVID-19 metric, main explanatory variable:					
	Reproduction rate	Incidence rate	Intensive care cases	Confirmed deaths	Confirmed cases
	(1)	(2)	(3)	(4)	(5)
Case numbers	-0.4855*** (-11.9271)	-0.5912*** (-16.4764)	-0.4247*** (-11.4550)	-0.4618*** (-12.1503)	-0.5606*** (-14.8308)
Controls	Yes	Yes	Yes	Yes	Yes
Scenario-dummies	Yes	Yes	Yes	Yes	Yes
Observations	1,165	1,190	1,186	1,211	1,150
Adjusted R ²	0.2742	0.3371	0.2400	0.2771	0.3024
Residual std. error	2.9931 (df = 1143)	2.9655 (df = 1168)	2.9821 (df = 1164)	3.0460 (df = 1189)	2.9792 (df = 1128)
F statistic	21.9398*** (df = 21; 1143)	29.7865*** (df = 21; 1168)	18.8224*** (df = 21; 1164)	23.0836*** (df = 21; 1189)	24.7171*** (df = 21; 1128)

<i>Dependent variable: Willingness to pay</i>					
COVID-19 metric, main explanatory variable:					
	Case-fatality ratio	Unrecorded cases	Probable cases	Hospital systems load	Traffic light model
	(6)	(7)	(8)	(9)	(10)
Case numbers	-0.3292*** (-7.5694)	-0.4898*** (-14.1085)	-0.4682*** (-14.7063)	-0.5001*** (-11.0083)	-0.2885*** (-10.2557)
Controls	Yes	Yes	Yes	Yes	Yes
Scenario-dummies	Yes	Yes	Yes	Yes	No
Observations	815	865	800	860	411
Adjusted R ²	0.2638	0.3637	0.3505	0.3071	0.2313
Residual std. error	3.2049 (df = 798)	2.8965 (df = 848)	2.9702 (df = 783)	3.0996 (df = 843)	2.2549 (df = 402)
F statistic	19.2303*** (df = 16; 798)	31.8623*** (df = 16; 848)	27.9517*** (df = 16; 783)	24.7992*** (df = 16; 843)	16.4242*** (df = 8; 402)

Regression coefficients are presented with *t*-statistics in parentheses. Standard errors are robust. *p<0.1; **p<0.05; ***p<0.01. Variable definitions can be found in Table A.1 in the Appendix.

(breakpoint: 1571). Panel D shows the case-fatality ratio. For values exceeding 3.12%, the WTP decreases linearly, in line with expectations. Probable and unrecorded cases as well as the traffic light model and the hospital systems load show a similar pattern.

Figure 2: Case numbers and consumers' willingness to pay: Breakpoint analyses



Covid Economics 83, 2 July 2021: 183-205

5 Discussion

The effort to control the spread of COVID-19 critically hinges on individuals' behavior in accordance to the recommended hygienic etiquette and social distancing guidelines. In this paper, we investigated how individuals react to different COVID-19 metrics and various realizations of these metrics. Using a survey experiment, we ask consumers how willing they are to pay for participating in common scenarios of daily life without being exposed to the risk of infection, given a certain level of the pandemic spread. In line with expectations, an increase in case numbers is associated with an increase in WTP, which we interpret as a sign of higher perceived infection risk. However, consistent with the “framing”-effect (Tversky and Kahneman, 1985) participants react with different sensitivities to different metrics. Participants react most sensitive to absolute numbers such as the incidence rate and confirmed cases, and least sensitive to the composite traffic light model and the case-fatality ratio. This may be explained by the fact that absolute numbers are generally easier to understand compared to probabilities and ratios, and thus invoke a swifter change in behavior (Gigerenzer and Hoffrage, 1995). At the same time, previous literature indicates that precise numbers are perceived as an indicator of higher competence in communication (Welsh et al., 2011) and that individuals are more likely to follow a precise advisor (see, e.g., Schultze and Loschelder, 2020). Thus, the comparably low sensitivities of the traffic light model may be explained by its three-dimensionality. While it has been introduced to condense various statistics into an overall indicator to facilitate navigating an excess of information on COVID-19 reporting, the aggregation of information may actually blur its meaningfulness.

We also show that perceived infection risk is not a linear function of case numbers. Instead, individuals tend to react less sensitive to certain values, once a particular threshold is exceeded. While, for example, this threshold seems to reflect the natural limit of one for the reproduction rate, the threshold for the incidence rate close to 50 is very likely explained by the political discussion that revolved around this value in the time the survey was conducted (see, e.g., Tversky and Kahneman, 1974, on the effect of anchoring). In

this aspect, our results are consistent with the “fatalism effect” found by Akesson et al. (2020), who find that the more infectious people believe that COVID-19 is, the less willing they are to take social distancing measures.

Thus, our findings have three important implications. First, policy interventions matter. Confirmed cases and incidence rate are co-dependent measures, yet they show distinctly different shapes. This difference is likely driven by the political focus on particular thresholds. Second, framing matters. People react differently to different COVID-19 metrics. Thus, communicators have to take the sensitivity of particular metrics into account. Choosing one metric over another may nudge individuals towards infection prevention behavior. Third, as individuals seem to become insensitive for high levels of some metrics, stricter policy interventions may become necessary under severe states of the pandemic.

Compliance with ethical standards

Funding Regina Ortmann gratefully acknowledges financial support by the German Research Foundation (DFG) - Collaborative Research Center (SFB/TRR) Project-ID 403041268 - TRR 266 Accounting for Transparency. Any errors, misrepresentations, and omissions are our own

Conflicts of interest All authors declare that they have no conflict of interest

Availability of data The data generated during the research process are available on request.

Ethics approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This study was reviewed and approved by the institutional ethics review board of a participating institution.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

- Akesson, Jesper, Sam Ashworth-Hayes, Robert Hahn, Robert D. Metcalfe, and Itzhak Rasooly, 2020, Fatalism, beliefs, and behaviors during the COVID-19 pandemic, NBER Working Papers 27245, National Bureau of Economic Research, Inc.
- Berkowitz, Leonard, and Louise R Daniels, 1964, Affecting the salience of the social responsibility norm: effects of past help on the response to dependency relationships., *The Journal of Abnormal and Social Psychology* 68, 275.
- Bierhoff, H-W, 2012, *SV-Skala der sozialen Verantwortung* (ZPID (Leibniz Institute for Psychology Information)–Testarchiv).
- Blackman, Allen, and Bridget Hoffmann, 2021, Diminishing returns: Nudging covid-19 prevention among colombian young adults, *IDB Working Paper Series* IDB-WP-1217.
- Chen, Daniel L., Martin Schonger, and Chris Wickens, 2016, otree—an open-source platform for laboratory, online, and field experiments, *Journal of Behavioral and Experimental Finance* 9, 88 – 97.
- Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G Wagner, 2011, Individual risk attitudes: Measurement, determinants, and behavioral consequences, *Journal of the European Economic Association* 9, 522–550.
- Durant, Lauren E., Michael P. Carey, and Kerstin E. E. Schroder, 2002, Effects of anonymity, gender, and erotophilia on the quality of data obtained from self-reports of socially sensitive behaviors, *Journal of Behavioral Medicine* 25, 439–467.
- Fan, Ying, A Yeşim Orhun, and Dana Turjeman, 2020, Heterogeneous actions, beliefs, constraints and risk tolerance during the covid-19 pandemic, Technical report, National Bureau of Economic Research.
- Gigerenzer, Gerd, Wolfgang Gaissmaier, Elke Kurz-Milcke, Lisa M Schwartz, and Steven Woloshin, 2007, Helping doctors and patients make sense of health statistics, *Psychological Science in the Public Interest* 8, 53–96.
- Gigerenzer, Gerd, and Ulrich Hoffrage, 1995, How to improve bayesian reasoning without instruction: frequency formats., *Psychological Review* 102, 684.
- Himmler, Sebastian, Job van Exel, Meg Perry-Duxbury, and Werner Brouwer, 2020, Willingness to pay for an early warning system for infectious diseases, *The European Journal of Health Economics* 21, 763–773.
- Kim, Jungkeun, Marilyn Giroux, Hector Gonzalez-Jimenez, Seongsoo Jang, Seongseop (Sam) Kim, Jooyoung Park, Jae-Eun Kim, Jacob C. Lee, and Yung Kyun Choi, 2020, Nudging to reduce the perceived threat of coronavirus and stockpiling intention, *Journal of Advertising* 49, 633–647.
- Levin, Irwin P., Sandra L. Schneider, and Gary J. Gaeth, 1998, All frames are not created equal: A typology and critical analysis of framing effects, *Organizational Behavior and Human Decision Processes* 76, 149–188.

- Li, Meng, and Gretchen B. Chapman, 2013, Nudge to health: Harnessing decision research to promote health behavior, *Social and Personality Psychology Compass* 7, 187–198.
- MacKinnon, James G, and Halbert White, 1985, Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties, *Journal of Econometrics* 29, 305 – 325.
- Muggeo, Vito MR, 2003, Estimating regression models with unknown break-points, *Statistics in medicine* 22, 3055–3071.
- Oosterhoff, Benjamin, Cara A Palmer, Jenna Wilson, and Natalie Shook, 2020, Adolescents' motivations to engage in social distancing during the covid-19 pandemic: associations with mental and social health, *Journal of Adolescent Health* 67, 179–185.
- Oppenheimer, Daniel M., Tom Meyvis, and Nicolas Davidenko, 2009, Instructional manipulation checks: Detecting satisficing to increase statistical power, *Journal of Experimental Social Psychology* 45, 867 – 872.
- Pearlin, Leonard I., 1961, The appeals of anonymity in questionnaire response, *The Public Opinion Quarterly* 25, 640–647.
- Perry-Duxbury, Meg, Job van Exel, and Werner Brouwer, 2019, How to value safety in economic evaluations in health care? a review of applications in different sectors, *The European Journal of Health Economics* 20, 1041–1061.
- Podsakoff, Philip M., Scott B. MacKenzie, Jeong-Yeon Lee, and Nathan P. Podsakoff, 2003, Common method biases in behavioral research: A critical review of the literature and recommended remedies., *Journal of Applied Psychology* 88, 879–903.
- Ries, Tonia E., Richard Edelman, Megan Van Someren, Lee Maicon, Michele Anderson, Jackie Cooper, Judy John, Kary Laskin, Ruth Warder, Lisa Ross, Nancy Jeffrey, Sarah Adkins, Esther Choi, David M. Bersoff, Antoine Harary, Louise Turner, Cody Armstrong, Giuseppe Bovenzi, Nick Maxwell, Abbey Fox, and John Zamites, 2020, 2020 Edelman Trust Barometer Special Report: Trust and the Coronavirus, Technical report, Edelman.
- Sasaki, Shusaku, Hirofumi Kurokawa, and Fumio Ohtake, 2020, Short-term responses to nudge-based messages for preventing the spread of covid-19 infection: Intention, behavior, and life satisfaction, *Osaka University Discussion Papers in Economics and Business* 20-11.
- Schultze, Thomas, and David D. Loschelder, 2020, How numeric advice precision affects advice taking, *Journal of Behavioral Decision Making* 34, 303–310.
- Shen, Luxi, and Christopher K. Hsee, 2017, Numerical nudging: Using an accelerating score to enhance performance, *Psychological Science* 28, 1077–1086.
- Slovic, Paul E, 2000, *The perception of risk* (Earthscan publications).
- Tversky, Amos, and Daniel Kahneman, 1974, Judgment under uncertainty: Heuristics and biases, *Science* 185, 1124–1131.

- Tversky, Amos, and Daniel Kahneman, 1985, *The Framing of Decisions and the Psychology of Choice*, 25–41 (Springer US, Boston, MA).
- Visschers, Vivianne HM, Ree M Meertens, Wim WF Passchier, and Nanne NK De Vries, 2009, Probability information in risk communication: a review of the research literature, *Risk Analysis: An International Journal* 29, 267–287.
- Welkenhuysen, Myriam, Gerry Evers-Kiebooms, and Gery d’Ydewalle, 2001, The language of uncertainty in genetic risk communication: framing and verbal versus numerical information, *Patient Education and Counseling* 43, 179–187.
- Welsh, Matthew, Daniel Navarro, and Steve Begg, 2011, Number preference, precision and implicit confidence, Proceedings of the 33rd Annual Meeting of the Cognitive Science Society, Cognitive Science Society (L. Carlson and C. Hoelscher and T. Shipley).

Table A.1: Variable definitions

Variable	Definition
<i>Dependent variables</i>	
Willingness to pay (WTP)	Monetary amount people are willing to pay to participate in leisure activities during the pandemic.
<i>COVID-19 metrics</i>	
Case-fatality ratio	Ratio of overall COVID-19 related deaths to confirmed cases.
Confirmed cases	Number of nationwide daily new infections.
Confirmed deaths	Number of individuals nationwide who died with or due to a COVID-19 infection in the past 24 hours.
Hospital systems load	Hospital systems load at participants area of residence.
Incidence rate	Rate of new COVID-19 cases within 7-days per 100,000 inhabitants in the respondent's area of residence.
Intensive care cases	Number of individuals nationwide currently requiring intensive medical care due to a COVID-19 infection.
Probable cases	Number of probable cases (per 100,000 population) in the area of residence. Suspicion of COVID-19 is established when individuals fulfil at least one of the following two constellations: i) Individuals with any symptoms consistent with COVID-19 AND contact with a confirmed case of COVID-19; ii) Occurrence of two or more cases of pneumonia in a medical facility, nursing home, or home for the elderly in which an epidemic link is likely or suspected, even in the absence of pathogen detection. Diagnostic clarification should be performed for these individuals.
Reproduction rate	The nationwide 7-day reproduction rate. It is the number of COVID-19 cases an infected person will cause during the infectious period. A value below 1 describes a decrease in case numbers, while a rate above 1 describes an increase in case numbers.
Traffic light model	The Corona traffic light risk assessment incorporates both spread risk (= risk to public health from the spread of COVID-19) and systemic risk (= risk of overloading the healthcare system with COVID-19 patients). Choices include red, yellow, and green.
Unrecorded cases	The suspected number of unrecorded cases of infected people (per 100,000 population) in the area of residence.
<i>Control variables</i>	
Age	6-level scale measuring the individuals' age.
Close contact job	Dummy variable that takes the value of one for individuals who work in a job with close contact to other people and zero otherwise.
Female	Dummy variable that takes the value of one for females and zero otherwise.
Protective measures	Perceived usefulness of the five most prominent government-mandated COVID-19 transmission and <i>protective measures</i> , the "AHA+L+A" guiding principles (distance, hygiene, masks, ventilation, and contact tracing app), on a 5-level Likert scale ranging from 0 to 4.
Risk group	Dummy variable that takes the value of one for individuals who consider themselves as belonging to a COVID-19 at-risk group and zero otherwise.

Continued on next page

Table A.1 – continued

Variable	Definition
Risk propensity	11-level scale measuring individual risk attitude based on Dohmen et al. (2011).
Social responsibility	8-item scale using a short version of the responsibility scale by Berkowitz and Daniels (1964). German translations are based on Bierhoff (2012).