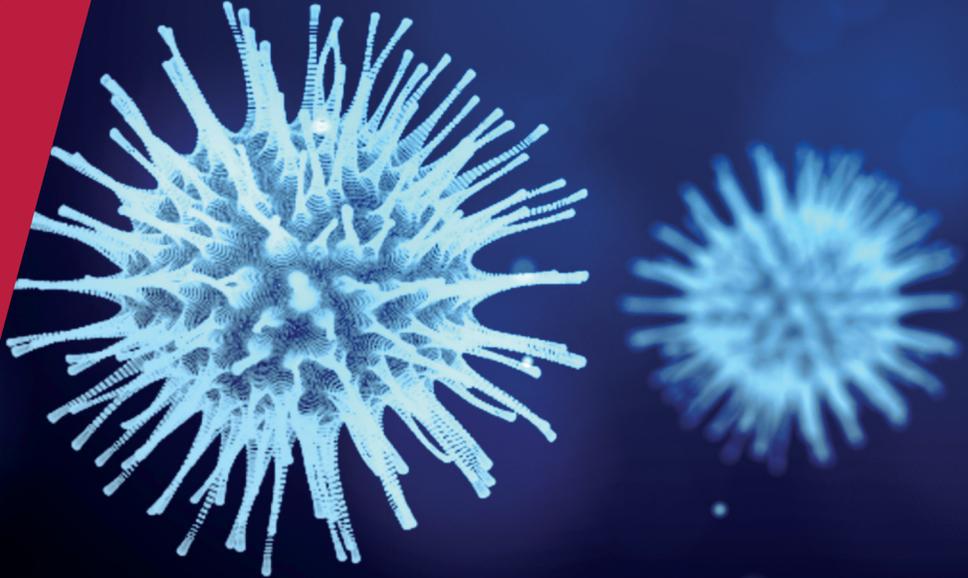


**CENTRE FOR
ECONOMIC
POLICY
RESEARCH**

CEPR PRESS



COVID ECONOMICS
VETTED AND REAL-TIME PAPERS

ISSUE 79
27 MAY 2021

EXPECTED STOCK VOLATILITY

Viral V. Acharya, Yang Liu
and Yunhui Zhao

**A GAME THEORETIC VIEW
OF COVID**

Santiago Forero-Alvarado, Nicolás
Moreno-Arias and Juan J. Ospina-Tejeiro

WHO TO VACCINATE FIRST?

Rikard Forslid and Mathias Herzing

PARENTS UNDER STRESS

Simone Schüller and Hannah S. Steinberg

GLOBAL STUDENT SURVEY

David A. Jaeger, Jaime Arellano-Bover,
Krzysztof Karbownik, Marta Martínez-
Matute, John M. Nunley, R. Alan Seals
et al.

FED COMMUNICATION

Jonathan Benchimol, Sophia Kazinnik
and Yossi Saadon

Covid Economics

Vetted and Real-Time Papers

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

Founder: Beatrice Weder di Mauro, President of CEPR

Editor: Charles Wyplosz, Graduate Institute Geneva and CEPR

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

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

The Centre for Economic Policy Research (CEPR)

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

Chair of the Board

Sir Charlie Bean

Founder and Honorary President

Richard Portes

President

Beatrice Weder di Mauro

Vice Presidents

Maristella Botticini

Ugo Panizza

Philippe Martin

Hélène Rey

Chief Executive Officer

Tessa Ogden

Editorial Board

Beatrice Weder di Mauro, CEPR

Charles Wyplosz, Graduate Institute Geneva and CEPR

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

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 79, 27 May 2021

Contents

COVID-19 containment measures and expected stock volatility: High-frequency evidence from selected advanced economies <i>Viral V. Acharya, Yang Liu and Yunhui Zhao</i>	1
Humans against virus or humans against humans: A game theory approach to the COVID-19 pandemic <i>Santiago Forero-Alvarado, Nicolás Moreno-Arias and Juan J. Ospina-Tejeiro</i>	41
Who to vaccinate first: Some important trade-offs <i>Rikard Forslid and Mathias Herzing</i>	93
Parents under stress: Evaluating emergency childcare policies during the first COVID-19 lockdown in Germany <i>Simone Schüller and Hannah S. Steinberg</i>	117
The Global COVID-19 Student Survey: First wave results <i>David A. Jaeger, Jaime Arellano-Bover, Krzysztof Karbownik, Marta Martínez-Matute, John M. Nunley, R. Alan Seals, et al.</i>	152
Federal Reserve communication and the COVID-19 pandemic <i>Jonathan Benchimol, Sophia Kazinnik and Yossi Saadon</i>	218

COVID-19 containment measures and expected stock volatility: High-frequency evidence from selected advanced economies¹

Viral V. Acharya,² Yang Liu³ and Yunhui Zhao⁴

Date submitted: 13 May 2021; Date accepted: 21 May 2021

We study the effect of COVID-19 containment measures on expected stock price volatility in some advanced economies, using event studies with hand-collected minute-level data and panel regressions with daily data. We find that six-month-ahead volatility indices dropped following announcements of initial or re-imposed lockdowns, and that they did not drop significantly following the easing of lockdowns. Such patterns are not as strong for three-month-ahead expected volatility and generally absent for one-month-ahead expected volatility. These results provide suggestive evidence for the existence of an intertemporal trade-off: although stringent containment measures cause short-term economic disruptions, they may reduce medium-term uncertainty (reflected in expected stock volatility) by boosting markets' confidence that the outbreak would be under control more quickly.

- 1 We are grateful for the helpful discussions with Alberto Behar (IMF, unless otherwise stated), Jorge Chan-Lau, Martin Cihak, Rupa Duttagupta, Alan Xiaochen Feng, Gita Gopinath, Burcu Hacibedel, Sandile Hlatshwayo, Anna Ilyina, Timothy Johnson (UIUC), Andras Komaromi, Romain Lafarguette, Wojciech Maliszewski, Jonathan Ostry, Chris Redl, Agustin Roitman, Eddy Tam (Oxford), Yannick Timmer, Wentao Xiong (Goldman Sachs), Daria Zakharova, Tao Zhang, and participants at the IMF interdepartmental surveillance meeting in October 2020 and at the SPARK series. We also thank William Kunxiang Diao for his excellent support on updating the volatility data, and Tiana Wang for her excellent research assistance. The views expressed here are those of the authors and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.
- 2 Professor, NYU Stern School of Business.
- 3 Data Scientist, International Monetary Fund.
- 4 Economist, International Monetary Fund.

Copyright: Viral V. Acharya, Yang Liu and Yunhui Zhao

1 Introduction

This paper aims at illustrating an economic benefit of COVID-19 containment measures reflected in lower expected stock price volatility. On such measures, many discussions focus disproportionately on their *humanitarian* benefit and characterize them as saving lives at the cost of sacrificing the economy or livelihoods. This perception has been the main reason for governments' reluctance to impose lockdowns and their rush to reopen the economies. However, despite the short-term economic disruptions, containment measures may also have a significant *economic* benefit, particularly over the medium term—they could help contain the outbreak, buy time for vaccination rollout and herd immunity, reduce uncertainty, and mitigate health/financial constraints.¹ These arguments are consistent with some market participants' views that investors' optimism can sustain only if they are confident that the outbreak is under control.² They are also in line with the views of some policymakers.³

However, there are at least two challenges for quantifying the economic benefits of containment measures. First, it is hard to distinguish between short-term costs and medium-term benefits of containment measures, given that the observed macroeconomic data reflect both the costs and benefits. Second, most macroeconomic data are available at relatively low frequency (monthly or quarterly), making the identification of the effects of containment measures difficult. To overcome these challenges, we proxy the “medium-term uncertainty” with the six-month-ahead stock price volatility indices implied by options prices; we then use the reduction in these indices to measure the economic benefit of containment measures. We also conduct the analysis separately for the initial tightening stage, the easing/reopening stage, and the retightening stage to account for the different impacts of containment measures at different stages. Specifically, two complementary approaches are employed in each stage.

First, using minute-level data, we conduct event studies for an *extreme* containment measure—lockdown. To do so, we take a deep dive into multiple information sources (such as English newspapers, local language newspapers, tweets, and government websites) and manually identify the minute when a COVID-related lockdown or reopening was announced. We then conduct event studies by comparing the post-announcement actual volatility indices with their counterfactuals. Due to limitations of the volatility data, only the US, Italy, Germany, and the eurozone are covered. To focus on systemic events only, we study the most significant

¹ Although containment measures may be less effective in countries with large informal sectors, the aforementioned benefit still exists in all countries and needs to be considered to properly assess the trade-off associated with containment measures.

² “US stocks in sharp late rally on hopes virus is slowing,” *Financial Times*, April 7, 2020.

³ One example is IMF Managing Director's [statement](#) that “*the faster the virus stops, the quicker and stronger the recovery will be.*”

lockdown/reopening announcements by the aforementioned countries/regions.⁴ Despite the small number of events, this approach has the advantage of mitigating the omitted variable bias commonly encountered in regression-based approaches. Various models are used to construct the counterfactuals, such as GARCH and EGARCH, and additional variables are controlled to account for other forces (e.g., fiscal stimulus) that may be at play at the time of the announcements.

Second, using daily data, we conduct panel regressions for *broader* containment measures (such as a moderate restriction on gathering) instead of a full-fledged lockdown. This approach regresses the daily volatility indices on a comprehensive stringency index of containment measures constructed by Oxford University for largely the same set of countries/regions as in the first approach. There are not many variations in the stringency index data due to the relatively infrequent policy measure changes, and there are many other forces that affect volatility. To address these concerns, we include the relevant stock price indices in the regressions in an attempt to control for the impacts of other driving forces in a parsimonious way. Since COVID-19 is a global shock, seemingly unrelated regressions are also conducted to account for the correlations among different countries and different volatility products.

Both approaches produce very similar results. First, during the initial tightening stage, stringent containment measures significantly reduce expected stock volatility, which directly supports our hypothesis stated above. Second, the easing of stringent containment measures is not associated with a significant reduction in expected volatility, contradicting the conventional wisdom (i.e., the easing of containment means less disruptions to the economy and lower uncertainty) and thus indirectly supporting our hypothesis. Third, during the retightening stage, more stringent containment policies are associated with lower expected volatility, although its statistical significance is lower than the initial tightening stage.⁵

In particular, the above results are the strongest for six-month-ahead expected volatility, not as strong for three-month-ahead expected volatility, and generally absent for one-month-ahead expected volatility (the left three panels in Figure 5). Although we do not study the volatility beyond the six-month horizon due to data limitations, the increasing significance of results over time seems informative. Taken together, the results provide suggestive evidence for the existence of an intertemporal trade-off: although stringent containment measures cause short-term economic

⁴ For the initial tightening stage, the events include (1) Trump's state-of-emergency declaration on March 13 that enabled states to impose lockdowns; (2) California's stay-at-home order on March 19, the very first state-wide lockdown in the US; (3) Italy's nationwide lockdown on March 9; (4) Germany's nationwide lockdown on March 16, which was widely regarded as a historic move by the press (for this event, we have data on both the domestic and eurozone-wide volatility); (5) France's nationwide lockdown on March 16.

⁵ This could reflect market participants' perception that the containment measures at the retightening stage are less stringent than the initial stage, which may be inadequate to contain the second waves and thus volatility did not drop as much.

disruptions, they may reduce medium-term uncertainty (reflected in expected stock volatility) by boosting markets' confidence that the outbreak would be under control more quickly.

The remainder of the paper is organized as follows. Section 2 reviews the related literature. Section 3 discusses the minute-level event studies, including the dataset, the methodology, and the results for the three stages. Section 4 discusses daily-level regressions. Section 5 concludes and discusses some policy implications.

2 Literature Review

Our paper relates to a growing literature on the effects of COVID-19 and containment measures. First, it relates to the literature on the COVID-era financial market responses. Beirne et al. (2020) find that emerging economies in Asia and Europe experienced the sharpest declines in stocks, bonds, and exchange rates due to COVID-19. Using data since January 1900, Baker et al. (2020) find that no previous infectious disease outbreak increased the US stock market volatility as forcefully as the COVID. Focusing on industry-specific realized volatility, Baek, Mohanty, and Glambsky (2020) find that changes in volatility are more sensitive to COVID news than to economic indicators. Using daily data and ARMA models, Cheng (2020) studies the futures of VIX (rather than the VIX itself, as we do) and finds that the VIX futures market underreacted to the growing risks of the pandemic during the early stages.⁶ Our paper complements these studies by focusing on the effect of containment measures and examining the minute-level, forward-looking volatility data right after a lockdown announcement, which is more likely to separate the effect of containment from that of other driving forces, such as the COVID outbreak itself.

Second, our paper relates to a large literature that points to a high economic *cost* of containment measures. For example, Deb et al. (2020a) find that containment measures are associated with a 15% decline in industrial production over a 30-day period. Kok (2020) finds that during the second quarter of 2020, “containment and closure policies” deducted about 8.6% (year-on-year) of GDP growth for advanced economies and 5.1% for Emerging Market and Developing Economies. However, some other studies suggest a mixed picture. Caselli et al. (2020) find that voluntary social distancing also contributed to short-term economic contractions. Goolsbee and Syverson (2020) find that legal shutdown orders account for only 7 of the 60 percentage-point decline of consumer visits to businesses, and that individual choices due to infection fears were a far more important factor. Chen et al. (2020) find that deterioration of economic conditions *preceded* the introduction of non-pharmaceutical interventions (NPIs). Aum, Lee, and Shin (2020b) find that at most half of the job losses in the US and the UK can be attributed to lockdowns. Amon, Ricco, and Smetters (2020) find that NPIs explain only about 15% of the decline in employment.

⁶ Jackwerth (2020) uses the derived distribution from option prices to discuss the market prediction of the future SP500 index in the COVID era (rather than the prediction of VIX as we do). A few other papers on stock market volatility include Zhang, Hu, and Ji (2020), Zaremba et al. (2020), and Haroon and Rizvi (2020).

Relatedly, in terms of the optimal lockdown policy, Alvarez, Argente, and Lippi (2020) models a planner who balances the fatality induced by the pandemic with the output costs of the lockdown policy. This is the standard life-livelihood trade-off, a framework adopted by a large number of studies. For example, under assumptions about the value of lives saved in the UK, Miles, Stedman, and Heald (2020) conclude that “the costs of continuing severe restrictions are so great relative to likely benefits in lives saved that a rapid easing in restrictions is now warranted.”⁷

Third, our paper relates to the literature on the *benefits* of containment measures. Deb et al. (2020b) find that such measures have been very effective in flattening the pandemic. Correia, Luck, and Verner (2020) analyze monthly data across US cities during the 1918 Flu Pandemic, and find that NPIs are associated with better economic outcomes in the medium term.⁸ Using a macroeconomic model calibrated to Korea and UK COVID dynamics, Aum, Lee, and Shin (2020a) find that a longer lockdown eventually mitigates the GDP loss, with a focus on the work-from-home channel, i.e., the lockdown lowers infections and induces people to switch from working from home (assumed to be less productive) to working on site.

Perhaps the two empirical papers most related to ours are by Sheridan et al. (2020) and Ashraf (2020). Using daily consumer spending data from a large bank in Scandinavia and exploiting an exogenous difference in COVID responses between Denmark and Sweden, Sheridan et al. (2020) find that social distancing laws may provide an economic benefit by reducing the economic activity of the low-risk population, lowering the overall prevalence of the virus in the society, and thus attenuating the COVID-induced drop in spending for high-risk individuals. Using daily stock market return data during January 22-April 17, 2020 from 77 countries, Ashraf (2020) finds that announcements of government social distancing measures have both a direct negative effect on stock market returns and an indirect positive effect through the reduction in COVID cases.

Our paper differs from these two in several dimensions. In terms of methodologies, in addition to using daily data, we employ minute-level event studies and account for other policies

⁷ Below are some other examples using this framework. Eichenbaum, Rebelo, and Trabandt (2020) find that the competitive equilibrium is not socially optimal due to externality, and that the best simple containment policy increases the severity of the recession but saves roughly half a million lives in the U.S. Jones, Philippon, and Venkateswaran (2020) find that private mitigation reduces the cumulative death rate by more than the planner does, albeit at the cost of a sharper drop in consumption. Hall, Jones, and Klenow (2020) estimate that the planner is willing to give up 41% of consumption for a full year to avoid the elevated mortality associated with the pandemic. Gourinchas (2020) concludes that “the measures that help solve the health crisis can make the economic crisis worse – at least in the short run.”

⁸ Relatedly, Fotiou and Lagerborg (2021) find that countries with previous SARS experience were able to both contain COVID-19 and mitigate lockdown-associated economic costs due to a “smart” containment strategy. Barro, Ursua, and Weng (2020) also quantify the medium-to-long-term effects by analyzing annual data for 48 countries. They find that the 1918 Flu Pandemic lowered real GDP by 6-8% in the typical country, which is suggested to be the upper bound of the effects of COVID-19.

(e.g., fiscal stimulus), providing a potentially cleaner identification. In terms of the scope, in addition to the initial lockdown stage, we also analyze the easing stage and the retightening stage.

3 Event Studies with Minute-Level Data

3.1 Data

Our data on the minute-level event times are hand-collected. Specifically, we take a deep dive into multiple information sources, such as English newspapers, local language newspapers, tweets of reporters, videos of the actual announcements, and government websites. We then manually identify the minute when a COVID-related lockdown or reopening was announced. In case the precise minute cannot be identified, we make our best estimate based on all available information. For example, the minute of France's reopening announcement is estimated using a three-step procedure (Appendix 1). For the announcements made outside of the trading hours, we treat the next opening minute as the event time; and to account for potentially higher fluctuations of the volatility indices right after the opening of markets, our counterfactual models have explicitly introduced a dummy variable for the first 30 minutes after the opening. The same treatment is applied to the last 30 minutes of the trading day.

As for the response variable of the event studies—medium-term expected volatility, we proxy it by the six-month-ahead options-based volatility indices. These indices are based on the core stock price index of the country/region (e.g., S&P 500 for the US), and they estimate the expected volatility by aggregating the weighted prices of the stock price index puts and calls over a wide range of strike prices (CBOE White Paper, 2019). In the case of the US, these are the six-month equivalents of the VIX Index, which measures the one-month-ahead expected volatility and is often referred to as the “fear gauge.” We choose three and six months because these horizons represent the “medium-term” to some extent and because data for longer horizons are not available in all countries we study (only the US has the one-year-ahead volatility index).

The minute-level data on expected volatility are from Bloomberg. Specifically, for the US events, we use the CBOE S&P 500 three-month and six-month expected volatility indices. For all events in Italy, Germany, and France, we use the Euro STOXX 50 expected volatility index, which is widely viewed as Europe's equivalent of the VIX in the US (See Smith, 2013, among others). For events in Germany, in addition to the eurozone-wide volatility index, we also use Germany's country-specific expected volatility index based on the DAX stock price index. However, no *intraday* expected volatility data are available for Italy, and no data *beyond one month* are available for France, so events in these two countries are only studied based on the eurozone-wide volatility index. The minute-level data on the underlying stock price indices from Bloomberg are also used in the construction of the counterfactual models.

Table 1 provides the summary statistics of the volatility and stock price index data. For all countries/regions, the data used in our event studies cover the business days from January 2, 2020

to October 29, 2020, around 210 days in total (the specific number of days varies slightly, depending on the particular country/region and on the maturity, i.e., six-month or three-month). To gauge the magnitude of the post-event change in volatility relative to the “usual” daily change, we also provide the mean of the daily changes across all days, where the daily change is defined as the highest volatility minus the lowest volatility observed during the day. Note that Table 1 does not distinguish between stages (initial tightening, easing, or retightening) because all stages use the same data to train the volatility prediction models.

Table 1. Summary Statistics of Event Study Data

	Number of obs	Mean	Sd	Number of days	Daily change mean	Daily change sd
Panel (A): Six-month volatility						
Volatility for S&P 500	84,849	31.1	8.3	210	1.9	2.1
Volatility for STOXX 50	105,647	29.5	9.9	213	1.9	2.2
Volatility for DAX	105,151	31.0	10.1	212	1.7	1.9
Panel (B): Three-month volatility						
Volatility for S&P 500	84,847	31.6	10.5	210	2.8	3.4
Volatility for STOXX 50	102,171	29.8	11.9	206	2.4	2.5
Volatility for DAX	101,663	30.0	11.7	205	2.0	2.2
Panel (C): One-month volatility						
Volatility for S&P 500	158,843	30.2	12.8	210	4.2	4.2
Volatility for STOXX 50	91,203	31.5	17.9	184	4.2	4.2
Volatility for DAX	90,889	33.0	17.4	183	4.1	4.0
Panel (D): Stock price indices						
S&P 500	85,050	3,136.0	284.7	210	52.3	38.4
STOXX 50	112,854	355.1	35.5	213	6.2	4.4
DAX	110,920	12,182.9	1,274.3	212	225.0	136.0

Sources: Bloomberg; Authors' calculations.⁹

3.2 Initial Tightening

In Spring 2020, triggered by the rapidly growing COVID cases, numerous Western countries announced strict nationwide lockdowns. Figure 1 presents the results for the six-month-ahead volatility indices, and the three-month results are available in the [Online Appendices](#) (the same comment applies to all the three-month results in subsequent sections unless otherwise stated).

⁹ The same sources apply to all figures and tables in Section 3.

In the figure, the solid black lines are the actual volatility, the solid grey lines are the counterfactuals, and the other two lines are the upper and lower bounds of the 90% confidence intervals. Take the event of Trump’s state-of-emergency declaration as an example. The six-month-ahead dropped sharply starting at 3:36 pm on March 13, the exact minute when Trump finished his remarks and started taking questions from the press.¹⁰ It then dropped by as much as 3.2% during one minute shortly after that. Moreover, in 15 minutes, it dropped by more than twice the average *daily* change (about 1.9), from a level of 45 at 3:36 pm to 41 at 3:51pm. Although this declaration was not a lockdown announcement *per se*, it signals that the federal government was serious about the situation and that harsh lockdowns by state governments would follow (which indeed happened).

Moreover, such drops lie below the lower bounds of the 90% confidence intervals of the counterfactual (expected) volatility paths. The same is true for Germany’s six-month volatility after its announcement of a historic national lockdown. Even though the actual (expected) volatility was rising, it still lay significantly below the mean counterfactual (which was also rising starting from a much higher level) and below the 90% confidence interval lower bound. These results suggest that, contrary to the widely-held belief, lockdown announcements may have decreased market participants’ perceptions of the six-month-ahead uncertainty.

We would like to make several comments on the methodology. First, in Figure 1, the model used to construct the counterfactual volatility is an ARIMA(1,1,1) model augmented with two additional predictors: the stock price index itself, and the GARCH-implied volatility. The ARIMA component mainly captures the persistence of *historical* patterns of volatility, while the stock price and the GARCH components mainly capture *new* information associated with the announcement. Our “augmented ARIMA approach” is in a similar spirit to the study by Engle and Gallo (2006), which adds a *one-month-ahead* forecast of MEM-implied volatility to an AR(1) model of VIX (whereas we add *time t* GARCH-implied actual volatility to an ARIMA(1,1,1) model of VIX; we use the *ex post* instead of the *ex ante* forecasted volatility to enhance the accuracy of the counterfactual). Our approach is also similar to the “factor model” in Fernandes, Medeiros, and Scharth (2014), who forecast the daily-level VIX with S&P 500 price and volume, an AR(1) component (i.e., last day VIX), and past 5-day, 10-day, 22-day, and 66-day VIX averages. In our case, the “factors” include S&P 500 price, the ARIMA(1,1,1) components, and a GARCH-implied volatility. Two other counterfactual models are also used, as discussed in the robustness check subsection.

Second, the stock price index itself is used to construct the counterfactual volatility because there are at least four forces associated with a lockdown announcement: (1) The announcement confirms that the outbreak was severe, which tends to decrease stock price and increase (short-

¹⁰ The full video for Trump’s state of emergency declaration is available [here](#), showing that he finished his remarks at 3 pm 35’40”.

and-medium-term) expected volatility; (2) Other stimulus policies (e.g., fiscal stimulus or monetary easing) may be announced at the same time (or the lockdown announcement signals the severity of the outbreak and makes the market to believe that policymakers will be more likely to pass these stimulus policies), which affects both stock price and expected volatility; (3) The lockdown causes *short-term* disruptions to the economy, which decreases stock price and increases volatility; (4) The lockdown may have a *medium-term* economic benefit through containing the outbreak and reducing expected volatility. Controlling for stock price allows us to somewhat proxy for Forces (1)-(3) and test the existence of Force (4), which is the focus of our paper.¹¹ Relatedly, it is not obvious that the lower volatility is due to (announced or expected) aggressive monetary easing: such easing could also be seen as the central bank running out of firepower, which tends to *increase* volatility. For example, on March 15 Sunday, the Fed surprised the market by cutting 125 bps to 0 and launching a massive \$700b QE. Some [news articles](#) believe this triggered the market's fear that all Fed's firepower had been used and was responsible for the massive stock price declines on March 16 (and may have led to the sharp increases in VIX in the morning of March 16).

Third, we use a 30-minute event window to control for confounding events. Although it appears short, 30 minutes are considered a relatively long window in intraday event study literature (Marshall, Nguyen, and Visaltanachoti, 2017). As McWilliams and Siegel (1997, p. 634) note: "The longer the event window, the more difficult it is for researchers to claim they have controlled for confounding events." In addition, as shown in Chordia, Roll, and Subrahmanyam (2005), among others, new information affecting the stock price is digested by the market within 5-60 minutes.

Fourth, to check whether this substantial drop of volatility is a usual pattern that occurs on most days or truly reflects the impact of the event, we further evaluate the performance of our counterfactual models across the entire sample. Figure 2 plots the average prediction errors of our counterfactual model for the six-month-ahead volatility at the same point of time on each day in our sample, where the prediction error is defined as the actual volatility minus the counterfactual. Across all minutes during our 30-minute event window, the mean prediction errors (for a given minute across all days) are very close to 0, suggesting that our counterfactual model is broadly unbiased. Moreover, the most negative prediction errors (i.e., the situations where the actual volatility is much lower than the counterfactual) are largely observed on the event days we are

¹¹ One may argue that there can be an endogeneity issue: On the one hand, the rapid increase in COVID cases induces policymakers to impose the lockdown; on the other hand, it also induces more market participants to believe that herd immunity will be achieved sooner due to the higher infections, which in turn tends to lower medium-term volatility. Hence, both the lockdown and the lower medium-term volatility are results of deterioration of COVID dynamics rather than the former causing the latter. However, very few countries (and market participants) believe that it is effective to achieve herd immunity through more infections. This is evidenced by the harsh criticism of Boris Johnson's earlier remarks and critical views on Sweden's initial "no-containment" strategy. In addition, the minute-level analysis can somewhat mitigate this endogeneity issue.

analyzing (i.e., the prediction errors on the event days fall below the 10th percentiles of the empirical distributions). This result reaffirms that these events indeed induced market participants to lower their volatility forecasts relative to the counterfactuals significantly.

Fifth and finally, the VIX (the one-month-ahead volatility index in the US) is widely regarded as the “fear gauge” by financial market participants, and the volatility indices we use are its counterparts for other maturities or in other countries. But to further check the macroeconomic relevance of these indices, we conduct a simple test by regressing the growth rate of the (normalized) purchasing manager index (PMI) on various (lagged) VIX measures. As shown in Table 2, VIX measures are negatively and significantly correlated with the growth rate of the one-month-ahead PMI, which provides some suggestive evidence that the (forward-looking) volatility indices we use are relevant to the macroeconomy and not just to financial markets.

Figure 1. Six-Month Volatility Indices (Initial Tightening)

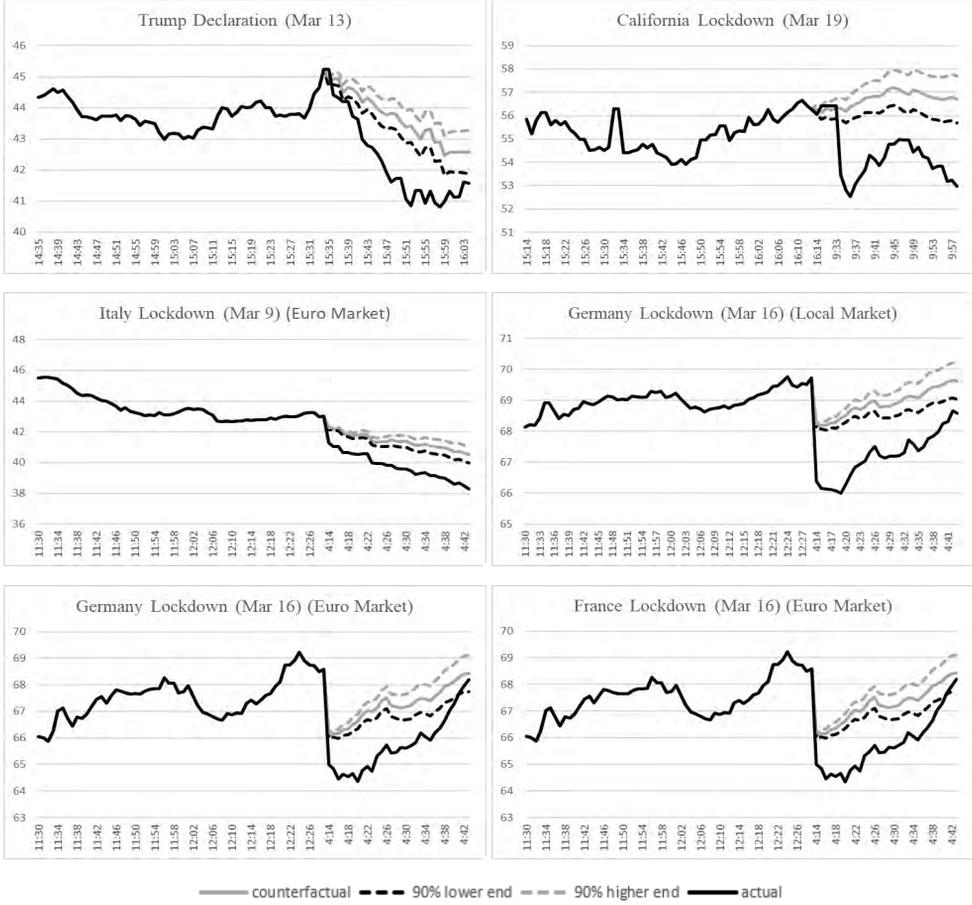
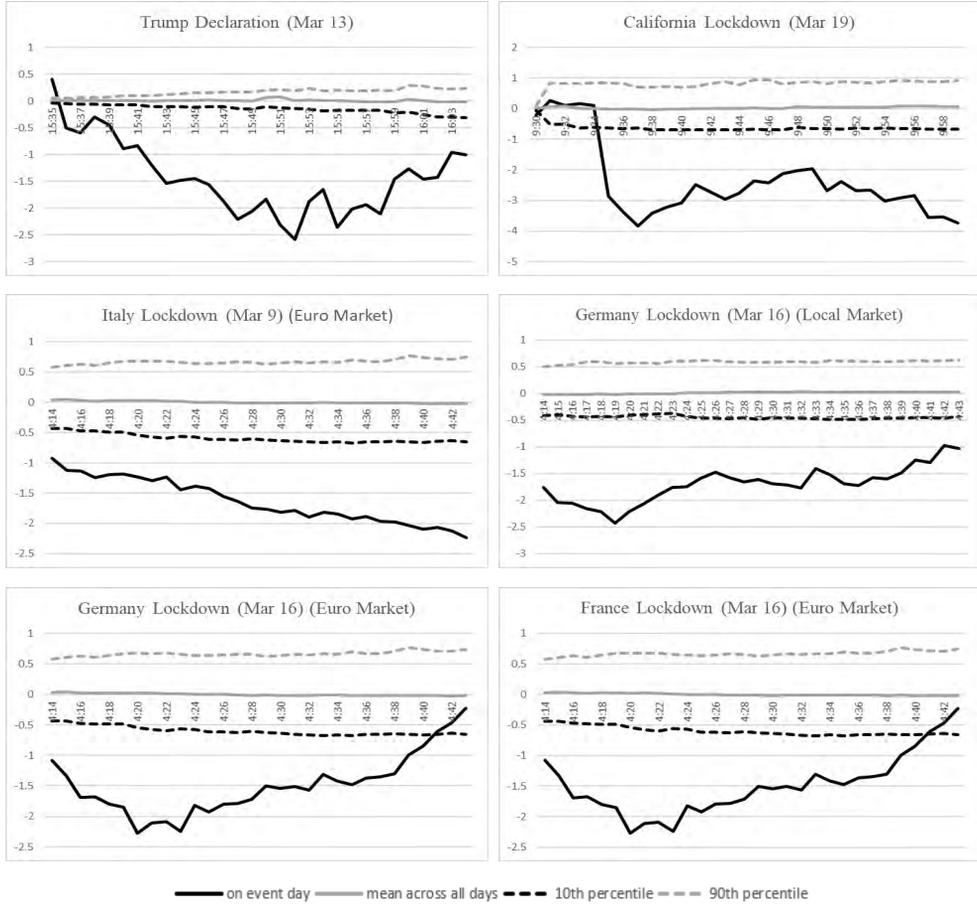


Figure 2. Prediction Errors for Six-Month Volatility Indices (Initial Tightening)



Covid Economics 79, 27 May 2021: 1-40

Table 2. Macroeconomic Relevance of Volatility Indices

	(1) Open	(2) Close	(3) High	(4) Low
L.VIX_Open	-3.368** (0.026)			
L.VIX_Close		-3.364** (0.027)		
L.VIX_High			-3.201** (0.021)	
L.VIX_Low				-3.486** (0.036)
Constant	258.717*** (0.000)	258.322*** (0.000)	258.872*** (0.000)	257.704*** (0.000)
Observations	119	119	119	119
R-squared	0.041	0.041	0.045	0.037

Notes: The horizon is Oct 2010 – Sep 2020; p-values are in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

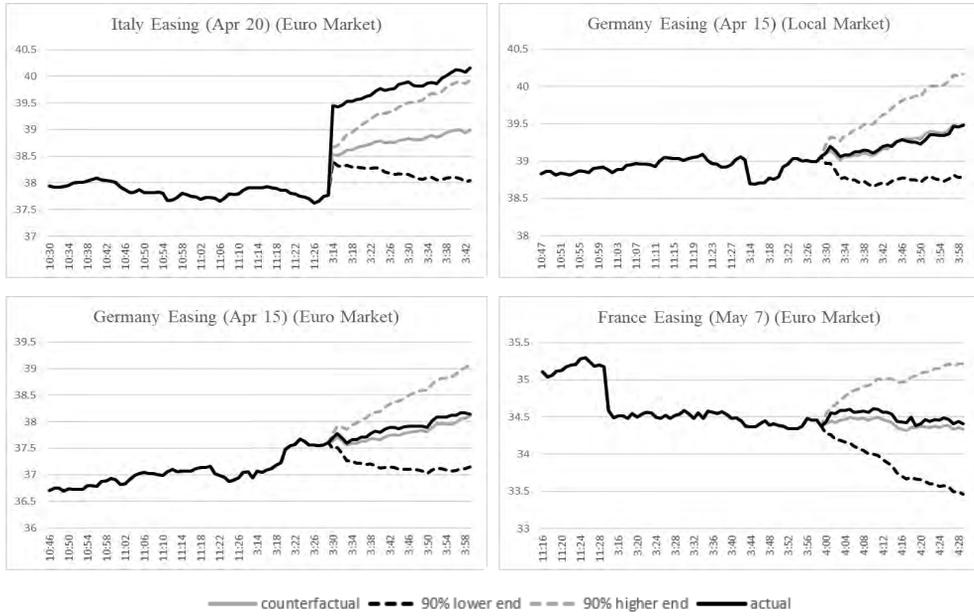
3.3 Easing/Reopening

We now turn to the event studies following the announcements of easing the lockdowns and reopening the economy, mostly in Summer 2020. The results for the six-month-ahead volatility are shown in Figure 3 (three-month results are in the [Online Appendices](#)). Since the US does not have a clear-cut easing/reopening date due to its gradual, state-operated reopening (and the difficulty of identifying the precise minute of the first reopening, which was by California), we do not include the US in the event studies for the easing stage.

As shown in Figure 3, for three out of the four events, the six-month-ahead volatility indices did not show statistically different paths from the counterfactuals following the easing announcements. And for the remaining event (Italy's reopening), the six-month-ahead volatility actually rose above the upper bound of the counterfactual's confidence interval. These results are in sharp contrast with conventional wisdom, which suggests that as the stringent containment measures are relaxed, there would be less disruptions to the economy, and thus the uncertainty would also be lower. Therefore, our results provide further suggestive evidence for the existence of the volatility-reducing effect of stringent containment measures emphasized in our paper: although the easing of containment measures provides immediate relief to the economy (which decreases uncertainty), it may raise concerns that the COVID outbreak might recur in the near future (which increases expected volatility).¹²

¹² One may argue that the lower volatility may be simply a result of reduced policy uncertainty rather than of the lockdown decision: even if the government instead announced that there would not be any lockdown, the lower policy uncertainty would still lead to lower volatility. It is indeed hard (if not impossible) to empirically rule out this argument because the counterfactual scenario suggested in the comment is not observed by definition. However, the reopening of an economy can be regarded as a “quasi counterfactual” experiment to test this. The

Figure 3. Six-Month Volatility Indices (Easing)



Similar to the case of initial lockdowns, we also evaluate the performance of our counterfactual models across the entire sample. Appendix Figure 1 plots the average prediction errors of our counterfactual model for the six-month-ahead volatility. Across all minutes during our 30-minute event window, the mean prediction errors (for a given minute across all days) are very close to 0, suggesting that our counterfactual model is also broadly unbiased for the easing stage. However, for three of the four events, the prediction errors on the event day mostly fall between the 10th and 90th percentiles of the empirical distribution of prediction errors across all days, suggesting that the events do not significantly lower the actual volatility relative to the counterfactual. Moreover, for Italy’s easing announcement, the prediction errors on the event day are among the top 10% most positive errors, suggesting that the actual volatility significantly rose above the counterfactual after Italy’s easing announcement. These results confirm the results presented in Figure 3.

reopening announcement also lowered policy uncertainty, but this announcement did not lead to lower volatility. As shown in Figure 3, for three out of the four events, the six-month-ahead volatility indices did not show statistically different paths from the counterfactuals; and for the remaining event (Italy’s reopening), the six-month-ahead volatility rose above the upper bound of the counterfactual’s confidence interval.

3.4 Retightening

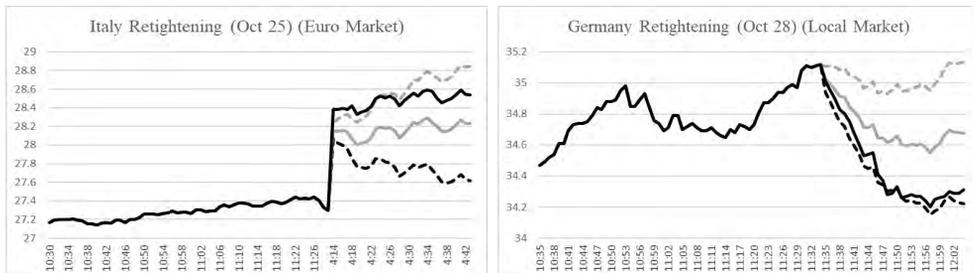
Finally, we discuss the event studies following the retightening announcements recently made in the context of new COVID waves. The results for the six-month-ahead expected volatility are shown in Figure 4. Also, because the US does not have a clear-cut retightening date yet, we do not include the US in the event studies for the retightening stage.

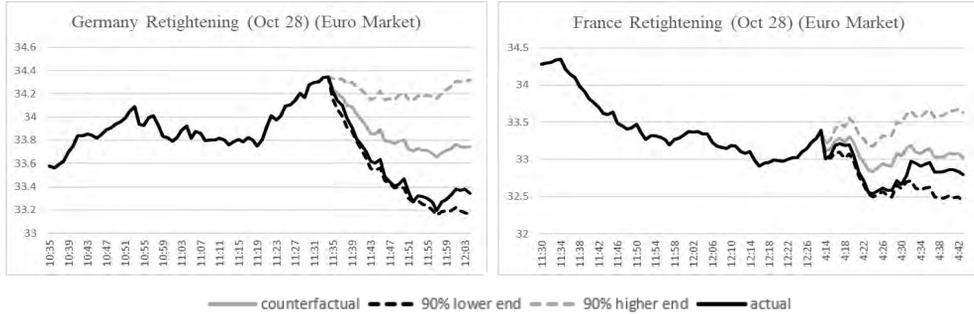
As shown in Figure 4, following Germany's retightening announcement around Eastern Time 11:35 am on October 28, the local and eurozone-wide expected volatility indices were significantly lower than the counterfactuals, with a significance level close to 10% (the actual paths almost overlapped the lower bounds of the counterfactuals' 90% confidence intervals). A similar pattern was observed following France's retightening announcement on the night of October 28, although the actual (expected) volatility was only significantly lower than the counterfactual for the first 19 minutes in the 30-minute event window and then rose to inside the 90% confidence interval.

The case of Italy's retightening announcement on the night of October 25 appears to display a more mixed pattern: expected volatility was slightly above the upper bound of the 90% confidence interval for the first 10 minutes and then dropped to inside the confidence interval. In addition, the average prediction errors (Appendix Figure 2 for six-month-ahead volatility) suggest that the deviations of the actual volatility from the counterfactuals on Italy's retightening day were within the range of the 10% and 90% percentiles of the empirical distribution across all days. Hence, Italy's retightening announcement was followed by neither a significantly higher nor lower volatility relative to the counterfactuals.

In sum, during the retightening stage, event studies show that the announcements of re-imposing lockdowns are still followed by somewhat significantly lower volatility, similar to the initial lockdowns. However, the statistical significance is lower than the initial tightening stage. There are multiple interpretations of these results. One interpretation is that they reflect market participants' perception that the governments' containment measures during the retightening stage are less stringent than the initial stage, which may be inadequate to contain the second waves and thus, volatility did not drop as much.

Figure 4. Six-Month Volatility Indices (Retightening)





3.5 Monotonicity Over Time

To shed more light on the intertemporal trade-off explained in the Introduction, we compare the responses of volatility indices across different maturities, i.e., the one-month-, three-month-, and six-month-ahead volatility indices. To ensure comparability, we use the same counterfactual model for all maturities, that is, an ARIMA(1,1,1) model augmented with the stock price index and the GARCH-implied volatility.

Figure 5 presents the comparison result following Germany's initial tightening on March 16, 2020 (for the volatility of its domestic stock index). As shown in the top left chart, the one-month-ahead volatility index actually jumped *above* the upper bound of the 90% confidence interval during the first half of the event window before falling within the interval. By contrast, the three-month-ahead volatility index dropped *below* the lower bound of the 90% confidence interval during the first half before falling within the interval, suggesting that the lockdown announcement significantly decreased the three-month-ahead volatility. The most significant response is displayed in the bottom left chart, where the six-month-ahead volatility index stayed below the lower bound of the 90% confidence interval throughout the entire event window. This monotonicity is confirmed by the three right-hand-side charts in Figure 5, which plot the prediction errors of the counterfactual model for the three maturities.

A similar pattern is observed following France's initial tightening on March 16, 2020 (Figure 6), Italy's initial tightening on March 9, 2020 (Appendix Figure 3), and Germany's *retightening* on October 28, 2020 (also for the volatility of its domestic stock index, Figure 7). For some other events studied in previous sections, including Trump's state-of-emergency declaration and California's initial lockdown during the initial tightening stage, the one-month-ahead volatility is still significantly lower after the announcement.¹³

¹³ Note that because the responses of volatility indices during the easing stage are not statistically significant, we do not compare the responses across different maturities.

In summary, we find suggestive evidence that that the results presented in previous sections are the strongest for six-month-ahead expected volatility; not as strong for three-month-ahead expected volatility; and generally absent for one-month-ahead expected volatility. This monotonicity provides suggestive evidence for the existence of the intertemporal trade-off associated with lockdown: The lockdown disrupts the economy, which increases volatility; But it contains the COVID outbreak, which decreases volatility. Both of these two forces are present not only in the medium term (six months) but also in the short term (one month). In general, the volatility-decreasing effect is more likely to dominate the volatility-increasing effect in the medium term, as suggested by the finding in this section. Although we do not study the volatility beyond the six-month horizon due to data limitations, the increasing significance of results over time seems informative.

Then how to explain the results where the one-month-ahead volatility is still significantly lower after the lockdown? The answer again lies in the intertemporal trade-off: since the observed volatility is a result of two countervailing forces, it is still possible that in some countries and for some events, the volatility-decreasing effect can already dominate the volatility-increasing effect even in the short term.

3.6 Robustness checks

To further validate our results, we conduct three sets of robustness checks. First, while constructing the counterfactual volatility, we drop the GARCH-implied volatility from the list of predictors. That is, we use an ARIMA model augmented with the stock price index as the only additional predictor. The results for six-month-ahead volatility are presented in Appendix Figures 3-5, with one figure for one stage (initial tightening, easing, and retightening). All results are very similar to the main results discussed above. Note that the alternative counterfactual models' empirical prediction error bands are not constructed due to the heavy computation burden (the construction for each event takes more than 5 hours).

Second, we replace the GARCH-implied volatility with the EGARCH-implied volatility in the counterfactual models. That is, we use an ARIMA model augmented with the stock price index and EGARCH-implied volatility as two additional predictors. The results are presented in Appendix Figures 6-8, with one figure for one stage.

Third, since the six-month-ahead volatility index overlaps the three-month-ahead volatility index for the first three months, we further decompose each index into non-overlapping indices. That is, we decompose the six-month-ahead volatility index into the three-month-ahead volatility and the volatility from Month 3 to Month 6. We then repeat the event studies for the initial tightening, easing, and retightening stages using these non-overlapping volatility indices as the dependent variables. The results are available in the [Online Appendices](#). All results in these robustness checks are very similar to the main results discussed above.

Figure 5. Volatility Responses Across Maturities: Germany’s Initial Tightening (Local Market)

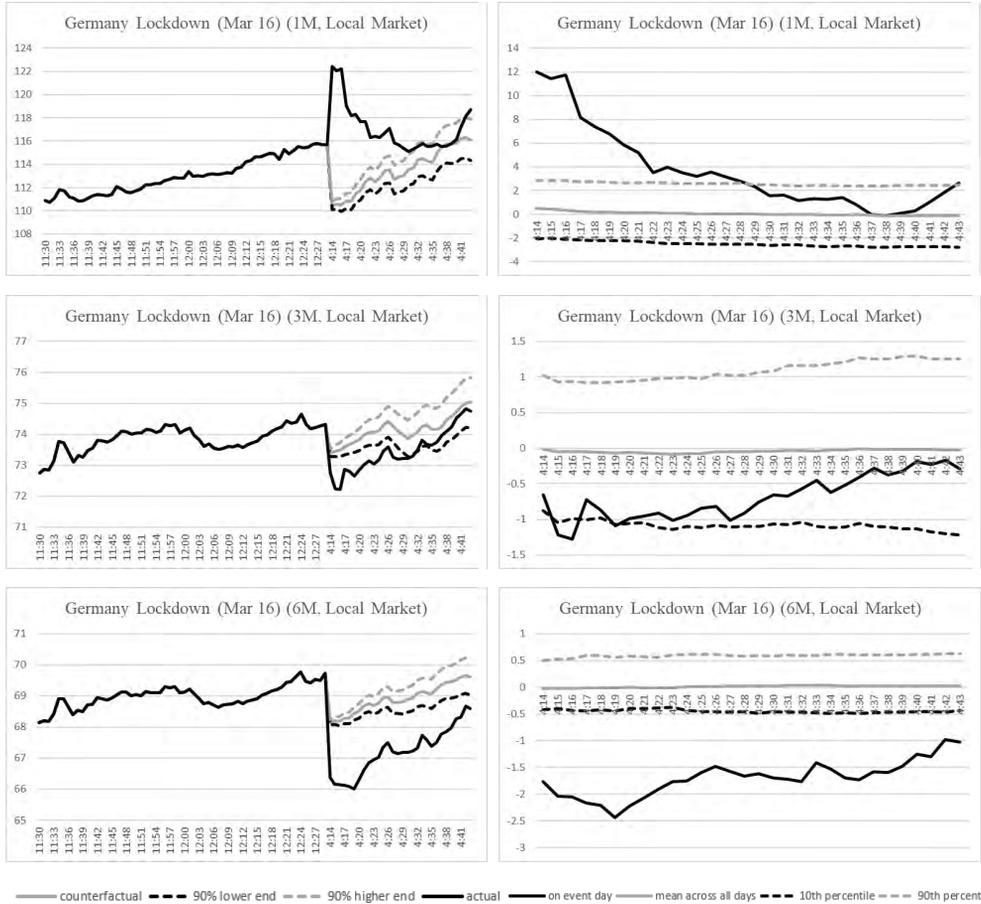
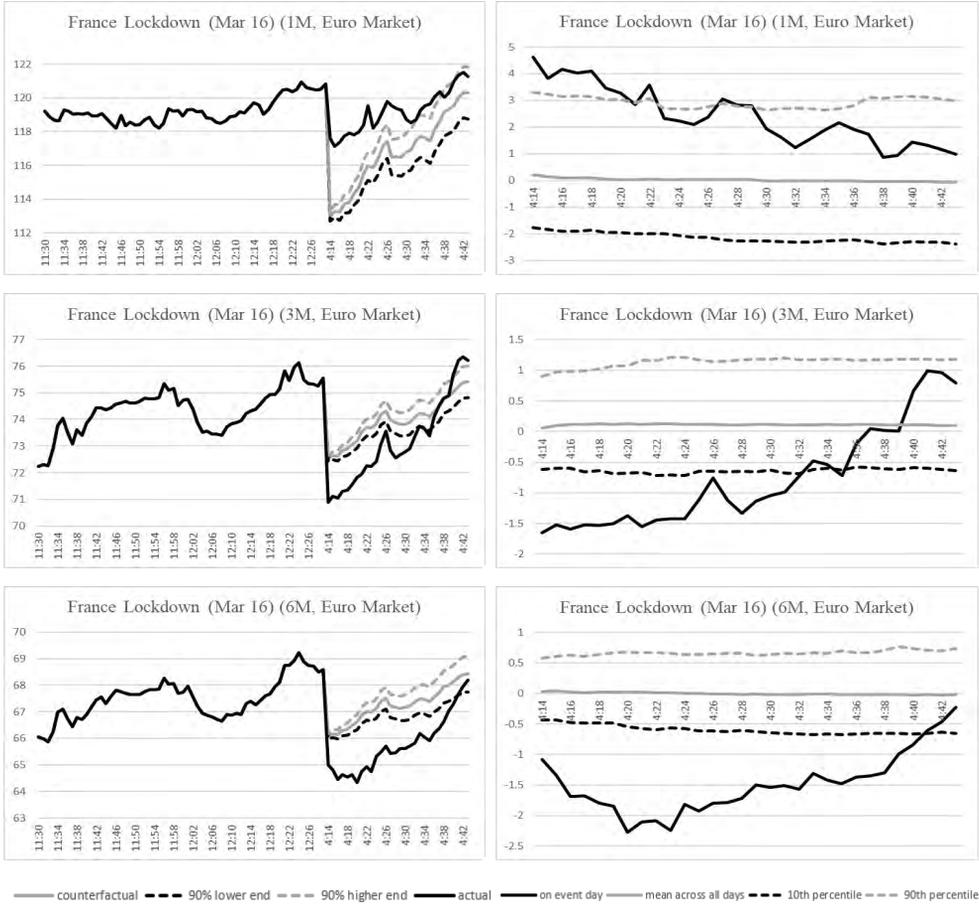
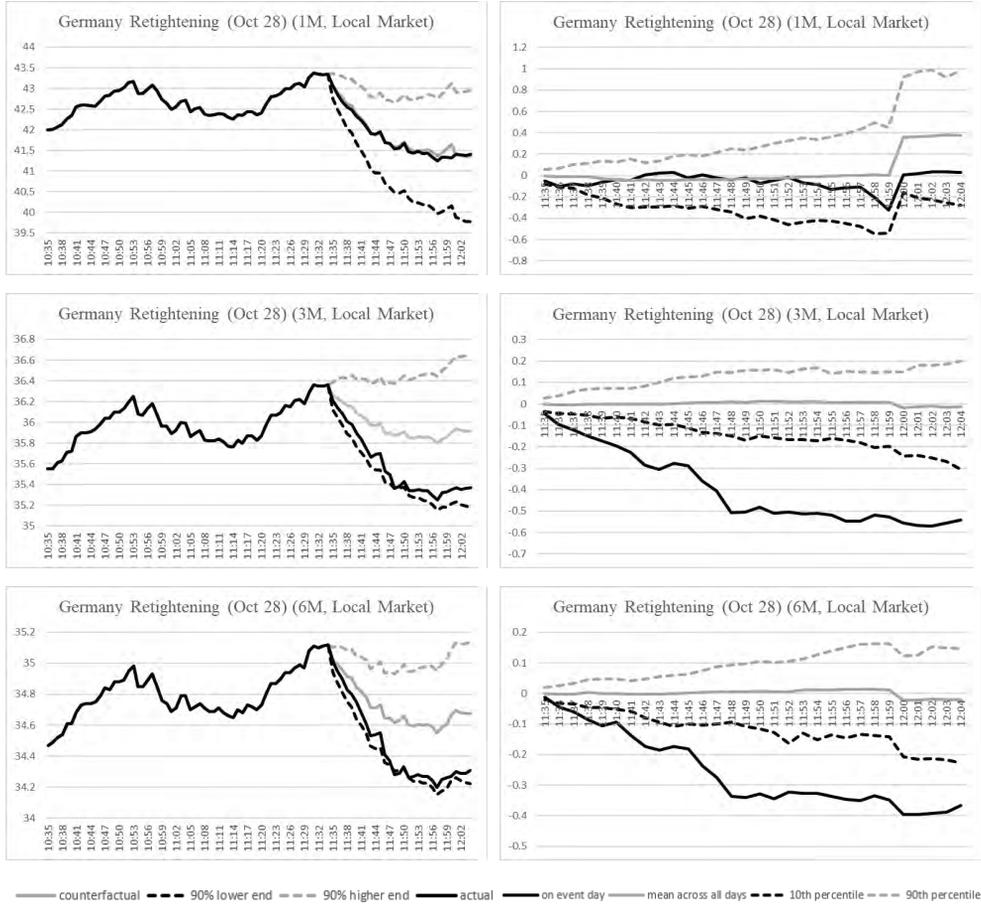


Figure 6. Volatility Responses Across Maturities: France's Initial Tightening



Covid Economics 79, 27 May 2021: 1-40

Figure 7. Volatility Responses Across Maturities: Germany's Retightening (Local Market)



4 Regressions with Daily Data

4.1 Data

The daily data used in the regression approach cover weekdays from January 3, 2020 to October 22, 2020, including the initial tightening, easing, and retightening stages. Due to limitations on the expected volatility data, the following five countries/regions are covered: the US, Italy, Germany, Euro Area, and the UK. Note that, unlike the event study approach, the regression approach uses the country-specific volatility index for Italy instead of the eurozone-wide volatility because the daily data for this volatility index is available. For the same reason, the UK is covered in the

regression approach, even though it is not in the event study approach. However, because France still does not have the daily data for its country-specific volatility index beyond the one-month horizon, it is not covered in the regression approach.

Daily data on the stringency index are from the widely-used Oxford COVID-19 Government Response Tracker (OxCGRT) database. The index provides a continuous measurement of the stringency of COVID containment and closure policies. It scales between 0 and 100, with 100 representing the most stringent measures. It is constructed based on eight indicators, including school closing, workplace closing, public events cancellation, restrictions on gatherings, public transport closure, stay-at-home requirements, restrictions on internal movement, and international travel controls.¹⁴ The same dataset also provides COVID case numbers. The stock price data are from Bloomberg.

The summary statistics are presented in Appendix Table 1 for each of the three stages. Note that rescaling is done to make the displayed coefficients more informative. Because of this, the units of the volatility percent change and of the stock price percent change are basis points (i.e., 1/100 percent), although the units of the COVID case percent change and stringency index remain as the percent. However, it can be shown that the interpretation of economic significance is invariant to the units of the variables and thus is not affected by the rescaling.

4.2 Initial Tightening

For each stage (initial tightening, easing, and retightening), we first conduct the regressions for the benchmark models that include stringency index, COVID case growth rate, and the interaction of the two as the only regressors (along with the constant term). We then add the stock price percent change (and its lag) and re-run the regressions. Finally, we obtain the “full” models after adding the intraday standard deviation of the stock price (and its lag) to capture forces that affect the *current* realized volatility (note that the dependent variables in the regressions are *forward-looking* expected volatility). The data samples used for the regressions in each stage are unbalanced panel datasets because different countries have different easing/reopening and retightening days, but all countries start the data from January 3, 2020 for the initial tightening stage.

The benchmark model results for the initial tightening stage are presented in Appendix Table 2 (only results for six-month volatility are shown; those for three-month are available in the [Online Appendices](#)). As the table shows, the interaction term is not statistically significant. However, the stringency index itself is statistically significant and negatively correlated with the percent changes of volatility, consistent with our main idea that containment measures help reduce uncertainties. Note that the R-squared's are low in the benchmark models, suggesting the possible existence of omitted variable bias.

¹⁴ More details are available [here](#).

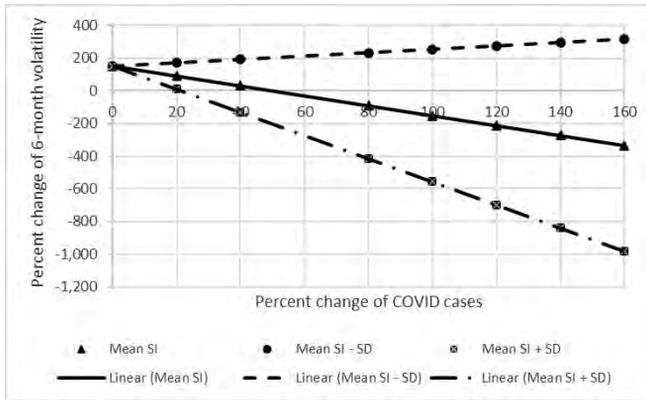
We then add the stock price percent change (and its lag). The results for the initial tightening stage are presented in Appendix Table 3. Now the stringency index itself becomes insignificant, and the interaction term becomes significantly negative, suggesting that the containment measures reduce expected volatility through the interaction with the outbreak dynamics. Importantly, the stock price percent change is highly significant, and the R-squared's have improved significantly, confirming the existence of omitted variable bias in the benchmark models.

Finally, we obtain the regression results for the initial tightening stage in the full models that include the intraday standard deviation of the stock price (and its lag). The results are presented in Table 3, which are similar to those in the models with stock price percent change. Specifically, we would like to highlight the following:

First, the stringency index itself is insignificant, and the interaction term is significantly negative. This implies that (a) the marginal effect of stringency index on expected volatility (equal to the coefficient of the interaction term, multiplied by the COVID case growth rate) is still negative, as the COVID case growth rate is positive; (b) containment measures reduce volatility mainly through the interaction with the outbreak dynamics—the more severe the outbreak is, the stronger this effect is; (c) containment measures mitigate the volatility-increasing effect of the COVID case growth,¹⁵ as illustrated in Figure 8: when the stringency index is low (equal to the sample mean minus one standard deviation), a higher COVID case growth is associated with a higher expected volatility (the dash line); but as the stringency index increases, e.g., to the sample mean (the solid line) or the sample mean plus one standard deviation (the dash-dot line), the volatility-increasing effect of COVID case growth is mitigated and ultimately reversed, possibly because the stringent containment measures have reduced infections and generated indirect economic benefit.

¹⁵ Note that the marginal effect on volatility equals the *positive* coefficient of the COVID case growth, plus the product of the *negative* coefficient of the interaction term and the (positive) stringency index. Hence, the COVID case growth has a positive marginal effect on volatility if the stringency index equals 0.

Figure 8. Interaction Between Stringent Containment Measures and COVID Case Growth



Note: SI = Stringency Index; All other regressors are evaluated at respective sample means.

Sources: Oxford; Authors' calculations.¹⁶

Second, as expected, the percent change of the stock price index is negatively correlated with both volatility indices in all specifications and acts as a control for other forces that drive the volatility. As for the standard deviation of the stock price, the lagged term is positively correlated with volatility, and the current term is negatively correlated. This seemingly counterintuitive result may be because the standard deviation is unable to capture the direction of stock price movement—a high standard deviation could mean either an increase in stock price (in which case volatility tends to be low) or a decrease in stock price (in which case volatility tends to be high).¹⁷

Third, the economic significance of the interaction term (in *italic>*) has the same order of magnitude as the stock price percent change. This reassures that the stringency index is as economically relevant as other forces (captured by the stock price percent change) in driving the expected volatility.

Table 3. Initial Tightening Stage Panel Regressions in the Full Model

	(1)	(2)	(3)	(4)	(5)	(6)
	SI	Cases	Interaction	SI	Cases	Interaction
	FE	FE	FE	RE	RE	RE
StringencyIndex	-1.819**		-0.527	-1.557*		-0.242
	(0.032)		(0.559)	(0.065)		(0.788)

¹⁶ The same sources apply to all figures and tables in Section 4.

¹⁷ Indeed, the models controlling for the standard deviation of the stock price (and not stock price percent change) have extremely low R-squared's (in the range of 3-7%), suggesting that this variable has low explanatory power for volatility.

Cases_pct		-0.181	1.731*		0.132	2.032**
		(0.817)	(0.063)		(0.866)	(0.029)
Cases_pct_SI			-0.114***			-0.112***
			(0.000)			(0.000)
			-0.182			-0.179
StockPrice_pct	-1.849***	-1.883***	-1.919***	-1.874***	-1.894***	-1.933***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
			-0.718			-0.723
Lag_StockPrice_pct	0.247**	0.207**	0.197**	0.199**	0.173*	0.156
	(0.013)	(0.039)	(0.048)	(0.043)	(0.082)	(0.117)
StockPrice_std	-2.165***	-2.385***	-2.007***	-2.780***	-2.930***	-2.639***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Lag_StockPrice_std	4.260***	4.085***	4.410***	3.694***	3.561***	3.802***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	396	396	396	396	396	396
R-squared	0.532	0.526	0.550			

Notes: (1) p-values are in parentheses. (2) FE = fixed effect; RE = random effect.

4.3 Easing/Reopening

The cutoff dates for the easing stage regressions are determined based on the Oxford stringency index. Note that to provide a comparison benchmark and obtain a sharper identification, the starting date used in the easing stage regressions is a few working days earlier than the actual easing/reopening day. For example, the stringency index shows that Germany eased on May 4, 2020, but we start the sample for Germany's easing stage regressions from April 24.

Regressions for the easing stage from the benchmark models and models with stock price percent change are also conducted, but the results are omitted for brevity.¹⁸ Similar to the initial tightening stage, the addition of stock price percent change (and its lag) also substantially increases the R-squared's. The results from the full models for the easing stage are presented in Panel (A) of Table 4, where the results on the lagged stock price percent change, the stock price standard deviation, and its lag are not reported.

¹⁸ These results are available upon request.

Table 4. Panel Regressions in the Full Model: Easing and Retightening Stages

	(1)	(2)	(3)	(4)	(5)	(6)
	SI	Cases	Interaction	SI	Cases	Interaction
	FE	FE	FE	RE	RE	RE
Panel (A): Easing stage						
StringencyIndex	-3.374 (0.124)		-4.461 (0.138)	-3.015 (0.110)		-3.166 (0.216)
Cases_pct		-24.827 (0.174)	-193.858 (0.310)		-22.657 (0.166)	-108.048 (0.528)
Cases_pct_SI			2.483 (0.353)			1.335 (0.583)
StockPrice_pct	-2.023*** (0.000)	-2.018*** (0.000)	-2.026*** (0.000)	-2.026*** (0.000)	-2.020*** (0.000)	-2.029*** (0.000)
Observations	257	257	257	257	257	257
R-squared	0.598	0.597	0.600			
Panel (B): Retightening stage						
StringencyIndex	2.890 (0.424)		5.875 (0.125)	-0.534 (0.684)		2.903 (0.136)
Cases_pct		2.299 (0.778)	128.279** (0.022)		3.033 (0.705)	130.986** (0.017)
Cases_pct_SI			-2.159** (0.024)			-2.206** (0.018)
			-0.652			-0.666
StockPrice_pct	-1.729*** (0.000)	-1.719*** (0.000)	-1.720*** (0.000)	-1.718*** (0.000)	-1.718*** (0.000)	-1.707*** (0.000)
			-0.520			-0.516
Observations	289	289	289	289	289	289
R-squared	0.538	0.537	0.547			

Notes: (1) p-values are in parentheses. (2) FE = fixed effect; RE = random effect.

As in the initial tightening stage, the stock price percent change is negatively correlated with expected volatility, and the effect is highly significant during the easing stage. A major difference is that during the easing stage, neither the containment measures nor their interactions with the COVID case growth are statistically significant, which suggests that the easing of stringent containment measures is *not* associated with a significant reduction in volatility. Conventional wisdom is that as the stringent containment measures are relaxed, there would be less disruptions to the economy, and thus the expected volatility and uncertainty would also be

lower. However, similar to the event study results, our regression results do not support this hypothesis. This, in turn, provides further suggestive evidence for the existence of the trade-off emphasized in our paper: although the easing of containment measures provides immediate relief to the economy (hence decreasing uncertainty perceived by the market), it may raise concerns that the COVID outbreak might recur in the near future (hence increasing uncertainty).

4.4 Retightening

The cutoff dates for the retightening stage regressions are generally determined based on the Oxford stringency index. But to obtain a sharper identification, in some cases we again skip the long “post-easing” period when the stringency index remained flat and low. For example, even though we end Italy’s easing stage on June 10, we skip the data in the next two months for Italy due to its flat and low stringency index; instead, we start Italy’s retightening stage on August 11, which is five weekdays before Italy publicly announced a reintroduction of restrictions on August 17 (see the announcement [here](#); we include the five extra weekdays to provide a comparison benchmark). Note that although the stringency index for the US did not show a retightening after its easing stage, we still include the US in the regression as a benchmark to help identify the effect of retightening by other countries.

Regressions for the retightening stage from the benchmark models and models with stock price percent change are also conducted, but the results are omitted for brevity.¹⁹ Similar to the initial tightening stage and the easing stage, the addition of stock price percent change also substantially increases the R-squared’s. The results from the full models for the retightening stage are presented in Panel (B) of Table 4.

As in the other two stages, the stock price percent change is negatively correlated with volatility, and the effect is highly significant during the retightening stage. A major finding is that during the retightening stage, more stringent containment measures are again associated with lower volatility, although its statistical significance is lower than the initial tightening stage. Specifically, for the six-month volatility, the p-values for the interaction term between stringency index and COVID case growth are around 2 percent in the retightening stage, compared with 0 percent in the initial tightening. And for three-month results (available in the [Online Appendices](#)), the p-values for the interaction term are above 20 percent in the retightening stage, compared with 1-2 percent in the initial tightening.

As discussed in the event study results, one interpretation is that these results reflect market participants’ perception that the governments’ containment measures during the retightening stage are less stringent than the initial stage, which may be perceived as inadequate to contain the second waves. As a result, volatility did not drop as much as in the initial tightening stage.

¹⁹ These results are available upon request.

4.5 Robustness checks

In addition to the various models presented above, we conduct two more sets of robustness checks. First, given that COVID-19 is a global shock that affects different countries and different volatility products, we conduct seemingly unrelated regressions (SUR) to account for the correlations among countries and volatility products. Doing so would require a balanced panel dataset, so we conduct the SUR for the whole sample only, without distinguishing among different stages (recall that different countries have different reopening and retightening dates, so distinguishing among different stages would result in unbalanced panel datasets). The results are presented in Appendix Table 3, which are very similar to the full model results during the initial tightening stage.

Second, as with the event studies, we also decompose each volatility index into non-overlapping indices. Specifically, we decompose the six-month-ahead volatility index into the volatility from the three-month-ahead volatility and the volatility from Month 3 to Month 6. We then repeat the regressions for the initial tightening, easing, and retightening stages using these non-overlapping volatility indices as the dependent variables (for the full models). The results are available in the [Online Appendices](#), which are again similar to the results in the corresponding stage.

5 Conclusion and Policy Implications

Using event studies with minute-level expected volatility data and panel regressions with daily data, we empirically show that COVID containment measures reduce six-month-ahead expected stock price volatility indices. This pattern is not as strong for the three-month-ahead expected volatility and generally absent for the one-month-ahead expected volatility. Our results provide some suggestive evidence that such measures may have an economic benefit of reducing medium-term uncertainty despite their short-term economic disruptions.

Future studies can explore further the channel through which containment measures reduce the expected volatility. To this end, one could analyze the responses of volatility in different sectors. If the contact-intensive sectors experienced a significantly larger drop in volatility, then this supports a real economy channel: containment measures would put the pandemic under control, which would be more beneficial to the contact-intensive sectors, leading to lower volatility in these sectors than in other sectors. Another caveat is that, due to data limitations, the number of events we study is relatively small, with an exclusive focus on advanced economies. Future studies can apply event studies to other measures of uncertainty or confidence in other types of economies, possibly at a daily or weekly frequency, given that it is hard to find minute-level data. Finally, since vaccination is also one containment measure, it is worth exploring the impact of positive vaccine-related news on the expected volatility.

Our results have some potential policy implications. First, on containment and reopening strategies, our results highlight that it is important to recognize the existence of a potential

economic benefit of containment measures, particularly when decisively implemented in advanced economies. Although containment may be less effective in emerging markets and low-income countries (e.g., due to large informal sectors), the benefit of reducing uncertainty may still exist in all countries and needs to be taken into account when assessing the trade-off associated with containment measures (as evidenced in China's experience). And in the context of local new COVID waves, the lockdowns can be localized and be combined with other containment measures such as mask wearing.

Second, on macroeconomic projections, ignoring this uncertainty-reducing benefit may lead to static projections. If one only considers the short-term economic disruptions of stringent containment measures while disregarding their medium-term benefit, macroeconomic projections would be overly conservative with containment measures or overly optimistic without them, distorting policy decisions.

References

- Alvarez, Fernando, David Argente, Francesco Lippi. 2020. "A Simple Planning Problem for COVID-19 Lockdown," Manuscript, University of Chicago.
- Arnon, Alexander, John Ricco, and Kent Smetters. 2020. "Epidemiological and Economic Effects of Lockdown," Brookings Papers on Economic Activity Conference Drafts.
- Ashraf, Badar Nadeem. 2020. "Economic Impact of Government Actions to Control COVID-19 Pandemic: Evidence from Financial Markets," *Journal of Behavioral and Experimental Finance*.
- Baek, Seungho, Sunil K. Mohanty, and Mina Glamboosky. 2020. "COVID-19 and Stock Market Volatility: An Industry Level Analysis," *Finance Research Letters*, <https://doi.org/10.1016/j.fl.2020.101748>
- Baker, Scott R., Nicholas Bloom, Steven J. Davis, Kyle Kost, Marco Sammon, and Tasaneeya Viratyosin. 2020. "The Unprecedented Stock Market Reaction to COVID-19," *Review of Asset Pricing Studies*, 0: 1-17, doi/10.1093/rapstu/raaa008/5873533
- Barro, Robert J., José F. Ursúa, and Joanna Weng. 2020. "The Coronavirus and the Great Influenza Pandemic: Lessons from the 'Spanish Flu' for the Coronavirus's Potential Effects on Mortality and Economic Activity," *NBER Working Paper No. 26866*.
- Beirne, John, Nuobu Renzhi, Eric Sugandi, and Ulrich Volz. 2020. "Financial Market and Capital Flow Dynamics During the COVID-19 Pandemic," Asian Development Bank Institute Working Paper No. 1158.
- Caselli, Francesca, Francesco Grigoli, Weicheng Lian, and Damiano Sandri. 2020. "Protecting Lives and Livelihoods with Early and Tight Lockdowns," IMF Working Paper No. 20/234.
- CBOE. 2019. "White Paper: Cboe Volatility Index," <https://cdn.cboe.com/resources/futures/vixwhite.pdf>.
- Chen, Sophia, Deniz Igan, Nicola Pierri, and Andrea F. Presbitero. 2020. "Tracking the Economic Impact of COVID-19 and Mitigation Policies in Europe and the United States," *COVID Economics*, 36: 1-24.
- Cheng, Ing-Haw. 2020. "Volatility Markets Underreacted to the Early Stages of the COVID-19 Pandemic," *Review of Asset Pricing Studies*, 10: 635-668.
- Chordia, Tarun, Richard Roll, and Avanidhar Subramanyam. 2005. "Evidence on the Speed of Convergence to Market Efficiency," *Journal of Financial Economics*, 76: 271-292.
- Correia, Sergio, Stephan Luck, and Emil Verner. 2020. "Pandemics Depress the Economy, Public Health Interventions Do Not: Evidence from the 1918 Flu," Working Paper.
- Deb, Pragyan, Davide Furceri, Jonathan D. Ostry, and Nour Tawk. 2020a. "The Economic Effects of COVID-19 Containment Measures," *COVID Economics*, 24: 32-75.
- Deb, Pragyan, Davide Furceri, Jonathan D. Ostry, and Nour Tawk. 2020b. "The Effect of Containment Measures on the COVID-19 Pandemic," *IMF Working Paper No. 20/159*.
- Eichenbaum, Martin, Sergio Rebelo, and Mathias Trabandt. "The Macroeconomics of Epidemics," March 2020. *Northwestern University Working Paper*.

- Fernandes, Marcelo, Marcelo C. Medeiros, and Marcel Scharth. 2014. "Modeling and Predicting the CBOE Market Volatility Index." *Journal of Banking and Finance*, 40: 1-10.
- Fotiou, Alexandra, and Andresa Lagerborg. 2021. "Smart Containment: Lessons from Countries with Past Experience," *COVID Economics*, 74: 94-141.
- Goolsbee, Austan, and Chad Syverson. 2020. "Fear, Lockdown, and Diversion: Comparing Drivers of Pandemic Economic Decline," NBER Working Paper No. 27432.
- Gourinchas, Pierre-Olivier. 2020. "Flattening the Pandemic and Recession Curves," <https://voxeu.org/article/flattening-pandemic-and-recession-curves>
- Hall, Robert E., Charles I. Jones, and Peter J. Klenow. 2020. "Trading Off Consumption and COVID-19 Deaths," *Stanford University Working Paper*.
- Haroon Omair, and Syed Aun R. Rizvi. 2020. "COVID-19: Media Coverage and Financial Markets Behavior—A Sectoral Inquiry," *Journal of Behavioral and Experimental Finance*, 27, <https://doi.org/10.1016/j.jbef.2020.100343>
- Jackwerth, Jens. 2020. "What Do Index Options Teach Us About COVID-19?" *Review of Asset Pricing Studies*, 10: 618-634.
- Jones, Callum, Thomas Philippon, and Venky Venkateswaran. 2020. "Optimal Mitigation Policies in A Pandemic," *NYU Working Paper*.
- Kok, Jason Lih Cheng. 2020. "Short-term Trade-off between Stringency and Economic Growth," *Covid Economics*, 60: 172-189.
- Marshall, Ben R., Nhut H. Nguyen, and Nuttawat Visaltanachoti. 2017. "A Note on Intraday Event Studies," Working Paper.
- McWilliams, Abigail., and Siegel, Donald. 1997. "Event Studies in Management Research: Theoretical and Empirical Issues," *The Academy of Management Journal*, 40(3): 626-657.
- Miles, David K., Michael Stedman, and Adrian H. Heald. 2020. "Stay at Home, Protect the National Health Service, Save Lives: A Cost Benefit Analysis of the Lockdown in the United Kingdom," *International Journal of Clinical Practice*.
- Sheridan, Adam, Asger Lau Andersen, Emil Toft Hansen, and Niels Johannesen. 2020. "Social Distancing Laws Cause Only Small Losses of Economic Activity During The COVID-19 Pandemic in Scandinavia," *PNAS* August 25, 2020 117 (34), <https://doi.org/10.1073/pnas.2010068117>
- Smith, Simon. February 27, 2013. "FTSE introduces UK and Italian equivalents of the VIX," *ETF Strategy*, <https://www.etfstrategy.com/ftse-introduces-uk-and-italian-equivalents-of-the-vix-the-ftse-implied-volatility-index-series-ivi-11335/>
- Zaremba, Adam, Renatas Kizys, David Y. Aharon, and Ender Demir. 2020. "Infected Markets: Novel Coronavirus, Government Interventions, and Stock Return Volatility around the Globe," *Finance Research Letters*, 35, <https://doi.org/10.1016/j.frl.2020.101597>
- Zhang, Dayong, Min Hu, and Qiang Ji. 2020. "Financial Markets under the Global Pandemic of COVID-19," *Finance Research Letters*, 35, <https://doi.org/10.1016/j.frl.2020.101528>

Appendices

Appendix 1. Estimating the Event Time: An Example

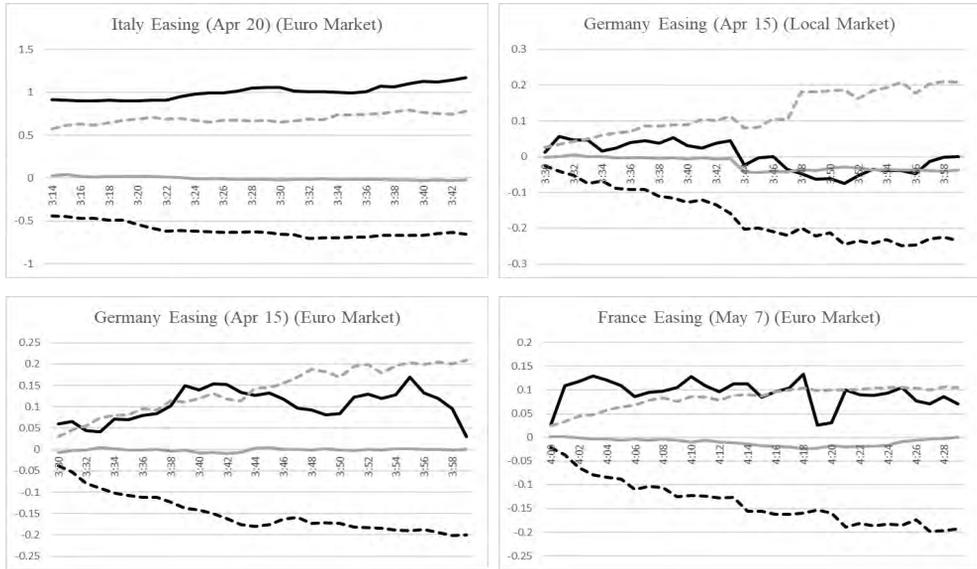
This appendix provides an example for estimating the event minute when this information is not readily available. The times for different events are estimated differently, and the following three-step procedure is used to estimate the time for France's reopening announcement.

Step 1: *Identifying the publication time of the relevant news article.* After an extensive search, we found that a French newspaper, France24, published an article on this event at 14:38 of July 5, 2020. See [this link](#).

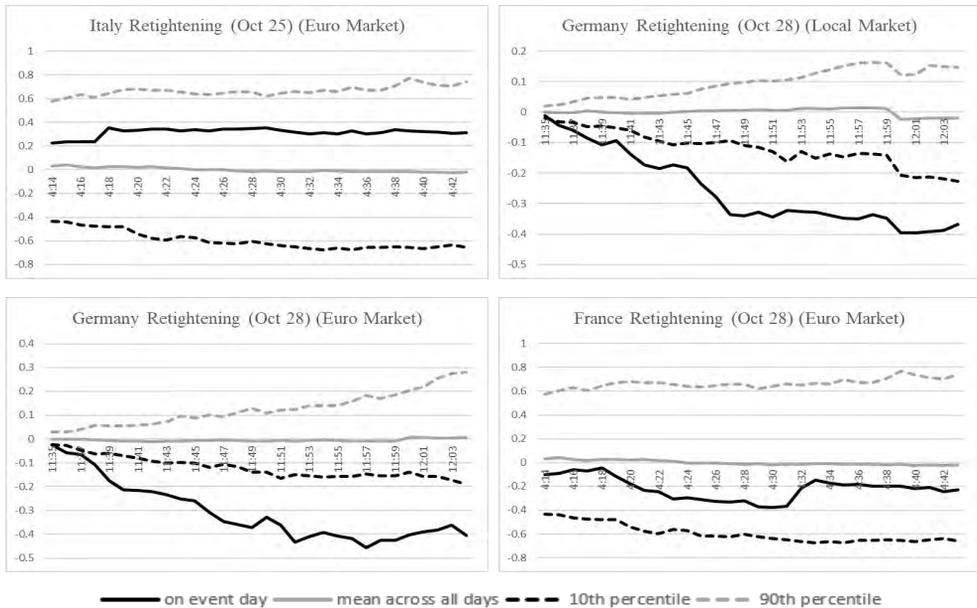
Step 2: *Confirming the time zone of the publication time.* We then checked another article by France24 published on the day when we were doing the search (November 12, 2020). It had already published an article about the US election at "11:33", when the actual time in Washington DC was only 8:39 am (Eastern time). This means that the time shown in France24's article is in French time.

Step 3: *Inferring the time of the announcement.* The article about France's easing (i.e., the article in Step 1) is long and may take some time to prepare, so it is hard to estimate the time of the announcement. However, the same article cited a reporter's tweet, which shows 10:15 am of May 7, 2020 (and it must be in French time, according to Step 1). Since sending a tweet takes only a few minutes, we can infer that the announcement time must be a few minutes before 10:15 am French time. In the end, we use 10 am French time of May 7 as our event time for this event, which is 4 am Eastern time (as shown in Figure 3).

Appendix Figure 1. Prediction Errors for Six-Month Volatility (Easing)

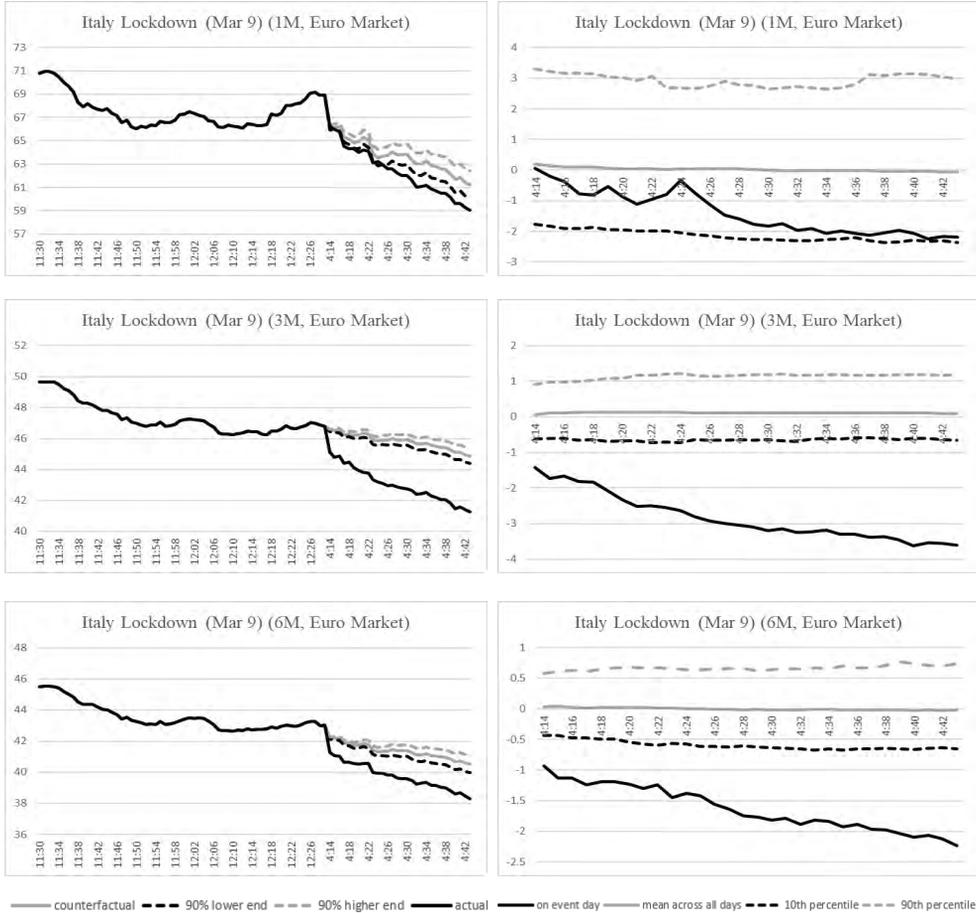


Appendix Figure 2. Prediction Errors for Six-Month Volatility (Retightening)



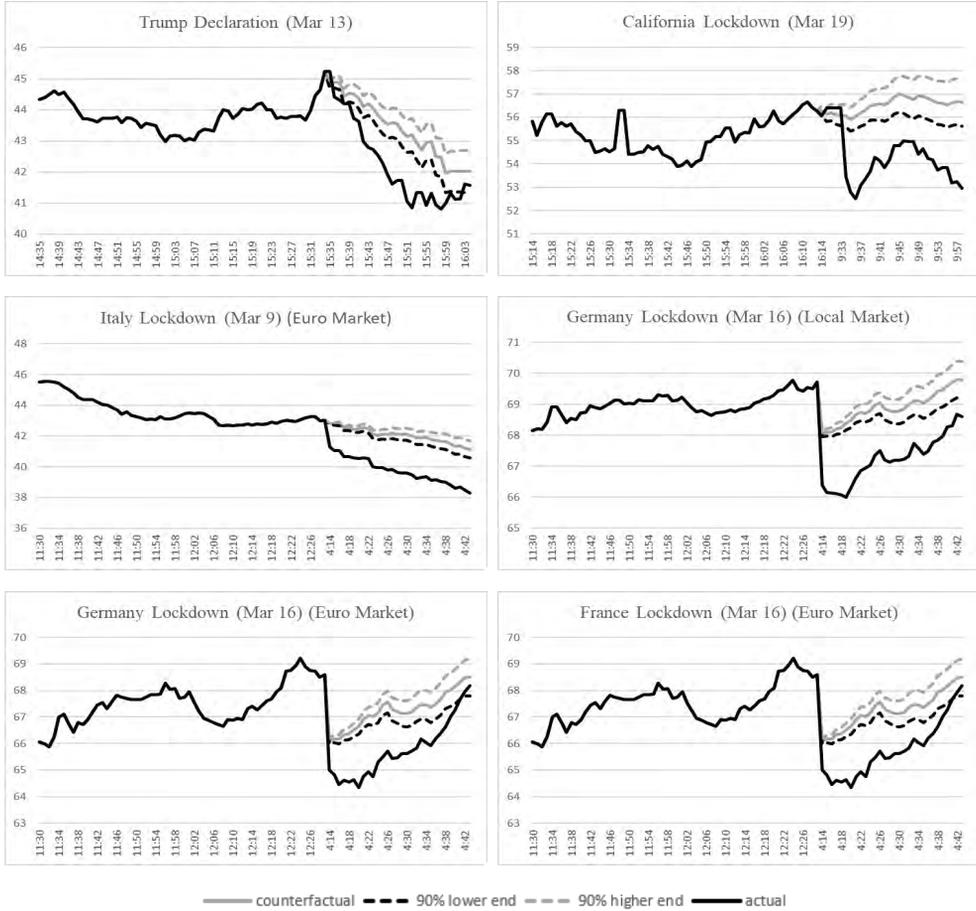
Covid Economics 79, 27 May 2021: 1-40

Appendix Figure 3. Volatility Responses Across Maturities: Italy's Initial Tightening

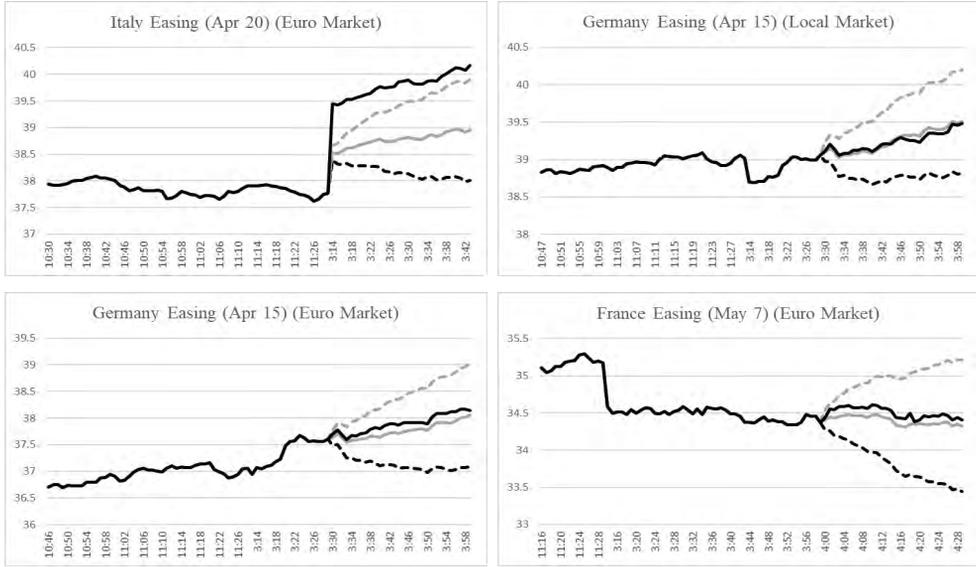


Covid Economics 79, 27 May 2021: 1-40

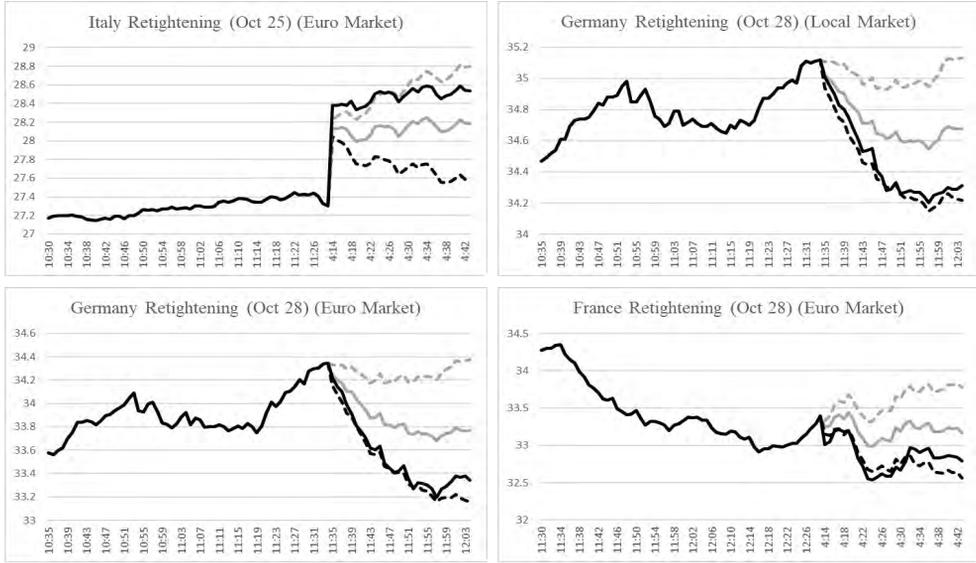
Appendix Figure 4. ARIMA Model: Six-Month Volatility Indices (Initial Tightening)



Appendix Figure 5. ARIMA Model: Six-Month Volatility Indices (Easing)



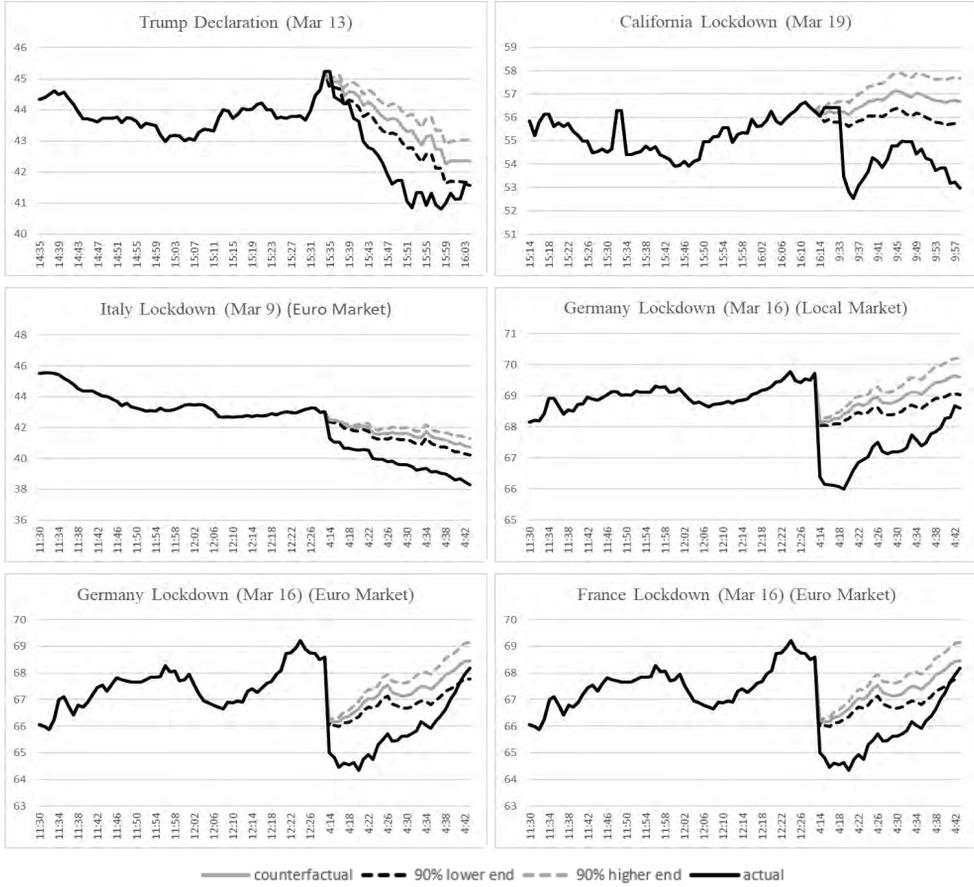
Appendix Figure 6. ARIMA Model: Six-Month Volatility Indices (Retightening)



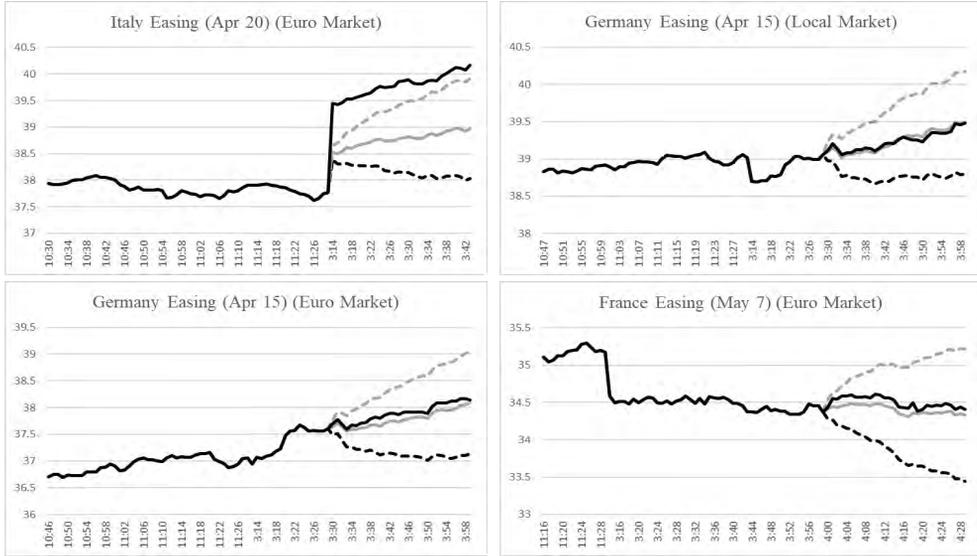
— counterfactual — — 90% lower end - - - 90% higher end — actual

Covid Economics 79, 27 May 2021: 1-40

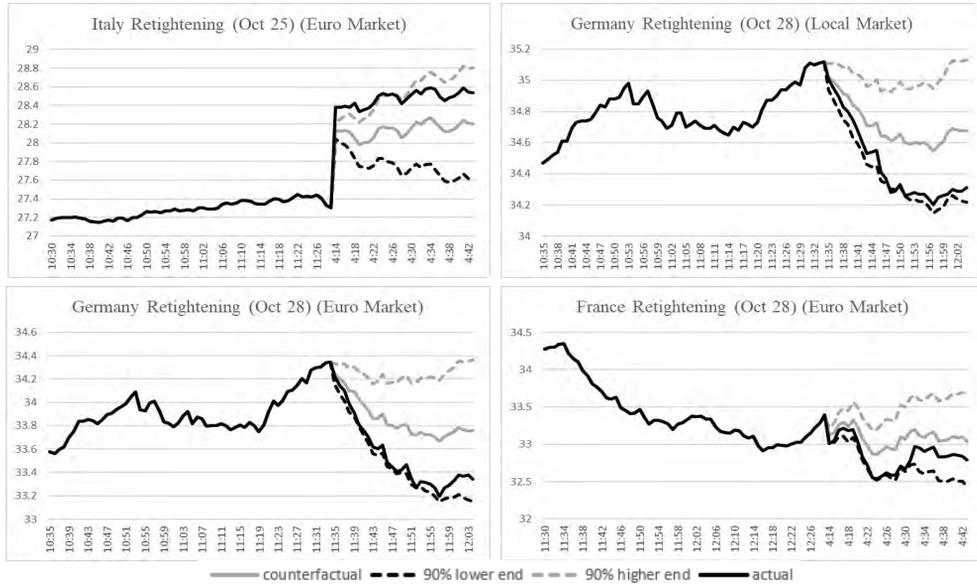
Appendix Figure 7. EGARCH Model: Six-Month Volatility Indices (Initial Tightening)



Appendix Figure 8. EGARCH Model: Six-Month Volatility Indices (Easing)



Appendix Figure 9. EGARCH Model: Six-Month Volatility Indices (Retightening)



Appendix Table 1. Summary Statistics of Panel Regression Data

	Number of obs	Mean	Sd	Min	Max
Panel (A): Initial tightening					
V_3M_pct	401	108.3	922.4	-4,881.7	4,013.7
V_6M_pct	401	96.5	791.8	-4,875.3	3,436.8
StringencyIndex	401	41.8	35.5	0.0	93.5
Cases_pct	401	18.2	37.4	0.0	400.0
Cases_pct_SI	401	660.3	1,263.5	0.0	11,174.8
StockPrice_pct	401	-23.4	296.3	-1,692.4	1,097.6
StockPrice_std	401	45.6	72.9	0.3	804.3
Panel (B): Easing tightening					
V_3M_pct	257	-11.3	625.0	-2,487.0	3,672.6
V_6M_pct	257	-16.0	450.1	-1,252.8	2,705.7
StringencyIndex	257	64.6	9.6	42.6	93.5
Cases_pct	257	1.0	1.1	0.1	7.5
Cases_pct_SI	257	66.9	81.9	3.4	519.1
StockPrice_pct	257	21.8	169.5	-589.4	567.3
StockPrice_std	257	30.5	34.4	0.4	153.4
Panel (C): Retightening stage					
V_3M_pct	289	-1.9	442.3	-2,195.0	2,440.0
V_6M_pct	289	-5.2	270.1	-730.6	1,553.5
StringencyIndex	289	57.8	8.6	43.5	69.9
Cases_pct	289	1.4	1.4	0.0	9.3
Cases_pct_SI	289	77.5	81.6	1.5	628.6
StockPrice_pct	289	1.2	114.4	-437.5	322.1
StockPrice_std	289	25.4	28.3	0.3	175.6

Notes: (1) pct = percent change; std = standard deviation; Cases_pct_SI is the interaction of COVID case percent change and stringency index. (2) Because of the rescaling, the units of the volatility percent change (e.g., V_3M_pct) and of the stock price percent change are basis point (i.e., 1/100 percent); the units of the Cases_pct and StringencyIndex remain as percent.

Appendix Table 2. Initial Tightening Stage Panel Regressions in the Benchmark Model

	(1)	(2)	(3)	(4)	(5)	(6)
	SI	Cases	Interaction	SI	Cases	Interaction
	FE	FE	FE	RE	RE	RE
StringencyIndex	-1.868 (0.101)		-1.462 (0.237)	-1.996* (0.072)		-1.515 (0.213)
Cases_pct		3.092*** (0.004)	3.358** (0.012)		3.173*** (0.002)	3.418*** (0.010)
Cases_pct_SI			-0.018 (0.659)			-0.019 (0.640)
Observations	401	401	401	401	401	401
R-squared	0.007	0.021	0.027			

Notes: (1) p-values are in parentheses. (2) FE = fixed effect; RE = random effect. (3) Constant not shown.

Appendix Table 3. Initial Tightening Stage Panel Regressions in the Stock Price Model

	(1)	(2)	(3)	(4)	(5)	(6)
	SI	Cases	Interaction	SI	Cases	Interaction
	FE	FE	FE	RE	RE	RE
StringencyIndex	-0.850 (0.321)		0.425 (0.650)	-0.998 (0.233)		0.354 (0.701)
Cases_pct		0.312 (0.706)	2.082** (0.037)		0.436 (0.595)	2.196** (0.025)
Cases_pct_SI			-0.099*** (0.002)			-0.101*** (0.001)
StockPrice_pct	-1.778*** (0.000)	-1.777*** (0.000)	-1.830*** (0.000)	-1.774*** (0.000)	-1.771*** (0.000)	-1.825*** (0.000)
Lag_StockPrice_pct	0.174* (0.086)	0.173* (0.094)	0.131 (0.206)	0.178* (0.077)	0.178* (0.082)	0.135 (0.188)
Observations	396	396	396	396	396	396
R-squared	0.464	0.463	0.478			

Notes: (1) p-values are in parentheses. (2) FE = fixed effect; RE = random effect. (3) Constant not shown.

Appendix Table 4. Seemingly Unrelated Regression Results (Full Sample, All Stages)

	(1)	(2)	(3)	(4)
	V_3M_pct	V_3M_pct	V_3M_pct	V_3M_pct
V_3M_pct				
StringencyIndex	-3.287*** (0.000)		-2.825*** (0.004)	-1.004 (0.183)
Cases_pct		4.448*** (0.000)	3.452*** (0.004)	1.942** (0.030)
Cases_pct_SI			0.025 (0.468)	-0.080*** (0.003)
StockPrice_pct				-2.234*** (0.000)
Lag_StockPrice_pct				-0.027 (0.743)
(mean) Std_SP				-0.837* (0.069)
Lag_StockPrice_std				1.678*** (0.000)
V_6M_pct				
StringencyIndex	-2.423*** (0.001)		-1.756** (0.027)	-0.485 (0.400)
Cases_pct		3.500*** (0.000)	3.457*** (0.000)	2.124*** (0.002)
Cases_pct_SI			-0.012 (0.660)	-0.097*** (0.000)
StockPrice_pct				-1.902*** (0.000)
Lag_StockPrice_pct				0.101 (0.109)
(mean) Std_SP				-2.060*** (0.000)
Lag_StockPrice_std				2.937*** (0.000)
Observations	924	924	924	919
R ²	0.014	0.026	0.034	0.469

Notes: (1) p-values are in parentheses. (2) FE = fixed effect; RE = random effect. (3) Constant not shown.

Humans against virus or humans against humans: A game theory approach to the COVID-19 pandemic¹

Santiago Forero-Alvarado,² Nicolás Moreno-Arias³ and Juan J. Ospina-Tejeiro⁴

Date submitted: 7 May 2021; Date accepted: 21 May 2021

Externalities and private information are key characteristics of an epidemic like the Covid-19 pandemic. We study the welfare costs stemming from the incomplete information environment that these characteristics foster. We develop a framework that embeds a game theory approach into a macro SIR model to analyze the role of information in determining the extent of the health-economy trade-off of a pandemic. We apply the model to the Covid-19 epidemic in the US and find that the costs of keeping health information private are between USD \$5.9\$ trillion and USD \$6.7\$ trillion. We then find an optimal policy of disclosure and divulgation that, combined with testing and containment measures, can improve welfare. Since it is private information about individuals' health what produces the greatest welfare losses, finding ways to make such information known as precisely as possible, would result in significantly fewer deaths and significantly higher economic activity.

- 1 We are grateful to Hernando Vargas, Juan Esteban Carranza, and Franz Hamann for their valuable comments on the paper. We also express special thanks to Marcela De Castro and Sara Naranjo for conversations, comments and suggestions that were relevant during the process. Sebastián Beltrán provided excellent research assistance.
- 2 Forecasting Professional at the Macroeconomic Modeling Department, Banco de la República.
- 3 Forecasting Specialized Professional at the Macroeconomic Modeling Department, Banco de la República.
- 4 Chief Officer for Monetary Policy and Economic Information, Banco de la República.

Copyright: Santiago Forero-Alvarado, Nicolás Moreno-Arias and Juan J. Ospina-Tejeiro

1 Introduction

A pandemic caused by a virus is a health shock that may induce consumers to reduce their activities to protect themselves and reduce the probability of contagion. It may also induce governments and policy makers to implement restrictive measures to slow down and control the spread of the virus. Both private and government responses produce a trade-off between economic and health outcomes (Eichenbaum et al. (2020a)). A characteristic of a pandemic like COVID-19, is that there are externalities (for example the infection externality), and information asymmetries (a person may know she is sick, but others may not), which sets up an environment of strategic interaction among consumers. Clearly, the fact that people lacks information about other people's health complicates the containment of the virus, magnifies the externalities and affects how people choose the extent of their social and economic activity.

However, concerns about the collection and use of private information about the health of individuals have been important in the efforts to control the COVID-19 disease, particularly the use of apps and technology to trace and track infected individuals. The following guidelines for contact tracing in the COVID-19 pandemic are provided by the The Centers for Disease Control and Prevention (CDC) in its website: "All public health staff involved in case investigation and contact tracing activities with access to such information should sign a confidentiality statement acknowledging the legal requirements not to disclose COVID-19 information. Efforts to locate and communicate with clients and close contacts must be carried out in a manner that preserves the confidentiality and privacy of all involved. This includes never revealing the name of the client to a close contact unless permission has been given (preferably in writing), and not giving confidential information to third parties (e.g., roommates, neighbors, family members)."¹

Given the importance of information and the limitations from privacy, in this paper we develop an analytical framework that combines a game theory set-up and the Macro-SIR model proposed in Eichenbaum et al. (2020a) to make explicit how information influences the spread of an epidemic and quantify its importance. We also extend the model to include asymptomatic infected people, an important characteristic of the COVID-19 pandemic as argued by Berger et al. (2020), and one that certainly entails a key source of information loss. Asymptomatic individuals increase infections, but as long as they do not die, taking them into account will change quantitatively the predictions of the model about deaths and the fall in economic activity with respect to a classic SIR. Our framework allows to understand and quantify the costs of privacy from a microeconomics perspective. It also provides a way to study how different degrees of information can determine how an infectious disease spreads and evolves over time, how it affects economic outcomes, and what the optimal mix of policy tools could be to reduce its negative effects.

As a case study, we apply the model to the US and analyze the recent COVID-19 crisis. We show that the

¹<https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/Confidentiality-Consent.html>

lack of both private and common information generates relevant welfare losses, albeit the greater losses are associated with the latter. Accordingly, we study and quantify the effects of a policy of disclosure and divulgation of private health information about individuals in alleviating the negative health and economic effects of a pandemic. We argue that disclosure and active divulgation of precise information about who is infected can have large welfare effects, especially when combined with the more traditional policy tools of testing and containment. We find that what we label the “Optimal Mix” of policies, calls for the use of containment and testing, but the welfare gains from these two policies are overshadowed by the gains from precise divulgation.

In our setup, we vary the information available to individuals. In a first case, which we call Total Incomplete Information (TII), agents do not know their own or other people’s health condition. In the second case, Partial Incomplete Information (PII), people know their own health condition but are ignorant of the infection status of others. Finally, in the third case, one of Complete Information (CI), everything is known. These different information worlds produce widely different welfare outcomes. Total Incomplete Information is of course the worst of the worlds, and it is arguably the closest to reality.

We study optimal policies in these different information worlds and how they allow us to improve outcomes and eventually go from one world to the next. We begin with containment policies, which have been the main policy tools used during the COVID-19 pandemic. We generally find that containment generates little welfare gains, and that the ability to make them targeted (so called conditional) and their optimal extent depends on the available information. In general, the scarcer information is, the more stringent and generalized optimal containments must be. The economic gains from general containment improve consumption in about USD 3.5 trillion when information is gathered and incorporated allowing for conditional quarantines. Despite this improvement, conditional containments are still insufficient to compensate the welfare losses from the disease under incomplete information.

We then study testing as a tool to gather information and produce better aggregate estimates of the extent of the infection in the economy. Testing allows people to learn about their health status. We show that testing alone improves welfare in a rather modest amount. The reason is that it is a double-edge sword. On the one hand testing improves private information about the disease, helping tested people improve their decision making. On the other hand, while testing improves aggregate information it also creates an information asymmetry, which, in the absence of any other policy, results in infected asymptomatic people not reducing consumption and work, and thus increasing the spread of the infection. Nevertheless, a combination of testing and targeted quarantines could generate better results as seen in Eichenbaum et al. (2020b).

We subsequently analyze a policy of disclosure and divulgation of private information about people’s health status at an individual level. Depending on the costs of making this information public and of people

being able to use it, such policy is the game changer. A disclosure and divulgation policy can be optimally used accompanied by policies of testing and containment. First, containment helps internalize the infection externality, which is still present in a world of complete information. Second, together with testing, divulging will flatten the infections curve with a much smaller economic downturn. Unlike containments and testing only, it does so by enabling mutually beneficial transactions with no contagion risk to take place normally. Ultimately, this relaxes the trade-off between economic activity and public health. We estimate the potential gains of frequently divulging precise information to fall between USD 5.9 trillion and USD 6.7 trillion in dollars of 2019.

Altogether, our paper illustrates that the COVID-19 crisis can be thought about as an information problem, rather than a problem that needs to be controlled through stringent containment. Consequently, an appropriate policy response to the epidemic should aim at closing information gaps. We are aware that a full disclosure of private information brings out several considerations about privacy rights, however, thinking about how to make precise information available is worthwhile. Paraphrasing Dr. Tedros Adhanom Ghebreyesus, the President of the WHO, when he recommends “Testing, Testing and Testing”, we encourage authorities to do “Divulgation, Divulgation and Divulgation”.

Our paper makes two contributions. First, a framework to analyze an epidemic and its economic consequences as the result of many strategic interactions where information is of the essence. An advantage of such framework is that information’s importance can be quantified to guide policy decisions. In this sense, our paper is broadly related to the literature that links epidemiological models and macroeconomics to think about optimal policies as in Eichenbaum et al. (2020a), Acemoglu et al. (2020), Rowthorn and Toxvaerd (2020), Alvarez et al. (2020), Jones et al. (2020), Farboodi et al. (2020), and Garriga et al. (2020). Our second contribution is to highlight the importance of disclosing and divulging more precise and disaggregated information about people’s health status, which can become a powerful policy tool to reduce the economic health trade-off of an epidemic like the COVID-19. Since we make more explicit the role of information in the analysis of pandemic dynamics, our work contributes more closely related to Argente et al. (2020), Eichenbaum et al. (2020b) and Berger et al. (2020). A key difference with these papers is that we micro-found how information affects economic decisions.

The rest of the paper is organized as follows. In Section 2 we describe the model and how it changes under different information contexts. In Section 3 we study the ability of containment policies to improve health and economic outcomes in the face of different information structures. Then, in Section 5, we analyze the effects of testing and disclosure and divulgation as policy tools that, by reducing the effects of information deficiencies and asymmetries, can potentially achieve better health and economic outcomes. In Section 6 we conduct some exercises to consider the possibility that beliefs are not “right”, which is a possibility in the

presence of poor information about the disease. Finally, Section 7 concludes.

2 Model

Our model builds on the framework of Eichenbaum et al. (2020a), in which we couple an economic structure together with the epidemiological model of Kermack and McKendrick (1927). We depart from Eichenbaum et al. (2020a) in two important ways. First, we extend the epidemiology block to include asymptomatic cases. Second, the economic structure is built using game theory. That approach is motivated by people's high interdependence when engaging in economic activities with contagion risk and the existence of information asymmetries. Together, these changes give information a key role in the model. This helps us incorporate and underscore the idea that people's economic choices and how they get infected critically hinges on health information available to them.

The economy is populated by three classes of agents: the government, a continuum of identical firms, and representative households. The government collects taxes from consumption and redistributes them among the population. Firms produce a consumption good, C_t , choosing how many hours of labor to hire, N_t , and using a linear production technology in order to maximize profits Π_t :

$$\Pi_t = AN_t - w_t N_t$$

Households make decisions on consumption and work hours in a strategic environment, which we model as a game. These decisions are strategic because the health status of individuals and the information that they have about it, matters for an economic transaction to happen and for the potential health consequences of the interaction.

2.1 Game Setup

The players of this game are households who choose how much to consume and work at every moment of time t to maximize utility, which takes the following form:

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

where i indexes the health status. Throughout the model section we will refer to players as agents. Given that the game is set up in an economy during an epidemic, agents know there is a risk of getting infected when interacting with others. Thus, their economic decisions become intertwined with the health status of

others and of themselves.

At the beginning of the game, Nature plays first and randomly chooses two agents (i, j) from the population P_t to interact with each other in an economic transaction.² Agents are *ex-ante* identical and they will differ only by their type, which is the health status they get assigned by Nature. These types are drawn from the set $T_i = \{S, I^E, I^A, R^E, R^A\}$, where S is Susceptible, I^E is Symptomatic Infected, I^A is Asymptomatic Infected, R^E is Symptomatic Recovered and R^A is Asymptomatic Recovered³. After Nature's move, the two players have to choose (actions) consumption c_t^i and work hours n_t^i simultaneously. In the baseline version of the model, the main difference between I^A and I^E individuals is that only the latter will have their productivity negatively affected by the shock.

Players' payoffs are given by value functions that depend on their types and actions. Clearly, being in a strategic environment means that players strategies will depend upon each player's information set, particularly what they learn about their health and others' health after Nature moves. In particular, we will study the game under three information worlds, in which we vary the information assumptions of the game: 1) Complete information (CI); 2) Partial incomplete information (PII); 3) Total incomplete information (TII). Since agents are *ex-ante* identical, hereafter we study the game only from player i 's perspective without loss of generality.

2.2 Complete Information

Our first world is one where players know their own type and the type of each player they face. We will call this world the Complete Information (CI) case. Even though this may be an unrealistic scenario it will serve as the ideal benchmark.

Given the game set up, player i 's strategy is contingent on both player's types. This can be described by a tuple of dimension 25 (all combinations of the 5 types in T_i), which contains, for each combination of player types, a pair of actions for consumption and hours worked. Since what is relevant for the consumption and labor decision is the chance of getting infected, to solve the game we group the subgames into two categories: 1) No contagion risk for player i and 2) Positive contagion risk for player i .

2.2.1 No Contagion Risk for player i

The interactions $T_i \times T_j$ where player i faces no risk of getting infected are:

²Throughout the paper, we assume all interactions occur only between two people at a time.

³Symptomatic infected are people who exhibit symptoms. We assume that these symptoms are observationally unique and thus, the virus cannot be confused with another disease. Symptomatic Recovered people got infected, had symptoms and recovered while Asymptomatic Recovered got infected, did not have symptoms and recovered

$$\{S\} \times \{S, R^E, R^A\} \cup \{I^E, I^A, R^E, R^A\} \times T_j.$$

Player i chooses consumption and hours worked to maximize her value function, subject only to her budget constraint. Since in these interactions player i can safely disregard player j 's type, she has a dominant strategy. Nonetheless, player i is affected by her own type through her budget constraint due to the productivity shock associated with the virus. In other words, player i 's dominant strategy will vary depending on her health status. We now find the strategies for the different cases.

Player i is Infected: $T_i = \{I^E, I^A\}$

Player i 's type is I^Z with $Z \in \{E, A\}$, and takes actions $(c_t^{I^{Z*}}, n_t^{I^{Z*}})$ by solving:

$$\begin{aligned} \max U_t^{I^Z} &= u(c_t^{I^Z}, n_t^{I^Z}) + \beta \left[(1 - \pi_d^Z - \pi_r^Z) U_{t+1}^{I^Z} + \pi_r^Z U_{t+1}^{R^Z} \right] \\ \text{s.t.} & (1 + \mu_t) c_t^{I^Z} = w_t \phi^{I^Z} n_t^{I^Z} + \Gamma_t \end{aligned}$$

where $u(\cdot)$ is the instant utility function, π_d^Z is the mortality rate, π_r^Z is the recovery rate, and $\phi^{I^Z} \in [0, 1]$ is a parameter that captures the fall in infected people's labor productivity. We assume that for Asymptomatic Infected (I^A) mortality rate is zero ($\pi_d^A = 0$) and that their productivity does not get affected $\phi^{I^A} = 1$. For the Symptomatic Infected (I^E), we have $\pi_d^E = \pi_d$ and $\phi^{I^E} = \phi^I$. The government enters the problem in the budget constraint through lump-sum transfers, Γ_t , and through a containment rate, μ_t , which affects consumption. It is worth noting that the value function reflects the assumption that the cost of death is the foregone lifetime utility.

Optimal levels of consumption and hours are such that:

$$\frac{\partial u(c_t^{I^{Z*}}, n_t^{I^{Z*}})}{\partial c_t^{I^{Z*}}} = \lambda_t^{I^Z} (1 + \mu_t) \tag{1}$$

$$\frac{\partial u(c_t^{I^{Z*}}, n_t^{I^{Z*}})}{\partial n_t^{I^{Z*}}} = -\lambda_t^{I^Z} w_t \phi^{I^Z} \tag{2}$$

Player i has Recovered $T_i = \{R^E, R^A\}$

In these cases the optimal choices $(c_t^{R^{Z*}}, n_t^{R^{Z*}})$ with $Z \in \{E, A\}$ come from the solution to the optimization problem:

$$\max U_t^{R^Z} = u(c_t^{R^Z}, n_t^{R^Z}) + \beta U_{t+1}^{R^Z}$$

$$s.t.(1 + \mu_t)c_t^{R^Z} = w_t n_t^{R^Z} + \Gamma_t \tag{3}$$

So that consumption and hours worked satisfy the optimality conditions:

$$\frac{\partial u(c_t^{R^Z}, n_t^{R^Z})}{\partial c_t^{R^Z}} = \lambda_t^{R^Z} (1 + \mu_t) \tag{4}$$

$$\frac{\partial u(c_t^{R^Z}, n_t^{R^Z})}{\partial n_t^{R^Z}} = -\lambda_t^{R^Z} w_t \tag{5}$$

Player i is Susceptible $T_i = S$

If player i is Susceptible of getting infected but player j 's type belongs to $\{S, R^E, R^A\}$, there is no risk of contagion. Then player i solves the optimization problem below to find her actions $(c_t^{S,NI}, n_t^{S,NI})$.

$$\begin{aligned} \max U_t^{S,NI} &= u(c_t^{S,NI}, n_t^{S,NI}) + \beta U_{t+1}^S \\ s.t.(1 + \mu_t)c_t^{S,NI} &= w_t n_t^{S,NI} + \Gamma_t \end{aligned} \tag{6}$$

With the optimal levels of consumption and hours worked satisfying:

$$\frac{\partial u(c_t^{S,NI}, n_t^{S,NI})}{\partial c_t^{S,NI}} = \lambda_t^{S,NI} (1 + \mu_t) \tag{7}$$

$$\frac{\partial u(c_t^{S,NI}, n_t^{S,NI})}{\partial n_t^{S,NI}} = -\lambda_t^{S,NI} w_t \tag{8}$$

2.2.2 Contagion Risk for Player i

Player i faces a risk of contagion as long as she is susceptible and Player j is of type I^Z with $Z \in \{E, A\}$.

In these interactions, player i will choose the pair $(c_t^{S,I^Z}, n_t^{S,I^Z})$ by solving:

$$\begin{aligned} \max U_t^{S,I^Z} &= u(c_t^{S,I^Z}, n_t^{S,I^Z}) + \beta [(1 - \tau_t^{I^Z}) U_{t+1}^S + \tau_t^{I^Z} U_{t+1}^{I^Z}] \\ s.t.(1 + \mu_t)c_t^{S,I^Z} &= w_t n_t^{S,I^Z} + \Gamma_t \end{aligned} \tag{9}$$

$$, \wedge, \tau_t^{I^Z} = \pi_1 c_t^{S,I^Z} c_t^{I^Z} + \pi_2 n_t^{S,I^Z} n_t^{I^Z} + \pi_3 \tag{10}$$

With $\tau_t^{I^Z}$ being the probability of Player i getting infected by Player j , π_1 the probability of getting infected from consumption interactions, π_2 the probability of getting infected from work interactions, and π_3 the probability of getting infected in any other way. Simultaneously, the Player j solves its own optimization problem and acts according to the pair $(c_t^{I^Z*}, n_t^{I^Z*})$. Thus, these actions influence Player i 's optimal decisions as follows:

$$\frac{\partial u(c_t^{S,I^Z*}, n_t^{S,I^Z*})}{\partial c_t^{S,I^Z*}} + \beta \pi_1 c_t^{I^Z*} (U_{t+1}^I - U_{t+1}^S) = \lambda_t^{S,I^Z} (1 + \mu_t) \tag{11}$$

$$\frac{\partial u(c_t^{S,I^Z*}, n_t^{S,I^Z*})}{\partial n_t^{S,I^Z*}} + \beta \pi_2 n_t^{I^Z*} (U_{t+1}^I - U_{t+1}^S) = -\lambda_t^{S,I^Z} w_t \tag{12}$$

2.2.3 Aggregates and Equilibrium

Given that the game is symmetric for players and that before Nature randomly selects their type they are identical, to find aggregates we can just aggregate over i . Aggregating over players i such that $T_i \in \{I, R\}$ yields the following aggregate value functions:

$$\begin{aligned} R_t U_t^R &= R_t^E U_t^{R^E} + R_t^A U_t^{R^A} \\ I_t U_t^I &= I_t^E U_t^{I^E} + I_t^A U_t^{I^A} \end{aligned}$$

Aggregate consumption and hours for the infected and recovered population have analogous expressions.

When Player i is Susceptible, her value function, consumption and work take into account Player j 's type, so we need to integrate over all other individuals in the population (P_t):

$$\begin{aligned} U_t^S &= \frac{1}{P_t} \int_0^{P_t} U_t^S(j) dj \\ &= \frac{1}{P_t} [(S_t + R_t) U_t^{S,NI} + I_t^E U_t^{S,I^E} + I_t^A U_t^{S,I^A}] \end{aligned}$$

And consumption and hours can be found similarly to get:

$$\begin{aligned} c_t^S &= \frac{1}{P_t} [(S_t + R_t) c_t^{S,NI} + I_t^E c_t^{S,I^E} + I_t^A c_t^{S,I^A}] \\ n_t^S &= \frac{1}{P_t} [(S_t + R_t) n_t^{S,NI} + I_t^E n_t^{S,I^E} + I_t^A n_t^{S,I^A}] \end{aligned}$$

Finally, Susceptible aggregates are just $S_t U_t^S$, $S_t c_t^S$, and $S_t n_t^S$.

Government

Government may collect taxes on consumption (μ_t) to disincentivize interactions. This will capture the effect of lockdowns. The government also makes transfers Γ_t to households. The government’s budget constraint is given by:

$$\mu_t(S_t c_t^S + I_t c_t^I + R_t c_t^R) = \Gamma_t(S_t + I_t + R_t) \tag{13}$$

Market clearing

Merging together consumers’ and government’s budget constraints and using the production function we obtain the market clearing condition in the good and services markets:

$$S_t c_t^S + I_t c_t^I + R_t c_t^R = AN_t \tag{14}$$

While market clearing in the labor market must satisfy:

$$S_t n_t^S + I_t^A n_t^{I^A} + I_t^E \phi^I n_t^{I^E} + R_t n_t^R = N_t \tag{15}$$

Population Dynamics

New infection cases T_t come from interactions between players i, j when there is risk of contagion:

$$\begin{aligned} T_t &= \int_0^{S_t} \int_0^{I_t^A} \tau_t^{I^A} dj di + \int_0^{S_t} \int_0^{I_t^E} \tau_t^{I^E} dj di \\ &= \pi_1 c_t^{S, I^A} S_t c_t^{I^A} I_t^A + \pi_2 n_t^{S, I^A} S_t n_t^{I^A} I_t^A + \pi_3 S_t I_t^A + \pi_1 c_t^{S, I^E} S_t c_t^{I^E} I_t^E + \pi_2 n_t^{S, I^E} S_t n_t^{I^E} I_t^E + \pi_3 S_t I_t^E \end{aligned}$$

Susceptible population evolves according to:

$$S_{t+1} = S_t - T_t$$

With the share of new infections that end up being asymptomatic given by χ^A , and the probability that an asymptomatic infected recovers given by π_r^A , total Asymptomatic Infected people in period $t + 1$ can be calculated as:

$$I_{t+1}^A = I_t^A + \chi^A T_t - \pi_r^A I_t^A$$

The number of Symptomatic Infected people in $t + 1$ will be equal to:

$$I_{t+1}^E = I_t^E + (1 - \chi^A) T_t - (\pi_r^E + \pi_d) I_t^E$$

where π_r^E and π_d are the probabilities of a symptomatic infected recovering and dying, respectively.

In period $t + 1$ the total infected, asymptomatic recovered, symptomatic recovered, and recovered populations are respectively:

$$I_{t+1} = I_{t+1}^A + I_{t+1}^E$$

$$R_{t+1}^A = R_t^A + \pi_r^A I_t^A$$

$$R_{t+1}^E = R_t^E + \pi_r^E I_t^E$$

$$R_{t+1} = R_{t+1}^A + R_{t+1}^E$$

Total deaths will accumulate over time according to:

$$D_{t+1} = D_t + \pi_d I_t^E$$

Finally, the economy's total population in $t + 1$ will be diminished by deaths occurred at time t :

$$P_{t+1} = P_t - \pi_d I_t^E$$

2.3 Partial incomplete information

In this second world, which we call Partial Incomplete Information (PII), every player knows her own type but ignores the type of others. Note that the number of subgames reduces to five in this information environment, because, for Player i , Player j 's type is actually one: unknown. Then a strategy for Player i is now a tuple of only five dimensions. In order to design a strategy Player i uses a Harsanyi prior F to assign probabilities to Player j 's potential types: p^S if $T_j = S$, p^{I^A} if it is I^A ($T_j = I^A$), p^{I^E} if it is I^E ($T_j = I^E$),

p^{RA} if it is R^A ($T_j = R^A$), and $p^{RE} = 1 - p^S - p^{IA} - p^{IE} - p^{RA}$ if it is R^E , $T_j = R^E$.

Notwithstanding the uncertainty about Player j 's type, optimization problems and solutions for subgames where Player $i \in \{I^E, I^A, R^E, R^A\}$ are identical to those already presented in the CI case (section 2.2), because the probability of getting infected is zero. Since this is not the case when Player i is Susceptible we write it here explicitly:

$$\begin{aligned} \max U_t^S &= u(c_t^S, n_t^S) + \beta \left[(1 - p_t^{IE} \tau_t^{IE} - p_t^{IA} \tau_t^{IA}) U_{t+1}^S + (p_t^{IE} \tau_t^{IE} + p_t^{IA} \tau_t^{IA}) U_{t+1}^I \right] \\ \text{s.t.} \quad &(1 + \mu_t) c_t^S = w_t n_t^S + \Gamma_t \\ &\tau_t^{IE} = \pi_1 c_t^S c_t^{IE} + \pi_2 n_t^S n_t^{IE} + \pi_3 \\ &\tau_t^{IA} = \pi_1 c_t^S c_t^{IA} + \pi_2 n_t^S n_t^{IA} + \pi_3 \end{aligned}$$

Strategically, Player i 's decisions take into account Player j 's best response and as such, her optimal consumption and hours worked are:

$$[c_t^{S*}] : \frac{\partial u(c_t^{S*}, n_t^{S*})}{\partial c_t^{S*}} + \beta \pi_1 (p_t^{IE} c_t^{IE*} + p_t^{IA} c_t^{IA*}) (U_{t+1}^I - U_{t+1}^S) = \lambda_t^S (1 + \mu_t)$$

$$[n_t^{S*}] : \frac{\partial u(c_t^{S*}, n_t^{S*})}{\partial n_t^{S*}} + \beta \pi_2 (p_t^{IE} n_t^{IE*} + p_t^{IA} n_t^{IA*}) (U_{t+1}^I - U_{t+1}^S) = -\lambda_t^S w_t$$

In terms of finding economic aggregates one can follow the same process as in the CI case (section 2.2). Nevertheless, note that in this case Susceptible people behave the same no matter who they interact with. Government budget constraint and market clearing conditions remain the same.

The total number of new infections is given by:

$$T_t = \int_0^{S_t} \int_0^{I_t^{IE}} \tau_t^{IE} dj di + \int_0^{S_t} \int_0^{I_t^{IA}} \tau_t^{IA} dj di = \pi_1 c_t^S S_t c_t^I I_t + \pi_2 n_t^S S_t n_t^I I_t + \pi_3 S_t I_t$$

while all other population dynamics behave as in the CI case (section 2.2).

It is worth noting that, the Macro-Sir Model in Eichenbaum et al. (2020a) is nested in our model. In

fact, it is a particular case of the PII world, in which there are no asymptomatic infections (i.e. $\chi^A = 0$).

2.3.1 Beliefs Dynamics

When information is partially incomplete, every interaction features an information asymmetry. However, we assume that this private information is collected by the government and made public as population aggregates. Later, we will explore the benefits from disclosing and divulging disaggregated information. For now, we assume all players can access this public aggregate information through government reports. Once they are informed, players go on to form their beliefs about the probabilities that the player they interact with is either I^E or I^A . Since population groups by health status changes over time, beliefs become dynamic:

$$\begin{aligned} p_t^{I^A} &= \frac{I_t^A}{P_t} \\ p_t^{I^E} &= \frac{I_t^E}{P_t} \end{aligned}$$

In section 6 we discuss in more detail the assumption that beliefs get these probabilities right.

2.4 Total incomplete information

The third world we study is one of Total Incomplete Information (TII), in which a player ignores other people's type and possibly her own type. In reality it is likely that an asymptomatic infected person does not know her health status. In our setup we assume this is the case. The uncertainty about one's own health status will affect consumption and work decisions and, in the aggregate, the pandemics dynamics will be different. The strategic behavior of the types Asymptomatic Infected, Asymptomatic Recovered and Susceptible will now be the same: in the absence of symptoms they behave as though there is always the risk of getting infected.

In the case of the Symptomatic Infected and Symptomatic Recovered we assume that because they exhibit or have exhibited symptoms, they do know their health status. The government can learn about it and publish aggregate statistics, but other players can not identify them individually. As they did in the previous two information cases, these types pick dominant strategies.

In this environment, the number of subgames becomes three. The two subgames that arise when Player i is I^E or R^E entail optimization problems and solutions that are the same as in the CI case (section 2.2). The subgame that arises when Player i is either susceptible, asymptomatic infected or asymptomatic recovered is the one we focus on now.

As in the PII case (section 2.3), Player i uses the Harsanyi prior F to assign probabilities to Player j 's possible types. To deal with the additional uncertainty about her own type, Player i when $T_i = \{S, R^A, I^A\}$ now employs another Harsanyi prior, G , that assigns probabilities: q^S if her type is S ($T_i = S$), q^{I^A} if it is I^A ($T_i = I^A$) and q^{R^A} if it is R^A ($T_i = R^A$). Player i will then solve (here A stands for asymptomatic including types S, R^A, I^A):

$$\begin{aligned} \max U_t^A &= q_t^S U_t^{A,S} + q_t^{I^A} U_t^{A,I^A} + q_t^{R^A} U_t^{A,R^A} \\ \text{s.t. } (1 + \mu_t) c_t^A &= w_t n_t^A + \Gamma_t \\ \wedge \tau_t^{I^E} &= \pi_1 c_t^A c_t^{I^E} + \pi_2 n_t^A n_t^{I^E} + \pi_3 \\ \wedge \tau_t^{I^A} &= \pi_1 c_t^A c_t^{I^A} + \pi_2 n_t^A n_t^{I^A} + \pi_3 \end{aligned}$$

With

$$\begin{aligned} U_t^{A,I^A} &= u(c_t^A, n_t^A) + \beta \left[(1 - \pi_r^A) U_{t+1}^{A,I^A} + \pi_r^A U_{t+1}^{A,R^A} \right] \\ U_t^{A,R^A} &= u(c_t^A, n_t^A) + \beta U_{t+1}^{A,R^A} \\ U_t^{A,S} &= u(c_t^A, n_t^A) + \beta \left[(1 - p_t^{I^E} \tau_t^{I^E} - p_t^{I^A} \tau_t^{I^A}) U_{t+1}^{A,S} + (p_t^{I^E} \tau_t^{I^E} + p_t^{I^A} \tau_t^{I^A}) U_{t+1}^I \right] \end{aligned}$$

Player i 's first-order conditions for consumption and hours worked are:

$$[c_t^A] : \frac{\partial u(c_t^{A^*}, n_t^{A^*})}{\partial c_t^{A^*}} + q_t^S \beta \pi_1 \left(p_t^{I^E} c_t^{I^E*} + p_t^{I^A} c_t^{I^A*} \right) \left(U_{t+1}^I - U_{t+1}^{A,S} \right) = \lambda_t^A (1 + \mu_t)$$

$$[n_t^A] : \frac{\partial u(c_t^{A^*}, n_t^{A^*})}{\partial n_t^{A^*}} + q_t^S \beta \pi_2 \left(p_t^{I^E} n_t^{I^E*} + p_t^{I^A} n_t^{I^A*} \right) \left(U_{t+1}^I - U_{t+1}^{A,S} \right) = -\lambda_t^A w_t$$

2.4.1 Aggregates and Equilibrium

Aggregate economic variables are obtained by a process analogous to that of the PII case (section 2.3).

Government

The new government budget constraint is:

$$\mu_t((S_t + I_t^A + R_t^A) c_t^A + I_t^E c_t^{I^E} + R_t^E c_t^{R^E}) = \Gamma_t(S_t + I_t + R_t) \tag{16}$$

Equilibrium

Market clearing conditions are:

$$(S_t + I_t^A + R_t^A) c_t^A + I_t^E c_t^{I^E} + R_t^E c_t^{R^E} = AN_t \tag{17}$$

$$S_t n_t^A + I_t^A n_t^A + I_t^E \phi^I n_t^{I^E} + R_t n_t^R = N_t \tag{18}$$

Population Dynamics

New infection cases are given by:

$$\begin{aligned} T_t &= \int_0^{S_t} \int_0^{I_t^E} \tau_t^{I^E} dj di + \int_0^{S_t} \int_0^{I_t^A} \tau_t^{I^A} dj di \\ &= \pi_1 c_t^A S_t c_t^I I_t + \pi_2 n_t^A S_t n_t^I I_t + \pi_3 S_t I_t \end{aligned}$$

The rest of the population dynamics remain unchanged relative to what was explained in the CI case (section 2.2).

2.4.2 Beliefs Dynamics

Contrary to the PII case (section 2.3), here players do not know their own health status when they have not exhibited symptoms. We will assume that all of those players believe they are susceptible, so that $q_t^S = 1$. The lack of private information for the asymptomatic also means that people cannot observe population aggregates about Asymptomatic Infected. Thus, players must form their beliefs about the probability of encountering Asymptomatic Infected players differently to how they did in the PII case (section 2.3). One can have beliefs that lie inside a neighborhood of the true probability:

$$p_t^{I^A} = \frac{I_t^A}{P_t} * (1 + \varepsilon_t^{I^A})$$

We will initially assume that this error in assessing the true probability, $\varepsilon_t^{I^A}$, is zero. This assumption can be relaxed, something we discuss in Section 6.

2.5 Calibration

Since our model follows the economic structure of Eichenbaum et al. (2020a), we take some parameters directly from them. Such is the case of A , θ and β . This parameters are set so that in the pre-epidemic

steady state the model is able to match some relevant economic statistics of the US economy.

Similarly, our calibration of the epidemiological parameters is also based in Eichenbaum et al. (2020a). We take the value of the parameter ϕ^I exactly from their model and maintain their assumptions about the herd immunity threshold (60% of initial population) and the time it takes for an infected person to either recover or die (14 days).

The other epidemiological parameters cannot have the same values because we incorporate asymptomatic infections. Nonetheless, we do use their method to calibrate such parameters. In particular, π_1, π_2, π_3 and π_d are set to match the same aggregate transmissions and mortality patterns as in Eichenbaum et al. (2020a). We assume that symptomatic and asymptomatic infected people share these transmission parameters. Finally, we use the 40% estimate of the CDC (2020) for the share of total infections that are asymptomatic and calibrate χ^A to match this. The table below summarizes the calibration used in our model.

Parameter	Value
A	39.835
β	$0.96^{\frac{1}{52}}$
θ	0.0013
ϕ^I	0.8
π_1	$7.8408e^{-8}$
π_2	$1.2442e^{-4}$
π_3	0.3902
π_d	0.0032
π_R^A	0.3889
π_R^E	0.3857
χ^A	0.3993

2.6 Welfare Analysis

In sections 2.2, 2.3 and 2.4, we study the influence of public and private information in players' decisions. In particular, we show how their optimal decisions on consumption and work are modified in response to changes in their information sets. This section evaluates the effects of such changes in the aggregate social welfare of the economy during a span of five years. Our simulation technique follows the algorithm exposed in Eichenbaum et al. (2020a), where value functions are iterated backwards and the epidemiological block forward. Given that we simulate our model in a deterministic fashion, the relevant welfare indicator is the weighted sum of the value function of each type of player, U_t , at the initial period:

$$U_t = S_t U_t^S + I_t U_t^I + R_t U_t^R$$

This indicator summarizes the two forces in action during the pandemic's evolution: economic activity, and people's health status and deaths. Table 1 contains the value of this indicator across the three cases considered so far. These results show that the Complete Information Case is our best possible scenario, followed by the Partial Incomplete Information and the Total Incomplete Information cases. Hence, one can see how losing information completeness gradually worsens welfare. This happens because poorer information prevents players from understanding the nature of their interactions and making the proper choices.

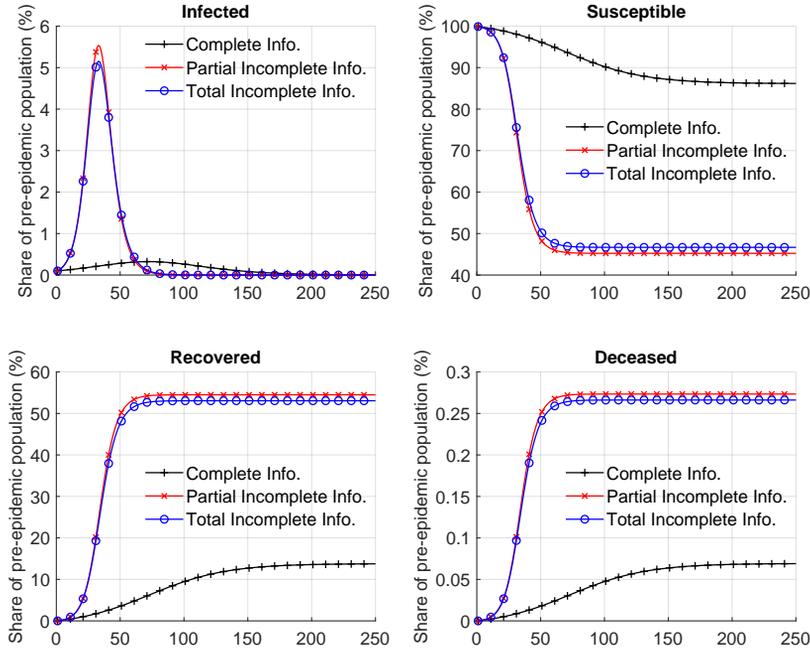
Furthermore, the welfare losses in each scenario can be understood through two different channels. First, the fall in consumption can be used as a proxy of the size of the recession induced by the epidemic. Our calibration implies that the pre-epidemic per capita annual consumption is 58.000 USD. We take this value and multiply it by the cumulative fall in aggregate consumption to obtain the monetary economic loss of the epidemic during the five years horizon. Second, our calibration also implies the statistical value of life is 9.3 million US dollars of 2019, so we can use this figure to quantify the costs of the deaths caused by the epidemic⁴. Let's recall that the cost of a death in the model is equal to the present value of foregone utility. For the CI case these losses are equal to 169.676.100.000 USD and 2.117.610.000.000 USD, respectively.

However, we are not only interested in quantifying the costs of the epidemic per se, but the costs or benefits of different information settings. Thus, we find that the absence of commonly known information about other players' health statuses (i.e. the PII case) causes additional losses of 1.025.808.300.000, due to lower economic activity, and of 6.273.036.000.000 associated with deaths. Analogously, we see that when private information is also absent (i.e. the TII case), the economy suffers an additional loss of 287.004.300.000 with respect to the PII case, while the cost of deaths actually decreases in 220.968.000.000 USD.

Figure 1 shows how the Complete Information case has a considerably lower number of deaths than the other two. This is because the public and total availability of private information about other's health status allows each player to reduce the intensity of interactions that bear contagion risk, flattening thusly the infections curve. Nevertheless, it is worth noting that this flat curve implies that Herd Immunity is not reached within the horizon considered, despite the epidemic is controlled and dies out as a consequence of the information completeness. The infection curves of the other two cases considered are quite similar, although it is slightly lower in the TII case. Here the difference is explained by the presence of a positive externality of losing private information about a player's type. Particularly, when asymptomatic infected players ignore

⁴The Unites States current GDP was 21.433.000.000.000 for 2019 <https://data.worldbank.org/country/united-states>

Figure 1: Population Dynamics



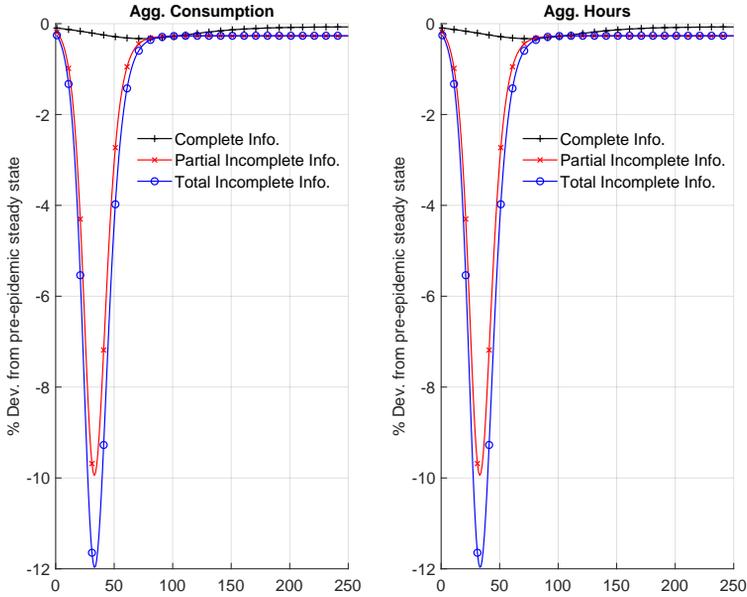
Covid Economics 79, 27 May 2021: 41-92

their type, they reduce their economic activity as a result of their false perception of being vulnerable to the virus, a behavior that ultimately reduces the propagation of the virus.

As we can see in Figure 2, the fall in economic activity is smallest in the CI case. In this scenario, susceptible agents only reduce their economic activity in risky interactions, which in turn implies a minor aggregate contraction. When susceptible agents are no longer able to make this distinction (i.e. the PII case) they reduce more aggressively their economic activity to avoid getting infected. Nonetheless, their efforts are not very effective and the infection curve rises considerably. This further decreases economic activity due to the greater aggregate loss of productivity when there are more symptomatic infected. Finally, aggregate economic variables also fall because deaths increase.

If agents are also unable to know their health status, the same channel that generated the positive externality in infections through asymptomatic infected, implies a negative effect on economic activity. This

Figure 2: Economic Aggregates



is magnified by the additional reaction of recovered asymptomatic agents that further reduces their economic activity as a result of their perception of contagion risk. The joint reaction of asymptomatic infected and recovered explain the greater fall in economic activity compared with the PII case.

3 Containment Measures

Containment measures, such as lockdowns, school closures, restrictions on gatherings, and other mobility restrictions, have been the primary policy intervention put in place by most countries in their attempt to limit the effects of the COVID-19 on fatalities and their health systems. The results have been heterogenous, with some countries apparently being more successful than others (Deb et al. (2020)).

These measures prevent many transactions from taking place, thereby reducing economic activity and creating a trade-off between economic and health outcomes. This trade-off is the motivation behind the literature that studies the interaction between an epidemic and the macroeconomy. Eichenbaum et al. (2020a) model a quarantine as a tax to consumption and find the optimal path for a simple containment policy in

which everybody is taxed, such that the benefits of lives saved outweigh the costs of worsening the recession. Eichenbaum et al. (2020b) study “smart” containment measures in which only sub-groups of the population are quarantined in the search for improving the health-economy trade-off. Acemoglu et al. (2020) and Berger et al. (2020) also study the gains from establishing quarantines for particular population groups.

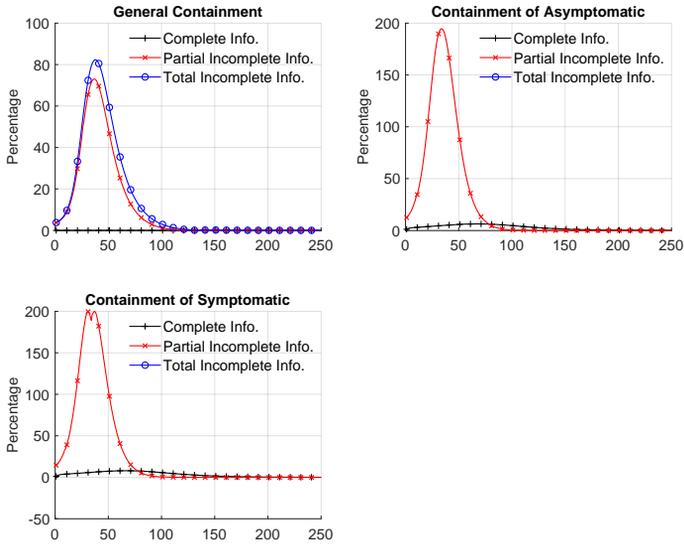
In this section we ask, how does information affect the ability of both general and conditional containment policies to improve the health-economy trade-off?

A general containment policy is one that applies to all the population. In the model it is instrumented through a consumption tax. The revenue raised by the government is then rebated back to all population groups by means of a lump-sum transfer. This type of containment is the one that is presented explicitly in the model of Section 2.

A conditional containment policy seeks to exploit information (kept private from other players) about the health status of people to establish focused quarantines, avoiding the confinement of people who have achieved immunity or who do not have the virus. Under this policy, only people who generate the negative externalities from contagion are put in lockdown. This is implemented in the model through a consumption tax rate μ_t^E for the symptomatic infected patients and μ_t^A for the asymptomatic infected people. At the same time, the government only makes transfers to people affected by the externality: the susceptible population in CI and PII cases. This conditional containment is in fact an imperfect compensation mechanism, given that the tax collection and the transfers do not occur by interaction, but by player’s type.

We now compare the level of containment that maximizes the discounted social welfare as defined in Section 2.6 for the different information cases. Figure 3 shows that in the Complete Information world optimal general containment is zero and there is only a small positive conditional containment. Notice that conditional containment is still desirable in this world, because the externality generated by infected people still exists.

Figure 3: Optimal Containment Policies Compared



Hence, in a world where information about the pandemic is fully available to everybody and everybody is able to process it efficiently, containment measures provide marginal gains. As shown in Table 1, this is true even if containments is conditional and its size is calibrated optimally, improving relative welfare to 0.0003%. The reason is that agents in the economy use the information to minimize market interactions where there is risk of contagion and engage normally in all other transactions. See Figures 4 and 5.

When information about people’s health status is kept private and agents cannot identify the contagion risk-free transactions, quarantine-type measures become optimal. As Table 1 shows, conditional containment policies yield better welfare outcomes vis-a-vis general containment, from -0.1741% to -0.164% relative welfare losses. This is the result of both higher aggregate consumption and hours worked (-1.14% vs. -4.97%). In turn, this is explained by 1) consumption of recovered patients does not fall; 2) due to the more-targeted transfers in conditional containment, susceptibles’ consumption does not fall as much; 3) despite the more pronounced decline in consumption of the infected, the flattening of the epidemic curve reduces the aggregate effect over time.

It is relevant to remark that conditional containment rates are orders of magnitude higher than those of general containment. To a large extent, this is due to the fact that asymptomatic infected people have

Figure 4: Population Dynamics - Comparison(CI)

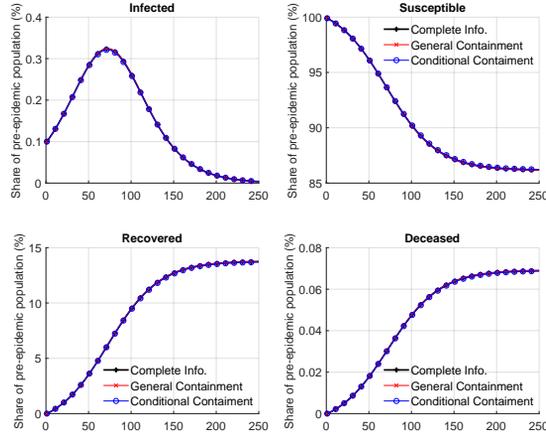
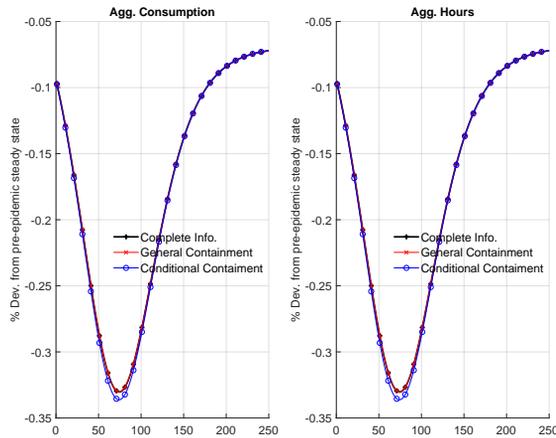


Figure 5: Economic Aggregates - Comparison(CI)



a dominant strategy, which is to engage in all possible transactions, posing a large negative externality on others. Moreover, this sort of containment allows the government to only impose a cost on those that are propagating the virus. In order to reduce the virus propagation, the conditional confinement rates must be high, reducing the consumption of those infected (see Figures 6 7). When there is common-knowledge, complete information, the negative effects of such strategy are attenuated since susceptibles are able to reduce

the intensity of interactions with asymptomatic patients.

Figure 6: Population Dynamics - Comparison(PII)

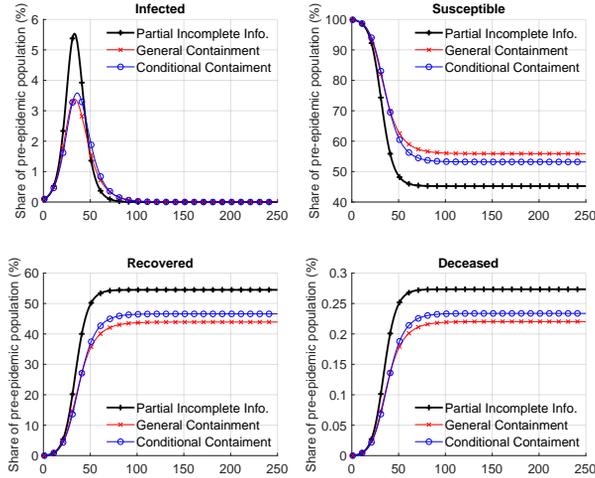
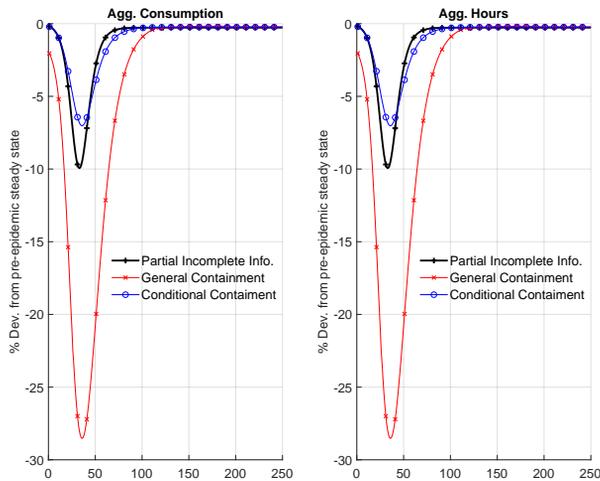
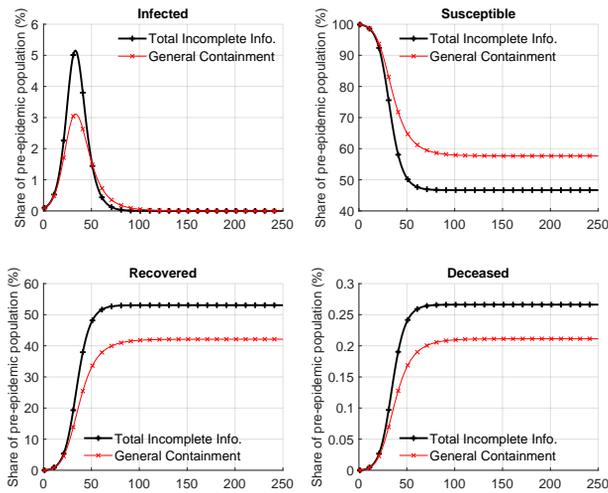


Figure 7: Economic Aggregates - Comparison(PII)



In the TII scenario, the unavailability of private information on health to the government and players makes it impossible to establish conditional containment measures. In this case policy makers are left with the option of general containment policies which, as the literature has shown, exacerbates the health economy trade-off: the reduction of contagion and fatalities comes at the expense of larger declines in economic activity as shown in Figures 8 and 9.

Figure 8: Population Dynamics - Comparison(TII)



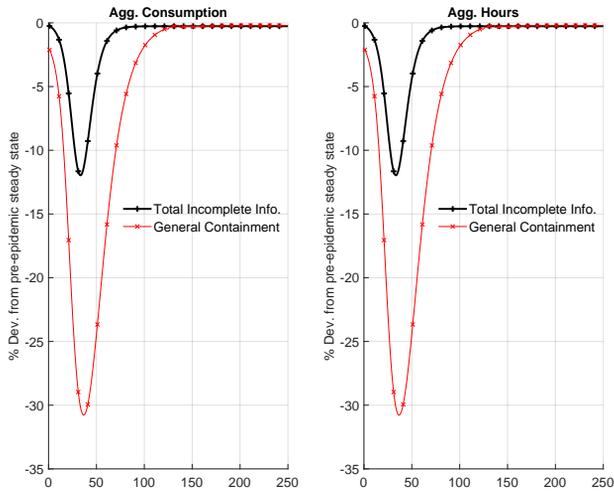
There are couple of points worth mentioning. First, Figure 3 shows that regardless of the type of containment being considered, the less information is available the more aggressive optimal containment measures must be. In other words, more complete information helps players choose better their interactions, therefore reducing the volume of hazardous interactions that need to be avoided or diminished through containments.

Finally, Table 1 reports the welfare losses from all confinement measures. Even though they allow to mitigate the effects from the externalities stemming from contagion and improve welfare, they still exhibit a large gap with respect to the ideal complete information world, with more infections, deaths, and larger reductions in economic activity. This is due to the fact that instead of relaxing the trade-off between economic and health outcomes, containment exploits the trade-off to control the infection. Even when there are conditional, confinements impose a consumption cost because agents are forced not to engage normally in transactions even if they pose no risk for themselves or others. In contrast, when all information about

Covid Economics 79, 27 May 2021: 41-92

each other's health is available to everybody the economic-health trade-off can be relaxed, so individuals are able to choose optimally the intensity of their interactions with each other and minimize thusly the risk of contagion without sacrificing their consumption.

Figure 9: Economic Aggregates - Comparison(TII)



4 Tables

Table 1: Welfare, economic and epidemiological results for Sections 2 and 3 - Information Scenarios and Containment Measures

This table summarizes the implications of the three information set-ups explained in Section 2 and the effects of the different containments considered in Section 3. These eight scenarios are simulated for 250 periods and analyzed through some indicators of welfare, economic activity, epidemiological dynamics and policy paths. The relative loss of aggregate welfare measures, in percentage points, the deviation of aggregate welfare under a certain specification with respect to the welfare of the Complete Information case. The maximum falls in aggregate consumption and aggregate hours are calculated relative to their pre-epidemic values and expressed in percentage points. The cumulative fall in aggregate consumption is the accumulation of all the foregone consumption during the simulation horizon, relative to a world where consumption remains all the time in its pre-epidemic value. The peak infection variable accounts for the total number of active infection cases at the height of the epidemic, as a percentage of the initial population. The final deaths and recoveries accumulate all the people that either died or recovered during the simulation horizon and express them as shares of the initial population. The containment measures show the maximum value of the consumption tax levied by the government for each type of containment.

	Complete Information (CI)	Partial Incomplete Information (PII)	Total Incomplete Information (TII)	CI General Containment	CI Conditional Containment	PII General Containment	PII Conditional Containment	TII General Containment
Relative loss of Aggregate Welfare	0	-0.2005	-0.2314	0	0.0003	-0.1741	-0.164	-0.1955
Max Fall in Aggregate Consumption %	-0.33	-9.94	-11.96	-0.33	-0.33	-28.53	-7.02	-30.78
Cumulative Fall in Aggregate Consumption %	-0.17	-1.24	-1.54	-0.17	-0.17	-4.97	-1.14	-5.88
Max Fall in Aggregate Hours %	-0.33	-9.94	-11.96	-0.33	-0.33	-28.53	-7.02	-30.78
Peak Infection %	0.32	5.53	5.15	0.32	0.32	3.37	3.58	3.11
Final Deaths %	0.06	0.27	0.26	0.06	0.06	0.22	0.23	0.21
Final Recoveries %	13.74	54.49	53.05	13.74	13.69	43.89	46.58	42.12
Peak of General Containment %	-	-	-	0	-	73.05	-	82.34
Peak of Symptomatic Containment %	-	-	-	-	7.89	-	199.98	-
Peak of Asymptomatic Containment %	-	-	-	-	6.28	-	194.58	-

Covid Economics 79, 27 May 2021: 41-92

5 Information Policy Tools

In the previous section we showed that optimal containment policies are not very effective in getting the economy close to a world of complete information in terms of welfare, deaths, and aggregate macroeconomic variables. In these section we consider policy tools that can actually fill the information gaps between the TII case and the first-best, so that the welfare gap between them closes.

In particular, we study two policy tools that can provide valuable information: testing and divulgation. Testing can fill the information gap that individuals have about their own health status. This information, gathered by health authorities, becomes privately known by the tested individual and disclosed at an aggregate level to all players. Divulgation makes this information publicly known, so that any given player can incorporate information on other people’s health in her decision making. As a matter of fact, a policy implemented along these lines was seen in South Korea, where authorities disclosed detailed information on infected people to manage the epidemic. More specifically, the mechanism consisted of an intensive use of text messages to disclose and propagate information about the health status of infected individuals and the places they had recently visited. Argente et al. (2020) examined the effects of such policy on the epidemiological dynamics in Seoul and found that people modified their commuting patterns in response to the information, which resulted in turn in a reduction in infections and deaths. One can think of divulgation as “painting people’s faces”. Here we think of divulgation as a tool to provide individuals with better information at the interaction level, so that by completing players’ information sets, we allow them to play out strategies where they can identify and engage normally in more mutually beneficial economic interactions (i.e. interactions with no contagion risk). The more information one provides, the lower the probability of engaging in risky interactions.

Starting from a world where information about health statuses is not common nor private, testing gets society closer to the PII case, whereas divulgation gets society closer to the CI case. We simplify the analysis by assuming that tests are performed only on asymptomatic people. In the model we do not test those who are already sick: we assume that the symptoms are enough to tell whether someone has the disease of interest. In this sense, testing serves the purpose of revealing private information to the agents about their own health status (type). With this in mind, the population subject to tests is given by $As_t - R_{t-1}^X - I_{t-1}^X$, where $As_t = S_t + I_t^A + R_t^A$ is the asymptomatic population at time t . From this population we subtract those asymptomatic who have recovered R_{t-1}^X because, due to the immunity assumption, once they know their type it will not change. Similarly, we also subtract people who were asymptomatic infected since the person will know her health status in the future from the recovery dynamics of the virus disease itself. A number of X_t tests are performed at random on this population, such that in expectation:

$$\begin{aligned} X_t &= X_t^S + X_t^I + X_t^R \\ &= X_t Prob(T_i = S) + X_t Prob(T_i = I^A) + X_t Prob(T_i = R^A) \\ &= X_t \frac{S_t}{As_t - R_{t-1}^X - I_{t-1}^X} + X_t \frac{I_t^A - I_{t-1}^X - \pi_r^A I_{t-1}^X}{As_t - R_{t-1}^X - I_{t-1}^X} + X_t \frac{R_t^A - R_{t-1}^X + \pi_r^A I_{t-1}^X}{As_t - R_{t-1}^X - I_{t-1}^X} \end{aligned}$$

The population dynamics of the groups that get to learn their type, that is, the tested asymptomatic recovered R_t^X ,

the tested asymptomatic infected I_t^X , and the tested susceptible S_t^X , are given by:

$$\begin{aligned} R_t^X &= R_{t-1}^X + X_t^R + \pi_r^A I_{t-1}^X \\ I_t^X &= I_{t-1}^X + X_t^I - \pi_r^A I_{t-1}^X \\ S_t^X &= X_t^S \end{aligned}$$

Once players get information about their own health through testing, an asymmetry of information arises and with it a rationale to make this information publicly known at an individual level. Divulgence is an instrument that theoretically gives the susceptibles the possibility to distinguish between contagion risky versus a contagion riskless interaction. However, since not everybody is tested and the policy maker does not know who all the susceptibles are, the information is aimed at two groups of people: the tested susceptible S_t^X and people who do not know their type $A_t^{NX} = S_t - S_t^X + I_t^A - I_t^X + R_t^A - R_t^X$. The divulgation mechanism consists in giving the available information to a number of people belonging to each of these groups and which we denote Z_t^{SX} and Z_t^A , respectively. The information that is revealed are: the symptomatic infected and recovered, the tested asymptomatic infected, the tested susceptibles and the tested recovered. The people who receive the information are selected randomly and its number is given in expectation by:

$$\begin{aligned} Z_t &= Z_t^{SX} + Z_t^A \\ &= Z_t^{SX} + Z_t^S + Z_t^I + Z_t^R \\ &= Z_t^{SX} + Z_t^A Prob(T_i = S) + Z_t^A Prob(T_i = I^A) + Z_t^A Prob(T_i = R^A) \\ &= Z_t^{SX} + Z_t^A \frac{S_t - S_t^X}{A_t^{NX}} + Z_t^A \frac{I_t^A - I_t^X}{A_t^{NX}} + Z_t^A \frac{R_t^A - R_t^X}{A_t^{NX}} \end{aligned}$$

where Z_t^S , Z_t^I y Z_t^R are the number of asymptomatic susceptibles, infected and recovered who do not know their type at time t and who are the receptors of the divulged private information.

In the model, the costs of testing and divulgation are financed through lump-sum taxes Γ_t^{Inf} levied on all agents in the economy, in such a way that:

$$\Gamma_t^{Inf} (S_t + I_t + R_t) = -mc_t^X X_t - mc_t^Z Z_t$$

The costs we consider have two components. The first component is the unit cost of a test (which we calibrate to be \$20). For simplicity, we abstract from any additional costs associated with testing. We also focus on the marginal cost, ignoring any initial investment required to set up testing infrastructure. The second component is the marginal cost of disclosing private information effectively at the individual level. This cost may include non-pecuniary costs such as ethical and regulatory restrictions on making public personal information, logistical and technological costs related to making the information available, and the capacity constraints that people may have when trying to process a high volume of information in an efficient way, such that they can use it to better choose how to act.

5.1 Modified Model

We now adjust the model to include the information tools introduced above. We do this for the TII case, so that this modified version can nest the three information cases explained in Section 2. When the cost of testing is zero, it is possible to get to the case of Partial Incomplete Information. Similarly, when the divulgation cost is zero and the information can be made available and processed perfectly, we can get to the Complete Information case.

In this version of the model we introduce new groups of agents. We will now distinguish the tested susceptibles, S_t^X , the tested asymptomatic infected, I_t^X , and the tested asymptomatic recovered, R_t^X . Additionally, due to divulgation, we will now consider four possible interaction cases for asymptomatic and tested susceptibles. These are 1) when they do not know the other's type; 2) when they know the other's type and there is no risk of contagion; 3) when they know the other's type is asymptomatic infected and there is risk of contagion; and 4) when they know the other's type is symptomatic infected and there is risk of contagion.

We now present the key elements of the game that change in this set-up due to the new types and interactions mentioned above. Nonetheless, one can notice that the symptomatic infected and symptomatic recovered face the same optimization problems than they did in Section 2.2. Similarly, tested asymptomatic infected and recovered behave as the asymptomatic infected and recovered of the CI case explained in Section 2.2.

5.1.1 Player i is asymptomatic

An asymptomatic player solves the following problem:

$$\begin{aligned} \max U_t^{A^j} &= q_t^S U_t^{S,A^j} + q_t^{I^A} U_t^{I^A,A^j} + q_t^{R^A} U_t^{R^A,A^j} \\ \text{s.t.} &(1 + \mu_t)c_t^{A^j} = w_t n_t^{A^j} + \Gamma_t + \Gamma_t^{Inf} \\ \wedge \tau_t^{I^E} &= \pi_1 c_t^{A^j} c_t^{I^E} + \pi_2 n_t^{A^j} n_t^{I^E} + \pi_3 \\ \wedge \tau_t^{I^A} &= \pi_1 c_t^{A^j} c_t^{I^A} + \pi_2 n_t^{A^j} n_t^{I^A} + \pi_3 \end{aligned}$$

With $J = \{U, NI, I^A, I^E\}$ indexing the different possibilities for information about player j 's type: unknown (U), known and no contagion-risk (NI), known and asymptomatic infected (I^A), and known and symptomatic infected (I^E). Also, the value functions on the right-hand side of the objective function are:

$$\begin{aligned} U_t^{I^A,A^j} &= u(c_t^{A^j}, n_t^{A^j}) + \beta \left[(1 - \pi_r^A) (U_{t+1}^{I^A,A^j}) + \pi_r^A U_{t+1}^{R^A,A^j} \right] \\ U_t^{R^A,A^j} &= u(c_t^{A^j}, n_t^{A^j}) + \beta U_{t+1}^{R^A,A^j} \end{aligned}$$

The problem above states that, despite varying with player j 's type, the total value function of an untested asymptomatic player weights the different value functions ($U_t^{S,A^j}, U_t^{I^A,A^j}, U_t^{R^A,A^j}$) according to her beliefs about her own health status ($q_t^S, q_t^{I^A}, q_t^{R^A}$).

However, note that the value functions when player i believes she is I^A or R^A will not change with player j 's type. Nevertheless, when player i believes she is susceptible, even if she does it to a small degree, her value function and the probabilities of contagion will be influenced by her risk perception; that is, by her information about player j 's type. See the Appendix (9) to explore the specifics of this problem.

Aggregation

The economic variables of the untested asymptomatic are given by two groups: those who received information about other players' types and those who did not. The former group is, in turn, subdivided into the different types she can encounter.

$$\begin{aligned}
 A_{st}U_t^A &= \frac{Z_t^A}{P_t} \int_0^{P_t} U_t^A(j) dj + (A_{st} - Z_t^A) U_t^{A^U} \\
 &= \frac{Z_t^A}{P_t} \left[(S_t^X + R_t^X + R_t^E) U_t^{A^{NI}} + I_t^X U_t^{A^{IA}} + I_t^E U_t^{A^{IE}} + A_t^{NX} U_t^{A^U} \right] \\
 &\quad + (A_{st} - Z_t^A) U_t^{A^U}
 \end{aligned}$$

Aggregation for consumption and hours worked follows this same procedure yielding:

$$\begin{aligned}
 A_{st}c_t^A &= \frac{Z_t^A}{P_t} \left[(S_t^X + R_t^X + R_t^E) c_t^{A^{NI}} + I_t^X c_t^{A^{IA}} + I_t^E c_t^{A^{IE}} + A_t^{NX} c_t^{A^U} \right] \\
 &\quad + (A_{st} - Z_t^A) c_t^{A^U}
 \end{aligned}$$

$$\begin{aligned}
 A_{st}n_t^A &= \frac{Z_t^A}{P_t} \left[(S_t^X + R_t^X + R_t^E) n_t^{A^{NI}} + I_t^X n_t^{A^{IA}} + I_t^E n_t^{A^{IE}} + A_t^{NX} n_t^{A^U} \right] \\
 &\quad + (A_{st} - Z_t^A) n_t^{A^U}
 \end{aligned}$$

The equations above show that the representative decisions of an untested asymptomatic player are influenced by the information sets that they get, which are improved by the divulgation mechanism, Z_t^A . From this, it follows that the decisions of this representative player will end up being a weighted average of the decisions taken when she has common information and when she does not. This helps to see divulgation as a tool that improves, in the average interaction, the information sets with which players choose their actions.

5.1.2 Player i is a tested susceptible

If player i is tested and knows she is susceptible, her optimization problem is:

Covid Economics 79, 27 May 2021: 41-92

$$\begin{aligned} & \max U_t^{S^{X,J}} \\ \text{s.t.} & (1 + \mu_t) c_t^{S^{X,J}} = w_t n_t^{S^{X,J}} + \Gamma_t + \Gamma_t^{Inf} \\ \wedge \tau_t^{I^E S^{X,J}} & = \pi_1 c_t^{S^{X,J}} c_t^{I^E} + \pi_2 n_t^{S^{X,J}} n_t^{I^E} + \pi_3 \\ \wedge \tau_t^{I^A S^{X,J}} & = \pi_1 c_t^{S^{X,J}} c_t^{I^A} + \pi_2 n_t^{S^{X,J}} n_t^{I^A} + \pi_3 \end{aligned}$$

Player i 's value function and contagion probabilities now change with the information she has on player j 's type. Hence, this player faces four interaction scenarios just as it occurred with the untested asymptomatic. See the Appendix (9) to explore the specifics of this problem.

Aggregation

The aggregate economic variables for those players who know themselves to be susceptibles is:

$$\begin{aligned} S_t^X U_t^{S^X} &= \frac{Z_t^{S^X}}{P_t} \left[(S_t^X + R_t^X + R_t^E) U_t^{S^{X,NI}} + I_t^X U_t^{S^{X,IA}} + I_t^E U_t^{S^{X,IE}} + A_t^{NX} U_t^{S^{X,U}} \right] \\ &+ (S_t^X - Z_t^{S^X}) U_t^{S^{X,U}} \\ S_t^X c_t^{S^X} &= \frac{Z_t^{S^X}}{P_t} \left[(S_t^X + R_t^X + R_t^E) c_t^{S^{X,NI}} + I_t^X c_t^{S^{X,IA}} + I_t^E c_t^{S^{X,IE}} + A_t^{NX} c_t^{S^{X,U}} \right] \\ &+ (S_t^X - Z_t^{S^X}) c_t^{S^{X,U}} \\ S_t^X n_t^{S^X} &= \frac{Z_t^{S^X}}{P_t} \left[(S_t^X + R_t^X + R_t^E) n_t^{S^{X,NI}} + I_t^X n_t^{S^{X,IA}} + I_t^E n_t^{S^{X,IE}} + A_t^{NX} n_t^{S^{X,U}} \right] \\ &+ (S_t^X - Z_t^{S^X}) n_t^{S^{X,U}} \end{aligned}$$

Note that, as seen in the aggregation of the untested asymptomatic players, here divulgation, $Z_t^{S^X}$, also acts as a tool that improves, in the average interaction of the tested symptomatic players, the information sets with which they choose their actions.

5.1.3 Final Aggregation and Market Clearing

The aggregation of the value functions for susceptibles, infected and recovered yields:

$$\begin{aligned} R_t U_t^R &= R_t^E U_t^{R^E} + (R_t^A - R_t^X) U_t^A + R_t^X U_t^{R^X} \\ I_t U_t^I &= I_t^E U_t^{I^E} + (I_t^A - I_t^X) U_t^A + I_t^X U_t^{I^X} \\ S_t U_t^S &= (S_t - S_t^X) U_t^A + S_t^X U_t^{S^X} \end{aligned}$$

Is easy to see that aggregate consumption and hours worked resemble this aggregate value functions.

Government

Government’s set-up looks a bit different in the modified version of the model, in view of the new policy tools. To finance testing and divulgation, the government levies a lump-sum tax on all players.

$$\mu_t(S_t c_t^S + I_t c_t^I + R_t c_t^R) = \Gamma_t(S_t + I_t + R_t) \tag{19}$$

$$-\Gamma_t^{Inf}(S_t + I_t + R_t) = mc_t^X X_t + mc_t^Z Z_t \tag{20}$$

Equilibrium

Adding up the budget constraints of the players that populate the economy and those of the government, we get that the aggregate budget constraint is:

$$S_t c_t^S + I_t c_t^I + R_t c_t^R = AN_t - mc_t^X X_t - mc_t^Z Z_t \tag{21}$$

The labor market clears so that:

$$S_t n_t^S + I_t^A n_t^{IA} + I_t^E \phi^I n_t^{IE} + R_t n_t^R = A_t N_t \tag{22}$$

New Population Dynamics

The total number of newly infected people at time t comes from all the interactions between players i and j that entangle a risk of contagion for either one:

$$\begin{aligned} T_t &= \int_0^{S_t} \int_0^{I_t} \pi_t(i, j) \, di \, dj \\ T_t &= (S_t - S_t^X - Z_t^S) \left(\pi_1 c_t^{AU} c_t^I I_t + \pi_2 n_t^{AU} n_t^I I_t + \pi_3 I_t \right) \\ &+ Z_t^S \left[I_t^A \left(\pi_1 c_t^{AIA} c_t^{IA} + \pi_2 n_t^{AIA} n_t^{IA} + \pi_3 \right) + I_t^E \left(\pi_1 c_t^{AIE} c_t^{IE} + \pi_2 n_t^{AIE} n_t^{IE} + \pi_3 \right) \right] \\ &+ (S_t^X - Z_t^{SX}) \left(\pi_1 c_t^{SXU} c_t^I I_t + \pi_2 n_t^{SXU} n_t^I I_t + \pi_3 I_t \right) \\ &+ Z_t^{SX} \left[I_t^A \left(\pi_1 c_t^{SXI^A} c_t^{IA} + \pi_2 n_t^{SXI^A} n_t^{IA} + \pi_3 \right) \right. \\ &\left. + I_t^E \left(\pi_1 c_t^{SXI^E} c_t^{IE} + \pi_2 n_t^{SXI^E} n_t^{IE} + \pi_3 \right) \right] \end{aligned}$$

The rest of the population dynamics are remained unchanged with respect to what was shown in Section 2.2.

5.1.4 Modified Beliefs

As in the case of Total Incomplete Information (Section 2.4), consumers form their beliefs over their own health status and the health status of people with whom they interact. However, now agents receive more information and use it to form their beliefs. The more information they receive, the closer their beliefs will be to the true probabilities.

On the one hand, individuals who have not been tested and have not experienced symptoms, form their beliefs using the aggregate information made publicly available by the policy makers. This aggregate information consists of the number

of tests performed, X_t , and their results by type: X_t^I , X_t^R y X_t^S . Under the assumption that agents in this economy know the testing technology, they use the tests results to gauge the probability of having a particular health status as follows:

$$q_t^{I^A} = \frac{X_t^I}{X_t} = \frac{I_t^X - I_{t-1}^X + \pi_r^A I_{t-1}^X}{S_t^X + I_t^X + R_t^X - I_{t-1}^X - R_{t-1}^X}$$

$$q_t^{R^A} = \frac{X_t^R}{X_t} = \frac{R_t^X - R_{t-1}^X - \pi_r^A I_{t-1}^X}{S_t^X + I_t^X + R_t^X - I_{t-1}^X - R_{t-1}^X}$$

$$q_t^S = \frac{X_t^S}{X_t} = \frac{S_t^X}{S_t^X + I_t^X + R_t^X - I_{t-1}^X - R_{t-1}^X}$$

Here we assume that in absence of testing all non-symptomatic players believe themselves to be susceptible (i.e. $q_t^S = 1$), like we did in the TII case (Section 2.4).

Additionally, by the assumption we have made throughout that aggregate information on symptomatic infected is publicly known, players form their beliefs as follows:

$$p_t^{I^E} = \frac{I_t^E}{P_t}$$

On the other hand, we assume that testing also affects the beliefs people have on the probability of meeting an asymptomatic infected person:

$$p_t^{I^A} = p_t^{I^A,PII} \frac{X_t}{A_{st} - R_{t-1}^X - I_{t-1}^X} + \left(1 - \frac{X_t}{A_{st} - R_{t-1}^X - I_{t-1}^X}\right) p_t^{I^A,TII}$$

where $p_t^{I^A,TII}$ are the beliefs under Total Incomplete Information (Section 2.4) and $p_t^{I^A,PII}$ are the beliefs under Partial Incomplete Information, just as they were shown in Section 2.3. Notice that under the assumption that $\varepsilon_t^{I^A} = 0$, which we also made in Section 2.4, this equation collapses to:

$$p_t^{I^A} = \frac{I_t^A}{P_t}$$

A more detailed discussion about this topic is provided in Section 6.

5.2 Testing

Testing is the first information instrument that one could consider to close the welfare gaps created by information incompleteness. Testing gives players private information about their own health and gives information to authorities that

is usually communicated to the public as aggregate numbers on the disease.

In order to gauge the impact of testing in closing the welfare gap, we run an exercise to find the optimal path of testing to maximize the discounted aggregate welfare of the economy. We do so assuming that no disaggregated, private information is revealed to the public. Through this exercise we go from the world of Total Incomplete Information to the world of Partial Incomplete Information in which there is perfect private information.

Figure 10: Population Dynamics - Testing

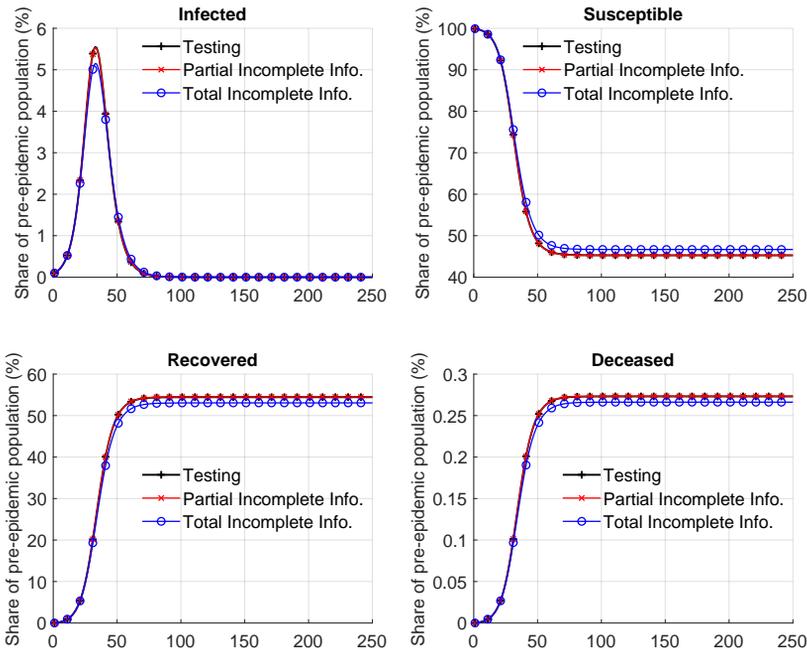
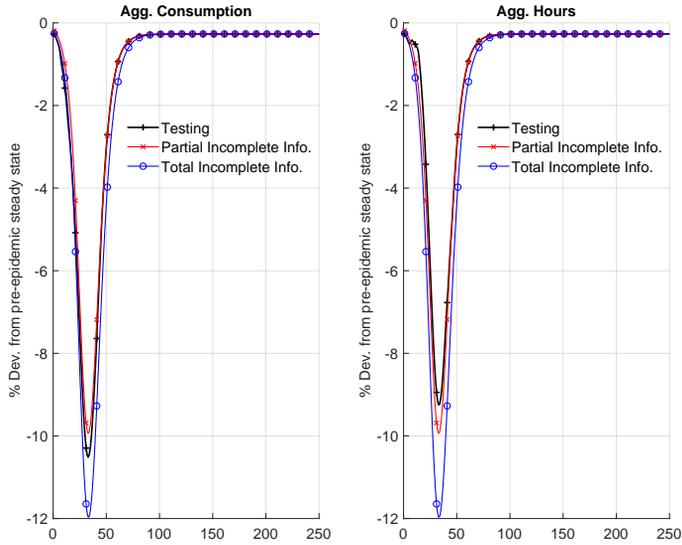


Figure 10 shows the evolution of the population by epidemiological status under the optimal testing policy. As shown, testing by itself makes the population outcomes worse. The reason is that testing creates an information asymmetry that exacerbates the negative externalities imposed by the infected and reduces the positive externalities imposed by the asymptomatic who are uncertain about their health status. Specifically, the asymptomatic infected have a dominant strategy in which they favor their economic decisions, engaging in more transactions and thereby pushing up the infections curve above that of the TII world. This affects the welfare of the susceptible population. Similarly, the recovered asymptomatic also have a dominant strategy favoring their economic activity. As Figure 11 shows that, under the baseline calibration and optimal testing, the increased economic interactions of these groups are enough to produce improved economic aggregates with respect to the TII world.

Covid Economics 79, 27 May 2021: 41-92

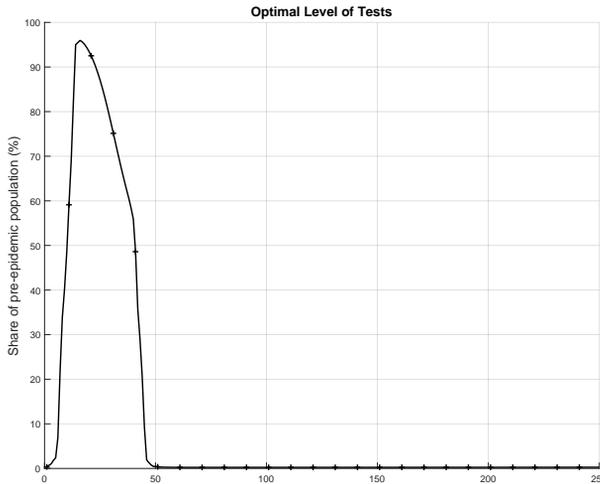
Figure 11: Economic Aggregates - Testing



Under the baseline calibration, welfare improves as the improvement in aggregate economic outcomes outweighs the negative effects of higher numbers of infections and deaths, as quantified in Table 2. Even though testing allows society to get closer (how close depends on the cost of testing) to the Partial Incomplete Information world, the gains from testing are modest vis-a-vis the losses stemming from the absence of common information. Despite this fact, it is important to recall that testing is a necessary step to implement a more detailed divulcation of health statuses.

When testing is the only policy in action, its optimal path exhibits an accelerated behavior in the first twenty weeks, since the marginal gains from obtaining the information are large as cases soar. At the peak, 96% of the initial population will be tested in a week. Once the epidemic starts to recede, the gains from testing diminish and in the optimal it falls to 0% by week 50. On average, over the 5-year horizon of analysis, 11% of the population is tested per week, a number that is similar to the one found in Eichenbaum et al. (2020b) (see Figure 12 and Table 2).

Figure 12: Optimal Tests - Testing



Covid Economics 79, 27 May 2021: 41-92

5.3 Testing and Divulcation

In the previous section we quantified the positive effects of testing on aggregate welfare in the economy, under the assumption that no personal, private health information was made public. This way of handling the information has been the norm in most countries.

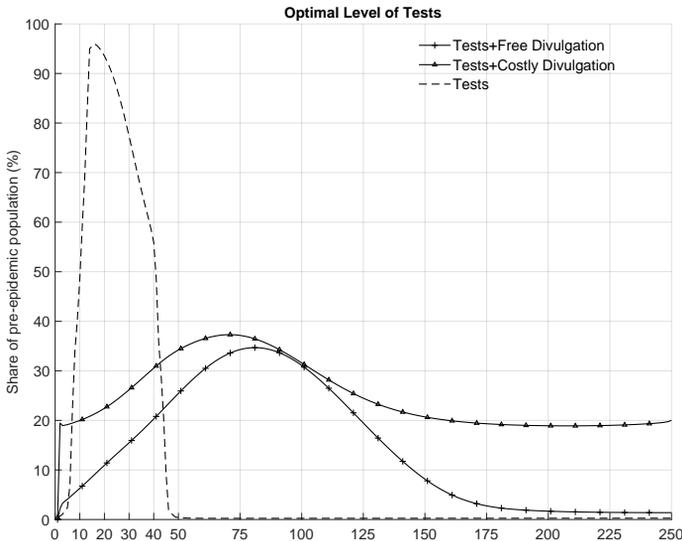
However, testing creates an information asymmetry and induces behavior that tempers its gains. Removing this asymmetry should be beneficial as the susceptible population could act optimally on it by reducing the intensity of those interactions where there is risk of contagion. As we show in Section 2.6, making all information public would imply gains of about \$6.3 trillion attributed to fewer deaths and of about \$1 trillion due to less pronounced falls in consumption. As found by Argente et al. (2020) in their study of the case of South Korea, the gains from making more detailed information publicly available are potentially large.

We now find the optimal paths for testing X_t and divulcation Z_t so as to maximize the discounted aggregate welfare of the economy. Since the information that is divulged to the public depends on the testing the optimization is performed over the two instruments simultaneously. A difficulty to produce an optimal path for divulcation is determining its marginal cost. This cost may include a variety of dimensions including some type of social cost due to loss of privacy, the cost/difficulty of processing large amounts of information due to some capacity constraint, or monetary costs. In order to illustrate the results we use two cases. In the first case divulcation has zero marginal cost and in the second one it has a cost of \$10, half the cost of testing. Together these two cases allow us to analyze the marginal effect on the health and economic dynamics of adding costs to divulcation. Such analysis gives us a more comprehensible understanding of the extent to which divulcation can relax the health-economy trade-off and improve welfare if some or all of the costs

mentioned are present. Table 2 contains the welfare calculations. Even though these calculations depend on the value of the marginal cost of divulgation, the economic benefits are so large that dwarf the cost. In other words, the cost of divulging would have to be too large in order not to do it. The margin to get positive benefits from making private health information public is given by the gap between the Complete Information and the Partial Incomplete Information worlds.

The large benefits from disclosing private information come from reducing the information asymmetries, which allows susceptible players to reduce intensity of risky interactions and dampen the effects of externalities. There are three sources of externalities: the symptomatic infected, I^E , who always know they have the virus and have no incentives to reduce their consumption, the non-tested asymptomatic infected, $I^A - I^X$, and the tested asymptomatic infected, I^X . The latter group learns about their health status once they get tested and modify their behavior imposing a negative externality on others. In the absence of testing, only the health status of those individuals who are symptomatic, either infected or recovered, could be potentially made public. Divulging private information from testing, makes testing more productive. At the beginning, as infections rise, it may be worth performing more tests as information becomes more valuable (see Figure 13). However, as Table 2 shows, the level of testing never reaches the levels of optimal testing of the no divulgation case, because the information that is released becomes very useful for players to make optimal decisions and avoid contagion. Some countries faced testing capacity constraints during the first stages of the COVID-19 pandemic. Divulgation lowers the capacity requirements and lowers the investment needed to set up such capacity.

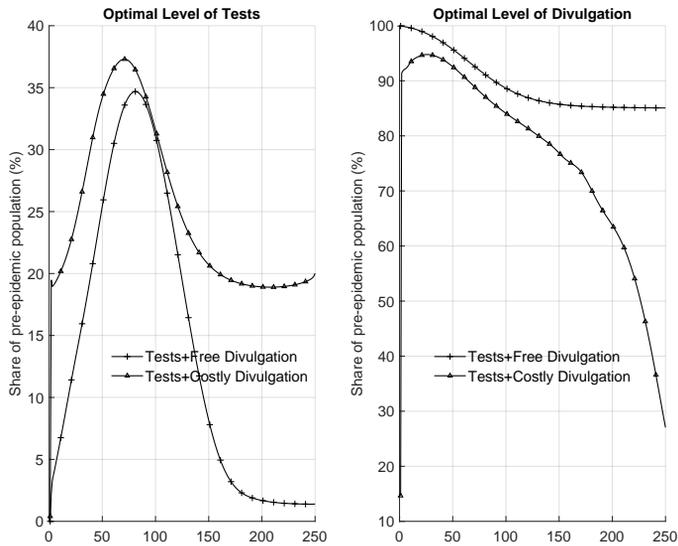
Figure 13: Optimal Tests - Testing vs Divuligation



Two more facts about testing and divulgation are illustrated in Figure 14. First, it shows that irrespective to its marginal cost, when divulgation is implemented, both testing and divulgation levels must be above zero during the entire

horizon. The reason for this is that divulging markedly flattens the infections curve and thus, herd immunity is never reached. If this is not attained, there is a latent risk of another outbreak. An alternative off-model benefit of this result is that an infection curve as flat as the one yielded by divuligation, buys authorities more time to find an effective vaccine or treatment with lower welfare losses.

Figure 14: Optimal Tests - Testing and Divuligation



Second, Figure 14 also shows that under costly divuligation there is a substitution between information policy tools. One reason for this substitution is that a different level of divuligation changes the infections curve and a higher peak of infections is reached faster. If there are more infections, testing brings larger gains and there is an incentive to increase it. Moreover, costly divuligation implies that the marginal benefit of divuliging is not necessarily greater than its marginal cost at all time. Notably, when there is not many people infected, revealing that information has low aggregate impacts, because there are not as much risky interactions but there are yet a lot of people to be informed.

Covid Economics 79, 27 May 2021: 41-92

Figure 15: Population Dynamics - Testing and Divulgence

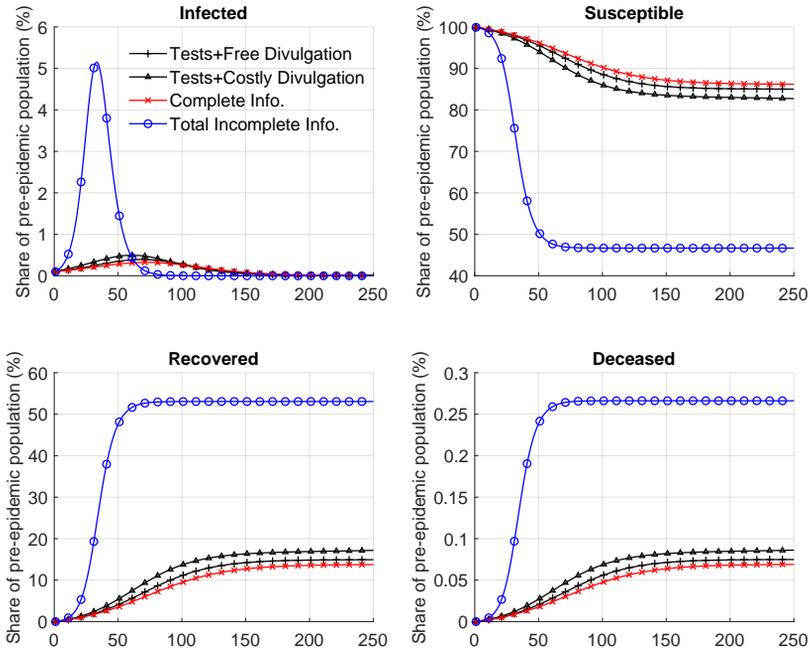
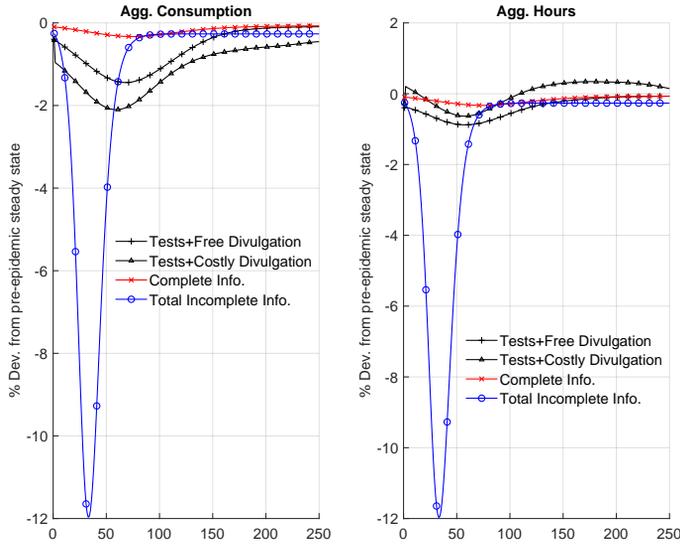


Figure 15 shows and quantifies the flattening of the infections curve and the reduction in deaths under the COVID-19 calibration for two levels of the divulgence marginal cost. At the same time, the publicly known information allows susceptible agents to carry out more (potentially all) of their contagion-risk free transactions, which results in higher consumption. Figure 16 shows substantially higher levels of economic activity and smoother dynamics. Using the results from Table 2, under the baseline calibration, the gains with respect to the case in which only testing is used are quantified to be between US\$216 billion and US \$670 billion coming from higher consumption and between US\$5.7 trillion and US \$6 trillion coming from fewer deaths, depending on the two different levels for the marginal cost of divulgence.

Covid Economics 79, 27 May 2021: 41-92

Figure 16: Economic Aggregates - Testing and Divuligation



The results aforementioned highlight the value of common information during a pandemic. It is worth mentioning that divuligation makes the pandemic last longer as the disease does not spread fast enough to reach the level of immunity as in the test-only case, but all outcomes exhibit smoother dynamics lowering the stress on testing facilities, health systems, and the economy.

5.4 The "Optimal Mix"

In this section we put together the three policy tools that we have considered and that are available to control the pandemic, and find its optimal combination. There are at least a couple of reasons to study the interaction of the three instruments.

First, the marginal cost of divuligation is hard to pin down and measure, which means that there is uncertainty as to how close divuligation can get us to the complete information world. One can end up close to the PII scenario, in which containment measures can generate significant welfare gains, as shown in Section 3.

Second, not only it is the case that, as long as there are information gaps, containment measures may have an important role to play, but there may also be uncertainty about the parameters that determine the power of different policies. For example, lower risk aversion would increase the marginal benefit of containment policies even in the presence of a high degree of divuligation⁵. This is true because even in an ideal scenario of full testing and divuligation, which is identical to the CI world, the infection externality persists. Recall that in Section 3, it is shown that even when there is complete

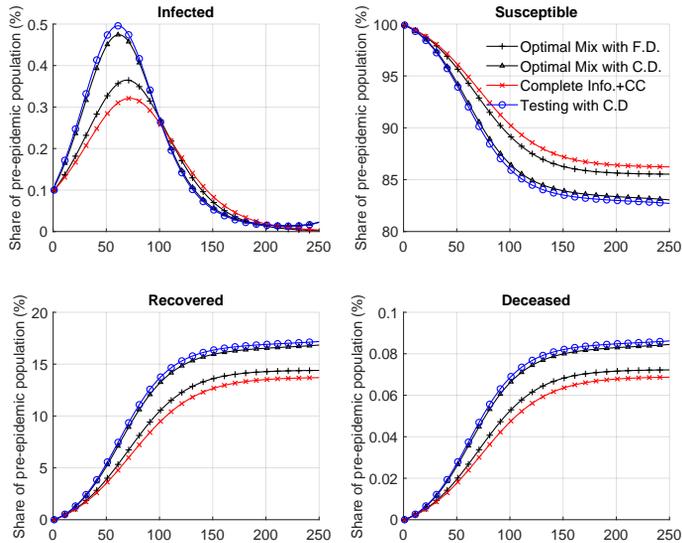
⁵The model could be potentially extended to study optimal policies in the presence of heterogeneity. For example, the existence of a mass of low risk aversion individuals may result in higher gains from combining instruments than the gains one gets under the average risk aversion. See Brotherhood et al. (2020)

information welfare can be improved through conditional containments. It is possible then to study in the modified model the optimal mix under different parameter configurations.

With this in mind, we performed two simulation exercises combining the three policy instruments optimally to maximize aggregate discounted welfare. We will find the “Optimal Mix” of policies for two different values of the marginal cost of divulgation. For this marginal cost we used the same values of the previous section. Regarding containment policies we only consider conditional containments because, as we showed in Section 3, general containment policies make sense only in the world of total incomplete information. There will be two conditional containment rates, one for symptomatic patients (μ_t^E) and one for asymptomatic infected that get identified by testing (μ_t^A). The containment scheme will still involve an imperfect compensation mechanism, since the revenue collected from those in lockdown is transferred to both the tested susceptible and the non-tested asymptomatic.

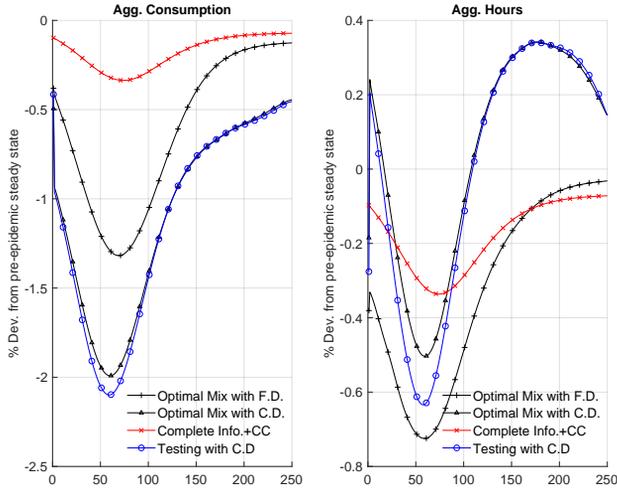
Under our baseline calibration, with costly divulgation, the gains from conditional containments are rather small, as shown in Table 2. The changes in both population outcomes (Figure 17) and aggregate economic variables (Figure 18) are marginal.

Figure 17: Population Dynamics - Optimal Mix



When the cost of divulgation falls to zero, there are important welfare gains but the outcome is still not quite the one of the CI case (see Table 2). Figure 19 illustrates the reason behind the small gains from adding containment policies to the mix. When divulgation is costly, the budget constraint of infected people becomes more stressed, which limits the room to impose more stringent lockdowns. Although optimal containment levels vary, the optimal paths for information

Figure 18: Economic Aggregates - Testing and Divulgate



instruments do not have significant changes because the tools are not substitutes.

Figure 19: Optimal Mixes

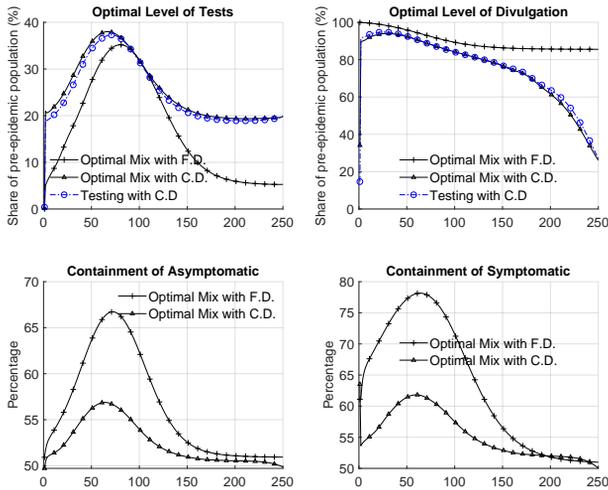


Table 2: Welfare, economic and epidemiological results for Section 4- Information Policy Tools

This table presents the simulation results for the modified model under different information and containment policies. These eight scenarios are simulated for 250 periods and analyzed through some indicators of welfare, economic activity, epidemiological dynamics and policy paths. The relative loss of aggregate welfare measures, in percentage points, the deviation of aggregate welfare under a certain specification with respect to the welfare of the Complete Information case. The maximum falls in aggregate consumption and aggregate hours are calculated relative to their pre-epidemic values and expressed in percentage points. The cumulative fall in aggregate consumption is the accumulation of all the foregone consumption during the simulation horizon, relative to a world where consumption remains all the time in its pre-epidemic value. The peak infection variable accounts for the total number of active infection cases at the height of the epidemic, as a percentage of the initial population. The final deaths and recoveries accumulate all the people that either died or recovered during the simulation horizon and express them as shares of the initial population. The containment measures show the maximum value of the consumption tax levied by the government for each type of containment. The information policies variables are expressed as a share of the initial population. The population that is informed is the people that acquire and incorporate the available private information of others health statuses. The averages are calculated for all of the 250 weeks simulated.

	Complete Information	Partial Incomplete Information	Total Incomplete Information	Testing	Testing Free Divuligation	Testing Costly Divuligation	Optimal Mix Free Divuligation	Optimal Mix Costly Divuligation
Relative loss of Aggregate Welfare	0	-0.2005	-0.2314	-0.2079	-0.0334	-0.0705	-0.0287	-0.0655
Max Fall in Aggregate Consumption %	-0.33	-9.94	-11.96	-10.51	-1.44	-2.09	-1.31	-1.99
Cumulative Fall in Aggregate Consumption %	-0.17	-1.24	-1.54	-1.34	-0.64	-1.12	-0.62	-1.08
Max Fall in Aggregate Hours %	-0.33	-9.94	-11.96	-9.24	-0.87	-0.63	-0.72	-0.50
Peak Infection %	0.32	5.53	5.15	5.54	0.39	0.49	0.36	0.47
Final Deaths %	0.06	0.27	0.26	0.27	0.07	0.08	0.07	0.08
Final Recoveries%	13.74	54.49	53.05	54.43	14.92	17.19	14.39	16.85
Peak of General Conatinent%	-	-	-	-	-	-	0	0
Peak of Symptomatic Containment %	-	-	-	-	-	-	78.15	63.51
Peak of Asymptomatic Containment %	-	-	-	-	-	-	66.72	56.88
Average % of Population Tested per Week	-	-	-	11.03	13.93	24.88	16.78	25.45
Max % of Population Tested per Week	-	-	-	95.96	34.68	42.26	35.18	38.00
Average % of Population Informed per Week	-	-	-	-	89.49	76.02	89.92	74.97
Max % of Population Informed per Week	-	-	-	-	99.93	94.76	99.93	94.00

6 Beliefs Discussion

In Sections 2.3.3, 2.4.3 and 4.1.4 we talked about players beliefs. Particularly, in Section 2.4.3 we made an explicit assumption about how players formed their beliefs in the absence of any information about asymptomatic population aggregates. The immediate consequence of this assumption is that players' beliefs are unbiased and reflect perfectly the real probabilities of facing asymptomatic infected players. Thus, the welfare loss between Partial Incomplete Information and Total Incomplete Information might be biased by this assumption. Nonetheless, our assumption is based in two factors: i) beliefs that are far from the real probabilities can induce instability in the equilibrium of this type of games⁶; ii) choosing the sign and magnitude of the bias will in any case be a difficult task, since there is no enough studies or information about it.

Moreover, this assumption was maintained in Section 4.1.4 when we introduced information policy tools, even though we introduced a new mechanism. Specifically, in this section we specified a technology of beliefs where private information attained by the government through testing and later disclosed through aggregates publication, helps to improve the precision with which players form their beliefs. However, this channel was not explored in such section, since the beliefs under Total Incomplete Information were already unbiased. Consequently, players were still able to know the real probabilities irrespective of the level of testing. This means that the welfare gains of testing could be biased.

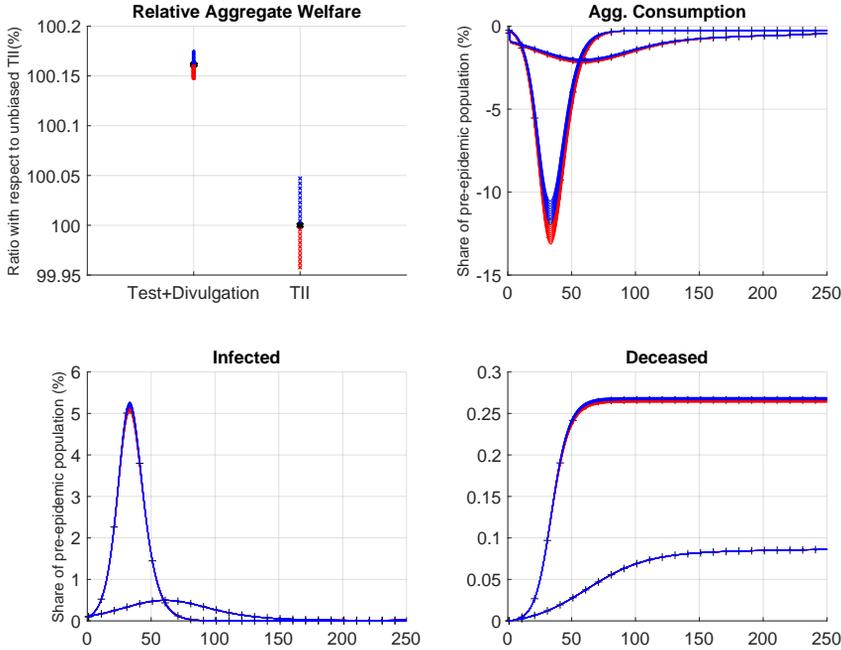
In the present section, we examine the welfare implications of deviating from this perfect beliefs formation assumption and analyze the mechanisms through which it can affect the model. For this purpose, we do two sensibility analyses, in which our benchmark scenario is the TII case. In both of them, we considered multiplicative constant biases present throughout the entire simulation horizon ranging between $[-0.5, 1.5]$ with a granularity of 0.05.

We **first** analyzed the sensibility of the results of the Testing and Costly Divulcation Scenario. For this purpose, we held constant the optimal paths of this scenario for both information policy tools across the different biases. This assumption enabled us to isolate the mechanism through which the bias distorts the model and its effects on welfare under two very different information environments.

We then did a **second** sensibility analysis in an information environment where there is no divulgation. However, we kept using the same path for testing as the one in the latter sensibility exercise. Removing divulgation in this fashion allowed us to analyze the marginal contribution of both testing and divulgation, while keeping the results as comparable as possible to the first sensibility analysis.

⁶This instability arises because the game equilibria will no longer be sequentially rational or Bayes consistent.

Figure 20: Beliefs Biases: TII vs Testing and Divulgateion

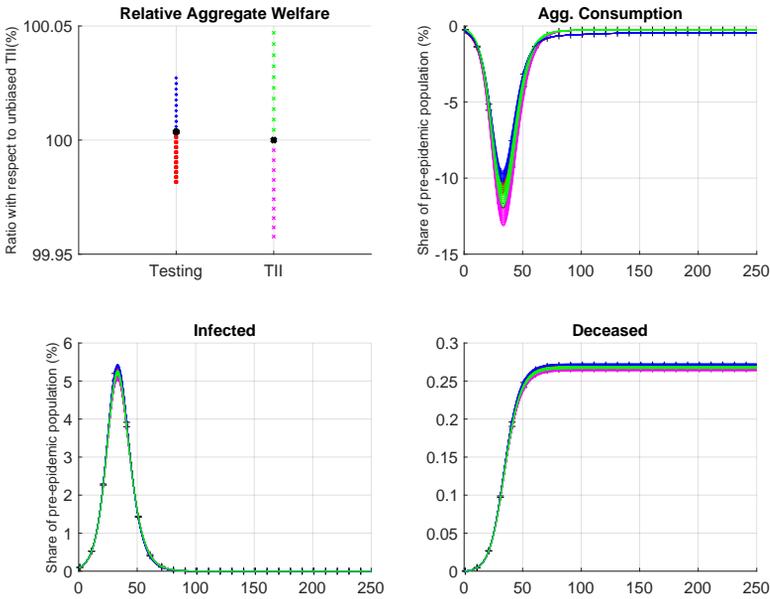


The results of this two exercises are shown in Figures 20 and 21, respectively. The main conclusions that arise in these graphs are:

1. A positive bias worsens welfare and a negative bias improves it. This holds in every information policy mix, even in total absence of testing and divulgation.
2. Regardless of the sign or magnitude of the bias, the divulgation improves welfare with respect to the TII case. This further supports the usefulness of this information policy tool.
3. When there is a positive bias, given a combination of information policy tools, the marginal welfare gain of such policy is greater than with no bias.
4. A positive bias of any magnitude does not seem to change the policy recommendation of testing to improve welfare. However, this might not be the case if the bias is negative, since if it is sufficiently large, such bias can entail welfare losses when testing. A note of caution about this conclusion is that,

although it is not possible to say that by itself testing is every time and everywhere desirable, it is possible to say that there cannot be any divulgation without testing.

Figure 21: Beliefs Biases: TII vs Testing



Covid Economics 79, 27 May 2021: 41-92

7 Conclusions

In this paper, we develop an analytical framework that combines a game theory set-up and the Macro-SIR model proposed in Eichenbaum et al. (2020a), to understand how information influences the spread of an epidemic and to quantify its importance for economic welfare. As a case study, we applied the model to analyze from its outbreak the COVID-19 epidemic in the US and show that the lack of both private and common information generates relevant welfare losses, albeit the greater losses are associated with the latter. Accordingly, we propose disclosure and divulgation as a novel policy tool to alleviate the consequences for society from publicly available disaggregated information scarcity. This policy can be combined optimally with other policies, such as testing and containments, obtaining great welfare gains. However, we are aware that privacy is important but its costs are much too great in the presence of externalities and this paper is useful for the privacy debate.

8 References

- Acemoglu, D., Chernozhukov, V., Werning, I., and Whinston, M. D. (2020). Optimal targeted lockdowns in a multi-group sir model. *NBER Working Paper*, 27102.
- Alvarez, F., Argente, D., and Lippi, F. (2020). A simple planning problem for covid-19 lockdown. *National Bureau of Economic Research*.
- Argente, D. O., Hsieh, C.-T., and Lee, M. (2020). The cost of privacy: Welfare effect of the disclosure of covid-19 cases. *National Bureau of Economic Research*.
- Atkeson, A. (2020). What will be the economic impact of covid-19 in the us? rough estimates of disease scenarios. *National Bureau of Economic Research*.
- Baker, S. R., Bloom, N., Davis, S. J., and Terry, S. J. (2020). Covid-induced economic uncertainty. *National Bureau of Economic Research*, 26983.
- Berger, D. W., Herkenhoff, K. F., and Mongey, S. (2020). An seir infectious disease model with testing and conditional quarantine. *National Bureau of Economic Research*, 26901.
- Brotherhood, L., Kircher, P., Santos, C., and Tertilt, M. (2020). An economic model of the covid-19 epidemic: The importance of testing and age-specific policies. *Institute of Labor Economics*.
- CDC (2020). Covid-19 pandemic planning scenarios. *Center for Disease Control*.
<https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html#table-1>.
- Deaton, A. (2021). Covid-19 and global income inequality. *National Bureau of Economic Research*.
- Deb, P., Furceri, D., Ostry, J., and Tawk, N. (2020). The effect of containment measures on the covid-19 pandemic. *International Monetary Fund*.
- Eichenbaum, M. S., Rebelo, S., and Trabandt, M. (2020a). The macroeconomics of epidemics. *National Bureau of Economic Research*, 26882.
- Eichenbaum, M. S., Rebelo, S., and Trabandt, M. (2020b). The macroeconomics of testing and quarantining. *National Bureau of Economic Research*, 27104.
- Farboodi, M., Jarosch, G., and Shimer, R. (2020). Internal and external effects of social distancing in a pandemic. *National Bureau of Economic Research*.
- Garriga, C., Manuelli, R., and Sanghi, S. (2020). Optimal management of an epidemic: An application to covid-19. a progress report. *National Bureau of Economic Research*.
- Jones, C., Philippon, T., and Venkateswaran, V. (2020). Optimal mitigation policies in a pandemic: Social distancing and working from home. *National Bureau of Economic Research*.

Kermack, W. O. and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115.

Rowthorn, R. and Toxvaerd, F. (2020). The optimal control of infectious diseases via prevention and treatment. *Cambridge-INET Working Paper*.

Stock, J. H. (2020). Data gaps and the policy response to the novel coronavirus. *National Bureau of Economic Research*, 26902.

9 Appendix A: Modified Model

9.1 Player i is asymptomatic

Player j type is unknown ($J = U$)

$$\begin{aligned}
 U_t^{S,A^U} &= u(c_t^{A^U}, n_t^{A^U}) + \beta \left[\left(1 - p_t^{I^E} \tau_t^{I^{E^U}} - p_t^{I^A} \tau_t^{I^{A^U}} \right) U_{t+1}^{S,A^U} \right. \\
 &\quad \left. + \left(p_t^{I^E} \tau_t^{I^{E^U}} + p_t^{I^A} \tau_t^{I^{A^U}} \right) U_{t+1}^I \right] \\
 \tau_t^{I^{E^U}} &= \pi_1 c_t^{A^U} c_t^{I^E} + \pi_2 n_t^{A^U} n_t^{I^E} + \pi_3 \\
 \tau_t^{I^{A^U}} &= \pi_1 c_t^{A^U} c_t^{I^A} + \pi_2 n_t^{A^U} n_t^{I^A} + \pi_3
 \end{aligned}$$

Then, Player i 's optimal decisions are:

$$[c_t^{A^U}] : \frac{\partial u(c_t^{A^U}, n_t^{A^U})}{\partial c_t^{A^U}} + q^S \beta \pi_1 \left(p_t^{I^E} c_t^{I^{E^*}} + p_t^{I^A} c_t^{I^{A^*}} \right) \left(U_{t+1}^I - U_{t+1}^{S,A^U} \right) = \lambda_t^{A^U} (1 + \mu_t)$$

$$[n_t^{A^U}] : \frac{\partial u(c_t^{A^U}, n_t^{A^U})}{\partial n_t^{A^U}} + q^S \beta \pi_2 \left(p_t^{I^E} n_t^{I^{E^*}} + p_t^{I^A} n_t^{I^{A^*}} \right) \left(U_{t+1}^I - U_{t+1}^{S,A^U} \right) = -\lambda_t^{A^U} w_t$$

Player j type is known and there is no contagion risk ($J = NI$)

$$U_t^{S,A^{NI}} = u(c_t^{A^{NI}}, n_t^{A^{NI}}) + \beta U_{t+1}^{S,A}$$

Then, Player i 's optimal decisions are:

$$[c_t^{A^{NI}}]: \frac{\partial u(c_t^{A^{NI*}}, n_t^{A^{NI*}})}{\partial c_t^{A^{NI*}}} = \lambda_t^{A^{NI}} (1 + \mu_t)$$

$$[n_t^{A^{NI}}]: \frac{\partial u(c_t^{A^{NI*}}, n_t^{A^{NI*}})}{\partial n_t^{A^{NI*}}} = -\lambda_t^{A^{NI}} w_t$$

Player j type is known and there is risk of contagion from I^A ($J = A$)

$$U_t^{S,A^{I^A}} = u(c_t^{A^{I^A}}, n_t^{A^{I^A}}) + \beta \left[(1 - \tau_t^{A^{I^A}}) U_{t+1}^{S,A} + \tau_t^{A^{I^A}} U_{t+1}^I \right]$$

$$\tau_t^{A^{I^A}} = \pi_1 c_t^{A^{I^A}} c_t^{I^A} + \pi_2 n_t^{A^{I^A}} n_t^{I^A} + \pi_3$$

Then, Player i 's optimal decisions are:

$$[c_t^{A^{I^A}}]: \frac{\partial u(c_t^{A^{I^A*}}, n_t^{A^{I^A*}})}{\partial c_t^{A^{I^A*}}} + \beta \pi_1 c_t^{I^A*} (U_{t+1}^I - U_{t+1}^{S,A}) = \lambda_t^{A^{I^A}} (1 + \mu_t)$$

$$[n_t^{A^{I^A}}]: \frac{\partial u(c_t^{A^{I^A*}}, n_t^{A^{I^A*}})}{\partial n_t^{A^{I^A*}}} + \beta \pi_2 n_t^{I^A*} (U_{t+1}^I - U_{t+1}^{S,A}) = -\lambda_t^{A^{I^A}} w_t$$

Player j type is known and there is contagion risk from I^E ($J = E$)

$$U_t^{S,A^{I^E}} = u(c_t^{A^{I^E}}, n_t^{A^{I^E}}) + \beta \left[(1 - \tau_t^{E^{I^E}}) U_{t+1}^{S,A} + \tau_t^{E^{I^E}} U_{t+1}^I \right]$$

$$\tau_t^{E^{I^E}} = \pi_1 c_t^{A^{I^E}} c_t^{I^E} + \pi_2 n_t^{A^{I^E}} n_t^{I^E} + \pi_3$$

Then, Player i 's optimal decisions are:

$$[c_t^{A^{IE}}]: \frac{\partial u(c_t^{A^{IE*}}, n_t^{A^{IE*}})}{\partial c_t^{A^{IE*}}} + \beta \pi_1 c_t^{IE*} (U_{t+1}^I - U_{t+1}^{S,A}) = \lambda_t^{A^{IE}} (1 + \mu_t)$$

$$[n_t^{A^{IE}}]: \frac{\partial u(c_t^{A^{IE*}}, n_t^{A^{IE*}})}{\partial n_t^{A^{IE*}}} + \beta \pi_2 n_t^{IE*} (U_{t+1}^I - U_{t+1}^{S,A}) = -\lambda_t^{A^{IE}} w_t$$

9.2 Player i is a tested susceptible

Player j type is unknown ($J = U$)

$$\begin{aligned} \max U_t^{S^{X,U}} = & u(c_t^{S^{X,U}}, n_t^{S^{X,U}}) + \beta \left[\left(1 - p_t^{IE} \tau_t^{IE^{S^{X,U}}} - p_t^{IA} \tau_t^{IA^{S^{X,U}}} \right) U_{t+1}^{S^{X,U}} \right. \\ & \left. + \left(p_t^{IE} \tau_t^{IE^{S^{X,U}}} + p_t^{IA} \tau_t^{IA^{S^{X,U}}} \right) U_{t+1}^I \right] \end{aligned}$$

$$\begin{aligned} s.a. (1 + \mu_t) c_t^{S^{X,U}} &= w_t n_t^{S^{X,U}} + \Gamma_t + \Gamma_t^{nf} \\ \wedge \tau_t^{IE^{S^{X,U}}} &= \pi_1 c_t^{S^{X,U}} c_t^{IE} + \pi_2 n_t^{S^{X,U}} n_t^{IE} + \pi_3 \\ \wedge \tau_t^{IA^{S^{X,U}}} &= \pi_1 c_t^{S^{X,U}} c_t^{IA} + \pi_2 n_t^{S^{X,U}} n_t^{IA} + \pi_3 \end{aligned}$$

Then, Player i 's optimal decisions are:

$$[c_t^{S^{X,U}}]: \frac{\partial u(c_t^{S^{X,U*}}, n_t^{S^{X,U*}})}{\partial c_t^{S^{X,U*}}} + \beta \pi_1 (p_t^{IE} c_t^{IE*} + p_t^{IA} c_t^{IA*}) (U_{t+1}^I - U_{t+1}^{S^{X,U}}) = \lambda_t^{S^{X,U}} (1 + \mu_t)$$

$$[n_t^{S^{X,U}}]: \frac{\partial u(c_t^{S^{X,U*}}, n_t^{S^{X,U*}})}{\partial n_t^{S^{X,U*}}} + \beta \pi_2 (p_t^{IE} n_t^{IE*} + p_t^{IA} n_t^{IA*}) (U_{t+1}^I - U_{t+1}^{S^{X,U}}) = -\lambda_t^{S^{X,U}} w_t$$

Player j type is known and there is no contagion risk ($J = NI$)

$$\begin{aligned} \max U_t^{S^{X,NI}} &= u(c_t^{S^{X,NI}}, n_t^{S^{X,NI}}) + \beta U_{t+1}^{S^{X,A}} \\ \text{s.a.} (1 + \mu_t)c_t^{S^{X,NI}} &= w_t n_t^{S^{X,NI}} + \Gamma_t + \Gamma_t^{Inf} \end{aligned}$$

Then, Player i 's optimal decisions are:

$$[c_t^{S^{X,NI}}]: \frac{\partial u(c_t^{S^{X,NI}*}, n_t^{S^{X,NI}*})}{\partial c_t^{S^{X,NI}*}} = \lambda_t^{S^{X,NI}} (1 + \mu_t)$$

$$[n_t^{S^{X,NI}}]: \frac{\partial u(c_t^{S^{X,NI}*}, n_t^{S^{X,NI}*})}{\partial n_t^{S^{X,NI}*}} = -\lambda_t^{S^{X,NI}} w_t$$

Player j is known and there is contagion risk from I^A ($J = A$)

$$\begin{aligned} \max U_t^{S^{X,IA}} &= u(c_t^{S^{X,IA}}, n_t^{S^{X,IA}}) + \beta \left[(1 - \tau_t^{IAS^{X,IA}}) U_{t+1}^{S^{X,A}} + \tau_t^{IAS^{X,IA}} U_{t+1}^I \right] \\ \text{s.a.} (1 + \mu_t)c_t^{S^{X,IA}} &= w_t n_t^{S^{X,IA}} + \Gamma_t + \Gamma_t^{Inf} \\ \wedge \tau_t^{IAS^{X,IA}} &= \pi_1 c_t^{S^{X,IA}} c_t^{IA} + \pi_2 n_t^{S^{X,IA}} n_t^{IA} + \pi_3 \end{aligned}$$

Optimal consumption and hours worked given by:

$$[c_t^{S^{X,IA}}]: \frac{\partial u(c_t^{S^{X,IA}*}, n_t^{S^{X,IA}*})}{\partial c_t^{S^{X,IA}*}} + \beta \pi_1 c_t^{IA*} (U_{t+1}^I - U_{t+1}^{S^{X,A}}) = \lambda_t^{S^{X,IA}} (1 + \mu_t)$$

$$[n_t^{S^{X,IA}}]: \frac{\partial u(c_t^{S^{X,IA}*}, n_t^{S^{X,IA}*})}{\partial n_t^{S^{X,IA}*}} + \beta \pi_2 n_t^{IA*} (U_{t+1}^I - U_{t+1}^{S^{X,A}}) = -\lambda_t^{S^{X,IA}} w_t$$

Player j type is know and there is contagion risk from I^E ($J = E$)

$$\begin{aligned} \max U_t^{S^{X,I^E}} &= u \left(c_t^{S^{X,I^E}}, n_t^{S^{X,I^E}} \right) + \beta \left[\left(1 - \tau_t^{I^{ES^{X,I^E}}} \right) U_{t+1}^{S^{X,A}} + \tau_t^{I^{ES^{X,I^E}}} U_{t+1}^I \right] \\ \text{s.a.} (1 + \mu_t) c_t^{S^{X,I^E}} &= w_t n_t^{S^{X,I^E}} + \Gamma_t + \Gamma_t^{Inf} \\ \wedge \tau_t^{I^{ES^{X,I^E}}} &= \pi_1 c_t^{S^{X,I^E}} c_t^{I^E} + \pi_2 n_t^{S^{X,I^E}} n_t^{I^E} + \pi_3 \end{aligned}$$

Optimal consumption and hours worked are given by:

$$[c_t^{S^{X,I^E}}] : \frac{\partial u \left(c_t^{S^{X,I^E}*}, n_t^{S^{X,I^E}*} \right)}{\partial c_t^{S^{X,I^E}*}} + \beta \pi_1 c_t^{I^E*} \left(U_{t+1}^I - U_{t+1}^{S^{X,A}} \right) = \lambda_t^{S^{X,I^E}} (1 + \mu_t)$$

$$[n_t^{S^{X,I^E}}] : \frac{\partial u \left(c_t^{S^{X,I^E}*}, n_t^{S^{X,I^E}*} \right)}{\partial n_t^{S^{X,I^E}*}} + \beta \pi_2 n_t^{I^E*} \left(U_{t+1}^I - U_{t+1}^{S^{X,A}} \right) = -\lambda_t^{S^{X,I^E}} w_t$$

Who to vaccinate first: Some important trade-offs¹

Rikard Forslid² and Mathias Herzing³

Date submitted: 21 May 2021; Date accepted: 24 May 2021

This paper models the current pandemic to analyze vaccination strategies in a setting with three age groups that differ with respect to their fatality rates. The model also accounts for heterogeneity in the transmission rates between and within these age groups. We compare the outcomes in terms of the total number of deceased, the total number of infected, the peak infection rate and the economic consequences. We find that fatalities are almost always minimized by first vaccinating the elderly, except when vaccination is slow and the general transmission rate is relatively low. In this case deaths are minimized by first vaccinating the middle-aged as this group is responsible for substantial spreading of the virus to the elderly. With regard to the other outcome variables it is always best to vaccinate the middle-aged group first. A trade-off may therefore emerge between reducing fatalities on the one hand and lowering the number of infected as well as maximizing the economic gains from vaccinations on the other hand.

1 Forslid is grateful for financial support from the Jan Wallander and Tom Hedelius Research Foundation and from the Swedish Research Council. Herzing is grateful for financial support from the Jan Wallander and Tom Hedelius Research Foundation.

2 Professor, Stockholm University and CEPR.

3 Associate Professor, Stockholm University.

Copyright: Rikard Forslid and Mathias Herzing

1 Introduction

The rapid development of vaccines against COVID-19 and the quick ramping up of production facilities has been unprecedented. However, vaccines also need to be distributed and administered to susceptible individuals. This is an important logistic challenge and also raises the question of whom to vaccinate first.¹ With the exception of front-line health care workers, most countries start with the older and most fragile part of the population. However, there is also the alternative strategy to first vaccinate a younger part of the population, which is primarily responsible for spreading the virus, in order to quickly reduce the number of infected. This reduction in the transmission of the infection could in principle result in fewer deaths in the long run also among the elderly.

This paper employs a SIR-model to examine the consequences of different vaccination strategies.² The focus is on outcomes in terms of the total number of deceased, the total number of infected, the peak infection rate, as well as economic consequences. In our model the population is divided into three groups. The first group consists of young individuals (below 20 years of age) that are very unlikely to die from the infection. The second group comprises working-age adults (20-59 years old) that have a slightly higher risk of dying. The third group consists of the elderly (60 years and older) who face a considerably higher fatality rate when being infected.

The transmission rates between and within these groups is an important factor for the spread of the infection. We will base our transmission rates on estimates from Wallinga et al. (2006), which use age-specific Dutch data on face-to-face conversations as a proxy for exposure to infectious respiratory-spread agents. Two important features stand out from this data. First, intra-group transmission rates are higher than transmission rates between age groups. Second, transmissions rates between middle-aged adults and the older group are considerably higher than between young individuals and the older group.³

We analyze six different vaccination strategies, under which the three population groups are vaccinated in sequence; since there are three groups there are six permutations. We have also considered alternative strategies, e.g. all groups being vaccinated simultaneously at rates in proportion to their share of the total population; none of these strategies generates an optimal outcomes in terms of any of the outcome measures that we focus upon.

To assess the implications of different vaccination strategies we will focus on the following outcome measures:

¹See e.g. Luyten et al. (2020) and Roope et al. (2020).

²This type of epidemiological model was introduced by Kermack and McKendrick (1927).

³An aspect not considered here is that transmission rates also differ among professions. Occupation-based infection risks are estimated by e.g. Babus et al. (2020).

(i) *The share of the population that will have deceased on day 730 after vaccinations have started.* Reducing the number of fatalities is obviously an important objective. The number of days is chosen to be equivalent to two years duration to make it possible to assess the impact also of low vaccination rates; in two years time the pandemic will have subsided even in the absence of a vaccine.

(ii) *The share of the population that will have been infected two years after vaccinations have started.* Keeping the total number of infected individuals low is also important; many surviving infected individuals have been ill for a long time and some suffer long-term consequences.

(iii) *The peak of the share of infectious individuals.* From a public health perspective it is desirable to keep the maximum number of infected persons low. In many countries the number of treated people has increased dramatically during the last weeks of 2020. It is therefore of interest to analyze which vaccination strategy dampens the peak infection rate most.

(iv) *Economic gains from vaccinations one year after their start.* We limit our analysis of economic consequences to one year, because these are almost entirely determined by the number of ill, which after one year will be negligible even in the absence of a vaccine.

Key parameters in our analysis are the efficacy of the vaccine (the share of vaccinated persons that become immune)⁴, the vaccination rate and the general transmission rate, which to varying degrees are policy parameters. The efficacy of the vaccine can be varied in a few discrete steps by the choice among existing vaccines against Covid-19. The vaccination rate can be increased by purchasing more doses and improving the implementation of a vaccine program. The general transmission rate can be influenced by the restrictions that a government imposes on the population.

We assume in our base case a vaccine efficacy of 0.9 in line with the reported levels of some of the vaccines against Covid-19 that have been developed so far. The recovery rate is assumed to be the same across the age groups and equal 0.2. However, as mentioned above, we allow for heterogeneity with respect to transmission rates within and between groups as well as age specific fatality rates. In our base case we assume a general transmission rate of 0.25, such that the infection reproduction number is given by $R_0 = 1.25$, which is roughly in line with current estimated levels for Sweden. To check the sensitivity of our results we consider the effects of changes in the efficacy rate as well as the transmission rates.

Our main results can be summarized as follows. First, the strategy of vaccinating the elderly first minimizes the fatality rates for most parameter value configurations. However, it is possible

⁴Hodgson et al. (2020) thoroughly discuss the definition of efficacy in relation to potential Covid-19 vaccines. As they point out, "many different endpoints are used in vaccine research to define efficacy depending on the pathogen, consequences of infection, and transmission dynamics." Furthermore, "outcomes might include reduction in infection (i.e., assessing sterilising immunity), severity of resultant clinical disease (i.e., assessing disease-modifying immunity), or duration of infectivity." Here we use the share of vaccinated persons that become immune for vaccine efficacy.

that vaccinating the middle-aged group first minimizes the fatality rate when the vaccination rate is low and the general transmission rate takes on values around 0.25, in particular at lower efficacy rates. Vaccinating the youngest group first always yields the highest number of deaths. Which group gets vaccinated secondly is less important, but the strategy of vaccinating the middle-aged group after the oldest group leads to the lowest fatality rates in the standard case with relatively high vaccination rates.

Second, the total number of infected persons (after two years) and the peak share of infectees are always minimized by first vaccinating the middle-aged. Since this age group is an important transmitter of the virus to all other age groups, this is the fastest way of eradicating the disease. Finally, the economic gains from vaccination are highest when the middle-aged group, which contributes most economically, gets vaccinated first. We also find that there are substantial economic gains from speeding up the vaccination campaign. For instance, we obtain a low benchmark for the gains from a campaign where the vaccination rate is doubled, such that it takes 168 rather than 325 days to vaccinate the entire susceptible population, of 0.28 billion USD (2.34 billion SEK) in the case of Sweden.

Hence, if the main policy objective is to reduce fatalities the elderly should be vaccinated first, followed by the middle-aged group. In contrast, when the aim is to minimize the total number of infected people or the peak share of infectees or to maximize economic output, the middle-aged group should be vaccinated first, while vaccinating the elderly first would yield the worst outcomes. We thus obtain a trade-off between minimizing fatalities and the other three outcome variables when it comes to the order of vaccination. Vaccinating the oldest group first implies less deaths, in particular at higher vaccination rates, but it comes at the cost of a higher share of the population becoming infected and therefore also a smaller economic gain from vaccinations. The negative consequence of not vaccinating the middle-aged group first are largest at intermediated vaccination rates.

There are a couple of highly related recent papers. Matrajt et al. (2020) analyze the optimal use of vaccine in an epidemiological model calibrated to U.S. demographics with 16 age groups that have different levels of susceptibility. Their main analysis assumes that vaccination has been carried out at the beginning of the simulations. The central result here is that deaths are minimized when vaccinating older people first, if the efficacy of a vaccine is between 10 and 50 per cent, but that it is optimal to switch to vaccinating younger persons first when the efficacy of the vaccine is above 60 per cent and there is enough vaccine to cover roughly half of the population. They also model a vaccination campaign where the entire population is vaccinated in 25, 50 or 101 weeks. Here they find that deaths are minimized by first vaccinating the elderly at low vaccination rates, but that vaccinating both old and young people is optimal at high vaccination rates. Although these results resemble ours, we find that it is optimal to start with the middle-aged group when vaccination is slow and the infection reproduction number is close to one (such that deaths are not minimized by starting with the oldest group

first). This different result is likely to be due to the fact that we are using a full matrix of estimated social interactions between age groups, whereby we account for there being relatively little transmission between the youngest and the oldest, most fragile age groups.

Moore et al. (2020) analyze the optimal sequence of Covid 19 vaccination in the UK in terms of deaths and quality adjusted life years. They divide the population into five age groups and use a social contact matrix between the groups based on UK data. They find that it is always optimal to target older age groups first. Three types of vaccine are analyzed: a vaccine that reduces susceptibility, one that reduces the probability of becoming symptomatic, and one that protects against symptoms becoming severe. The bulk of their analysis is based on the assumption that the vaccine can be instantaneously administered, but they also simulate a case where the speed of vaccine deployment is varied. In contrast to Matrajt et al. (2020) and the present study they find that the optimal ordering of age groups is unaffected by the speed of vaccine deployment.

Bartsch et al. (2020) calibrate a model of the spread of SARS-CoV-2 in the U.S. to identify the vaccine efficacy thresholds, above which vaccination could extinguish an ongoing wave of the pandemic across a range of possible scenarios. Contrary to our paper they do not analyze a vaccination campaign or consider a population divided into age groups with different mortality and transmission rates.

Vellodi and Weiss (2020) analyze optimal vaccination in a model without infection dynamics where individuals are randomly matched. Agents differ in exposure vulnerability and they may voluntarily chose to self-isolate. They find that it is optimal to first vaccinate individuals with an intermediate risk of severe illness.

Gollier (2021) analyzes the welfare consequences of different vaccination strategies using an epidemiological model, which is calibrated on French data and contains three age groups. In a setting with two countries it is shown that the observed vaccine nationalism, where rich countries prioritize to fully vaccinate their own population before exporting any vaccine, could increase the global death toll by 20 per cent.

Our paper is also related to a recent publication by Britton et al. (2020) where population heterogeneity is accounted for to assess herd immunity. The analysis focuses on four cases, where the population is either homogeneous, or is categorized by age cohorts but not by activity levels, or is categorized by different activity levels but not by age, or is categorized both by age and activity levels. Here, we introduce vaccination in a modified version of the case with age cohorts (using three age groups).

2 The Model

We employ a modified SIR-model, where there are three groups of individuals (A , B and C) with different characteristics. In each group $X \in \{A, B, C\}$ there are six categories of individuals: susceptible persons (S_X) who have never been exposed to the virus; infectious persons (I_X); recovered persons (R_X) who are no longer infectious and have developed resistance to the virus; deceased persons (D_X); vaccinated persons who are immune (V_X^{im}); and vaccinated persons who are still susceptible (V_X^s).

The dynamics in a SIR-model depend on the recovery rate and the transmission rate. We use a uniform recovery rate γ across all groups. Assuming that infectees are, on average, sick for five days, implies that $\gamma = 0.2$.

In standard pandemic models the transmission rate β is homogenous across the entire population, i.e. the rate at which a susceptible individual becomes infected by infectious individuals is βI , where I is the total number (or share) of infectious persons. Here, we instead assume that the rate of transmission varies across different segments of the population, which has been analyzed e.g. in Britton et al. (2020). We employ a simple modification of this approach. Groups A , B and C correspond to age cohorts consisting of young persons (below 20 years of age), middle-aged persons (20 to 59 years of age) and old persons (above 60 years of age), respectively. The shares of these three groups are 0.25, 0.5 and 0.25, roughly corresponding to Swedish population data.

To assess the evolution of the pandemic it is crucial to capture differences in social contact patterns within the population. Unfortunately there is a lack of detailed data on interactions between different age groups. A notable exception is the study by Wallinga et al. (2006), which uses age-specific Dutch data on face-to-face conversations as a proxy for exposure to infectious respiratory-spread agents. They obtain normalized age-specific contact rates for six cohorts (1-5, 6-12, 13-19, 20-39, 40-59 and 60-). We use the same data, but reduce the number of cohorts to three (1-19, 20-59 and 60-), to obtain the following transmission rates β_{XY} between an infected individual in group Y and a susceptible person in group X , given a general transmission rate of $\beta = 0.25$ across the entire population.⁵

⁵More specifically, we use the normalized age-specific contact rates (after correction for reciprocity) of Appendix Table 2 in Wallinga et al. (2006) to calculate the total number of reported weekly contacts for every age group (1-5, 6-12, 13-19, 20-39, 40-59, 60-). Next, we use these numbers to calculate the total number of reported weekly contacts for our cohorts (1-19, 20-59, 60-), which we then, following Diekmann et al. (1990), transform into a next-generation transmission matrix that is adjusted so that its largest eigenvalue equals $R_0 = 1.25$, implying a general transmission rate of $\beta = 0.25$ among the entire population given that $\gamma = 0.2$.

β_{XY}	1-19 (A)	20-59 (B)	60- (C)
1-19 (A)	0.5184	0.1907	0.0690
20-59 (B)	0.1907	0.3371	0.1510
60- (C)	0.0690	0.1510	0.2945

The values for the general transmission rate and the recovery rate imply that the reproduction rate is 1.25, which is roughly in line with the estimated levels for Sweden by the Public Health Agency of Sweden.⁶ This relatively low reproduction number is a result of the restrictions on public life that have been implemented.

Data on deaths due to Covid-19 clearly reveal a fatality rate that increases sharply with age. A meta-analysis by Levin et al. (2020) provides estimates of infection fatality rates (i.e. the likelihood of dying from Covid-19 among those infected by the virus) for different cohorts. On the basis of these estimates and Swedish population data for 2019, provided by Statistics Sweden, we obtain the following probabilities of dying per day of being infected: $\delta_A = 0.00000396017$, $\delta_B = 0.000268962$ and $\delta_C = 0.010573046$.

A vaccination program is introduced, such that susceptible persons are vaccinated at rate u . The vaccination rate is u_A , u_B and $u_C = u - u_A - u_B$ for groups A , B and C , respectively; while u is assumed to be constant over time (as long as there are still susceptible persons in the population), u_A , u_B and u_C change over time in accordance with the chosen vaccination strategy. For example, vaccination strategy ABC implies that $u_A = u$ and $u_B = u_C = 0$ until $S_A = 0$ (i.e. until all susceptible group A individuals have been vaccinated), whereafter $u_B = u$ and $u_A = u_C = 0$ until $S_B = 0$, followed by $u_C = u$ and $u_A = u_B = 0$ until $S_C = 0$ and the vaccination campaign ends.

Given a vaccination rate u_X of susceptibles in group X and a vaccine efficacy $e \in (0, 1]$ the number of immune vaccinated individuals increases by eu_X and the number of vaccinated individuals that remain susceptible increases by $(1 - e)u_X$ in group X per day.⁷ For each group

⁶The reproduction number for Covid 19 for Sweden has hovered between 1 and 1.5 since September 2020 according to estimates by the Public Health Agency of Sweden that publish the current reproduction number on their home page: <https://www.folkhalsomyndigheten.se/smittskydd-beredskap/utbrott/aktuella-utbrott/covid-19/statistik-och-analyser/analys-och-prognoser/>

⁷We thus assume that immunity lasts for the time interval analyzed here.

X the dynamics of the pandemic can be described as follows:

$$\begin{aligned}\dot{S}_X &= -(\beta_{XA}I_A + \beta_{XB}I_B + \beta_{XC}I_C)S_X - u_X, \\ \dot{I}_X &= (\beta_{XA}I_A + \beta_{XB}I_B + \beta_{XC}I_C)S_X + (\beta_{XA}I_A + \beta_{XB}I_B + \beta_{XC}I_C)V_X^s - \gamma I_X - \delta_X I_X, \\ \dot{R}_X &= \gamma I_X, \\ \dot{D}_X &= \delta_X I_X, \\ \dot{V}_X^{im} &= eu_X, \\ \dot{V}_X^s &= (1 - e)u_X - (\beta_{XA}I_A + \beta_{XB}I_B + \beta_{XC}I_C)V_X^s.\end{aligned}$$

For simplicity it will be assumed that $S_A, I_A, R_A, D_A, V_A^{im}, V_A^s, S_B, I_B, R_B, D_B, V_B^{im}, V_B^s, S_C, I_C, R_C, D_C, V_C^{im}$ and V_C^s represent shares of the population, i.e. $S_A(t) + I_A(t) + R_A(t) + D_A(t) + V_A^{im}(t) + V_A^s(t) + S_B(t) + I_B(t) + R_B(t) + D_B(t) + V_B^{im}(t) + V_B^s(t) + S_C(t) + I_C(t) + R_C(t) + D_C(t) + V_C^{im}(t) + V_C^s(t) = 1$ at any point in time t , where day 1 is the first day that people start being vaccinated. We assume that at day 0, the number of deceased individuals is zero, i.e. $D_A(0) = D_B(0) = D_C(0) = 0$; in terms of how vaccination strategies affect outcomes the number of those who have already died from Covid-19 is of no importance. Rather, our focus is on how many more fatalities there will be under different vaccination schemes.

As mentioned above, it is difficult to assess exactly how many have already been infected, as there have been many asymptomatic cases or cases with very light symptoms, where it was never established whether these were due to the Corona virus or not. In our calibration we assume that an equal share of 0.1 in all groups belong to the category of recovered people, implying that $R_A(0) = 0.025$, $R_B(0) = 0.05$ and $R_C(0) = 0.025$. Data on new infections per day suggest that by the end of December 2020 (when the first doses of vaccine were administered in Sweden and many other countries) about 0.3 per cent of the Swedish population was infectious. However, although testing capacity has increased considerably, there might still be many undiscovered Covid-19 cases. We therefore assume that a share of 0.005 in all groups are infected on day 0, i.e. $I_A(0) = 0.00125$, $I_B(0) = 0.0025$ and $I_C(0) = 0.00125$. Hence, the share of susceptibles is 0.895 in all groups, such that $S_A(0) = 0.22375$, $S_B(0) = 0.4475$ and $S_C(0) = 0.22375$.

While we allow for the possibility of vaccinated persons not being immune (whenever the efficacy of the vaccine is below one), we do not account for recovered persons becoming infected again. In light of reports of people having become infected more than once, this may be a strong assumption. However, the number of persons having been infected by the Covid-19 virus twice is still very low, suggesting that recovery provides immunity at least in the short run. A vaccination program is likely to lead to the pandemic having run its course in the not-too-distant future, such that the number of those reinfected will probably still be very low. Currently lack of data makes it hard to make a meaningful assessment of the reinfection rate.

To evaluate the implications of different vaccination strategies we will focus on the following measures:

- (i) The share of the population that will have deceased two years after vaccinations have started.
- (ii) The share of the population that will have been infected two years after vaccinations have started.
- (iii) The peak of the share of infectious individuals.
- (iv) Economic gains from vaccinations one year after their start, measured as the percentage gain in output in relation to one year's output in the absence of a vaccine.

Productivity has been normalized such that non-infectious (and non-deceased) group B individuals have a productivity of 1 per day in the presence of the current pandemic. While many infectees only suffer light symptoms and may still be able to work from home we make the simplifying assumption that productivity is zero for all infectious persons. For non-infected (and non-deceased) individuals it is assumed that productivity is 0 per group A individual and day, while it is 0.1 per group C individual and day. Hence, no young person contributes to output, while old people make a contribution roughly corresponding to the number of 60-64 year olds in relation to the number of group B individuals.

Normalized total output at any day t is given by

$$Y(t) = S_B(t) + R_B(t) + V_B^{im}(t) + V_B^s(t) + 0.1 [S_C(t) + R_C(t) + V_C^{im}(t) + V_C^s(t)].$$

To assess the economic consequences of the pandemic, $Y = \sum_{t=1}^{365} Y(t)$ will be measured.⁸ As a benchmark we use the outcome in the absence of a vaccine, such that we are able to calculate the economic gain during one year in relation to the vaccination rate.

3 Simulations

We simulate the outcomes for six different vaccination strategies (ABC , ACB , BAC , BCA , CAB , and CBA), according to which susceptible individuals in the three age groups are vaccinated in sequence, for vaccination rates between 0 and 0.01.⁹ The upper bound would imply that the entire population would be vaccinated in less than one hundred days, which would be hard to implement in most countries. In our simulations we assume a general transmission rate of $\beta = 0.25$ and a vaccine efficacy of $e = 0.9$. The sensitivity of our results with regard to changes in these parameter values is examined in section 4.

⁸The economic outcomes in year 2 after the start of vaccinations are hardly affected by the vaccination rate. We therefore restrict our analysis to the effects during the first year of the vaccination campaign.

⁹We have also simulated alternative strategies, such as all groups being vaccinated simultaneously at rates in proportion to their share of the total population have. None of these strategies have generated optimal outcomes in terms of any of the outcome measures that we focus upon.

3.1 Fatalities

The total number of fatalities decreases sharply in the vaccination rate, from 1294 deaths per million in the absence of a vaccination program to 215-382 fatalities per million for $u = 0.01$, as shown in Figure 1. Hence, rapidly creating vaccination capacities is crucial to keep the number of deceased as low as possible.

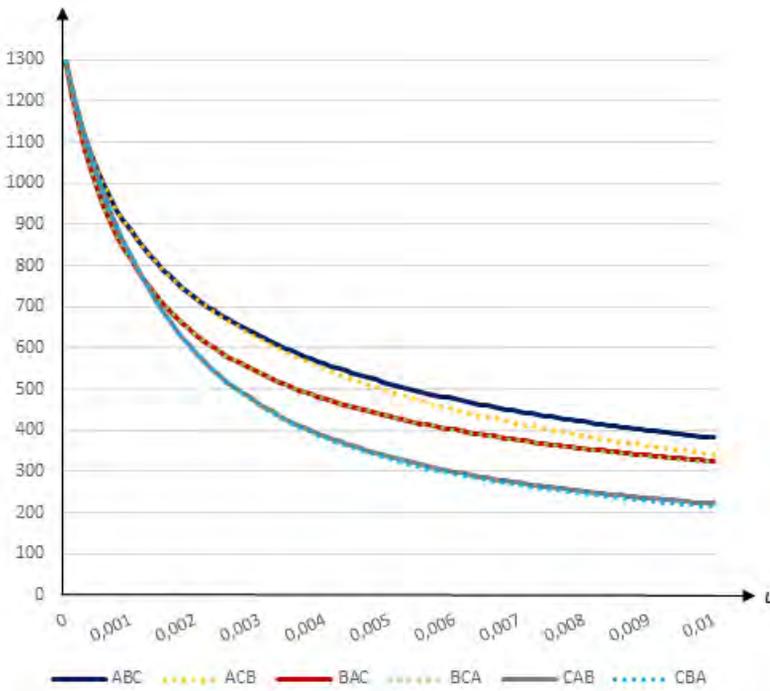


Figure 1. The number of deceased (per million) in relation to the vaccination rate when $\beta = 0.25$ and $e = 0.9$.

Our simulations also reveal significant differences between the strategies under consideration. Figure 1 shows that vaccinating group A (the young) first leads to most fatalities. However, which strategy minimizes fatalities crucially depends on the vaccination rate. To illustrate this, Figure 2 zooms in on vaccination rates below 0.002.

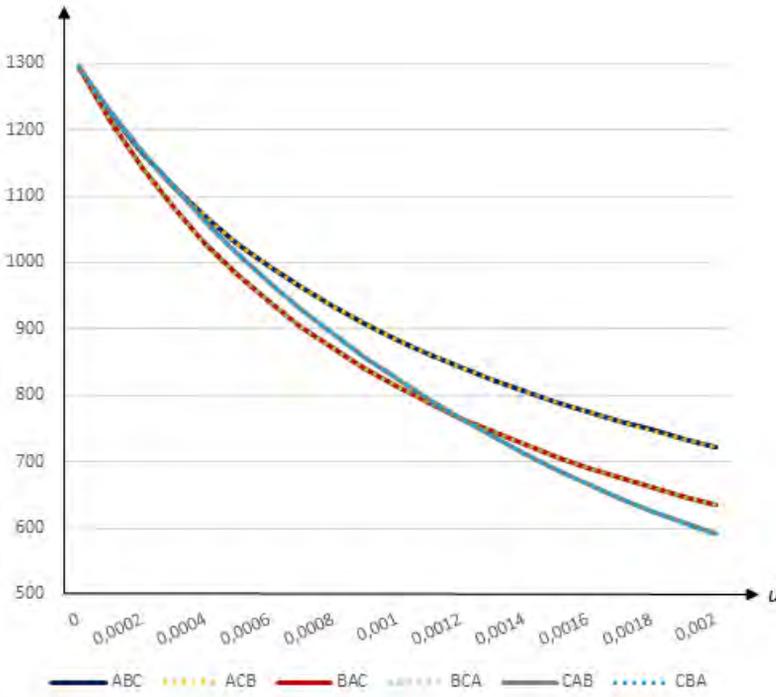


Figure 2. The number of deceased (per million) in relation to the vaccination rate when $\beta = 0.25$ and $e = 0.9$

Figure 2 shows that first vaccinating the elderly (group *C*) leads to the lowest total number of deaths for sufficiently high vaccination rates. In contrast, fatalities are minimized by first vaccinating group *B* (the middle-aged) for vaccination rates below 0.00117 (implying that it would take 699 days to cover the entire susceptible population under strategies *BAC* and *BCA*). When vaccination proceeds slowly it is more important to dampen the spread of the virus to protect the elderly. Under these conditions the total number of deaths are minimized by first vaccinating the middle aged group, which is crucial for the transmission of the virus across age groups.

Figure 2 also reveals that it is of little importance which group gets vaccinated secondly when the vaccination rate is low. However, at higher vaccination rates there are differences in outcomes depending on which group gets vaccinated secondly, as shown by Figure 3, which illustrates outcomes for vaccination rates above 0.005. In particular, there is a substantial and increasing difference in outcomes between strategies *ABC* and *ACB*, but there is also a non-negligible difference between strategies *CAB* and *CBA*. Vaccinating group *C* first and group *B*

secondly is the optimal strategy if the objective is to minimize fatalities at all vaccination rates above 0.00117.(i.e. for vaccination campaigns of less than 673 days duration under strategy *CBA*).

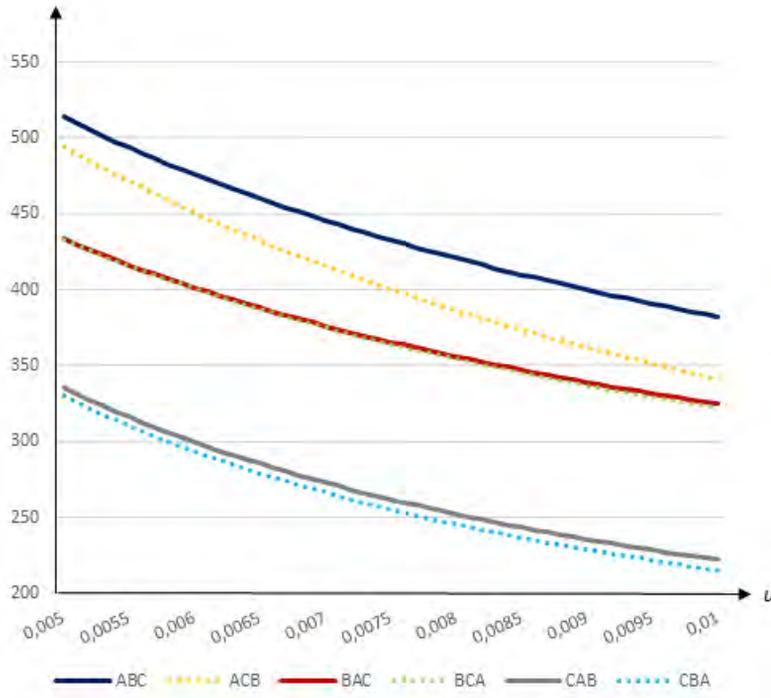


Figure 3. The number of deceased (per million) in relation to the vaccination rate when $\beta = 0.25$ and $e = 0.9$

3.2 Total number of infected

The share of the population that will have been infected by the virus decreases substantially in the vaccination rate, from 23.9 per cent for $u = 0$ to 12.9-14.2 per cent for $u = 0.01$, as shown by Figure 4.

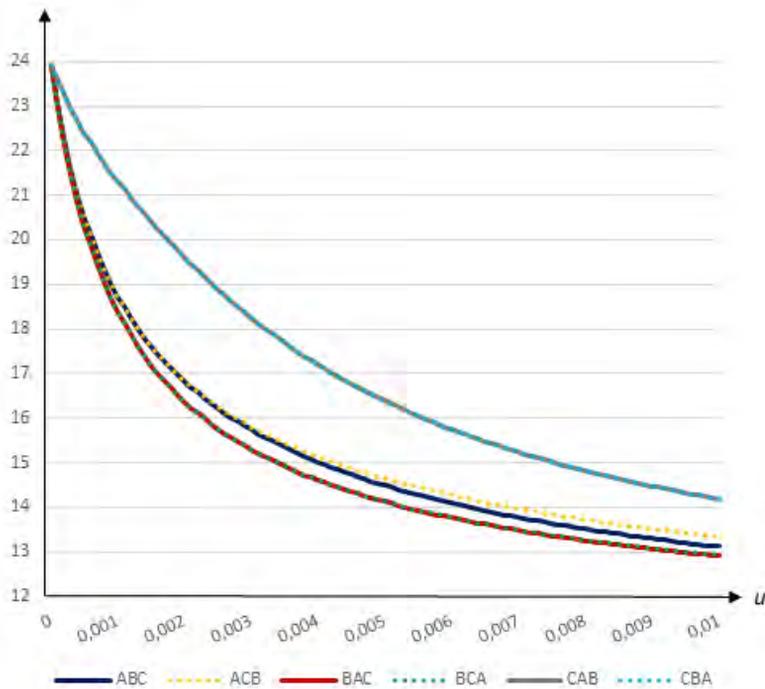


Figure 4. The total share of infected people (per cent of population) in relation to the vaccination rate when $\beta = 0.25$ and $e = 0.9$

It turns out that vaccinating group *C* first clearly yields the highest total number of infected people, while vaccinating group *B* first leads to the lowest total number of infected persons at all vaccination rates. This is due to group *B* having high rates of transmission to all groups. The difference between vaccinating group *B* first and vaccinating group *C* first is largest at intermediate vaccination rates. For example, given a vaccination rate of 0.0025 (such that it would take 325-337 days to vaccinate the entire population), 18.9 per cent of the population will have been infected under strategy *CBA*, whereas 15.8 per cent of the population will have been infected under strategy *BAC*. Figure 4 also shows that what matters for the total share of infected people is which group gets vaccinated first, while it is of little importance which group gets vaccinated secondly.

3.3 The peak share of infectees

Given that the vaccination program is implemented in the midst of the pandemic (the total share of infected people is 0.5 per cent on day 0) and the transmission rates are relatively low

due to the imposed restrictions, the peak rate of infection does not increase dramatically even in the absence of vaccinations. The day when the infection rate peaks occurs relatively early (on day 38 for $u = 0$ and in less than a week for $u = 0.01$). Nevertheless, there are differences in the peak infection rate with respect to which strategy is chosen, as shown in Figure 5.

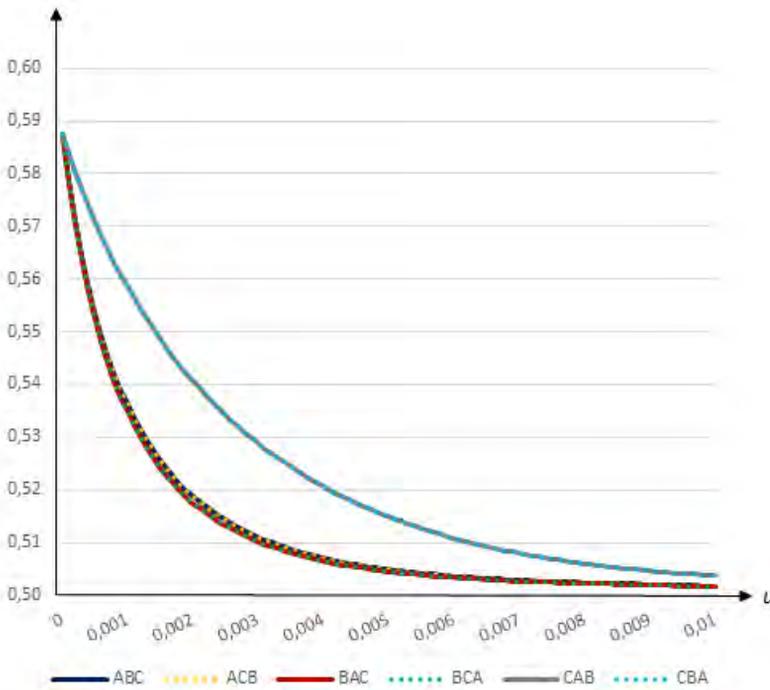


Figure 5. The peak share of infected people (per cent of population) in relation to the vaccination rate when $\beta = 0.25$ and $e = 0.9$

The peak share of infectees decreases in the vaccination rate, from 0.587 per cent for $u = 0$ to 0.502-0.504 per cent for $u = 0.01$. Vaccinating group B first yields the best outcome, but the difference to vaccinating group A first is very small. In contrast, vaccinating group C first leads to a somewhat higher peak infection rate, especially at intermediate vaccination rates, but the difference compared to when group B gets vaccinated first is never higher than 0.0245 per cent, which implies a difference of about 2500 cases in Sweden (population 10.4 million in 2020).¹⁰

¹⁰ As will be demonstrated in section 4.2, a higher transmission rate will increase this difference substantially.

3.4 Economic gains from vaccination

Economic gains increase in the vaccination rate, and they reach 0.16-0.20 per cent of GDP for $u = 0.01$, as shown in Figure 6.

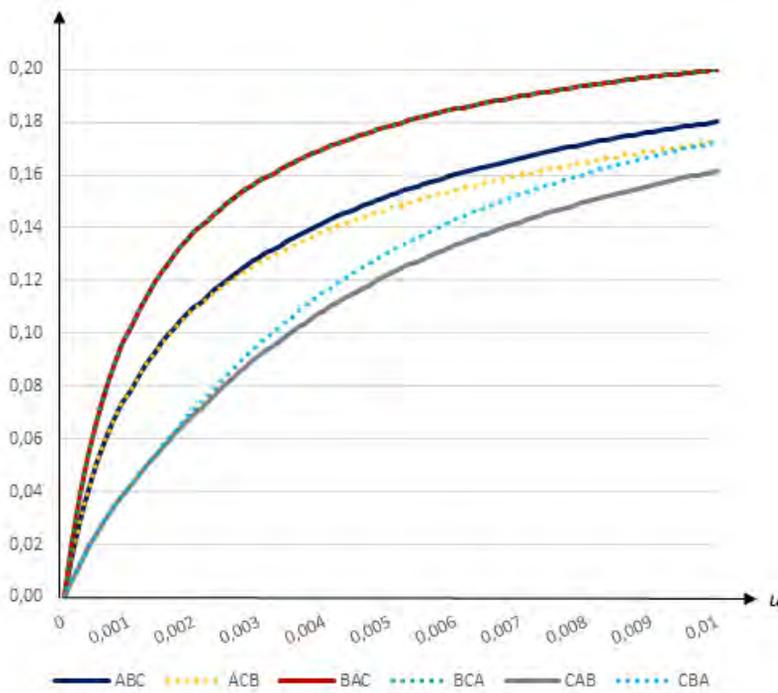


Figure 6. The economic gain from vaccinations (per cent) in relation to the vaccination rate when $\beta = 0.25$ and $e = 0.9$

Our estimates of economic gains are probably close to a lower bound since neither economic multipliers nor the cost of health care for those infected are taken into account, although we also do not consider the costs of administrating vaccinations. Nevertheless, our simulations suggest that there are substantial economic gains from increasing the vaccination rate. For instance, an acceleration of the vaccination rate from $u = 0.0025$ (lasting 325 days to cover the entire susceptible population) to $u = 0.005$ (lasting 168 days) under the fatality-minimizing strategy *CBA* would lead to a gain of roughly 0.28 billion USD (2.34 billion SEK) for Sweden.¹¹ Moreover, this increase in the vaccination rate would lead to a decrease in fatalities of 186 per million (1932 fewer deaths in Sweden) and a decrease of 2.51 per cent in the total share of

¹¹Sweden had a GDP of about 5000 billion SEK or 600 billion USD in 2020.

infected people (261 040 fewer cases in Sweden), while also slightly reducing the peak infection rate.

With respect to the economic outcomes, vaccinating group *B* first yields the highest gains and implementing strategy *CAB* leads to the smallest gains from vaccinations. Especially at intermediate vaccination rates the differences in economic outcomes between the strategies are considerable. For example, given a vaccination rate of $u = 0.0025$, the economic gain generated by switching from the fatality-minimizing strategy *CBA* to vaccinating group *B* first is about 0.39 billion USD (3.25 billion SEK) for Sweden. In addition, such a switch would imply a reduction in the total share of infected people of 3.13 per cent (about 325 000 more cases in Sweden); however, it would also lead to an increase in fatalities by 64 per million (668 cases in Sweden) as more old persons would become exposed to the virus.

3.5 Summary

The simulations where $\beta = 0.25$ and $e = 0.9$ provide clear-cut results. What matters most to outcomes is which group gets vaccinated first, while it is much less important which group gets vaccinated next.

If the main policy objective is to reduce fatalities, it is optimal to start vaccinating group *C*, followed by group *B*, for sufficiently high vaccination rates (such that it takes less than 700 days to cover the entire susceptible population). However, at lower vaccination rates it is optimal to start vaccinating the middle-aged (group *B*) first in order to minimize the total number of deaths. When the aim is to minimize the total number of infected people or the peak share of infectees or to maximize economic output, group *B* should be vaccinated first, while vaccinating group *C* would yield the worst outcomes, at all vaccination rates.

If capacities for swift vaccinations of the population are limited it is thus unambiguously optimal to start vaccinating group *B* first. In contrast, we obtain a trade-off between minimizing fatalities and the other three outcome variables at higher vaccination rates. Vaccinating group *C* implies less deaths, but it comes at the cost of a higher share of the population becoming infected and therefore also a smaller economic gain from vaccinations. The negative consequences of vaccinating group *C* rather than group *B* first are largest at intermediate vaccination rates.

4 Sensitivity analysis

In what follows we examine how sensitive the above results are to changes in the parameter values for the efficacy of the vaccine and the general transmission rate.

4.1 Vaccine efficacy

Naturally, a higher efficacy of the vaccine is associated with a lower fatality rate, a lower total number of infected people and a lower peak share of infectees, as well as larger economic gains from vaccinations for all strategies under consideration. For example, when $\beta = 0.25$ and $u = 0.0025$ (such that it would take 321-332 days to vaccinate all susceptible persons) the fatality rate is 728-800 per million people for $e = 0.5$, which decreases to 476-648 per million people for $e = 1$. The total share of infected persons is 17.4-20.2 per cent for $e = 0.5$, which is reduced to 15.5-18.6 per cent for $e = 1$. The peak share of infectees is 0.53-0.55 per cent of GDP for $e = 0.5$, which decreases to 0.51-0.53 per cent for $e = 1$. Finally, the economic gain is 0.06-0.12 per cent for $e = 0.5$, which increases to 0.09-0.15 per cent for $e = 1$. Although we obtain improvements with respect to all our four measures by increasing the efficacy, the reduction in the peak infection rate is only marginal, which is due to the fact that vaccinations are initiated at a time with relatively low transmission rates.

The effects of a lower vaccine efficacy resembles in some ways the effect of a lower vaccination rate. The threshold value for u , below which first vaccinating the middle aged (group B) minimizes fatalities, decreases in the efficacy rate; it is 0.002475 (such that it would take 336 days to vaccinate the entire susceptible population under strategies BAC and BCA) when $e = 0.5$ and 0.00105 (such that it would take more than two years to vaccinate all susceptible persons under all vaccination strategies) when $e = 1$. The effects of doubling the vaccination rate rather than the vaccine efficacy. For $e = 0.5$ and $u = 0.005$ the fatality rate is 519-620 per million people, the total share of infected persons is 15.5-17.7 per cent, the peak infection rate is 0.51-0.53 per cent and the economic gains are 0.10-0.15 per cent. Hence, we obtain similar, but slightly stronger, improvements in terms of our outcome measures when increasing the vaccination rate instead of the vaccine efficacy.

4.2 Transmission rates

If the transmission rates increase uniformly, this will obviously lead to more fatalities, a higher total share of infected persons and a higher peak number of infectees at all vaccination rates for all strategies. Not surprisingly, the economic gain from vaccinations also increases as transmission rates increase. We obtain qualitatively similar results regarding which strategies are best in terms of our outcome measures, with one important exception. For general transmission rates below 0.205 and above 0.271, it is optimal to implement strategy CBA to minimize fatalities at all vaccination rates.

The gains from increasing the vaccination rate in terms of the four outcome measures become more pronounced for higher transmission rates. In what follows we present simulation results when $\beta = 0.3$ (implying a R_0 -value of 1.5) and $e = 0.9$. Figure 7 illustrates how the fatality rate decreases from 3383 per million for $u = 0$ to 357-826 per million for $u = 0.01$. Clearly

higher rates of transmission make increasing vaccination capacities more urgent to avoid a high number of deaths. As noted above, group *C* should be vaccinated first at all vaccination rates when the virus spreads quickly to reduce the number of deaths among the elderly.

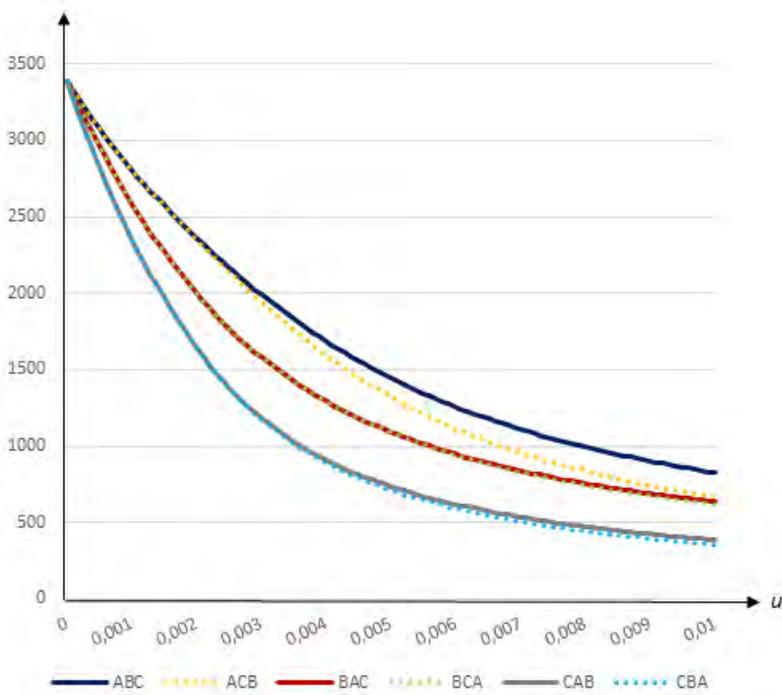


Figure 7. The number of deceased (per million) in relation to the vaccination rate when $\beta = 0.3$ and $e = 0.9$.

The total share of infected persons decreases substantially in the vaccination rate, from 44.6 per cent for $u = 0$ to 15.5-19.8 per cent for $u = 0.01$, as shown by Figure 8. Also in terms of the total number of infected people increases in the vaccination rate lead to substantially better outcomes. The difference between vaccinating group *B* rather than group *C* first becomes larger when the virus spreads more quickly, in particular at intermediate vaccination rates.

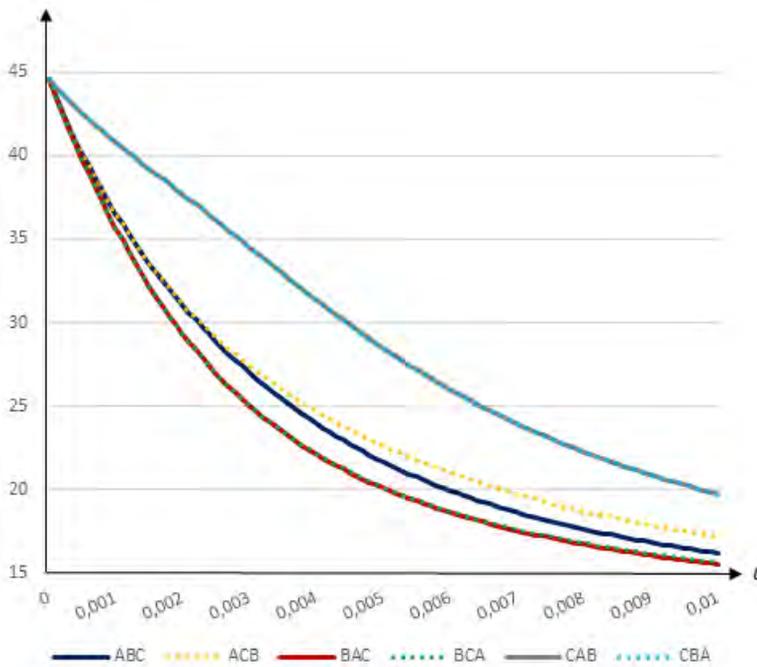


Figure 8. The total share of infected people (per cent of population) in relation to the vaccination rate when $\beta = 0.3$ and $e = 0.9$

The impact of an increasing vaccination rate on the peak infection rate becomes much stronger when the general transmission rate increases. Figure 9 shows how the peak share of infected people decreases from 2.49 per cent for $u = 0$ to 0.74-1.12 per cent for $u = 0.01$. It also clearly demonstrates that vaccinating group *B* rather than group *C* first yields substantially better outcomes, especially at intermediate vaccination rates, which is a crucial factor to account for to avoid the health care system becoming overwhelmed in case the virus spreads quickly.

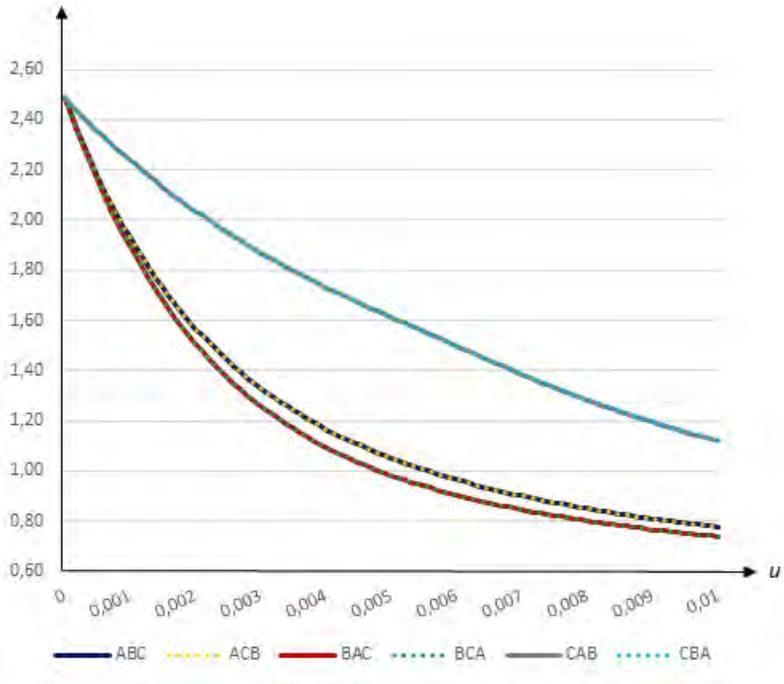


Figure 9. The peak share of infected people (per cent of population) in relation to the vaccination rate when $\beta = 0.3$ and $e = 0.9$

Economic gains increase in the vaccination rate, and they reach 0.41-0.53 % for $u = 0.01$, as shown in Figure 10. The results in section 3.4 are confirmed. Vaccinating the middle-aged first yield the best economic outcomes, particularly at intermediate vaccination rates.

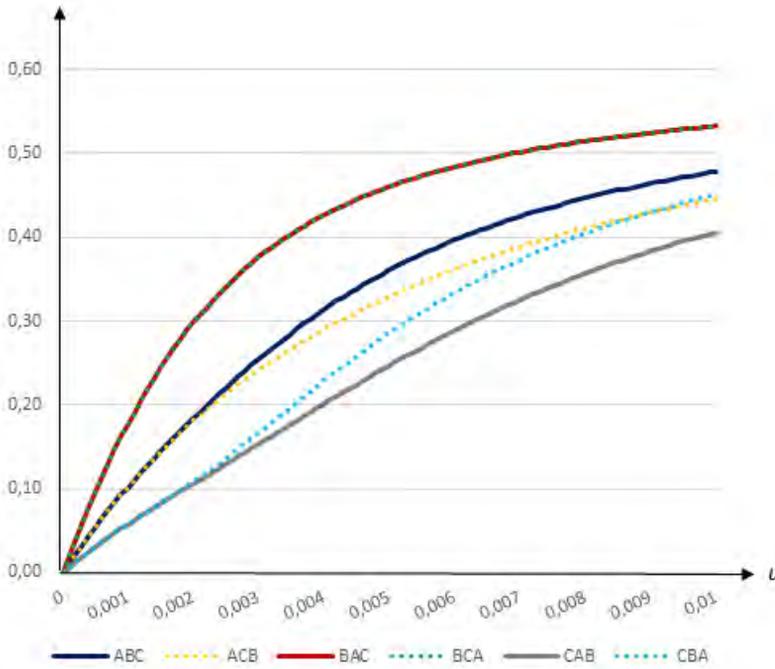


Figure 10. Economic gains of vaccination

To summarize, although the general pattern becomes more pronounced, it remains qualitatively the same for higher transmission rates. Fatalities are substantially lower when group *C* gets vaccinated first, while vaccinating group *B* first leads to considerably lower total and peak shares of infectees, as well as higher economic gains from vaccinations. Thus, the trade-off between these two alternative approaches becomes more apparent.

For example, given a vaccination rate of $u = 0.0025$ (implying that it would take 256 days to cover all susceptible persons) and an efficacy of $e = 0.9$, the fatality-minimizing strategy *CBA* leads to 1382 deaths per million, 36.1 per cent of the population becoming infected, a peak infection rate of 1.95 per cent and economic gains of 0.14 per cent, while implementing strategy *BAC* (which would last for 293 days) would lead to 1771 deaths per million, 27.0 per cent of the population becoming infected, a peak infection rate of 1.37 per cent and economic gains of 0.34 per cent. In the Swedish case, choosing strategy *CBA* rather than strategy *BAC* would imply 4042 fewer deaths, but almost one million more infected persons, a peak number

of infected people about 60 000 higher and foregone economic gains of 1.23 billion dollars (10.2 billion SEK).

The gains from accelerating the administration of vaccines are more substantial at higher transmission rates. By doubling the vaccination rate to $u = 0.005$ (such that it would take 144 days to vaccinate all susceptible persons) when implementing strategy *CBA*, fatalities would be reduced by 700 per million, the total share of infected persons would decrease by 7.7 per cent, the peak infection rate would fall by 0.34 per cent and economic gains would be 0.15 per cent higher. In the Swedish case this would imply almost 700 fewer deaths, about 800 000 fewer infected persons, a peak number of infected about 35 000 lower and economic gains of 0.9 billion USD (7.47 billion SEK).

5 Conclusions

We analyze a vaccine campaign against Covid 19 in a stylized model with three age groups that are roughly calibrated to Swedish demographic data. The age groups differ with respect to their fatality rates. Crucially, we also account for heterogeneity in contact patterns within and between age groups, such that the transmission parameters are specific to each pair of age groups.

A vaccine campaign can either prioritize the most fragile part of the population to protect them from the infection or prioritize to quickly eradicate the infection, in which case age groups with high transmission rates should be vaccinated first. We show that fatalities are almost always minimized by first vaccinating the elderly, followed by the middle-aged group. However, for some combinations of low vaccination rates and low transmission rates (e.g. due to restrictions) deaths are minimized by first vaccinating the middle-aged group; the lower is the efficacy of the vaccine, the wider is the range of vaccinations rates, for which this is true. This is due to a strong decrease in the spread of the infection, as the middle-aged have high transmission rates within their own group as well as to the other age groups. A policy implication for countries where vaccinations cannot progress at a high rate and the vaccine efficacy is not so high might therefore be to impose further restrictions in order to protect the elderly and to start vaccinating the working-age population first. Thereby deaths would be minimized, while at the same time the spreading of the disease would be countered most efficiently.

In terms of other outcome measures such as the total number and the peak number of infected persons it is always best to start vaccinating the middle-aged group first, because this groups is driving the infection through its many social contacts with the other age groups. Vaccinating the young first is never optimal. This group has a very low fatality rate, and the fact that intra-group transmission rates are high is of less importance, because the transmission rate to other groups, in particular the elderly, is relatively low.

When it comes to the economic gains from vaccinations it is always best to start vaccinating the middle-aged group first as this group has the highest productivity. We also demonstrate that there are very substantial economic gains, in addition to the health benefits, from a speedy vaccination campaign. In our model we obtain a low benchmark for the gains from doubling the vaccination rate, such that covering the entire susceptible population would take 168 rather than 325 days, of 0.28 billion USD (2.34 billion SEK) in the case of Sweden.

6 References

Babus, A., S. Das, and S. Lee, 2020, "The optimal allocation of Covid-19 vaccines", *CEPR Covid Economics* 44, 47-71.

Bartsch, S.M., K.J. O'Shea, M.C. Ferguson, M.E. Bottazzi, P.T. Wedlock, U. Strych, J.A. McKinnell, S.S. Siegmund, S.N. Cox, P.J. Hotez, and B.Y. Lee, 2020, "Vaccine efficacy needed for a COVID-19 coronavirus vaccine to prevent or stop an epidemic as the sole intervention", *American Journal of Preventive Medicine* 59(4), 493-503.

Britton, T., F. Ball, and P. Trapman, 2020, "A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2", *Science* 369, 846-849.

Diekmann, O., J.A.P. Heesterbeek, and J.A.J. Metz, 1990, "On the definition and the computation of the basic reproduction number R_0 in models for infectious diseases in heterogeneous populations", *Journal of Mathematical Biology* 28(4), 365-382.

Gollier, C., 2021, "The welfare cost of vaccine misallocation, delays and nationalism", forthcoming in the *Journal of Benefit-Cost Analysis*, republished in *CEPR Covid Economics* 74, 1-24.

Hodgson, S.H., K. Mansatta, G. Mallett, V. Harris, K.R.W. Emary, and A.J. Pollard, 2020, "What defines an efficacious COVID-19 vaccine? A review of the challenges assessing the clinical efficacy of vaccines against SARS-CoV-2", *The Lancet Infectious Diseases* doi: [https://doi.org/10.1016/S1473-3099\(20\)30773-8](https://doi.org/10.1016/S1473-3099(20)30773-8).

Kermack, W.O., and A.G. McKendrick, 1927, "A contribution to the mathematical theory of epidemics", *Proceedings of the Royal Society A. Mathematical, Physical and Engineering Sciences* 115 (772), 700-721.

Levin, A.T., W.P. Hanage, N. Owusu-Boaitey, K. B. Cochran, S.P. Walsh, and G. Meyerowitz-Katz, 2020, "Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications", *European Journal of Epidemiology* 35, 1123-1138.

Luyten, J., S. Tubeuf, and R. Kessels, 2020, "Who should get it first? Public preferences for distributing a COVID-19 vaccine", *CEPR Covid Economics* 57, 1-19.

Matrajt, L., J. Eaton, T. Leung, and E.R. Brown, 2020, "Vaccine optimization for COVID-19, who to vaccinate first?", *medRxiv preprint* doi: <https://doi.org/10.1101/2020.08.14.20175257>.

Moore, S., E.M. Hill, L. Dyson, M.J. Tildesley, and M.J. Keeling, 2020, “Modelling optimal vaccination strategy for SARS-CoV-2 in the UK”, *medRxiv preprint* doi: <https://doi.org/10.1101/2020.09.22.20194183>.

Roope, L.S.J., J. Buckell, F. Becker, P. Candio, M. Violato, J. L. Sindelar, A. Barnett, R. Duch, and P.M. Clarke, 2020, “How should a safe and effective COVID-19 vaccine be allocated? Health economists need to be ready to take the baton”, *PharmacoEconomics-Open* 4(4), 557-561.

Vellodi N. and J.Weiss, 2021, “Optimal vaccine policies: Spillovers and incentives”, *CEPR Covid Economics* 65, 1-46.

Wallinga, J., P. Teunis, and M. Kretzschmar, 2006, “Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents”, *American Journal of Epidemiology* 164(10), 936-944.

Parents under stress: Evaluating emergency childcare policies during the first COVID-19 lockdown in Germany¹

Simone Schüller² and Hannah S. Steinberg³

Date submitted: 6 May 2021; Date accepted: 13 May 2021

What are the effects of school and daycare facility closures during the COVID-19 pandemic on parental well-being and parenting behavior? Can emergency childcare policies during a pandemic mitigate increases in parental stress and negative parenting behavior? To answer these questions, this study leverages cross-state variation in emergency childcare eligibility rules during the first COVID-19 lockdown in Germany and draws on unique data from the 2019 and 2020 waves of the German AID:A family panel. Employing a DDD and IV approach we identify medium-term ITT and LATE effects and find that while emergency care policies did not considerably affect parents' life satisfaction, partnership satisfaction or mental health, they have been effective in diminishing harsh parenting behavior. We find partly gendered effects, specifically on paternal parenting behavior. Our results suggest that decreasing parental well-being likely constitutes a general effect of the pandemic, whereas the observed increase in negative and potentially harmful parenting behavior is largely directly caused by school and daycare facility closures.

1 We thank Stefan Bauernschuster, Christina Boll, Vincenzo Galasso, Mariana Grgic, Susanne Kuger, Ulrich Pötter, Claudia Zerle-Elsässer, our colleagues at DJI and participants of the 2021 CESifo Area Conference on Public Economics, participants of the DJI online conference "Aufwachsen in Deutschland 2019. Alltagswelten von Kindern, Jugendlichen und Familien" and the colloquium of the Bamberg Graduate School of Social Science for valuable discussions and suggestions. We are grateful to Kirsty Weiler for editorial advice, Stephan Heuberger for research assistance, and the Corona-KiTa-Study Team for providing us with information on utilized daycare capacity data from the DJI-RKI (2020) report. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Declarations of interest: none.

2 Researcher, German Youth Institute (DJI) and CESifo, IZA, FBK-IRVAPP

3 Researcher, German Youth Institute (DJI).

Copyright: Simone Schüller and Hannah S. Steinberg

1 Introduction

Confronted with nationwide closures of schools and daycare facilities due to the COVID-19 pandemic in spring 2020, many parents in Germany suddenly had to overhaul their work-care-arrangements. Only a very limited number of parents who between them had a specific constellation of systemically relevant occupations were granted access to emergency childcare (henceforth referred to as EC). In this paper, we exploit plausibly exogenous variation in eligibility rules across German federal states to evaluate the effect of emergency childcare policies on parental well-being and parenting behavior during the pandemic. Closed schools and childcare facilities, work from home arrangements, social distancing policies, and financial and health-related worries during the first lockdown created a stressful environment for families. These circumstances are likely to increase parenting stress, which in turn, might negatively influence parenting behavior (Abidin, 1992; Jackson and Choi 2018). Previous research indicates a positive association between negative parenting behavior and higher levels of children externalizing and internalizing problems, even if negative parenting behavior occurs infrequently (Pinquart, 2017). Moreover, harsh parenting has also been shown to be a risk factor for child abuse and neglect (Lee, Grogan-Kaylor, and Berger, 2014).

We employ a difference-in-difference-in-differences (DDD, or: triple differences) design combined with instrumental variable (IV) estimation to identify intention-to-treat (ITT) as well as local average treatment effects (LATE). We leverage cross-state variation in EC eligibility rules during the first COVID-19 lockdown in Germany, and can hence compare outcome changes of systemically relevant parents *with* EC access to outcome changes of equally systemically relevant parents in other states *without* EC access. This controls for the fact that systemically relevant parents (nurses, doctors and other key workers) are specifically affected in pandemic times, irrespective of EC. This cross-state comparison is based on the assumption that, were it not for EC policies, outcome changes for systemically relevant parents would have been similar across federal states. However, federal states were differently affected by the pandemic and hence reacted differently. To control for state-specific shocks that may have affected parents, we hence

additionally use outcome changes for parents with no systemically relevant occupation between them, resulting in a triple-difference design.

We draw on unique data from two waves of a German family panel “Growing up in Germany: Everyday Worlds” (AID:A) that surveyed families in 2019 and 4 to 5 months after the first COVID-19 lockdown in 2020. Based on a sample of 646 parents, we find that EC was not able to permanently shelter families from a considerable reduction in parental well-being. However, the provision of EC was effective in diminishing increases in harsh parenting in terms of ‘becoming angry’ among the EC-eligible parents. This effect of EC is evident several months post-lockdown, more pronounced in families with children of preschool-age or younger, and completely cancels out increases in harsh parenting among compliers. Furthermore, for fathers only, we find that EC prevented decreases in positive parenting behavior (child-centered communication), and increases in harsh parenting in terms of ‘punishing harder than merited’. Evaluating medium-term effects, we likely measure lower bounds of immediate effects during the lockdown and identify the persistent component of the overall impact.

Our results disentangle effects caused directly by school and daycare closures from general effects of the pandemic since by studying families that use childcare, we are able to compare families that experienced a complete disruption of external childcare provision with those that, thanks to their access to EC, did not. This comparison is most meaningful for children in daycare where EC was quantitatively and qualitatively more equivalent to the pre-pandemic situation than EC for school-children. We find that, while decreasing parental well-being appears to be a general pandemic effect rather than a specific effect of the closures, the observed increase in negative and potentially harmful parenting behavior is largely directly caused by school and daycare closures.

These findings contribute to the growing body of empirical literature on how the COVID-19 pandemic affects families and family well-being¹. We add to the existing literature by providing a first rigorous

¹ See e.g. Huebener *et al.* (2021) or Möhring *et al.* (2020) for empirical evidence on Germany and Prime, Wade, and Brown (2020) for a literature review on the possible consequences of the COVID-19 pandemic on well-being of families and children.

evaluation of EC policies, evaluating mid-term (by August, September 2020) rather than immediate effects, by exploiting intra-individual variation in a range of parental well-being and parenting behavior indicators. Furthermore our analysis relates to previous literature evaluating expansions in early childhood education and care (ECEC) on maternal labor market participation, parental well-being and child development (e.g. Bauernschuster and Schlotter, 2015; Schmitz, 2020; van Huizen and Plantenga, 2018). We contribute to this strand of literature by evaluating the effects of a temporary disruption in external childcare provision.

The paper is structured as follows: Section 2 describes the EC policies that were in place during the first COVID-19 lockdown in 2020 in the German federal states; Section 3 introduces the data used and describes sample selection; Section 4 presents the identification strategy; and Section 5 reports the main results, as well as complier analysis and robustness checks. The final section concludes.

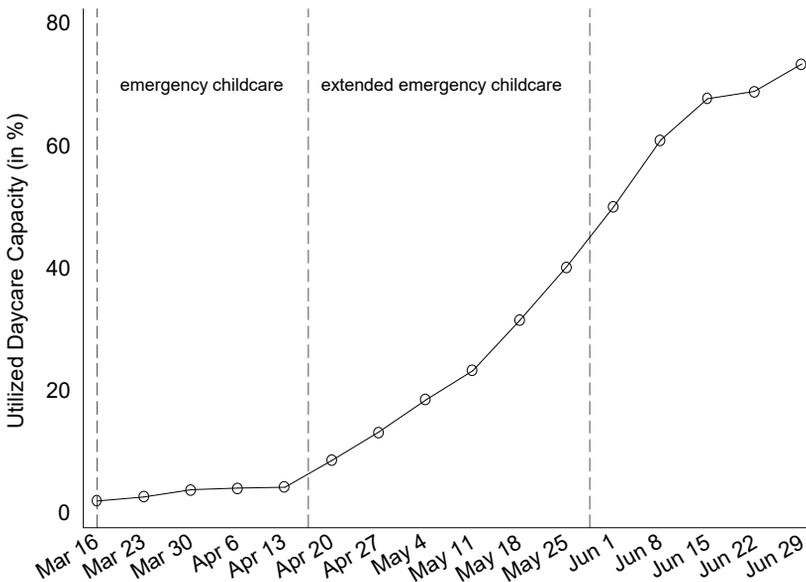
2 Emergency Childcare Policies in the German Federal States in 2020

In Germany, childcare options and their take-up depend heavily on the child's age. Attendance rates are lowest for children under three. In 2019, only 34 percent of under-threes attended a childcare facility, with significant differences between East and West Germany, but also between urban and more rural regions. In contrast, daycare usage from the age of three is almost universal: daycare coverage for children aged three to five was over 90 percent in 2019 (BMFSFJ, 2019). Finally, by age six, 64 percent attend school, and reaching almost 100 percent by age seven (Statistisches Bundesamt, 2018).

During the nation-wide lockdown between mid-March and mid-April 2020, school and daycare closures were mandated in all German federal states. Moreover, to prevent the spread of COVID-19, and especially to protect the elderly, parents were encouraged not to rely on friends, neighbors, and grandparents for childcare support. Thus, many parents in Germany suddenly had to overhaul their work-care-arrangements and provide home-schooling on their own. However, a small number of parents with a specific constellation of systemically relevant occupations were granted access to EC. In the period between mid-March and mid-April 2020, all federal states provided "emergency childcare" based on parents' occupational systemic

relevance. Subsequently —according to a mutually agreed upon framework for the stepwise opening process (JFMK, 2020)—there was a phase of gradual re-opening where emergency childcare was steadily extended (“extended emergency childcare”). By June 2020, most federal states then switched to (restricted) normal operations of daycare facilities.

Figure 1. Utilized childcare capacity in Germany during the first COVID-19 lockdown in early 2020 and subsequent re-opening



Source: DJI-RKI (2020); own calculations.

Note: Utilized daycare capacity represents the share of children that are currently in daycare among those children that were registered in daycare by March 2020. DJI-RKI (2020) reports these shares weekly by federal state, based on communications from the respective federal state ministries; we subsequently aggregate those shares to the national level. Not all federal states report utilized capacities every week (week of Mar 16: $N=9$, Mar 23-30: $N=13$, Apr 6-13: $N=14$, Apr 20-June 1: $N=15$, June 8-15: $N=12$, June 22: $N=11$, June 29: $N=7$). There is no information available on the federal state of Baden-Württemberg throughout. For six federal states, these data also include after-school childcare for school children. We define the timing of transition from emergency childcare to extended emergency childcare and from extended emergency childcare to the phase of (restricted) normal operation as the week where more than five observed federal states switch status, based on information from DJI-RKI (2020, Table 1).

Figure 1 depicts the utilized childcare capacity during the first COVID-19 lockdown and the subsequent re-opening phase based on data from the “*Corona-KiTa-Studie*” (DJI and RKI, 2020), whereby weekly utilized childcare capacities represent the share of children that were in childcare compared to the total of children registered for daycare by March 2020.² In the initial phase of the lockdown with EC—between mid-March and mid-April—on average 3 percent of the childcare capacities were utilized while subsequently, during the phase of extended EC, on average 27 percent were utilized.

In nearly all states (with exception of Hamburg and Saarland) parents’ occupational systemic relevance was a crucial factor for EC eligibility during the “emergency care” phase of facility closures.³ Systemically relevant occupations were defined as either occupations in the health and care sector (such as physicians, nursing staff or laboratory assistants) or occupations needed to maintain the infrastructure (such as in the energy or water industries, transportation, alimentation or public safety). Note that there is some variation, since German federal states applied stricter or looser definitions of systemically relevant occupations (Deutscher Bundestag, 2020; Blum and Dobrotić, 2021). Moreover, since in Germany, federal state and county-level authorities are responsible for education and social services, regulations regarding EC eligibility varied across federal states. Table 1 provides information about these differences in EC eligibility rules (see Figure A.1 in the Appendix for a graphical display). While in some states both parents had to work in a systemically relevant occupation to be eligible for EC (*2-parent rule*), in other states only one parent had to (*1-parent rule*). Additionally, in Bremen, Bavaria, Saxony and Schleswig-Holstein, a “*mixed rule*” was applied according to which, to gain access to EC parents had to be either both in a systemically relevant occupation, or at least one parent had to work in an occupation in the health and care sector. In the federal states of Hamburg and Saarland, parents were encouraged to keep their children at home, but access to EC was not further regulated. Furthermore, while in some states the child’s age was also a limitation factor in whether it could be placed in EC, in others child’s grade determined the relevant upper limit.

² For six federal states, these data also include after-school childcare for school children.

³ To be eligible for EC, employers had to confirm systemic relevance.

Table 1. Emergency Childcare Policies in the German Federal States during the first COVID-19 Lockdown

	Eligibility rules based on parents' occupations	Eligibility limit according to age or grade of children
Baden-Württemberg	2-parent rule	up to 6th grade
Bavaria	mixed rule	up to 6th grade
Berlin	1-parent rule	up to 6th grade
Brandenburg	2-parent rule	no restrictions
Bremen	mixed rule	up to 8th grade
Hamburg	all access	up to age 14
Hesse, link1 , link2	1-parent rule	up to 6th grade
Lower Saxony	1-parent rule	up to 8th grade
Mecklenburg-Vorpommern	2-parent rule	up to 6th grade
North Rhine-Westphalia	2-parent rule	up to 6th grade
Rhineland-Palatinate	1-parent rule	up to 7th grade
Saarland	all access	up to age 12
Saxony	mixed rule	up to 4th grade
Saxony-Anhalt	2-parent rule	up to age 11
Schleswig-Holstein	mixed rule	up to 6th grade
Thuringia	2-parent rule	up to 6th grade

Sources: Decrees or corresponding press releases by the respective federal state (see hyperlinks in the first column).

Note: “2-parent rule”: both parents have to work in a systemically relevant occupation to be eligible for EC. “1-parent rule”: at least one parent has to work in a systemically relevant occupation to be EC eligible. “Mixed rule”: parents have to either work both in a systemically relevant occupation or at least one parent works in an occupation in the health and care sector to gain access to EC. “All access”: parents are encouraged to keep their children at home, but access to EC is not further regulated. Lower Saxony and Schleswig-Holstein adjusted their regulations after the first week of lockdown. We employ the adjusted regulations. The eligibility rules summarized here concern the first phase of emergency childcare. Eligibility expanded in the subsequent phase of extended emergency childcare, whereby all federal states except Thuringia (where the 2-parent rule remained in place) applied the 1-parent rule.

Our analysis aims at identifying the impact of EC policies during the acute phase of school and daycare closures between mid-March and mid-April, exploiting the differences in eligibility rules across federal states as exogenous variation. Our empirical strategy abstracts from the effects of the subsequent provision of “extended emergency childcare” during the re-opening phase following the first COVID-19 lockdown, which is a key feature of our analysis. We feel it is important to focus on parents that had access to EC from the beginning of the lockdown, and hence experienced a significantly smaller disruption of daycare provision in comparison to all other parents that had children in daycare pre-pandemic (including those that utilized extended emergency childcare after two months of childcare at home). In view of the fact that in many cases emergency childcare did not provide full-time daycare, our estimates likely represent the lower bounds of the true effects.

3 Data and Sample Selection

Our analysis uses data from the 2019 wave of the AID:A family panel that surveyed about 6.000 households on living conditions of children, youth, young adults and parents. We combine this data with information from the “AID:A Corona Add-on” study, which re-interviewed about 780 households in August and September 2020 on their current living conditions, as well as their circumstances during the first COVID-19 lockdown in Germany in March and April 2020. Importantly for the purpose of our study, parents of children of pre-school age or younger were asked about their utilization of emergency childcare during the lockdown. Moreover, parents were also asked about their occupational “systemic relevance” and, crucially for our identification strategy, the type of systemic relevance (health-related or not). That is, respondents were not directly asked about their occupational systemic relevance, but whether they “work in the health or care sector” and whether they “work in a sector that is prescribed systemic importance, such as, for example, energy and water supply, transportation, alimentation or public security”. If partner information on systemic relevance is missing, we impute it via the partner’s occupation stated in 2019, based on the classification of systemic relevance employed in Koebe et al. (2020), which links up with occupations at the 3-digit KldB 2010 level.⁴ This information allows us to determine parents’ eligibility for EC according to the official rules of the respective federal states of residence (see Section 2 and Table 1), irrespective of their reported utilization of EC. Note that the information on parents’ occupational systemic relevance is crucial for our ability to identify effects of emergency childcare abstracting from effects of extended emergency childcare.

Note also, that our measure of EC utilization does not distinguish between utilization during the immediate “emergency childcare” phase or the subsequent phase of “extended emergency childcare”. Only by instrumenting EC utilization with EC eligibility, can we tease out the local average treatment effects of EC

⁴ A robustness check presented in Table A.6 in the Appendix is based on an alternative regression-based imputation employing federal state and occupation fixed effects (KldB 2010, 3-digit level).

use during the “emergency childcare” phase. We assume implicitly in our analysis that compliers in the IV analysis identify EC-eligible parents that actually utilized emergency childcare in the acute lockdown and not only extended emergency childcare in the subsequent phase of re-opening.

Estimating intention-to-treat effects on parental well-being and parenting behavior, we consider 636 parents from 482 two-parent families with at least one child below the age of 12⁵ that was either in school or in external daycare when the pandemic hit Germany (Sample A). This sample is restricted to parents for whom we have full information on the main outcomes, and families for whom we have either information on occupational systemic relevance or occupational classification for both parents. Additionally, 16 families from federal states without within-state variation of EC access (Hamburg and Saarland according to Table 1) have been excluded, since they do not contribute to our identifying variation.⁶

To estimate local average treatment effects, we consider a subset of 319 parents from 227 families with at least one child of pre-school age or younger (Sample B). We focus only on this group since unfortunately information on EC utilization was not collected for school children and is hence not available for the full sample (Sample A). Table 2 reports summary statistics for both samples.

About 54 percent of parents in our sample have no systemically relevant occupation in the parental couple and are thus not eligible for EC. Conversely, roughly 46 percent of parents⁷ have a specific constellation of systemically relevant occupations between them. However, only 23.4 (24.1) percent of parents in Sample A (B) are eligible for EC, with their specific constellation of systemic relevance in the parental couple matching the EC eligibility rule applied in their federal state of residence. The EC utilization rate, at 26.3 percent in Sample B is significantly higher than the observed eligibility rate, since the

⁵ Children below the age of 12 are defined as necessitous of childcare according to the Infection Protection Act (§56, Abs.1a). We exclude 24 single-parent households since some federal states applied particular EC eligibility rules to single parents, which we cannot examine based on low observation numbers.

⁶ Our main results also hold without imposing this restriction (see Table A5 in the Appendix).

⁷ About 32 percent of parents in Sample A work themselves in a systemically relevant occupation. 57 percent are employed in a health-related systemically relevant occupation, with “medical and health care occupations” as the dominant occupation category. 18 percent are employed in a non-health-related occupation of systemic relevance, with “occupations in business management and organization”, “occupations in teaching and training” and “occupations in education and social work, housekeeping, and theology” as the most common occupational classifications.

information on utilization does not distinguish between utilization during the period of “emergency childcare” (mid-March to mid-April 2020) and utilization during the subsequent period of “extended emergency childcare” (mid-April to end-May 2020) during which utilization rates increased significantly (see Section 2, Figure 1).

Table 2. Summary Statistics

	Sample A: Families with children below age 12		Sample B: Families with children of preschool-age or younger	
	Mean	SD	Mean	SD
<i>Outcome variables – parental well-being (1 "not at all satisfied" to 6 "very satisfied")</i>				
Life satisfaction 2019	5.074	0.775	5.107	0.778
Life satisfaction 2020	4.857	0.949	4.859	0.969
Δ life satisfaction	-0.217	0.949	-0.248	0.951
Partnership satisfaction 2019	5.072	1.004	5.103	0.974
Partnership satisfaction 2020	4.943	1.075	5.028	0.919
Δ Partnership satisfaction	-0.129	0.918	-0.075	0.908
WHO-5 2019 (index 0-100)	60.151	15.963	58.520	15.997
WHO-5 2020 (index 0-100)	59.818	18.129	57.969	17.888
Δ WHO-5	-0.333	18.022	-0.552	18.510
<i>Outcome variables – harsh parenting (1 "never" to 6 "(almost) always")</i>				
‘Become Angry’ 2019	2.525	0.941	2.586	0.977
‘Become Angry’ 2020	2.907	0.942	2.947	0.911
Δ ‘Become Angry’	0.382	0.971	0.361	0.948
‘Punish Harder’ 2019	1.788	0.782	1.796	0.764
‘Punish Harder’ 2020	1.808	0.794	1.859	0.794
Δ ‘Punish Harder’	0.020	0.816	0.063	0.787
<i>Treatment variables</i>				
Eligibility for emergency childcare	0.234	0.424	0.241	0.429
Usage of (extended) emergency childcare			0.263	0.441
<i>Systemic relevance constellation</i>				
No parent systemically relevant	0.538	0.499	0.539	0.499
One parent systemically relevant, <u>not</u> health-related	0.226	0.419	0.185	0.389
One parent systemically relevant, health-related	0.108	0.311	0.119	0.324
Both parents systemically relevant	0.127	0.334	0.157	0.364
<i>Complier characteristics</i>				
Up to 50,000 inhabitants	0.701	0.458	0.639	0.481
At least one parent holds university degree	0.411	0.492	0.436	0.497
Mother’s weekly working hours (2019)	19.236	15.047	17.896	15.918
Age youngest child in household	5.127	3.213	3.242	1.577
<i>N</i> (Nr. of parents)	636		319	

Source: AID:A 2019, AID:A Corona Add-on 2020; own calculations.

Notes: The World Health Organization Well-Being Index (WHO-5) is a five-item measure of self-reported current mental well-being (WHO 1998). The resulting index ranges from 0 “absence of well-being” to 100 “maximal well-being” (see Topp et al., 2015). ‘Become Angry’: “How frequently does the following occur? I quickly become angry when my child(ren) don’t do as I say”. ‘Punish Harder’: “How frequently does the following occur? I punish my child(ren) harder than what they merit.”

Covid Economics 79, 27 May 2021: 117-151

The information displayed in Table 2 also shows that, on average, all employed measures of parental well-being and parenting behavior worsened from 2019 to 2020. Life satisfaction decreased by on average 0.22 (0.25) points, partnership satisfaction by 0.13 (0.08) points on a 6-point scale from 1 “*not at all*” to 6 “*very satisfied*” for Sample A (B). Observed decreases in the WHO-5 Well-Being Index⁸ amount to on average 0.33 points on a 100-point scale for Sample A, and 0.55 for Sample B. Harsh parenting behavior increased in its frequency on average by 0.38 (0.36) points on a 6-point scale with respect to ‘becoming angry’, and less strongly in terms of ‘punishing harder’ with an average increase of 0.02 (0.06) points for Sample A (B). The latter measures of negative parenting behavior stem from the survey questions “*How frequently does the following occur? I quickly become angry when my child(ren) don’t do as I say*”, and “*I punish my child(ren) harder than what they merit*,” with answer categories ranging from 1 “*never*” to 6 “*(almost) always*”.

4 Empirical Strategy

The source of exogenous variation underlying our identification strategy mainly comes from the cross-state variation in EC eligibility rules during the first COVID-19 lockdown in Germany. Our identification strategy compares groups of parents with the same constellations of occupational systemic relevance, which differ in their EC eligibility due to variations in EC eligibility rules across federal states. We define EC-eligible parents as the “treatment group”, and parents that are in some constellation of systemically relevant occupation but who are not eligible for EC as the “control group”. For a more robust analysis, we add parents without any systemic relevant occupation in the parental couple as a further control group. Altogether, this leads to a difference-in-difference-in-differences design (Wooldridge, 2010, p.151). Subsequently, we employ an instrumental variable approach instrumenting EC utilization with EC eligibility to estimate local average treatment effects.

⁸ The World Health Organisation Well-Being Index (WHO-5) is a five-item measure of self-reported current mental well-being (WHO 1998). The resulting index ranges from 0 “*absence of well-being*” to 100 “*maximal well-being*” (see Topp et al., 2015).

4.1 Regression Model of Intention-to-Treat Effects

To estimate the intention-to-treat effects of EC, that is, the effects of EC eligibility on parental well-being and negative parenting behavior, we use the following model in a triple differences setup:

$$(1) \quad y_{irst} = \gamma_{st} + \alpha_{rt} + \theta_{rs} + \beta \text{Eligibility}_{rst} + \varepsilon_{irst},$$

where y_{irst} is the outcome of interest (parental well-being or parenting behavior, respectively) observed for parent i of systemic relevance (SR) constellation group r resident in state s in period t (with $t = [2019; 2020]$). The three dimensions of state (s), time (t) and SR constellation group (r) allow us to control non-parametrically for state-specific shocks (γ_{st}), interactions of SR constellation group and time effects (α_{rt}), as well as state-specific effects of SR constellation groups (θ_{rs}).

Eligibility_{rst} is our treatment variable, and results from the entanglement of the four SR constellation groups, the three state-specific eligibility rule types, and a period effect. Specifically, in federal states with a “1-parent rule”, all constellations except “no parent SR” are EC eligible; in federal states that apply a “2-parent rule” only the constellation of “both parents SR” is EC eligible; in federal states with a mixed rule, the constellations of “both parents SR”, as well as the constellation “one parent SR, health related” are eligible for EC. Our coefficient of primary interest β is a difference-in-difference-in-differences type estimator. This parameter is identified through (1) cross-sectional variation across states with different EC eligibility rules (with EC-eligible parents as the treatment group and parents in similar SR constellations who are not EC-eligible due to different state rules as a control group), (2) temporal variation in parents’ average outcome levels between the survey waves 2019 and 2020 (with the untreated year 2019 as control), and (3) temporal variation within states (with parents without a systemically relevant occupation as the “within-state” control group).

Throughout the analysis, all standard errors are clustered at the household level and are robust to heteroscedasticity. We show results from individual fixed effects regressions that additionally control for

unobserved time-invariant factors, and improve precision with respect to pooled OLS.⁹ Note that all within-individual time-invariant factors (including θ_{rs}) are controlled for by the individual fixed effects.

For completeness, we additionally report the difference-in-differences (or: double difference) results from the subsample of parents with some kind of SR constellation. That is, for the double-difference setup, we exclude the data for parents without any occupational SR in the parental couple, and we estimate equation (2), in which states are again subscripted with s , SR constellation groups with r , and time period with t :

$$(2) \quad y_{irst} = \gamma_s + \alpha_t + \theta_r + \beta \text{Eligibility}_{rst} + \varepsilon_{irst}.$$

The double-difference estimator assumes that, were it not for differences in EC eligibility rules, outcome changes for parents of the same SR constellation group would have been similar across federal states. However, there might well be state-specific period effects in the context of the COVID-19 pandemic and associated state-level policy measures that may have affected parental stress irrespective of EC policies. A neat way to account for state-specific shocks is in fact to use other groups that are not directly affected by EC policies in either state (such as parents without any SR occupation in the parental couple) as an additional control group in a triple-difference setup, as outlined above. Outcome changes in this group, which is unaffected by the policy of interest, are then presumed to reflect region-specific period effects.

The causal interpretation of the intention-to-treat effects in this (and any DD or DDD) setting hinges on the common trend assumption. However, we cannot investigate pre-trends since the AID:A family panel only started in 2019. Nevertheless, it seems plausible to assume that EC eligibility, i.e., a combination of a certain occupational SR combination in the parental couple and the federal state of residence, was of no importance pre-pandemic. To assess the legitimacy of this assumption, i.e., the exogeneity of EC eligibility, we regress EC eligibility on a variety of family sociodemographic characteristics. As expected, none of them appears to be statistically significantly associated with EC eligibility status (see Table A.1 in the Appendix).

⁹ Individual fixed effects help reduce the variance of ε_{irst} and hence the standard errors of the estimate of β .

4.2 Regression Model of Local Average Treatment Effects

To directly examine the causal effects of EC utilization on parental well-being and negative parenting behavior, we exploit the fact that access to EC was only possible when parents met the state-specific eligibility rules. We employ a first-differences model in the first stage, which—in two period panel models—is numerically equivalent to an individual fixed effects model. With respect to the regression model employed to estimate intention-to-treat effects, we exclude the second-level interactions (γ_{st} , α_{rt} and θ_{rs}) to yield a powerful first stage (Pischke, 2007, p.16). Overall, causal identification in this setup stems from the instrumental variable rather than the triple-difference approach. The analysis can be represented by the following system of equations:

$$(3) \quad \Delta y_{irst} = \vartheta + \sigma Utilization_{irst} + \Delta \varepsilon_{irst},$$

with the first stage given by:

$$(4) \quad Utilization_{irst} = \omega + \delta Eligibility_{irst} + \Delta \epsilon_{irst},$$

where $Utilization_{irst}$ is a dummy variable equal to 1 if parent i in federal state s reports having utilized EC during the first COVID-19 lockdown in Germany in spring 2020 (and 0 otherwise). The constant terms ϑ and ω represent first differences of the time effect. $Eligibility_{irst}$ serves as an instrument for parents' actual usage of EC. While exogeneity of the instrument (EC eligibility) is sufficient for a causal interpretation of the intention-to-treat effects from equations (1) and (2), IV estimation of equations (3) and (4) require the additional assumption that EC eligibility affects parental well-being and parenting behavior only through the actual utilization of EC, and not directly in any other way. In the context of the first COVID-19 lockdown, this assumption appears rather plausible: while systemic relevance *per se* might have been associated with factors that also influenced parental well-being and stress-levels (such as work in the health sector, potential exemptions from curfews and work-from-home-orders, or augmented infection risk exposure), the differences in EC eligibility regulations concerning the SR constellation in parental couples allow us to explicitly control for such differences.

5 Results and Discussion

5.1 Intention-to-Treat Effects

We use model (1) to estimate the intention-to-treat effects of the provision of EC on parental well-being and parenting behavior outcomes. These estimates inform about EC provision in its acute phase, when the occupation-based eligibility rules described in Section 2 (Table 1) were in place, abstracting from the effects of subsequently extended emergency childcare.

Table 3 presents the main results with respect to parental well-being measures and indicators of negative parenting behavior. For each outcome, we present the estimated coefficient β —based on individual fixed effects regressions—as the double-difference estimator in Panel A, and the triple-difference estimator in Panel B of Table 3. Panel C reports a gender interaction of the triple-difference estimator to investigate treatment heterogeneity for mothers and fathers.

The double-difference results in Panel A indicate that among parents with at least one SR occupation among the parental couple, EC-eligible and non-eligible parents experienced similar decreases in parental well-being between 2019 and 2020. With respect to negative parenting behavior, it appears that while non-EC-eligible parents report strong increases in “harsh parenting” in terms of “quickly becoming angry if children don’t do as I say”, the EC-eligible are significantly less prone to such increases. These effects are marginally statistically significant. However, they cannot isolate the causal effect of EC eligibility, as there may have been other state-specific shocks to parental stress levels (e.g., due to other COVID-19 measures at the state-level or regional infection dynamics). We hence augment the double-difference model to examine this possibility by taking advantage of the fact that parents without any occupational systemic relevance were not granted access to EC in the acute lockdown period. This allows us to use outcome changes for this group to control for unobserved state-specific shocks via a triple-difference technique.

The triple-difference results presented in Panel B of Table 3 confirm that overall, access to EC did not considerably affect parental life satisfaction or partnership satisfaction. However, we estimate a positive effect on well-being according to the WHO-5 index of about 7.3 points on a 100-point scale, which is

statistically significant at the 5-percent level. To classify the effect size, we compare it to the gender difference between mothers and fathers, which amounts to 3.0 points in the year 2019. Hence, our estimated effect is 2.4 times as large as the average gender difference in the WHO-5 index.

Table 3. Intention-To-Treat Effects on Parental Well-Being and Negative Parenting Behavior

	Parental Well-Being			Negative Parenting Behavior	
	Life Satisfaction (1)	Partnership Satisfaction (2)	WHO-5 (3)	'Become Angry' (4)	'Punish Harder' (5)
<i>Panel A: Double Difference</i>					
Year = 2020	-0.145*	-0.166**	-1.545	0.441***	0.090
	(0.084)	(0.079)	(1.697)	(0.080)	(0.076)
Eligibility × Year = 2020	-0.144	0.018	1.786	-0.180*	-0.117
	(0.116)	(0.114)	(2.226)	(0.107)	(0.097)
Constant	5.102***	5.133***	60.803***	2.503***	1.782***
	(0.029)	(0.029)	(0.555)	(0.027)	(0.024)
Individual FE	yes	yes	yes	yes	yes
N	588	588	588	588	588
Nr. of individuals	294	294	294	294	294
Nr. of households	217	217	217	217	217
<i>Panel B: Triple Difference</i>					
Eligibility × Year = 2020	0.141	-0.021	7.317**	-0.407**	-0.111
	(0.188)	(0.159)	(3.463)	(0.175)	(0.161)
SR constellation group × Year = 2020	yes	yes	yes	yes	yes
State × Year = 2020	yes	yes	yes	yes	yes
Individual FE	yes	yes	yes	yes	yes
N	1,272	1,272	1,272	1,272	1,272
Nr. of individuals	636	636	636	636	636
Nr. of households	482	482	482	482	482
<i>Panel C: Triple Difference with Gender Interaction</i>					
Eligibility × Year = 2020	0.142	0.009	6.936*	-0.360*	0.002
	(0.196)	(0.165)	(3.584)	(0.187)	(0.156)
Father × Eligibility × Year = 2020	-0.004	-0.083	1.070	-0.130	-0.316**
	(0.159)	(0.131)	(2.751)	(0.137)	(0.127)
SR constellation group × Year = 2020	yes	yes	yes	yes	yes
State × Year = 2020	yes	yes	yes	yes	yes
Individual FE	yes	yes	yes	yes	yes
N	1,272	1,272	1,272	1,272	1,272
Nr. of individuals	636	636	636	636	636
Nr. of households	482	482	482	482	482

Source: AID:A 2019, AID:A Corona Add-on 2020; own calculations.

Notes: Life satisfaction, as well as partnership satisfaction, are measured on a 6-point scale ranging from 1 "not at all satisfied" to 6 "very satisfied". The WHO-5 Well-being Index ranges from 0 "absence of well-being" to 100 "maximal well-being". 'Become Angry': "How frequently does the following occur? I quickly become angry when my child(ren) don't do as I say" (1 "never" to 6 "(almost) always"). 'Punish Harder': "How frequently does the following occur? I punish my child(ren) harder than what they merit." (1 "never" to 6 "(almost) always"). Systemic relevance (SR) constellation groups: (a) no parent SR, (b) one parent SR, not health-related, (c) one parent SR, health related, (d) both parents SR. Cluster-robust standard errors at household level. *** p < 0.01, ** p < 0.05, * p < 0.1.

Covid Economics 79, 27 May 2021: 117-151

With respect to parenting behavior, it appears that EC was effective at preventing increases in “harsh parenting” in terms of ‘becoming angry’. The estimated effects are statistically significant at the 5-percent level, and indicate that being EC-eligible decreases the frequency of ‘becoming angry’ by 0.407 points on a 6-point scale. To classify the effect size, we again compare it to the average gender difference between mothers and fathers (0.115 points in the year 2019). Our estimated effect is 3.5 times as large as the average gender difference for ‘becoming angry’.

To compare effect sizes between WHO-5 well-being, which is an index from 0 to 100, and harsh parenting in terms of ‘becoming angry’, which is measured on a 6-point scale, we compute effects on standardized outcomes (see Table A.2 in the Appendix). It turns out that both effects are in fact rather similar in size: 0.429 and 0.423 of one standard deviation for WHO-5 well-being and ‘becoming angry’, respectively. Note again that these are mid-term effects, measured as of September/August 2020. Immediate effects during the acute lockdown in March/April 2020 are likely to have even been stronger.

Gender interactions presented in Panel C of Table 3 reveal statistically significant gender differences only for negative parenting behavior with respect to ‘punishing harder’. Here, EC appears to have affected fathers only, in that it decreased the frequency of ‘punishing children harder than merited’ by about 0.316 points on a 6-point scale (or by 0.401 of one standard deviation, according to Table A.2 in the Appendix). Interestingly, we find a similar gender pattern when investigating positive parenting behavior in terms of child-centered communication (see Table A.3 in the Appendix).¹⁰ It is only fathers that appear to respond to EC with increased frequencies of ‘speaking with the child about his/her experiences’ (by 0.336 points on a 6-point scale) and ‘speaking with the child about things that annoy or burden him/her’ (by 0.554 points on a 6-point scale), while control-group parents show significant decreases in the frequency of child-centered communication.

¹⁰ Note that survey items on positive parenting behavior were child-specific in 2019 and collected only for children above the age of two. For this reason, we report results on this subsample and take the mean across children to make these observations comparable to the 2020 survey where these items were parent-specific. The survey questions read as follows: “How frequently does the following occur? I speak with the child about his/her experiences” (1 “never” to 6 “(almost) always”), “How frequently does the following occur? I speak with the child about things that annoy or burden him/her” (1 “never” to 6 “(almost) always”).

Tables A.4 to A.6 in the Appendix provide three types of robustness checks for our main outcomes. First, we reproduce the triple-difference results of Panel B of Table 3 (and additionally the gender interaction for the outcome ‘punish harder’) based on weighted regressions employing a combination of AID:A design weights at the household level and “staying probability” weights at the individual level (Table A.4 in the Appendix). Second, we re-run the analyses on a sample that includes the two federal states that applied an “all access” EC eligibility rule in the acute lockdown (Hamburg and Saarland), which we exclude in our main analysis sample (see Table A.5 in the Appendix). Third, we show results based on an alternative imputation of partners’ occupational systemic relevance, which is regression-based, employing federal state and occupation fixed effects (see Table A.6 in the Appendix). Overall, the results on harsh parenting behavior remain (sometimes marginally) statistically significant, and are qualitatively similar to our preferred estimates. This is also the case for estimates with respect to WHO-5 well-being, except for the weighted regressions, where the effect becomes statistically insignificant (see Table A.4 in the Appendix).

Given the categorical nature of the dependent variable ‘becoming angry’, we also investigate which frequency categories are most affected. That is, does the EC effect largely stem from changes in modest frequencies in ‘becoming angry’ or rather from frequency changes among already somewhat ‘angry’ parents? Table A.7 in the Appendix reports our key triple-difference results on ‘becoming angry’ from Panel B of Table 3 with respect to dichotomized outcome variables indicating different groupings of the frequency categories: the lowest frequency (“*never*”), the two lowest frequencies (“*never*” and “*seldom*”), the three highest frequencies (“*often*”, “*very often*” and “(*almost*) *always*”), and the two highest frequencies (“*very often*” and “(*almost*) *always*”).¹¹ The overall effect appears to largely originate in movements from the upper two frequency categories “*very often*” and “(*almost*) *always*” toward the category “*often*”, as well as movements from “*sometimes*” toward the lowest two frequencies of “*never*” or “*seldom*” becoming angry.

¹¹ There are too few observations in the highest frequency category “(*almost*) *always*” to allow its separate investigation.

5.1 Local Average Treatment Effects

The intention-to-treat results show how EC availability during the first COVID-19 lockdown in Germany affected parental well-being and negative parenting behavior. To interpret these results, it is—as a first step—important to understand the pattern of EC take-up. First, we quantify the relationship between EC availability and EC utilization by estimating the first stage model (4) based on Sample B.¹² We estimate the coefficient on eligibility (δ) to be about .32 with a standard error of .08. This estimate implies a 10 percentage point increase in EC eligibility among parents induces (an additional) 3.2 percent of parents to take up EC. To roughly understand what type of parents utilize EC when they are eligible (compliers), we estimate equation (4) separately for different types of parents. We partition Sample B sequentially by regional population size, parental education, mothers' labor market involvement, and age of the youngest child—that is, we split the sample in two, with one part including values equal to or below the median, and another for values above the median.

Column (1) of Table A.8 in the Appendix displays the median value of each characteristic. Column (2) reports the proportion of the sample that falls above the respective median value. Columns (3) and (4) show the distribution of compliers across the two subgroups (below or equal to the median, and above median) for each characteristic. Following Akerman, Gaarder, and Mogstad (2015), the proportion of the compliers of a given type is calculated as the ratio of δ for that subgroup to the δ in the overall sample, multiplied by the proportion of the sample in the respective subgroup.

We see that EC is not randomly adopted within the group of eligible parents. Compliers are slightly underrepresented among families that live in relatively more urban areas (with more than 50,000 inhabitants), and strongly underrepresented among parents with a university degree, in parental couples where mothers work more than 20 hours a week, and in families with relatively older children (i.e., the youngest child is above the age of three). The underrepresentation of compliers among the high-educated might in part be explained by the fact that the feasibility of working-from-home strongly increases with an

¹² Note that we now restrict the sample to families with children of preschool age or younger, since there is no information on EC utilization for school children in the data.

academic degree (Alipour, Falck and Schüller, 2020). This conjecture is corroborated by the finding that compliers are also underrepresented in families where mothers did work more than 50 percent of their work time remotely during the lockdown in March/April 2020.

In the following, we focus on the outcomes of harsh parenting behavior, where the intention-to-treat regressions yield robustly statistically significant “reduced form” effects.¹³ We deem the results with respect to WHO-5 well-being not entirely robust due to the lack of statistical significance in the double-difference as well as in the weighted regressions. In fact, also the LATE effects are not statistically significant for WHO-5 (see Table A.9 in the Appendix).

Invoking the exclusion restriction, we estimate how the utilization of EC affects the incidence of harsh parenting behavior among compliers. We approach the presentation of the LATE effects in a stepwise manner, taking the ITT effects presented in Table 3 as a starting point. Column (1) of Table 4 repeats the intention-to-treat effects reported in Table 3, as resulting from a first-differences regression and now without second-level interactions. In comparison with Table 3, the ITT effects without second-level interactions are about half the size, but are still statistically significant at the 5-percent level. Column (2) of Table 4 reports intention-to-treat effects estimated in the same way using Sample B instead of Sample A. ITT effects appear to increase in size and in statistical significance when families with relatively older children (i.e., with the youngest child being of school-age) are excluded from the sample. The specification without second-level interactions allows for a neat comparison of outcome changes between EC-eligible and non-eligible parents, since the estimated constant now indicates the average outcome changes for non-eligible parents. Hence, we can observe that the frequency of ‘becoming angry’ significantly increases by about 0.419 (0.426) points in the 6-point scale between 2019 and 2020 for non-EC-eligible parents, and that this increase is effectively reduced by about half (by 0.175 and by 0.270 points respectively for Sample A and B) for EC-eligible parents. Instead, with respect to ‘punishing harder’, we observe much smaller increases for the non-EC-eligible parents (0.035 and 0.095), which are not statistically significant for

¹³ We report LATE estimates on the remaining outcomes in Table A.9 in the Appendix. They are statistically insignificant throughout. There are also no significant gender differences.

Sample A, and only marginally statistically significant at the 10-percent level for Sample B. There are also no considerable differences in outcome changes for EC-eligible parents.

Column (3) of Table 4 presents estimates of equation (2) based on OLS. The OLS estimate is informative regarding the correlation between EC utilization and our outcome of interest, without any distinction made between EC utilization during the acute lockdown and utilization of *extended* EC during the subsequent re-opening. Interestingly, in this case, there appears to be no statistically significant association between EC utilization (including during extended EC) and the frequencies of harsh parenting behavior.

Table 4. Intention-To-Treat and LATE Effects on Negative Parenting Behavior for Parents with Children of Preschool Age or Younger. First-Differences Estimation.

<i>Panel A</i>		First Difference: 'Become Angry'				
	(1)	(2)	(3)	(4)	(5)	
	Red. Form (ITT)	Red. Form (ITT)	OLS	IV(LATE)	IV (LATE)	
Eligibility	-0.175*	-0.270**				
	(0.085)	(0.112)				
Usage			0.060	-0.842**	-0.816*	
			(0.118)	(0.426)	(0.463)	
Father × Usage					-0.071	
					(0.410)	
Constant	0.419***	0.426***	0.345***	0.582***	0.581***	
	(0.046)	(0.062)	(0.062)	(0.128)	(0.129)	
<i>N</i>	636	319	319	319	319	
Rkf	.	.	.	16.15	8.40	
Sample	A	B	B	B	B	
<i>Panel B</i>		First Difference: 'Punish Harder'				
	(1)	(2)	(3)	(4)	(5)	
	Red. Form (ITT)	Red. Form (ITT)	OLS	IV(LATE)	IV (LATE)	
Eligibility	-0.062	-0.134				
	(0.072)	(0.090)				
Usage			0.044	-0.418	-0.132	
			(0.101)	(0.313)	(0.289)	
Father × Usage					-0.795**	
					(0.343)	
Constant	0.035	0.095*	0.051	0.173*	0.162	
	(0.039)	(0.057)	(0.056)	(0.104)	(0.104)	
<i>N</i>	636	319	319	319	319	
Rkf	.	.	.	16.15	8.40	
Sample	A	B	B	B	B	

Source: AID:A 2019, AID:A Corona Add-on 2020; own calculations.

Notes: 'Become Angry': "How frequently does the following occur? I quickly become angry when my child(ren) don't do as I say" (1 "never" to 6 "(almost) always"). 'Punish Harder': "How frequently does the following occur? I punish my child(ren) harder than what they merit." (1 "never" to 6 "(almost) always"). Columns 4 and 5: EC usage is instrumented by eligibility to EC during the acute lockdown based on parental SR constellation. Rkf: Kleibergen-Paap rk Wald F statistic. Sample A: families with children below age 12. Sample B: families with children of preschool-age or younger. Cluster-robust standard errors at household level.

*** p < 0.01, ** p < 0.05, * p < 0.1.

Column (4) of Table 4 reports estimates based on IV estimation of equations (3) and (4). In contrast to the OLS estimate, with IV we estimate the effect of EC utilization on *compliers*, i.e., parents who utilize EC due to their SR-constellation based eligibility status. These compliers are most likely parents that had already taken up emergency care by the beginning of the lockdown. In turn, non-EC eligible parents that took up EC during the phase of extended EC are not compliers in this setup. The first stage is strong, with a Kleibergen-Paap rk Wald F-statistic of 16.15 on the excluded instrument, which means weak instrument bias is not a concern. The IV estimate with respect to ‘becoming angry’ is statistically significant at the 5-percent level, and suggests that EC utilization due to eligibility based on parents’ occupational systemic relevance was effective at preventing increases in negative and potentially harmful parenting behavior that would have happened in the absence of EC. Specifically, EC utilization reduced the frequency of ‘becoming angry’ by almost one point on the 6-point scale (0.842). As expected, the effect size is considerably larger among compliers than among all EC-eligible parents, where eligibility is associated with a 0.270-point lower frequency of ‘becoming angry’ (see Column 2). It also becomes evident that EC utilization among the EC-eligible can entirely prevent the increase in the frequency of ‘becoming angry’ that non-EC eligible parents experienced (0.582 points on the 6-point scale). In contrast, with respect to ‘punishing harder’, there appears to be no significant effect of EC utilization on compliers.

Column (5) finally reports on an investigation of potential gender differences in the IV estimates. It turns out that while there is no significant difference in EC effects on ‘becoming angry’ between fathers and mothers, the interaction with parental gender reveals that EC utilization significantly reduced the frequency of ‘punish harder’ for fathers, but not for mothers. The decrease in fathers’ harsh parenting in terms of ‘punishing harder’ is statistically significant at the 5-percent level, and is of substantial size (0.795 points on the 6-point scale).

6 Concluding Remarks

Having carried out a first rigorous evaluation of emergency childcare (EC) policies during the first COVID-19 lockdown in Germany in early 2020, we find that EC was not able to permanently shelter families from a considerable reduction in parental well-being. However, the provision of EC was effective in diminishing increases in harsh parenting among EC-eligible parents. This effect of EC is more pronounced in families with children of preschool-age or younger, and completely cancels out increases in harsh parenting among compliers.

We evaluate effects 4 to 5 months after the first COVID-19 lockdown in Germany, and hence provide evidence on the medium-term consequences, rather than the immediate impact of emergency childcare. Further research is needed to assess the mechanisms behind the gendered impact of EC on ‘punishing harder’ and child-centered communication, where we find effects exclusively for fathers.

Overall, our results disentangle effects caused directly by school and daycare closures from general effects of the pandemic, since among families with childcare usage we compare those who experienced or—due to EC—did not experience a complete disruption of external childcare provision. Thus, we conclude that, while decreasing parental well-being is likely to be a general pandemic effect rather than a specific effect of the closures, the observed increase in negative and potentially harmful parenting behavior is largely directly caused by school and daycare closures.

An important limitation of our study is that—given the data at hand—we can only provide somewhat isolated effects on parental well-being and parenting behavior, and not a comprehensive view of the impacts of school and daycare closures. To draw meaningful policy conclusions, impacts on e.g. long-term child development, health risks for parents and children, or the rate of new infections (see e.g. Dehning *et al.*, 2020; Brauner *et al.*, 2021) must be additionally considered.

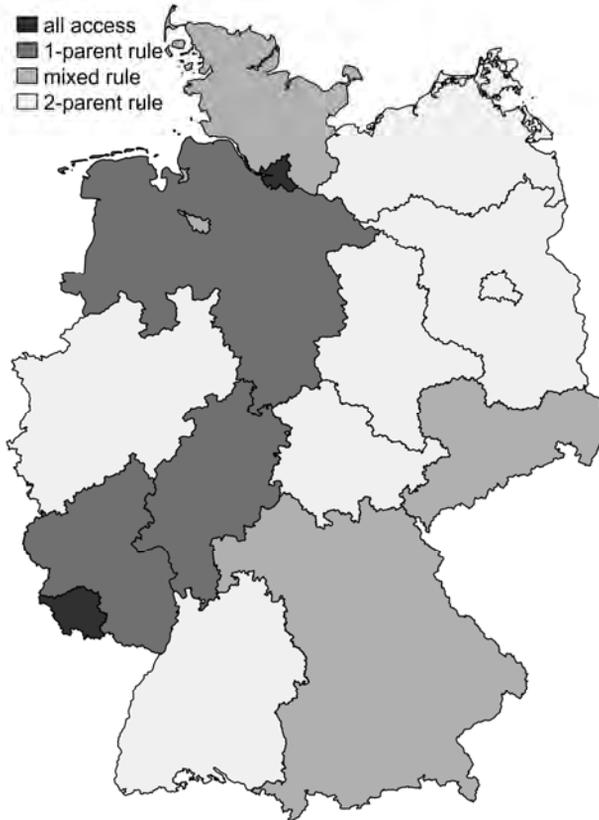
References

- Abidin, R. R. (1992). The determinants of parenting behavior. *Journal of Clinical Child Psychology* 21(4), 407–412.
- Akerman, A., I. Gaarder, and M. Mogstad (2015). The Skill Complementarity of Broadband Internet. *Quarterly Journal of Economics* 130(4), 1781–1824.
- Alipour, J.-V., O. Falck, and Schüller, S. (2020). Germany's Capacity to Work from Home. CESifo Working Paper No. 8227.
- Bauernschuster, S. and Schlotter, M. (2015). Public childcare and mother's labor supply. Evidence from two quasi-experiments. *Journal of Public Economics*, 123, 1–6.
- Blum, S., and Dobrotić, I. (2021). Childcare-policy responses in the COVID-19 pandemic: unpacking cross-country variation. *European Societies* 23(S1), 545–563.
- Brauner, J.M., Mindermann, S., Sharma, M., Johnston, D., Salvatier, J., Gavenčiak, T., Stephenson, A.B., Leech, G., Altman, G., Mikulik, V., Norman, A.J., Teperowski Monrad, J., Besiroglu, T., Ge, H., Hartwick, M.A., The, Y.W., Chindelevitch, L., Gal, Y., and Kulveit, J. (2021). Inferring the effectiveness of government interventions against COVID-19. *Science*, 371(6531).
- Bundesministerium für Familie, Senioren, Frauen und Jugend—BMFSFJ (2020). Kindertagesbetreuung Kompakt. Ausbaustand und Bedarf 2019 – Ausgabe 05a Kinder bis zum Schuleintritt. Retrieved from: <https://www.bmfsfj.de/blob/156672/aba616b5c3fc1cb9bd52e41aec73d246/kindertagesbetreuung-kompakt-ausbaustand-und-bedarf-2019-ausgabe05a-data.pdf> (last accessed on 18.02.2021).
- Dehning, J., Zierenberg, J., Spitzner, F.P., Wibral, M., Neto, J.P., Wilczek, M., and V. Priesemann (2020). Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science*, 369(6500).
- Deutscher Bundestag (2020). Medien und Journalisten im Kontext von „Kritischer Infrastruktur“ und „Systemrelevanz“. Sachstand WD 10 – 3000 – 016/20 des Wissenschaftlichen Dienstes, Fachbereich WD 10: Kultur, Medien und Sport. Retrieved from: <https://www.bundestag.de/resource/blob/691856/9f7d74b63d25edcac723005261905d88/WD-10-016-20-pdf-data.pdf> (last accessed on 16.03.2021).
- DJI and RKI (2020). *Monatsbericht der Corona-KiTa-Studie Ausgabe 03/2020*, Juli 2020.
- Huebener, M., Waights, S., Spiess, C.K., Siegel, N.A., and Wagner, G.G. (2021). Parental Well-Being in Times of COVID-19 in Germany. *Review of Economics of the Household*, 19, 91–122.
- Jackson, A. P., and Choi, J.-K. (2018). Parenting stress, harsh parenting, and Children's behavior. *Journal of Family Medicine & Community Health* 5(3), 10.
- Jugend- und Familienministerkonferenz—JFMK (2020). Gemeinsamer Rahmen der Länder für einen stufenweisen Prozess zur Öffnung der Kindertagesbetreuungsangebote von der Notbetreuung hin zum Regelbetrieb im Kontext der Corona-Pandemie. Beschluss der Jugend- und Familienministerkonferenz (JFMK) gemeinsam mit der Bundesministerin für Familie, Senioren, Frauen und Jugend vom 28.04.2020.

- Koebe, J., Samtleben, C., Schrenker, A., and Zucco, A. (2020). Systemrelevant und dennoch kaum anerkannt: Das Lohn- und Prestigeniveau unverzichtbarer Berufe in Zeiten von Corona, DIW aktuell, Nr. 28 — 24. März 2020.
- Lee, S. J., Grogan-Kaylor, A., and Berger, L.M. (2014). Parental spanking of 1-year-old children and subsequent child protective services involvement. *Child Abuse and Neglect* 38(5), 875–883.
- Möhring, K., Naumann, E., Reifenscheid, M. Wenz, A., Rettig, T., Krieger, U., Friedel, S., Finkel, M., Cornesse, C. and Bloom, A.G. (2020). The COVID-19 Pandemic and Subjective Well-being: Longitudinal Evidence on Satisfaction with Work and Family. *European Societies* 23(1), 601–617.
- Pinquart, M. (2017). Associations of parenting dimensions and styles with externalizing problems of children and adolescents: An updated meta-analysis. *Developmental Psychology* 53(5), 873–932.
- Pischke, J.-S. (2007). Empirical Methods in Applied Economics. Fixed Effects, Differences-in-differences, and Panel Data. Lecture Notes, London School of Economics. Retrieved from: http://econ.lse.ac.uk/staff/spischke/ec524/evaluation4_07.pdf (last accessed on 03.03.2021)
- Prime, H., Wade, M., and Browne, D.T. (2020). Risk and Resilience in Family Well-being during the COVID-19 Pandemic. *American Psychologist*, 75(5), 631–643.
- Schmitz, S. (2020). The Impact of Publicly Funded Childcare on Parental Well-Being: Evidence from Cut-Off Rules. *European Journal of Population* 36, 171–196.
- Statistisches Bundesamt (2018). Schulen auf einen Blick. Retrieved from: <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Schulen/Publikationen/Downloads-Schulen/broschuere-schulen-blick-0110018189004.pdf?blob=publicationFile>. (last accessed on 19.03.2021)
- Topp C.W., Østergaard S.D., Søndergaard S., and Bech. P. (2015). The WHO-5 Well-Being Index: A Systematic Review of the Literature. *Psychotherapy and Psychosomatics*, 84(3), 167–176.
- van Huizen, T. and Plantenga, J. (2018). Do children benefit from universal early childhood education and care? A meta-analysis of evidence from natural experiments. *Economics of Education Research* 66, 206–222.
- Wooldridge, J.M. (2010). *Econometric Analysis of Cross Section and Panel Data. Second Edition*. MIT Press.
- World Health Organization (1998). *Info Package: Mastering Depression in Primary Care, Version 2.2*. WHO, Regional Office for Europe. Psychiatric Research Unit, Frederiksborg.

Appendix

Figure A.1. Emergency Childcare Policies in the German Federal States during the first COVID-19 Lockdown



Sources: Decrees or corresponding press releases by the respective federal state (see Table 1).

Covid Economics 79, 27 May 2021: 117-151

Table A.1. Exogeneity of EC Eligibility

	EC Eligibility
<i>Q2 equivalence household income (ref.)</i>	
Q2 equivalence household income	-0.061 (0.053)
Q3 equivalence household income	0.019 (0.059)
Q4 equivalence household income	-0.012 (0.064)
Age mother	0.004 (0.006)
Age father	-0.005 (0.004)
Parent with university degree	0.052 (0.045)
≤ 50.000 inhabitants	-0.031 (0.046)
Migration background	-0.032 (0.051)
More than 1 child in hh	-0.009 (0.042)
Mean age children	0.000 (0.009)
Share male children in hh	0.012 (0.049)
At least one room per child	0.022 (0.057)
Constant	0.290 (0.177)
<i>N</i>	478

Source: AID:A 2019, AID:A Corona Add-on 2020; own calculations.

Notes: Cluster-robust standard errors at household level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.2. Triple Difference. Intention-To-Treat Effects on Parental Well-Being and Negative Parenting Behavior. Standardized Outcomes.

	Parental Well-Being			Negative Parenting Behavior		
	Life Satisfaction (1)	Partnership Satisfaction (2)	WHO-5 (3)	'Become Angry' (4)	'Punish Harder' (5)	(6)
Eligibility × Year = 2020	0.161 (0.215)	-0.020 (0.153)	0.429** (0.203)	-0.423** (0.182)	-0.141 (0.205)	0.002 (0.198)
Father × Eligibility × Year = 2020						-0.401** (0.161)
SR constellation group × Year = 2020	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
State × Year = 2020	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
Individual FE	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>N</i>	1,272	1,272	1,272	1,272	1,272	1,272
Nr. of individuals	636	636	636	636	636	636
Nr. of households	482	482	482	482	482	482

Source: AID:A 2019, AID:A Corona Add-on 2020; own calculations.

Notes: All outcomes are z-score rescaled to have a mean of zero and a standard deviation of one. Systemic relevance (SR) constellation groups: (a) no parent SR, (b) one parent SR, not health-related, (c) one parent SR, health related, (d) both parents SR. Cluster-robust standard errors at household level. *** p < 0.01, ** p < 0.05, * p < 0.1.

Covid Economics 79, 27 May 2021: 117-151

Table A.3. Intention-To-Treat Effects on Positive Parenting Behavior (Child-Centered Communication)

	Double Difference		Triple Difference			
	'Speak about Experiences,	'Speak about Annoyances'	'Speak about Experiences,		'Speak about Annoyances'	
	(1)	(2)	(3)	(4)	(5)	(6)
Year = 2020	-0.161** (0.072)	-0.320*** (0.090)				
Eligibility × Year = 2020	0.155 (0.111)	0.083 (0.146)	0.056 (0.172)	-0.056 (0.175)	0.117 (0.243)	-0.066 (0.237)
Father × Eligibility × Year = 2020				0.336** (0.156)		0.554** (0.246)
Constant	5.430*** (0.021)	5.258*** (0.027)	5.432*** (0.030)	5.430*** (0.021)	5.258*** (0.039)	5.257*** (0.027)
SR constellation group × Year = 2020	<i>no</i>	<i>no</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
State × Year = 2020	<i>no</i>	<i>no</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
Individual FE	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>N</i>	552	552	1,186	1,186	1,185	1,185
Nr. of individuals	294	294	636	636	636	636
Nr. of households	217	217	482	482	482	482

Source: AID:A 2019, AID:A Corona Add-on 2020; own calculations.

Notes: 'Speak about Experiences': "How frequently does the following occur? I speak with the child about his/her experiences" (1 "never" to 6 "(almost) always"). 'Speak about Annoyances': "How frequently does the following occur? I speak with the child about things that annoy or burden him/her" (1 "never" to 6 "(almost) always"). AID:A 2019 surveyed both items child-specific for children above age two, whereas these items were surveyed parent-specific in the AID:A Corona Add-on 2020. We employ the mean across children for t = 2019. Systemic relevance (SR) constellation groups: (a) no parent SR, (b) one parent SR, not health-related, (c) one parent SR, health related, (d) both parents SR. Cluster-robust standard errors at household level. *** p < 0.01, ** p < 0.05, * p < 0.1.

Table A.4. Triple Difference. Intention-To-Treat Effects on Parental Well-Being and Negative Parenting Behavior. Weighted Regressions

	Parental Well-Being			Negative Parenting Behavior		
	Life Satisfaction (1)	Partnership Satisfaction (2)	WHO-5 (3)	'Become Angry' (4)	'Punish Harder' (5)	(6)
Eligibility × Year = 2020	0.310 (0.210)	0.096 (0.173)	3.710 (3.633)	-0.688*** (0.223)	-0.162 (0.179)	-0.084 (0.181)
Father × Eligibility × Year = 2020						-0.254* (0.134)
SR constellation group × Year = 2020	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
State × Year = 2020	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
Individual FE	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>N</i>	1,272	1,272	1,272	1,272	1,272	1,272
Nr. of individuals	636	636	636	636	636	636
Nr. of households	482	482	482	482	482	482

Source: AID:A 2019, AID:A Corona Add-on 2020; own calculations.

Notes: Results of weighted regressions based on a combination of AID:A design weights at the household level and “staying probability” weights at the individual level. Life satisfaction, as well as partnership satisfaction, are measured on a 6-point scale ranging from 1 “not at all satisfied” to 6 “very satisfied”. The WHO-5 Well-being Index ranges from 0 “absence of well-being” to 100 “maximal well-being”. ‘Become Angry’: “How frequently does the following occur? I quickly become angry when my child(ren) don’t do as I say” (1 “never” to 6 “(almost) always”). ‘Punish Harder’: “How frequently does the following occur? I punish my child(ren) harder than what they merit.” (1 “never” to 6 “(almost) always”). Systemic relevance (SR) constellation groups: (a) no parent SR, (b) one parent SR, not health-related, (c) one parent SR, health related, (d) both parents. Cluster-robust standard errors at household level. *** p < 0.01, ** p < 0.05, * p < 0.1.

Covid Economics 79, 27 May 2021: 117-151

Table A.5. Triple Difference. Intention-To-Treat Effects on Parental Well-Being and Negative Parenting Behavior. Including “All-Access” States.

	Parental Well-Being			Negative Parenting Behavior		
	Life Satisfaction (1)	Partnership Satisfaction (2)	WHO-5 (3)	‘Become Angry’ (4)	‘Punish Harder’ (5)	(6)
Eligibility × Year = 2020	0.171 (0.184)	-0.058 (0.156)	7.204** (3.359)	-0.336* (0.180)	-0.083 (0.158)	0.036 (0.153)
Father × Eligibility × Year = 2020						-0.343*** (0.123)
SR constellation group × Year = 2020	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
State × Year = 2020	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
Individual FE	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>N</i>	1,318	1,318	1,318	1,318	1,318	1,318
Nr. of individuals	659	659	659	659	659	659
Nr. of households	498	498	498	498	498	498

Source: AID:A 2019, AID:A Corona Add-on 2020; own calculations.

Notes: This table displays results including observations from the federal states Hamburg and Saarland. Life satisfaction, as well as partnership satisfaction, are measured on a 6-point scale ranging from 1 "not at all satisfied" to 6 "very satisfied". The WHO-5 Well-being Index ranges from 0 "absence of well-being" to 100 "maximal well-being". 'Become Angry': "How frequently does the following occur? I quickly become angry when my child(ren) don't do as I say" (1 "never" to 6 "(almost) always"). 'Punish Harder': "How frequently does the following occur? I punish my child(ren) harder than what they merit." (1 "never" to 6 "(almost) always"). Systemic relevance (SR) constellation groups: (a) no parent SR, (b) one parent SR, not health-related, (c) one parent SR, health related, (d) both parents. Cluster-robust standard errors at household level.

*** p < 0.01, ** p < 0.05, * p < 0.1.

Table A.6. Triple Difference. Intention-To-Treat Effects on Parental Well-Being and Negative Parenting Behavior. Regression-based Imputation.

	Parental Well-Being			Negative Parenting Behavior		
	Life Satisfaction (1)	Partnership Satisfaction (2)	WHO-5 (3)	'Become Angry' (4)	'Punish Harder' (5)	(6)
Eligibility × Year = 2020	-0.035 (0.193)	0.082 (0.181)	6.078* (3.586)	-0.351* (0.179)	-0.068 (0.171)	0.064 (0.161)
SR constellation group × Year = 2020	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
State × Year = 2020	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
Individual FE	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>N</i>	1,272	1,272	1,272	1,272	1,272	1,272
Nr. of individuals	636	636	636	636	636	636
Nr. of households	482	482	482	482	482	482

Source: AID:A 2019, AID:A Corona Add-on 2020; own calculations.

Notes: If missing, partner's SR status is predicted by regressing individual health-related SR (or non-health related SR, respectively) on occupation (KldB 2010, 3-digit) as stated in 2019 and state fixed effects. Life satisfaction, as well as partnership satisfaction, are measured on a 6-point scale ranging from 1 "not at all satisfied" to 6 "very satisfied". The WHO-5 Well-being Index ranges from 0 "absence of well-being" to 100 "maximal well-being". 'Become Angry': "How frequently does the following occur? I quickly become angry when my child(ren) don't do as I say" (1 "never" to 6 "(almost) always"). 'Punish Harder': "How frequently does the following occur? I punish my child(ren) harder than what they merit." (1 "never" to 6 "(almost) always"). Systemic relevance (SR) constellation groups: (a) no parent SR, (b) one parent SR, not health-related, (c) one parent SR, health related, (d) both parents. Cluster-robust standard errors at household level. *** p < 0.01, ** p < 0.05, * p < 0.1.

Covid Economics 79, 27 May 2021: 117-151

Table A.7. Triple Difference. Intention-To-Treat on ‘Becoming Angry’. Dichotomized Outcome.

<i>Panel A: I = “never”; 0 otherwise</i>	
Eligibility × Year = 2020	0.081 (0.067)
<i>N</i>	1,272
<i>Panel B: I = “seldom” or “never”; 0 otherwise</i>	
Eligibility × Year = 2020	0.161* (0.097)
<i>N</i>	1,272
<i>Panel C: I = “often”, “very often” or “(almost) always”; 0 otherwise</i>	
Eligibility × Year = 2020	-0.037 (0.088)
<i>N</i>	1,272
<i>Panel D: I = “very often” or “(almost) always”; 0 otherwise</i>	
Eligibility × Year = 2020	-0.092** (0.037)
<i>N</i>	1,272

Source: AID:A 2019, AID:A Corona Add-on 2020; own calculations.

Notes: Each panel represents results of separate regressions. All regressions include fixed effects as reported in Panel B of Table 3. ‘Become Angry’: “How frequently does the following occur? I quickly become angry when my child(ren) don’t do as I say.” Cluster-robust standard errors at household level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.8. Complier Analysis

	Median	Sample share > median	Proportion of compliers:	
	(1)	(2)	≤ median	> median
	(1)	(2)	(3)	(4)
Population size of residence municipality (7 categories)	4 “20,000–50,000 inh.”	0.36	0.66	0.34
At least one parent holds university degree (0/1)	0	0.44	0.72	0.31
Mother’s weekly working hours (2019)	20	0.41	0.72	0.26
Age youngest child in household	3.17	0.48	0.65	0.37
Mother more than 50% remote work (2020)	0	0.30	0.75	0.24

Source: AID:A 2019, AID:A Corona Add-on 2020; own calculations.

Notes: We partition the IV sample (Sample B: families with a child of preschool-age or younger) sequentially by regional population size, parental education, mother’s labor market involvement, and age of the youngest child (above and below median of each characteristic). Column (1) displays the median value of each characteristic. Column (2) reports the proportion of the sample that falls above the respective median value. Columns (3) and (4) show the distribution of compliers across the two subgroups (below or equal to the median and above median) for each characteristic. The proportion of compliers of a given type is calculated as the ratio of $\hat{\delta}$ for that subgroup to the $\hat{\delta}$ in the overall IV sample (Sample B), multiplied by the proportion of the sample in the respective subgroup.

Table A.9. LATE Effects on Parental Well-Being for Parents with Children of Preschool Age or Younger. First-Differences Estimation.

	First Differences:					
	Life Satisfaction		Partnership Satisfaction		WHO-5	
	(1)	(2)	(3)	(4)	(5)	(6)
Usage	0.110 (0.402)	0.049 (0.456)	0.309 (0.446)	0.355 (0.478)	0.133 (7.656)	3.578 (7.897)
Father × Usage		0.171 (0.407)		-0.128 (0.360)		-9.593 (7.055)
Constant	-0.277** (0.122)	-0.275** (0.123)	-0.157 (0.125)	-0.158 (0.125)	-0.587 (2.433)	-0.712 (2.422)
<i>N</i>	319	319	319	319	319	319
Rkf	16.15	8.40	16.15	8.40	16.15	8.40
Sample	B	B	B	B	B	B

Source: AID:A 2019, AID:A Corona Add-on 2020; own calculations.

Notes: EC usage is instrumented by eligibility to EC during the acute lockdown based on parental SR constellation. Life satisfaction, as well as partnership satisfaction, are measured on a 6-point scale ranging from 1 "not at all satisfied" to 6 "very satisfied". The WHO-5 Well-being Index ranges from 0 "absence of well-being" to 100 "maximal well-being". Cluster-robust standard errors at household level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The Global COVID-19 Student Survey: First wave results¹

David A. Jaeger,² Jaime Arellano-Bover,³ Krzysztof Karbownik,⁴ Marta Martínez-Matute,⁵ John M. Nunley,⁶ R. Alan Seals,⁷ Miquel Almunia,⁸ Mackenzie Alston,⁹ Sascha O. Becker,¹⁰ Pilar Beneito,¹¹ René Böheim,¹² José E. Boscá,¹³ Jessica H. Brown,¹⁴ Simon Chang,¹⁵ Deborah Cobb-Clark,¹⁶ Shooshan Danagoulian,¹⁷ Sandra Donnally,¹⁸ Marissa Eckrote-Nordland,¹⁹ Lúdia Farré,²⁰ Javier Ferri,²¹ Margherita Fort,²² Jane Cooley Fruewirth,²³ Rebecca Gelding,²⁴ Allen C. Goodman,²⁵ Melanie Guldi,²⁶ Simone Häckl,²⁷ Janet Hankin,²⁸ Scott Imberman,²⁹ Joanna Lahey,³⁰ Joan Llull,³¹ Hani Mansour,³² Isaac McFarlin,³³ Jaakko Meriläinen,³⁴ Tove Mörtlund,³⁵ Martin Nybom,³⁶ Stephen D. O'Connell,³⁷ Rupert Sausgruber,³⁸ Amy Ellen Schwarz,³⁹ Jan Stuhler,⁴⁰ Petra Thiemann,⁴¹ Roel van Veldhuizen,⁴² Marianne Wannamaker,⁴³ Maria Zhu⁴⁴

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

University students have been particularly affected by the COVID-19 pandemic. We present results from the first wave of the Global COVID-19 Student Survey, which was administered at 28 universities in the United States, Spain, Australia, Sweden, Austria, Italy, and Mexico between April

1 This research has been reviewed and approved by the National Bureau of Economic Research Institutional Review Board, IRB Ref #20_097. David A. Jaeger thanks IZA for financial support and Kristen Kenny at the NBER for her tireless efforts with IRB Review. John Nunley thanks Natalie Solverson, Director of Institutional Research, at University of Wisconsin–La Crosse. Joan Llull acknowledges financial support by the European Research Council (ERC), through Starting Grant n. 804989, by the Generalitat de Catalunya (2017-SGR-1765), and by the Spanish Ministry of Science and Innovation through grant PGC2018-094364-B-I00 and through the Severo Ochoa Program for Centers of Excellence in RD (CEX2019-000915-S).

2 Professor, University of St Andrews, CESifo, IZA, and CEPR. 3 Postdoctoral Associate, Yale University and IZA. 4 Assistant Professor, Emory University, CESifo, IZA, and NBER. 5 Assistant Professor, Universidad Autónoma de Madrid and IZA. 6 Professor, University of Wisconsin-La Crosse. 7 Associate Professor, Auburn University. 8 Professor, Colegio Universitario de Estudios Financieros (CUNEF) and CEPR. 9 Assistant Professor, Florida State University. 10 Professor, Monash University, University of Warwick, CESifo, IZA, and CEPR. 11 Associate Professor, Universitat de València. 12 Professor, Johannes Kepler Universität Linz and IZA. 13 Professor, Universitat de València and FEDEA. 14 Assistant Professor, University of South Carolina. 15 Senior Lecturer, University of Western Australia, GLO, and IZA. 16 Professor, University of Sydney and IZA. 17 Assistant Professor, Wayne State University. 18 PhD Student, Lunds Universitet. 19 PhD Student, Michigan State University. 20 Associate Professor, Universitat de Barcelona and IZA. 21 Professor, Universitat de València and FEDEA. 22 Professor, Università di Bologna, FBK-IRVAPP, CESifo, IZA, and CEPR. 23 Associate Professor, University of North Carolina-Chapel Hill. 24 Project Officer, University of Sydney. 25 Professor, Wayne State University. 26 Assistant Professor, University of Central Florida. 27 Assistant Professor, Universitetet di Stavanger. 28 Professor, Wayne State University. 29 Professor, Michigan State University, CESifo, and NBER. 30 Associate Professor, Texas A&M University and NBER. 31 Associate Professor, MOVE, Universitat Autònoma de Barcelona, and Barcelona GSE. 32 Associate Professor, University of Colorado Denver, CESifo, and IZA. 33 Assistant Professor, University of Florida and NBER. 34 Assistant Professor, Instituto Tecnológico Autónomo de México (ITAM). 35 Analyst, IFAU. 36 Associate Professor, IFAU and Uppsala Universitet. 37 Assistant Professor, Emory University and IZA. 38 Professor, Wirtschaft universität Wien. 39 Professor, Syracuse University. 40 Associate Professor, Universidad Carlos III de Madrid, IZA, and CEPR. 41 Assistant Professor, Lunds Universitet and IZA. 42 Assistant Professor, Lunds Universitet. 43 Associate Professor, University of Tennessee-Knoxville, IZA, and NBER. 44 Assistant Professor, Syracuse University.

Copyright: David A. Jaeger, Jaime Arellano-Bover, Krzysztof Karbownik, Marta Martínez-Matute, John M. Nunley, R. Alan Seals, et al.

and October 2020. The survey addresses contemporaneous outcomes and future expectations regarding three fundamental aspects of students' lives in the pandemic: the labor market, education, and health. We document the differential responses of students as a function of their country of residence, parental income, gender, and for the US their race.

The COVID-19 pandemic has affected virtually every aspect of life in most countries. Education at all levels has been particularly disrupted, with formal instruction either ceasing or moving online, often for months at a time. In the spring of 2020, most university administrators faced difficult decisions regarding whether to move students out of university accommodations and whether and how to move instruction online, with concerns about student experience and whether students might abandon their university altogether. Students confronted health and well-being concerns, uncertainty regarding their immediate educational future, as well as parental job loss or loss of income, and their own future labor market prospects.

A extensive literature has emerged that documents the changes brought about by the pandemic. For current students and recent graduates, the consequences of the sudden transition to remote instruction (Blaskó et al., 2021) and remote work (Barrero et al., 2021) are likely to persist (and not be understood) for many years. Research from the United States also shows the pandemic altered student expectations for their careers and earnings (e.g., Aucejo et al., 2020), as well as their relative valuations of the college experience (e.g., Aucejo et al., 2021). The economic shutdowns and social-distancing protocols of pandemic life have also had disproportionate effects on women (e.g., Alon et al., 2021; Albanesi and Kim, 2021). The COVID-19 pandemic has refocused and magnified racial/ethnic inequality in the United States (e.g., Polyakova et al., 2021; Wrigley-Field, 2020) and in Europe (e.g., Razai et al., 2021; Shaaban et al., 2020). Recent papers have also documented the pandemic's effect on student stress and wellbeing (e.g., Aucejo et al., 2020; Rodríguez-Planas, 2020; Browning et al., 2021; Logel et al., 2021).

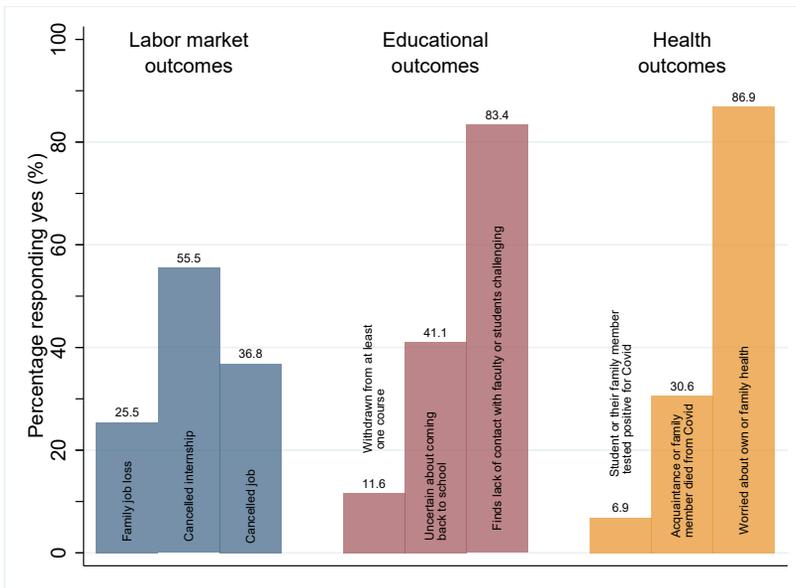
To measure college students' reactions to the various crises presented by the COVID-19 pandemic, we created the Global COVID-19 Student Survey (subsequently GC19SS). The goal of the GC19SS was to capture, on a global scale, how students were coping with the unprecedented (in their lifetimes) disruptions. By necessity working within a short time frame, the survey was written, IRB permission obtained, and the survey fielded at 28 large, mostly public, universities in the United States, Australia, Austria, Italy, Mexico, Spain, and Sweden beginning in late April 2020. This paper reports the basic first-wave results of the GC19SS.

The survey addresses three fundamental aspects of students' lives in the pandemic: their current and future academic situation, their current health and well-being (including that of their families), and their perceptions about their future labor market preferences and success. Labor market questions refer to job loss, students' labor market activity, preferences for positive job characteristics and willingness to accept negative ones, and earnings expectations at ages 30 and 45. Questions on educational outcomes concern contemporaneous learning, time allocation to class work, and future schooling plans. Health-related questions gather information on COVID-19 incidence and mental health issues related to the pandemic.

Figure 1 provides a broad summary of the survey's findings on labor market, educational, and health outcomes. The main message of Figure 1 is that COVID-19 has deeply

affected a generation of university students across the globe. Pooling all respondents together, Figure 1 shows that 26% of students had a family member experience job loss, 56% of those who had internship plans for the summer of 2020 had them cancelled, and 37% of those who had a job offer had it cancelled. With respect to education, 12% of students withdrew from at least one course, 41% were uncertain about coming back to school in the fall of 2020, and 83% expressed that the lack of contact with faculty or other students was challenging. At a time when testing was still not widespread, 7% students experienced a positive test for COVID-19 either personally or in their family, 31% had a family member or acquaintance die from COVID-19, and 87% were worried about their health or that of their family members.

Figure 1: Labor, educational, and health consequences of COVID-19 pandemic



The first three bars of this figure summarize labor market outcomes, the next three bars summarize educational outcomes, and the last three bars summarize health outcomes. Sample sizes differ by question. They are 28263, 12026, 1015, 36415, 29687, 34552, 28263, 26859, 32053 for bars one to nine, respectively. When it comes to the labor market outcomes both internship and job cancellations (bars two and three) are conditional on having been offered a job or planning an internship.

In the remainder of the paper, we present more detailed results on the three broad areas of labor market, education, and health outcomes. For each outcome, we document heterogeneous responses according to students' country of residence, parental income, gender, and, for US respondents, race/ethnicity. In our view, a key strength of the GC19SS is the ability to document the experiences of university students—and how they differ across types of students—in a manner that is consistent and comparable across countries and institutions.

Covid Economics 79, 27 May 2021: 152-217

1 Data

1.1 Survey instrument and data collection

The survey instrument for the first wave of the GC19SS was developed in late March and early April of 2020 by a small subset of the research team. The goals in designing the instrument were to gauge the impact of the developing pandemic on students academic experience and well-being, their expectations about the future job market and how those had been affected by the pandemic, and a set of demographic and preference questions. One of the guiding principles in designing the survey, to the extent possible, was to use questions that had been used previously or concurrently in other surveys, particularly the US Census (for demographic information), the International Survey on Coronavirus (Fetzer et al., 2020), and the the Global Preference Survey (Falk et al., 2018). This allows comparability of responses in the GC19SS to other surveys and data sources. To facilitate follow-ups, students were asked to provide an email address. IRB approval for the survey instrument was received from the NBER on 17 April 2020. An example of the US version of the survey instrument is included as Appendix B.

The survey was first designed in English to be appropriate for the United States, and then was translated for use in other countries. Questions were adapted to be appropriate for the context in each country. For example, questions regarding employer-provided health insurance are not relevant in some countries such as Sweden. Questions that refer to income levels (both family income and prospective income at ages 30 and 45 for the survey respondents) were designed to be comparable across countries, using as reference the same quantiles from each country's income distribution. Education categories were adopted from standard surveys in each country rather than trying to shoehorn responses into categories relevant for the US

Research partners were successfully solicited at (mostly) large public universities in the United States, Spain, Australia, Austria, Sweden, Italy, and Mexico.¹ Universities either ceded human subjects authority to the NBER or subjected the survey to IRB/ethics board review. In addition, approval to use student email addresses was received at each university in the survey.

The GC19SS is administered using the Qualtrics platform. Students were contacted through email in all cases, either directly through Qualtrics (when universities provided us with a list of email addresses) or by receiving an email from the university's administration with a link to the survey. In most cases, reminder emails were sent to students at various intervals after the initial solicitation. Response rates varied by university, but were usually close to 10-12 percent. Typically, just less than half of those who responded provided email addresses for subsequent follow-up. All identifying information was removed from the data before analysis.

¹Our IRB agreement prevents us from identifying at which universities the survey was administered. This was an intentional choice designed to increase the likelihood that administrators would approve the survey at their university.

1.2 Sample

Our sample includes data from 7 countries and 28 universities. We gathered information for 14 schools in the US, 5 schools in Spain, 3 schools in Australia, 2 schools in Sweden and Austria, and 1 school in both Italy and Mexico. Our full sample size contains 39,172 unique students but not all their responses are complete. Throughout the analysis we utilize maximum available samples for each question of interest, and we report these in figures' notes.

In the full sample, 54 percent of students come from the US, followed by Spain at 17 percent, Australia at 13 percent, and Italy at 11 percent. The remaining countries contribute less than 2 percent of the full sample each due to their smaller educational markets. We observe 25 percent of males, 54 percent of females, and 22 percent of students who do not report their gender. Similarly, in the US, we miss racial information for about 28 percent of respondents. As noted below these missing data issues are due to positioning of the demographic questions in the survey document. Excluding these missing values, which we do whenever we split the sample by either gender or race and ethnicity in the US, results in a sample with 69 percent of females and 31 percent males. This is not surprising given that in all countries considered in these survey females are over-represented among college enrollees. For example, this ratio is approximately 60 to 40 in the US and 58 to 42 in Australia and Sweden. In the US sample, the racial-ethnic percentages are 50 percent White, 5 percent Black, 7 percent Asian, and 10 percent Hispanic. Irrespective of the exact characteristics our conclusions remain very similar if we re-weight the results with racial and gender composition of all students enrolled in universities considered in our study. Finally, income information is not reported by about 12 percent of students in our sample.

2 Empirical approach

We document the findings of GC19SS across three broad topics: those related to the labor market (contemporaneous outcomes and future prospects), education, and health. For each of these sets of outcomes, we document heterogeneous student responses, stratifying the data along four dimensions: country, parental income, gender, and, for US respondents, their race/ethnicity.

Results by country. We report results from the survey separately for each of the seven countries in the sample: Australia, Austria, Italy, Mexico, Spain, Sweden, and the United States. Asking comparable questions to undergraduate students across countries is a strength of GC19SS and sheds light on how the pandemic affected university students in different parts of the world. When comparing results across countries, however, it should be kept in mind that Mexico respondents are students from a single elite institution who are likely not representative of the broader population of Mexican undergraduate students. For most countries our sample includes multiple universities: three in Australia, two in Austria, five in Spain, two in Sweden, and fourteen in the United States. For Italy, our sample also only includes

one university but, contrary to the Mexican case, this is a large public institution.

Parental income differences. We document results by students' socioeconomic backgrounds by asking about respondents' parental income in a comparable way across countries, using common percentiles of each country's household income distribution. We group students by the quintiles of the household income distribution to which their parents' belong. These analyses complement, from an international perspective, existing evidence showing how the pandemic has disproportionately negatively affected workers and households with lower incomes (e.g. Chetty et al., 2020).

Gender differences. Across most institutions in our sample, women are a majority of undergraduates, and existing evidence suggests asymmetric impacts of the pandemic on men and women (Alon et al., 2020).² We believe this makes understanding disparate impacts of the pandemic by gender on the population of undergraduate students particularly important. We note that the survey instrument asked about gender towards the the end, which in turn resulted in missing gender information for 22% of respondents. Our analyses by gender are thus carried out on the remaining 78% of the sample.³

Racial differences in the US. Lastly, we separately examine respondents from US institutions and document heterogeneous results by race/ethnicity for Whites, Blacks, Asians, and Hispanics. Several reports indicate that racial minorities have been most severely affected in the US (e.g., Couch et al., 2020; Hardy and Logan, 2020). Our analyses by race/ethnicity for US respondents contribute to understanding the degree to which these disparities extend to the population of undergraduate students. Similarly as with gender, due to its location within the survey, we note that race is missing for 28% of respondents so our analyses by race are carried out on the remaining 72% of US respondents.⁴

We use graphs to report the majority of our results, showing differences in mean responses by group, for each of the four above-mentioned dimensions of heterogeneity we consider. Additionally, for differences across parental income, gender, and race/ethnicity, we estimate different versions of the following linear regression:

$$y_i = \Gamma_{g(i)} + X_i' \beta + \Phi_{u(i)} + \varepsilon_i, \quad (1)$$

²Overrepresentation of women among university students is consistent with national and international statistics (e.g., UNESCO, 2012). Nonetheless, in all universities considered here, female students were more likely to participate in and finish the survey. This female-favorable gap ranged across institutions from 2.5 to 28 percentage points. We computed results presented in all figures and tables re-weighting for the share of women at a particular institutions which yielded almost identical results. For brevity we do not present the re-weighted results, but they are available upon request.

³When we re-weight the sample to be representative of actual gender composition of the university, we assign weight of one to students who did not respond to the question concerning their gender.

⁴The Pearson correlation between missing race/ethnicity and missing gender in the United States is 0.84.

where y_i is an outcome of interest of student i , $\Gamma_{g(i)}$ are categorical dummies for each of the relevant dimensions of heterogeneity g (i.e., parental income quintile, gender, or race/ethnicity), X_i are student baseline covariates (gender, field of study, and university year), and $\Phi_{u(i)}$ are university fixed effects.

We estimate versions of (1) which sequentially include i) only $\Gamma_{g(i)}$; ii) $\Gamma_{g(i)}$ and X_i ; and iii) $\Gamma_{g(i)}$, X_i , and $\Phi_{u(i)}$. The first specification simply tests for the statistical significance of the raw mean differences across parental income, gender, and race/ethnicity that we present graphically. The second specification checks whether such differences remain when holding constant basic demographics and student characteristics. The third specification further asks whether such differences arise when comparing students within the same university. The last specification, which includes university fixed effects, implicitly controls for country fixed effects, and, since the survey was fielded at slightly different time at different institutions, timing of the survey.

We focus our main results on showing unconditional means which we report graphically in the main text. Additionally, we present tables of estimates based on equation (1) in Appendix A. As it turns out, most of the differences we emphasize across parental income, gender, and race/ethnicity remain when controlling for student covariates and university fixed effects. In the main text, each figure showing unconditional means references its corresponding regression table.

3 Results

In this section, we report GC19SS findings on three broad set of students' outcomes related to the labor market, education, and health.

3.1 Labor market outcomes

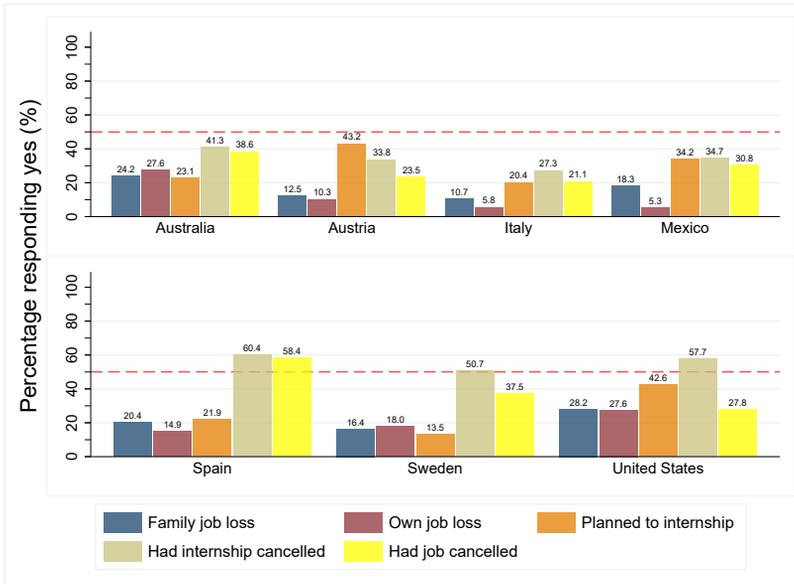
We analyze multiple outcomes related to the labor market: job loss, students' labor market activity, future career considerations, willingness to accept negative job characteristics after graduation, and earnings expectations at ages 30 and 45.

3.1.1 Job loss

We document the intensity of job loss experienced by university students showing the rates of job loss of an immediate family member, own job loss, canceled internships, and canceled job offers.

Results by country. Figure 2 shows how job loss intensity varied across countries. The US had the highest rate of family job loss, with 28% of respondents having one or more immediate family members lose their job. This number was equal to 11% in Italy, 13% in Austria, 16% in Sweden, 18% in Mexico, 20% in Spain, and 24% in Australia. US and Australian students were also the most likely to report having lost an existing job themselves (28% in both countries). By contrast, only 5% and 6% of Mexico and Italy respondents, respectively,

Figure 2: Job loss measures, by country



Note: This figure presents mean values of responses to the following five questions: (1) One or more of my immediate family members (parents, siblings, partner) has lost their job (navy bars); (2) I have lost a job (maroon bars); (3) Before COVID-19 pandemic, were you planning on doing an internship at any time between May 2020 and August 2020 (orange bars); (4) My internship got cancelled (khaki bars); (5) Conditional on having a job offer was it withdrawn or cancelled (yellow bars). The responses are stratified by country. Top panel presents results for Australia, Austria, Italy, and Mexico while bottom panel presents results for Spain, Sweden, and the United States. Sample sizes differ by question and country. These are, respectively for questions (1) to (5): for Australia 3645, 3645, 4799, 1092, 44; for Austria 320, 320, 507, 219, 17; for Italy 2435, 2435, 3886, 780, 38; for Mexico 525, 525, 593, 202, 39; for Spain 5311, 5311, 6442, 1385, 267; for Sweden 377, 377, 565, 75, 16; for the United States 15650, 15650, 19557, 8273, 594.

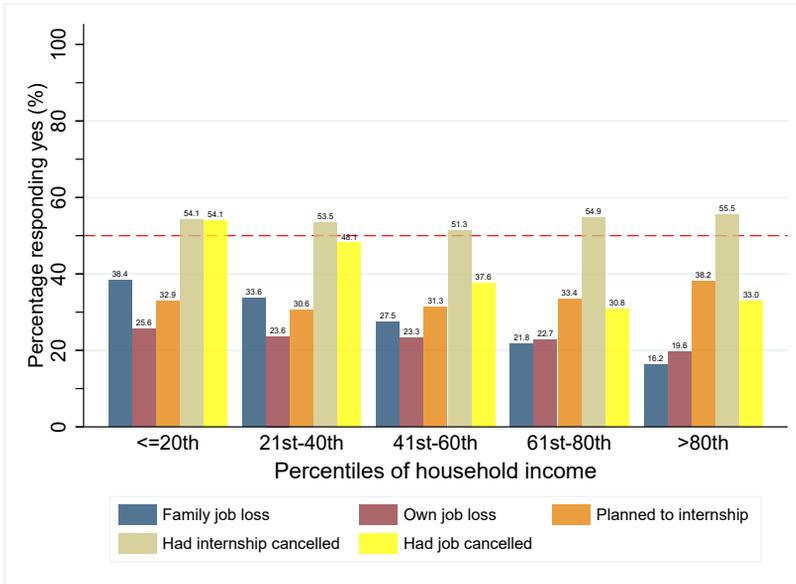
lost a job.⁵ Cancellation of internships planned for May–August 2020 was commonplace: 27% of planned internships in Italy, 34% in Austria, 35% in Mexico, 41% in Australia, 51% in Sweden, 58% in US, and 60% in Spain were cancelled.⁶ Lastly, the withdrawal or cancellation of existing job offers also occurred at high rates across the seven countries. As with summer internships, the extent of job-offer retractions was highest in Spain where cancellations reached 58%. Job offers in the US were cancelled to a lesser extent than internships (28% canceled), and students in Italy experienced the least job cancellations at 21%.

Parental income differences. Figure 3 shows that job loss events were not uniformly distributed across respondents of different socioeconomic backgrounds. Among respondents with parents in the bottom income quintile, 38% had an immediate family member experience job loss. The rate was more than halved (16%) for students with parents in the top quintile of the earnings distribution. Own job loss was also negatively related to parental

⁵The probabilities of reporting job loss do not condition on having a job at the beginning of the pandemic.

⁶Figure 2 also shows that the existence of internship plans varied across countries, from a low of 14% of respondents in Sweden to a high of 43% in Austria and US.

Figure 3: Job loss measures, by parental income



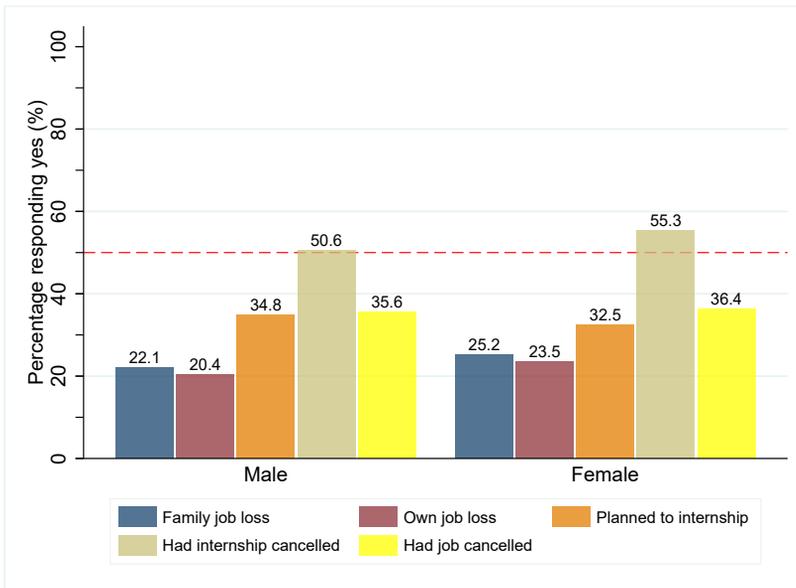
Note: This figure presents mean values of responses to the following five questions: (1) One or more of my immediate family members (parents, siblings, partner) has lost their job (navy bars); (2) I have lost a job (maroon bars); (3) Before COVID-19 pandemic, were you planning on doing an internship at any time between May 2020 and August 2020 (orange bars); (4) My internship got cancelled (khaki bars); (5) Conditional on having a job offer was it withdrawn or cancelled (yellow bars). The responses are stratified by student’s household (parents) quintile which is country-specific based on national income distribution. Sample sizes differ by question and quintile. These are, respectively for questions (1) to (5): for bottom quintile 1637, 1637, 1803, 593, 37; for 21st-40th percentile 2897, 2897, 3219, 981, 81; for 41st-60th percentile 3953, 3953, 4482, 1397, 101; for 61st-80th percentile 5327, 5327, 6063, 2020, 172; for top quintile 7808, 7808, 8861, 3365, 291. Equivalent regression analyses with and without controls are presented in panel A of Table A1.

income: 26% of bottom-quintile students experienced it, compared to 20% in the top quintile. Interestingly, internship cancellation rates are similar across parental income groups, ranging between 51% and 56% percent.⁷ In contrast, job offer withdrawals and cancellations were negatively correlated with parental income: 54% of those who had a standing job offer and parents in the bottom income quintile lost the offer, while the corresponding number was 33% for students with parents with incomes in the top quintile.

Gender differences. Figure 4 shows somewhat higher incidence of job loss measures among women. The probability of family job loss is higher for female (25%) than for male (22%) students, as well as for own job loss (24% for women and 20% for men). At the same time, internship cancellation rates were 55% for women and 51% for men. By contrast, job offer withdrawals were quite similar for women and men, with both probabilities equal to about 36%.

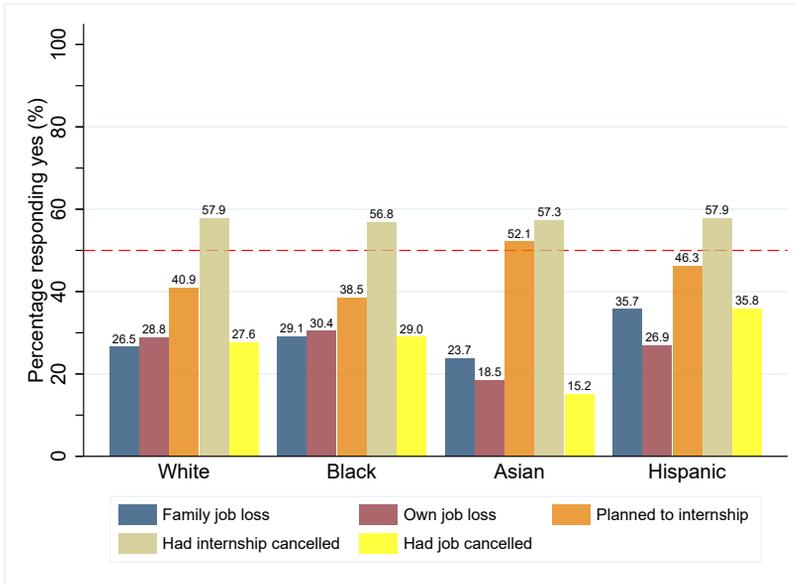
⁷Planning to do an internship to begin with was more common among top-quintile students (38% vs. 31%–33% among the other groups).

Figure 4: Job loss measures, by gender



Note: This figure presents mean values of responses to the following five questions: (1) One or more of my immediate family members (parents, siblings, partner) has lost their job (navy bars); (2) I have lost a job (maroon bars); (3) Before COVID-19 pandemic, were you planning on doing an internship at any time between May 2020 and August 2020 (orange bars); (4) My internship got cancelled (khaki bars); (5) Conditional on having a job offer was it withdrawn or cancelled (yellow bars). The responses are stratified by gender. Sample sizes differ by question and gender. These are, respectively for questions (1) to (5): for males 8123, 8123, 9582, 3322, 289; and for females 18707, 18707, 20953, 6783, 583. Equivalent regression analyses with and without controls are presented in panel B of Table A1.

Figure 5: Job loss measures, by race/ethnicity (US only)



Note: This figure presents mean values of responses to the following five questions: (1) One or more of my immediate family members (parents, siblings, partner) has lost their job (navy bars); (2) I have lost a job (maroon bars); (3) Before COVID-19 pandemic, were you planning on doing an internship at any time between May 2020 and August 2020 (orange bars); (4) My internship got cancelled (khaki bars); (5) Conditional on having a job offer was it withdrawn or cancelled (yellow bars). The responses are stratified by race/ethnicity for the United States only. Sample sizes differ by question as well as race/ethnicity. These are, respectively for questions (1) to (5): for Whites 9395, 9395, 10527, 4295, 384; for Blacks 959, 959, 1004, 387, 31; for Asians 1436, 1436, 1554, 808, 33; and for Hispanics 1933, 1933, 2006, 926, 53. Equivalent regression analyses with and without controls are presented in panel C of Table A1.

Racial differences in the US. Figure 5 documents job loss separately for Whites, Blacks, Asians, and Hispanics, among respondents from US institutions. Hispanics and Blacks experienced greater family job loss (36% and 29%, respectively) compared to Asians and Whites (24% and 27%, respectively). Blacks and Whites were the most likely to experience own job loss (30% and 29%, respectively), compared to 27% of Hispanics, and 19% of Asians. Blacks and Whites were least likely to have planned a summer internship (39% and 41%, respectively), whereas Asian and Hispanic students were considerably more likely to have planned summer internships (52% and 46%, respectively). Internship cancellation was similarly likely across groups (between 57%–58%), but job offer withdrawals disproportionately affected Hispanic respondents (36%) compared to Blacks, Whites, and Asians (29%, 28%, and 15%, respectively).

3.1.2 Student's labor market activity

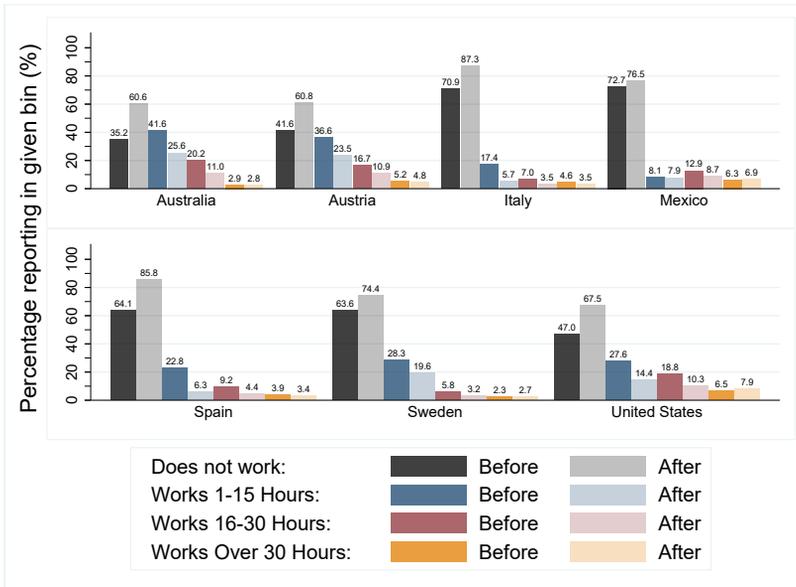
We documented extensive margin labor market responses above but how did COVID-19 affect students' current engagement with the labor market on the intensive margin? To answer this question, we document changes in the distribution of working hours, before and after the pandemic started.

Results by country. Figure 6 shows distributions of hours of work, before and after the start of the pandemic, for each of the countries in our sample. The share of students who work a positive number of hours per week varied across countries before the pandemic: 65% of respondents in Australia worked, 52% in Austria, 53% in the US, 36% in Spain and Sweden, 29% in Italy, and 27% in Mexico. Across all countries, however, we see marked increases in the fraction of students who report working zero hours after the start of the pandemic (e.g., from 47% to 68% in US, and from 64% to 86% in Spain). These increases in the share working zero hours are accompanied by substantial decreases in the fractions of those who work between 1–15 hours and 16–30 hours. The fraction working full-time (over 30 hours) remained quite similar across countries with the exception of the US (where it increased from 6.5% to 7.9%) and Italy (where it decreased from 4.6% to 3.5%).

Parental income differences. Figure 7 shows working hours by parental income quintiles. Before the pandemic, top-quintile students were least likely to work, especially 16 or more hours (14% of them did, compared to 18%–21% among the other groups). After the pandemic started, however, top-quintile students were working 16 or more hours at similar rates as the other students (11% vs. between 10%–13%). Overall, the share working zero hours increased substantially across all income quintiles, but the magnitude of the change before and after the pandemic was less extensive for the richest students.

Gender differences. Figure 8 shows that, before the pandemic, men were less likely to work than women (45% vs. 50%, respectively). After the start of the pandemic, the share not working increased for both men and women. But the extent of the increase was larger for women (16 percentage points for men, and 23 percentage points for women). Men and women worked full-time at similar rates before the pandemic (5.8% and 5.3%, respectively) but afterwards men were more likely to do so (6.7%) than women (5.4%).

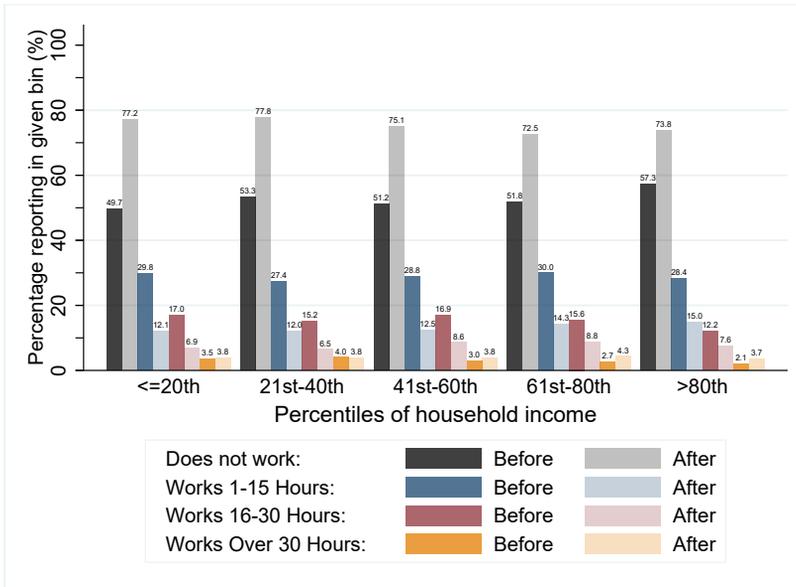
Figure 6: Student’s labor market activity before and after pandemic start, by country



Note: This figure presents mean values of responses to questions regarding employment of students prior to and since COVID-19 pandemic. The exact questions were “Before [Since] the COVID-19 pandemic, did [do] you work for pay (including work-study) while pursuing your studies?”. Respondents had multiple options including: “No, not at all”, 5-hour intervals above zero (e.g., “About 1-5 hours per week”), up to “More than 40 hours per week”. We aggregated these responses to dichotomous scale of four variables depicted in this figure. Variables are multiplied by 100 and sum to 100 within a question. Black bars represent no work, navy bars represent working between 1 and 15 hours per week, maroon bars represent working 16 to 30 hours per week, and orange bars represent working more than 30 hours per week. Solid bars are for work situation before while faded bars are for work situation after the start of COVID-19 pandemic. Sample is divided by country. Top panel presents results for Australia, Austria, Italy, and Mexico while bottom panel presents results for Spain, Sweden, and the United States. Sample sizes are 4758 for Australia, 503 for Austria, 3844 for Italy, 583 for Mexico, 6412 for Spain, 555 for Sweden, and 19359 for the United States.

Covid Economics 79, 27 May 2021: 152-217

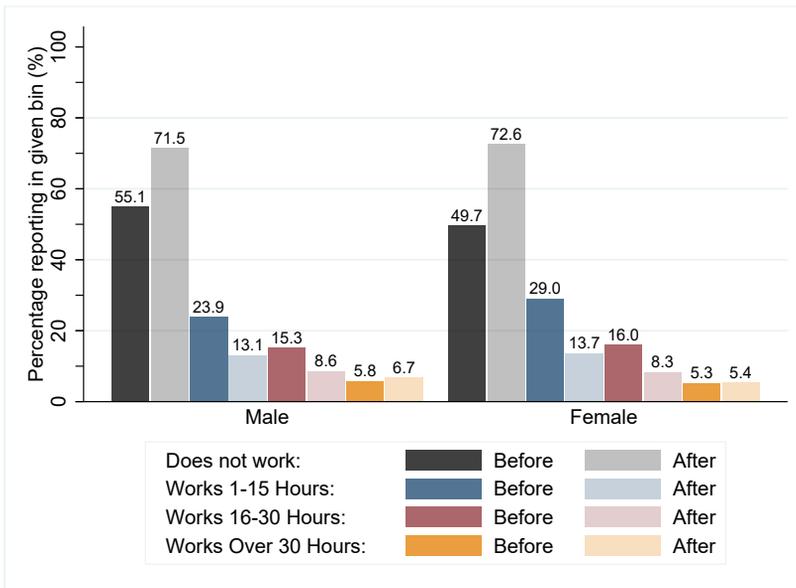
Figure 7: Student’s labor market activity before and after pandemic start, by parental income



Note: This figure presents mean values of responses to questions regarding employment of students prior to and since COVID-19 pandemic. The exact questions were “Before [Since] the COVID-19 pandemic, did [do] you work for pay (including work-study) while pursuing your studies?”. Respondents had multiple options including: “No, not at all”, 5-hour intervals above zero (e.g., “About 1-5 hours per week”), up to “More than 40 hours per week”. We aggregated these responses to dichotomous scale of four variables depicted in this figure. Variables are multiplied by 100 and sum to 100 within a question. Black bars represent no work, navy bars represent working between 1 and 15 hours per week, maroon bars represent working 16 to 30 hours per week, and orange bars represent working more than 30 hours per week. Solid bars are for work situation before while faded bars are for work situation after the start of COVID-19 pandemic. The responses are stratified by student’s household (parents) quintile which is country-specific based on national income distribution. Sample sizes are 1785 for bottom quintile, 3193 for 21st to 40th percentile, 4458 for 41st to 60th percentile, 6024 for 61st to 80th percentile, and 8789 for top quintile.

Covid Economics 79, 27 May 2021: 152-217

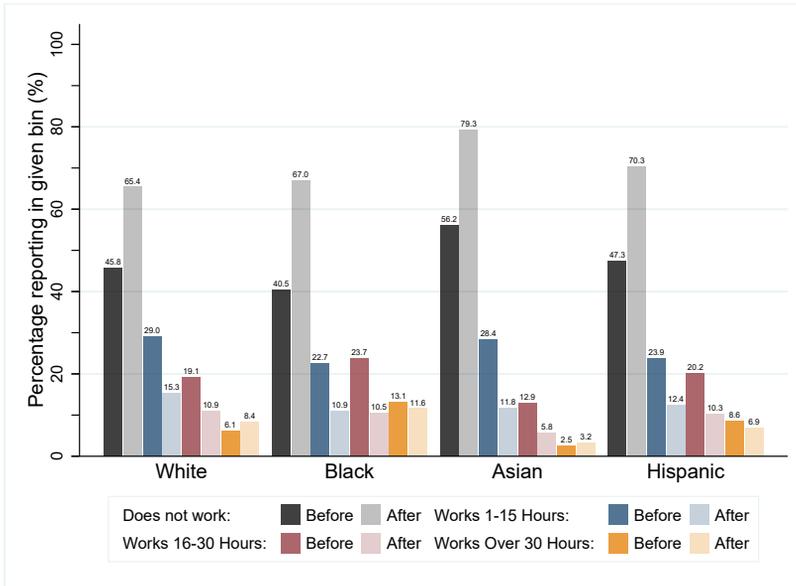
Figure 8: Student’s labor market activity before and after pandemic start, by gender



Note: This figure presents mean values of responses to questions regarding employment of students prior to and since COVID-19 pandemic. The exact questions were “Before [Since] the COVID-19 pandemic, did [do] you work for pay (including work-study) while pursuing your studies?”. Respondents had multiple options including: “No, not at all”, 5-hour intervals above zero (e.g., “About 1-5 hours per week”), up to “More than 40 hours per week”. We aggregated these responses to dichotomous scale of four variables depicted in this figure. Variables are multiplied by 100 and sum to 100 within a question. Black bars represent no work, navy bars represent working between 1 and 15 hours per week, maroon bars represent working 16 to 30 hours per week, and orange bars represent working more than 30 hours per week. Solid bars are for work situation before while faded bars are for work situation after the start of COVID-19 pandemic. The responses are stratified by student’s gender. Sample sizes are 9497 for males and 20790 for females.

Covid Economics 79, 27 May 2021: 152-217

Figure 9: Student’s labor market activity before and after pandemic start, by race/ethnicity (US only)



Note: This figure presents mean values of responses to questions regarding employment of students prior to and since COVID-19 pandemic. The exact questions were “Before [Since] the COVID-19 pandemic, did [do] you work for pay (including work-study) while pursuing your studies?”. Respondents had multiple options including: “No, not at all”, 5-hour intervals above zero (e.g., “About 1-5 hours per week”), up to “More than 40 hours per week”. We aggregated these responses to dichotomous scale of four variables depicted in this figure. Variables are multiplied by 100 and sum to 100 within a question. Black bars represent no work, navy bars represent working between 1 and 15 hours per week, maroon bars represent working 16 to 30 hours per week, and orange bars represent working more than 30 hours per week. Solid bars are for work situation before while faded bars are for work situation after the start of COVID-19 pandemic. The responses are stratified by race/ethnicity and gender for the United States only. Sample sizes are 10454, 992, 1530, and 1994, for Whites, Blacks, Asians, and Hispanics, respectively.

Racial differences in the US. Figure 9 shows changes in hours of work for US respondents by race/ethnicity. Asian students were the least likely to work any number of hours before the pandemic (44% of them did) compared to the three other groups (ranging between 53%–59%). Full-time work (over 30 hours) was quite uncommon before the pandemic for Asians (2.5%), and somewhat more common for Whites (6.1%) or Hispanics (8.6%) and especially much more common for Blacks (13.1%). After the start of the pandemic, zero hours of work substantially increased for all racial groups, and the share working 1–15 or 16–30 hours similarly decreased for all. Whites’ and Asians’ probability of working full time increased with the pandemic (from 6.1% to 8.4%, and from 2.5% to 3.2%, respectively), while for Blacks’ and Hispanics’ equivalent probabilities somewhat decreased (from 13% to 12%, and from 8.6% to 6.9%, respectively).

Covid Economics 79, 27 May 2021: 152-217

3.1.3 Career considerations

How has the pandemic affected the importance that undergraduates attach to future job and career characteristics? We now examine to what degree students consider that several (positive) career considerations have become more important as a result of the pandemic.

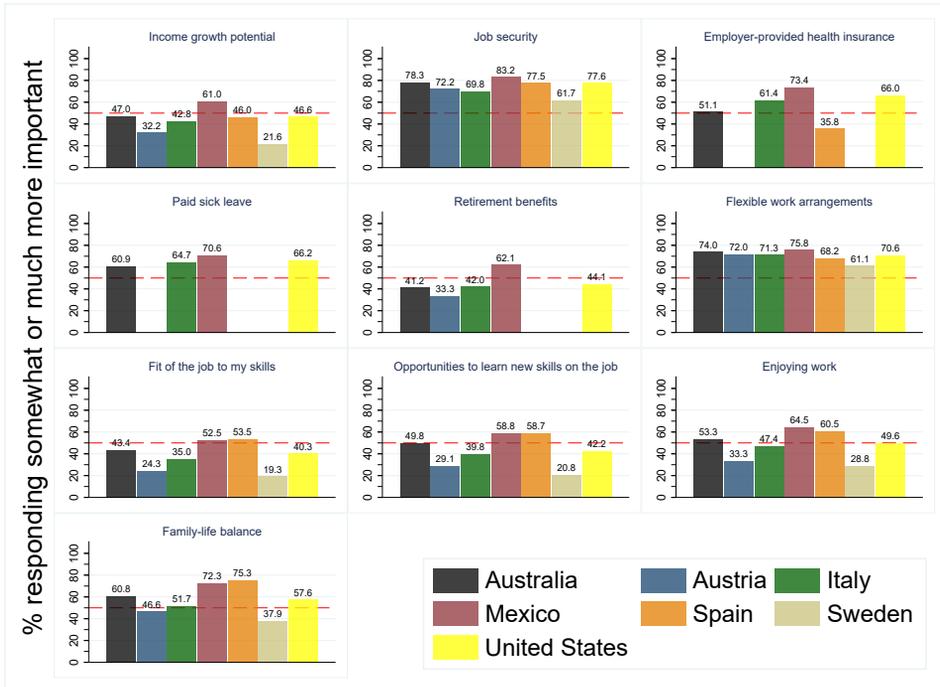
Results by country. Figure 10 plots the fraction of respondents in each country who respond that a given career consideration has become somewhat or much more important as a result of the pandemic. The job aspects that respondents feel have become particularly more important are job security and flexible work arrangements. Greater importance of job security was reported by over 60% of respondents in all countries, ranging from 62% in Sweden to 83% in Mexico. Likewise, flexible work arrangements were also reported by over 60% of respondents in all countries. Paid sick leave was also reported by large fractions of students in the countries that the survey offered as an option: Australia (61%), Italy (65%), Mexico (71%), and US (66%).

Students in Mexico and Spain reported opportunities to learn new skills on the job (59% in both cases) and the fit of the job to existing skills (53% and 54%, respectively) had become more important. In Australia and US, 40%–50% of students contend that the fit of the job to existing skills and/or opportunities to learn new skills are increasingly important. Respondents in Italy, Austria, and Sweden were the least likely to believe job fit to their skills and/or opportunities to learn new skills had become more important in response to the pandemic, with about 20%–40% reporting greater importance.

Other job aspects the survey inquired about include employer-provided health insurance, income growth potential, retirement benefits, enjoying work, and family-life balance. Figure 10 shows that respondents selecting these aspects as being more important due to the pandemic vary but, generally speaking, students in Mexico and Spain are the most likely to select these options, followed by those in the US and Australia, those in Italy, and then those in Austria and Sweden.

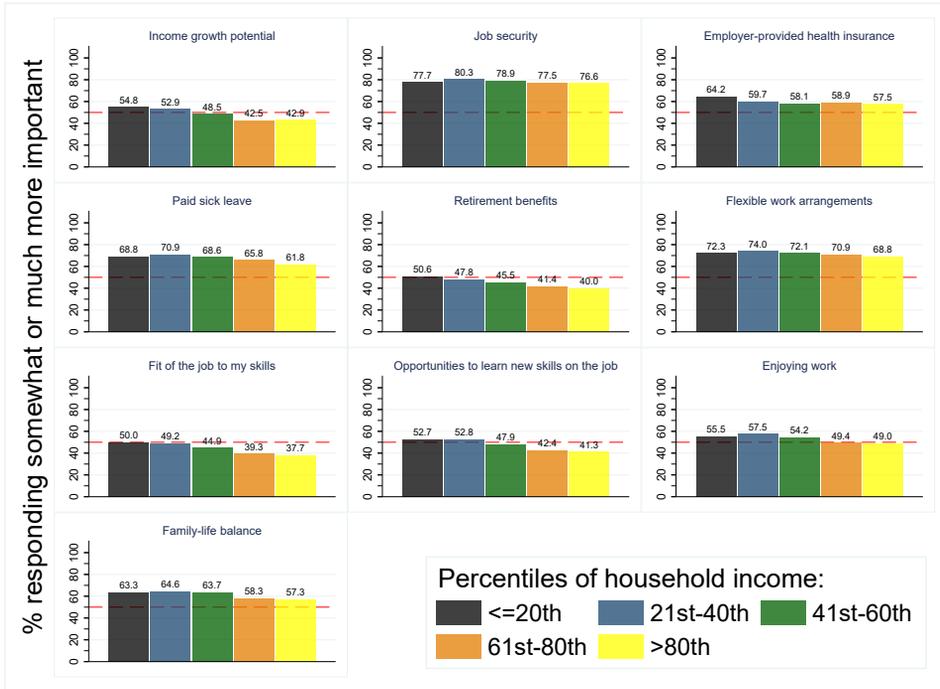
Parental income differences. Figure 11 shows a marked parental income gradient for most career considerations, with wealthier students less likely to report positive job characteristics have become more important as a result of the pandemic. Comparing students with parents in the bottom vs. top quintiles, income growth potential has become more important for 55% vs. 43%, respectively; employer-provided health insurance for 64% vs. 58%; paid sick leave for 69% vs. 62%; retirement benefits for 51% vs. 40%; flexible work arrangements for 72% vs. 69%; fit of job to skills for 50% vs. 38%; opportunities to learn new skills on the job for 53% vs. 41%; enjoying work for 56% vs. 49%; and family-life balance for 63% vs. 57%. Job security is the one characteristic that students from all income backgrounds feel similarly about, with close to 80% of all groups claiming the attribute has become more important.

Figure 10: Career considerations as a result of the pandemic, by country



Note: This figure presents mean values of responses to question regarding career considerations. The exact question was worded as "Below are some things that might be important when choosing a career. As a result of the COVID-19 pandemic, how has their importance to you changed?" with the following options: (1) income growth potential, (2) job security, (3) employer-provided health insurance (not asked in Austria and Sweden), (4) paid sick leave (not asked in Austria, Spain, and Sweden), (5) retirement benefits (not asked in Spain and Sweden), (6) flexible work arrangements (for example: working from home, telecommuting), (7) fit of the job to my skills, (8) opportunity to learn new skills on the job, (9) enjoying work, and (10) family-life balance. Each option is depicted as a separate panel and the responses are stratified by country. Black bars are for Australia, navy for Austria, green for Italy, maroon for Mexico, orange for Spain, khaki for Sweden, and yellow for the United States. Sample sizes are 4202 for Australia, 457 for Austria, 3678 for Italy, 541 for Mexico, 5543 for Spain, 514 for Sweden, and 17384 for the United States.

Figure 11: Career considerations as a result of the pandemic, by parental income



Note: This figure presents mean values of responses to question regarding career considerations. The exact question was worded as “Below are some things that might be important when choosing a career. As a result of the COVID-19 pandemic, how has their importance to you changed?” with the following options: (1) income growth potential, (2) job security, (3) employer-provided health insurance (not asked in Austria and Sweden), (4) paid sick leave (not asked in Austria, Spain, and Sweden), (5) retirement benefits (not asked in Spain and Sweden), (6) flexible work arrangements (for example: working from home, telecommuting), (7) fit of the job to my skills, (8) opportunity to learn new skills on the job, (9) enjoying work, and (10) family-life balance. Each option is depicted as a separate panel and the responses are stratified by student’s household (parents) quintile which is country-specific based on national income distribution. Black bars are for bottom quintile, navy for 21st to 40th percentile, green for 41st to 60th percentile, orange for 61st to 80th percentile, and yellow for top quintile. Sample sizes across quintiles are 1770, 3178, 4414, 5993, and 8750 for questions (1), (2), (6), (7), (8), (9), (10); they are 1734, 3118, 4338, 5897, and 8559 for question (3); they are 1481, 2271, 3324, 4917, and 7397 for question (4); and they are 1495, 2312, 3377, 4975, 7534 for question (5). Equivalent regression analyses with and without controls are presented in panel A of Tables A2 and A3.

Gender differences. Figure 12 shows that women are more likely than men to place increased importance on positive career characteristics as a result of the pandemic. Gender differences are especially pronounced for paid sick leave (70% of women vs. 55% of men), employer-provided health insurance (62% vs. 51%), family-life balance (63% vs. 53%), flexible work arrangements (74% vs. 64%), and job security (80% vs. 69%). Women were also more likely than men to assign greater importance on income growth potential (48% vs. 41%, respectively), retirement benefits (46% vs. 37%), fit of the job to existing skills (43% vs. 38%), opportunities to learn new skills (47% vs. 41%), and enjoying work (53% vs. 48%).

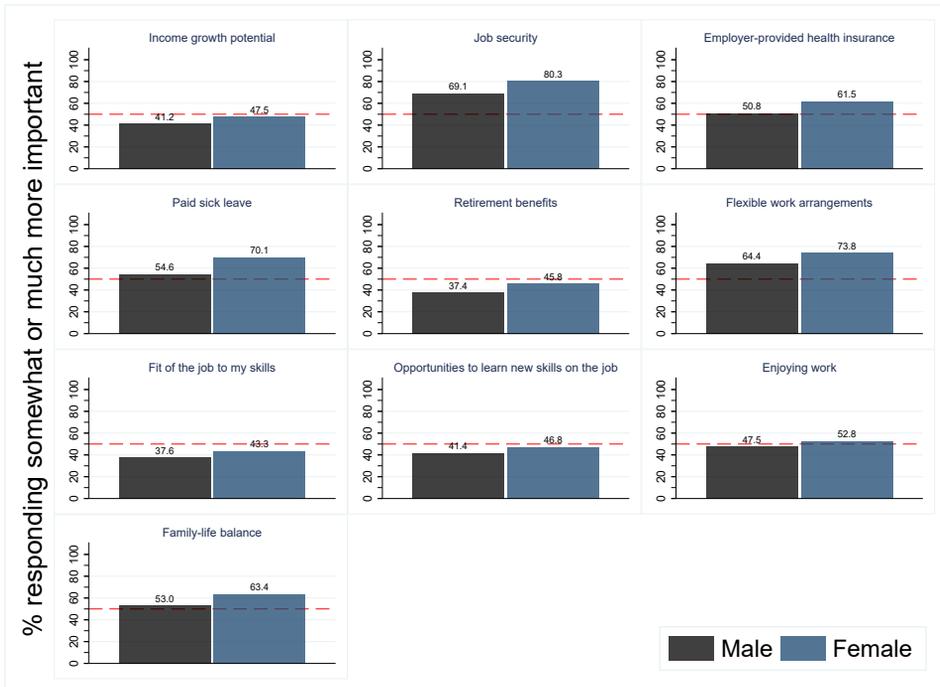
Racial differences in the US. Figure 13 shows that White students were less likely than members of the other racial/ethnic groups to assign increased importance to positive career characteristics. The greatest gap arises with income growth potential, where 42% of Whites assigned increased importance compared to 51% of Asians, 56% of Hispanics, and 62% of Blacks. Job security was assigned increased importance at high rates for all groups. Blacks (83%) and Hispanics (82%) are considerably more likely to consider job security as more important than Asians (78%) and Whites (77%), however. Employer-provided health insurance and paid sick leave were assigned increased importance by between 71%–75% of Asians, Hispanics, and Blacks, compared to 64% of Whites. The remaining career considerations—retirement benefits, flexible work arrangements, job fit to existing skills, opportunities to learn new skills, enjoying work, and family-life balance—all show similar patterns: highest assignment of increased importance among Blacks, followed by similar rates among Asians and Hispanics, and significantly lower rates for Whites.

3.1.4 Willingness to accept negative job characteristics

We now ask whether the pandemic has made students more willing to accept negative job characteristics after graduating. We consider four dimensions: working part-time, working at a job for which the student is overqualified, doing an unpaid internship, and working for the minimum wage.

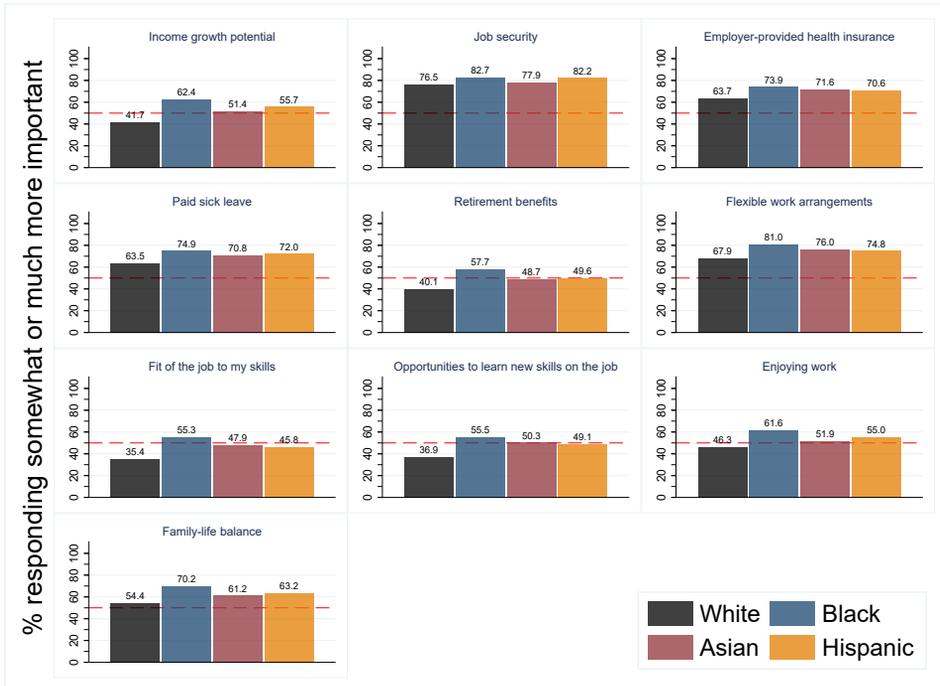
Results by country. Figure 14 shows the fraction of respondents whom the pandemic has made somewhat or much more willing to accept negative job aspects after graduation. Across all countries and all negative job aspects, less than 35% of students reported being more willing to work with such conditions. Between 21%–34% are more willing to work part-time, with the maximum fraction occurring in Mexico (34%), Spain (33%), and Australia (33%), and the minimum in Italy (21%). The fractions being more willing to be overqualified are somewhat similar, with students in Spain, Australia, and US being the most willing (34%, 32%, and 30%, respectively). Students were generally not willing to hold an unpaid internship after graduation (9% in Austria and up to 24% in Australia) or work for minimum wages (4% in Mexico and up to 20% in Spain).

Figure 12: Career considerations as a result of the pandemic, by gender



Note: This figure presents mean values of responses to question regarding career considerations. The exact question was worded as "Below are some things that might be important when choosing a career. As a result of the COVID-19 pandemic, how has their importance to you changed?" with the following options: (1) income growth potential, (2) job security, (3) employer-provided health insurance (not asked in Austria and Sweden), (4) paid sick leave (not asked in Austria, Spain, and Sweden), (5) retirement benefits (not asked in Spain and Sweden), (6) flexible work arrangements (for example: working from home, telecommuting), (7) fit of the job to my skills, (8) opportunity to learn new skills on the job, (9) enjoying work, and (10) family-life balance. Each option is depicted as a separate panel and the responses are stratified by gender. Black bars are for males while navy bars are for females. Sample sizes for males and females respectively are 9412 and 20579 for questions (1), (2), (6), (7), (8), (9), (10); they are 9085 and 19980 for question (3), they are 7565 and 16347 for question (4); and they are 7723 and 16628 for question (5). Equivalent regression analyses with and without controls are presented in panel B of Tables A2 and A3.

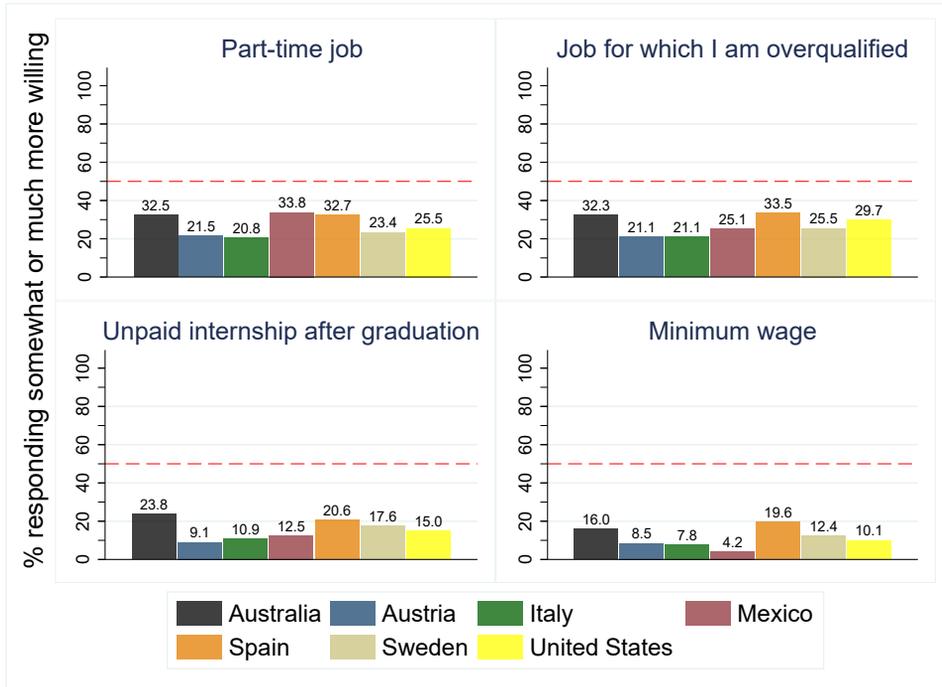
Figure 13: Career considerations as a result of the pandemic, by race/ethnicity (US only)



Note: This figure presents mean values of responses to question regarding career considerations. The exact question was worded as "Below are some things that might be important when choosing a career. As a result of the COVID-19 pandemic, how has their importance to you changed?" with the following options: (1) income growth potential, (2) job security, (3) employer-provided health insurance (not asked in Austria and Sweden), (4) paid sick leave (not asked in Austria, Spain, and Sweden), (5) retirement benefits (not asked in Spain and Sweden), (6) flexible work arrangements (for example: working from home, telecommuting), (7) fit of the job to my skills, (8) opportunity to learn new skills on the job, (9) enjoying work, and (10) family-life balance. Each option is depicted as a separate panel and the responses are stratified by race/ethnicity for the United States only. Black bars are for Whites, navy bars are for Blacks, maroon bars are for Asians, and orange bars are for Hispanics. Sample sizes for Whites, Blacks, Asians, and Hispanics respectively are 10407, 979, 1522 and 1971. Equivalent regression analyses with and without controls are presented in panel C of Tables A2 and A3.

Covid Economics 79, 27 May 2021: 152-217

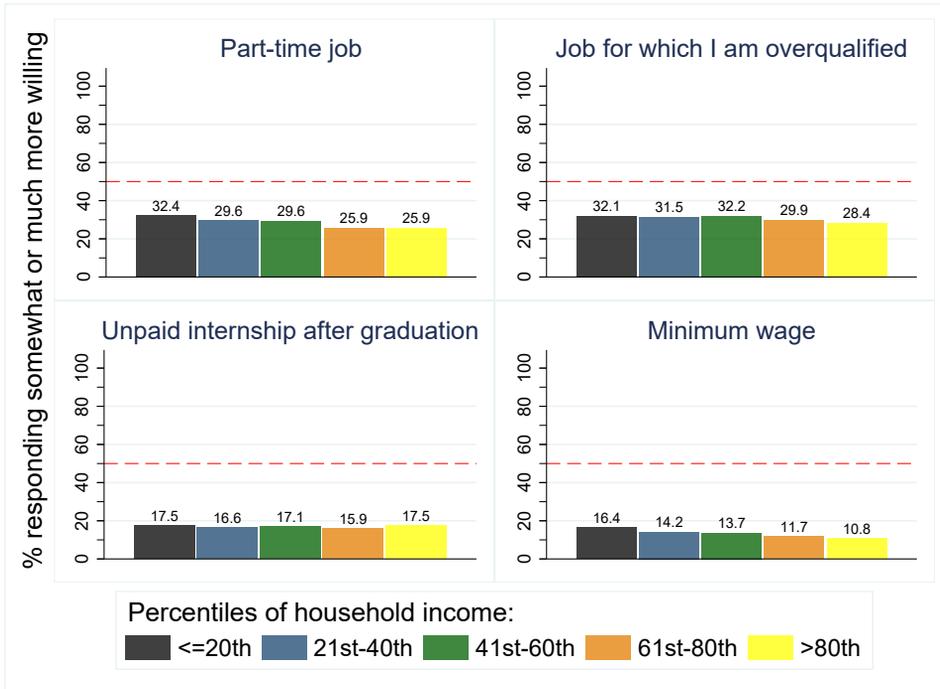
Figure 14: Greater willingness to accept negative job characteristics as a result of the pandemic, by country



Note: This figure presents mean values of responses to question regarding career compromises. The exact question was worded as “Think about the job market in the first two years after you complete your current degree. As a result of the COVID-19 pandemic, how has your willingness to work in jobs with the characteristics listed below changed?” with the following options: (1) work in a part-time job, (2) work in a job for which I am overqualified, (3) take an unpaid internship after graduation, and (4) work for minimum wage. Each option is depicted as a separate panel and the responses are stratified by country. Black bars are for Australia, navy for Austria, green for Italy, maroon for Mexico, orange for Spain, khaki for Sweden, and yellow for the United States. Sample sizes are 4223 for Australia, 460 for Austria, 3677 for Italy, 542 for Mexico, 5545 for Spain, 518 for Sweden, and 17420 for the United States.

Covid Economics 79, 27 May 2021: 152-217

Figure 15: Greater willingness to accept negative job characteristics as a result of the pandemic, by parental income

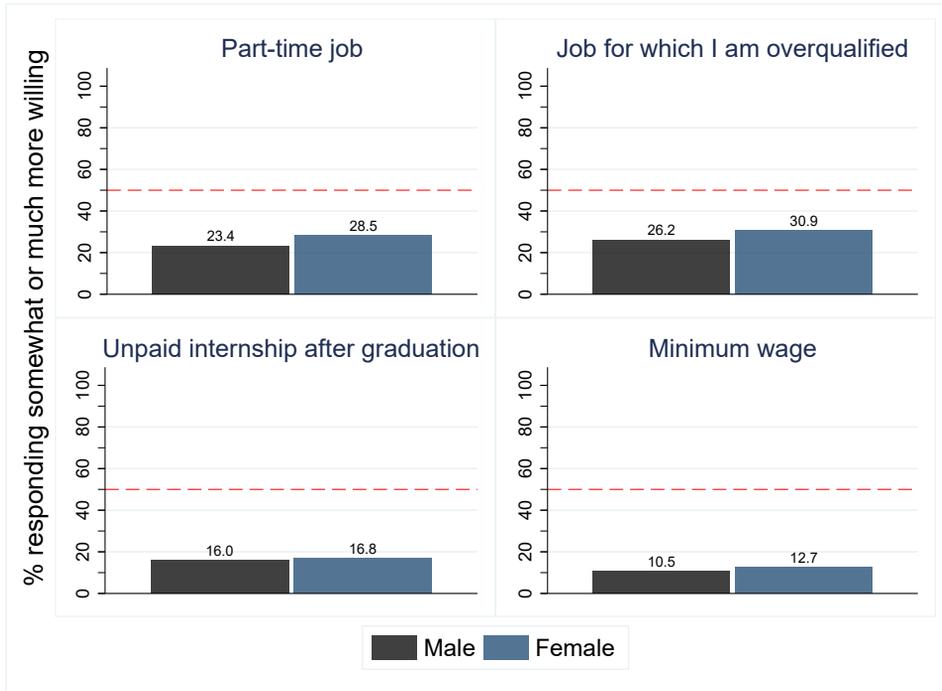


Note: This figure presents mean values of responses to question regarding career compromises. The exact question was worded as “Think about the job market in the first two years after you complete your current degree. As a result of the COVID-19 pandemic, how has your willingness to work in jobs with the characteristics listed below changed?” with the following options: (1) work in a part-time job, (2) work in a job for which I am overqualified, (3) take an unpaid internship after graduation, and (4) work for minimum wage. Each option is depicted as a separate panel and the responses are stratified by student’s household (parents) quintile which is country-specific based on national income distribution. Black bars are for bottom quintile, navy for 21st to 40th percentile, green for 41st to 60th percentile, orange for 61st to 80th percentile, and yellow for top quintile. Sample sizes across quintiles are 1769, 3184, 4440, 6013, and 8796. Equivalent regression analyses with and without controls are presented in panel A of Table A4.

Parental income differences. Figure 15 shows that students with wealthier parents are less likely to report that the pandemic has made them more willing to accept negative job characteristics. Comparing students with parents in the bottom vs. top quintiles, the fraction who have become more willing to have a part-time job is 32% vs. 26%, respectively; a job for which they are overqualified is 32% vs. 28%; and work for minimum wage is 16% vs. 11%. Having an unpaid internship after graduation is the one characteristic that students from all backgrounds feel similarly about, with between 16%–18% of all groups saying they have become more willing to take such a position.

Covid Economics 79, 27 May 2021: 152-217

Figure 16: Greater willingness to accept negative job characteristics as a result of the pandemic, by gender



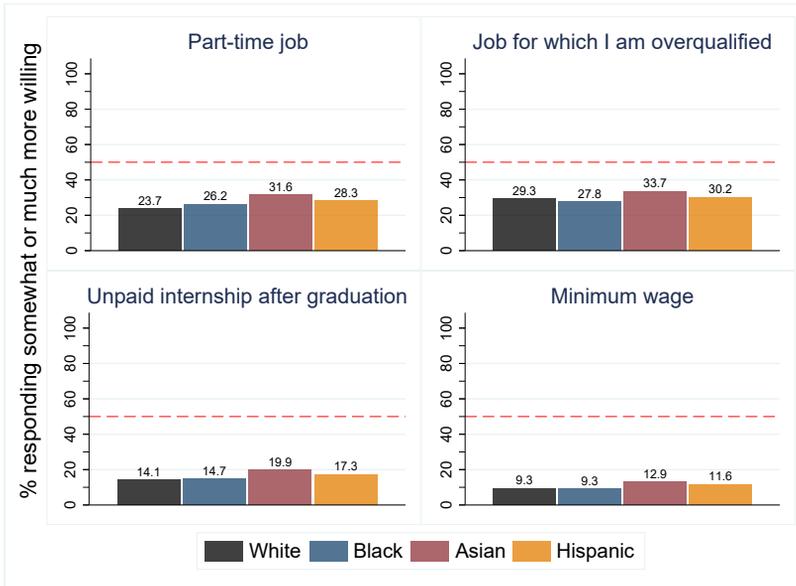
Note: This figure presents mean values of responses to question regarding career compromises. The exact question was worded as “Think about the job market in the first two years after you complete your current degree. As a result of the COVID-19 pandemic, how has your willingness to work in jobs with the characteristics listed below changed?” with the following options: (1) work in a part-time job, (2) work in a job for which I am overqualified, (3) take an unpaid internship after graduation, and (4) work for minimum wage. Each option is depicted as a separate panel and the responses are stratified by gender. Black bars are for males while navy bars are for females. Sample sizes for males and females respectively are 9463 and 20634. Equivalent regression analyses with and without controls are presented in panel B of Table A4.

Gender differences. Figure 16 shows that women are slightly more willing to accept negative job characteristics after graduation as a result of the pandemic. This is true for part-time jobs (29% of women vs. 23% of men), being overqualified (31% vs. 26%), doing an unpaid internship (17% vs. 16%), and working for minimum wage (13% vs. 11%).

Racial differences in the US. Figure 17 shows how changes in the willingness to accept negative job characteristics after graduation as a result of the pandemic vary by race/ethnicity in the US. Asians and Hispanics are the most willing to accept negative characteristics, while Blacks and Whites are the least willing to accept negative job characteristics. Among Asians, 32% would be more willing to work part-time, 34% would be more willing to be overqualified, 20% to do an unpaid internship, and 13% to work for minimum wage. Among Hispanics, 28% would be more willing to work part-time, 30% to be overqualified, 17% to do an unpaid internship, and 12% to work for minimum wage. Among Blacks, 26% would be

Covid Economics 79, 27 May 2021: 152-217

Figure 17: Greater willingness to accept negative job characteristics as a result of the pandemic, by race/ethnicity (US only)



Note: This figure presents mean values of responses to question regarding career compromises. The exact question was worded as “Think about the job market in the first two years after you complete your current degree. As a result of the COVID-19 pandemic, how has your willingness to work in jobs with the characteristics listed below changed?” with the following options: (1) work in a part-time job, (2) work in a job for which I am overqualified, (3) take an unpaid internship after graduation, and (4) work for minimum wage. Each option is depicted as a separate panel and the responses are stratified by race/ethnicity for the United States only. Black bars are for Whites, navy bars are for Blacks, maroon bars are for Asians, and orange bars are for Hispanics. Sample sizes for Whites, Blacks, Asians, and Hispanics respectively are 10431, 981, 1530 and 1987. Equivalent regression analyses with and without controls are presented in panel C of Table A4.

more willing to work part-time, 28% to be overqualified, 15% to do an unpaid internship, and 9% to work for minimum wage. Lastly, among Whites, 24% would be more willing to work part-time, 29% to be overqualified, 14% to do an unpaid internship, and 9% to work for minimum wage.

3.1.5 Earnings expectations at ages 30 and 45

What are students’ earnings expectations in the long term, and how do they vary across groups? Survey respondents were presented with the contemporaneous average earnings of 30- and 45-year-olds in their country who hold a college degree, and then asked about their expected earnings at those same ages. We summarize this information by showing the share of respondents who reported earnings expectations that are greater (in real terms) than the average presented to them.

Results by country. Figure 18 shows how the fraction who reported higher-than-average expected earnings varies across countries. Almost all respondents in Mexico reported expected earnings greater than the Mexican college average for both ages (98%). In contrast, only 38% and 33% of students in Spain expected earnings greater than the average at 30 and 45, respectively. Like in Spain, a common theme across countries is that a greater share of students expected higher-than-average earnings at 30 compared to 45: 48% and 44% in Australia, 62% and 44% in Italy, 71% and 64% in Sweden, and 65% and 60% in US. The exception is Austria, where 43% and 57% expect greater-than-average earnings at ages 30 and 45, respectively.

Parental income differences. Figure 19 shows earnings expectations at ages 30 and 45, by parental income quintile. Compared to the bottom three quintiles, students with parents in the top two quintiles are more likely to expect greater-than-average earnings at 30 and 45. Between 49%–50% of students in the bottom three quintiles expect greater-than-average earnings at 30, while this number is equal to 57% of students with parents in the fourth quintile, and 72% with parents in the fifth quintile. The levels for greater-than-average earnings at age 45 are somewhat lower, but the relative patterns are similar.

Gender differences. Figure 20 shows men have much higher earnings expectations than women: 67% of men and 54% of women expect greater-than-average earnings at age 30, and 64% of men and 46% of women do so at age 45.

Racial differences in the US. Figure 21 shows that, in the US, Asian students have the greatest earnings expectations. At age 30, 64% of Whites, 65% of Blacks, 67% of Hispanics, and 72% of Asians expect greater-than-average earnings. At age 45, the corresponding fractions for Whites, Blacks, Hispanics, and Asians are, respectively, 58%, 63%, 63%, and 70%.

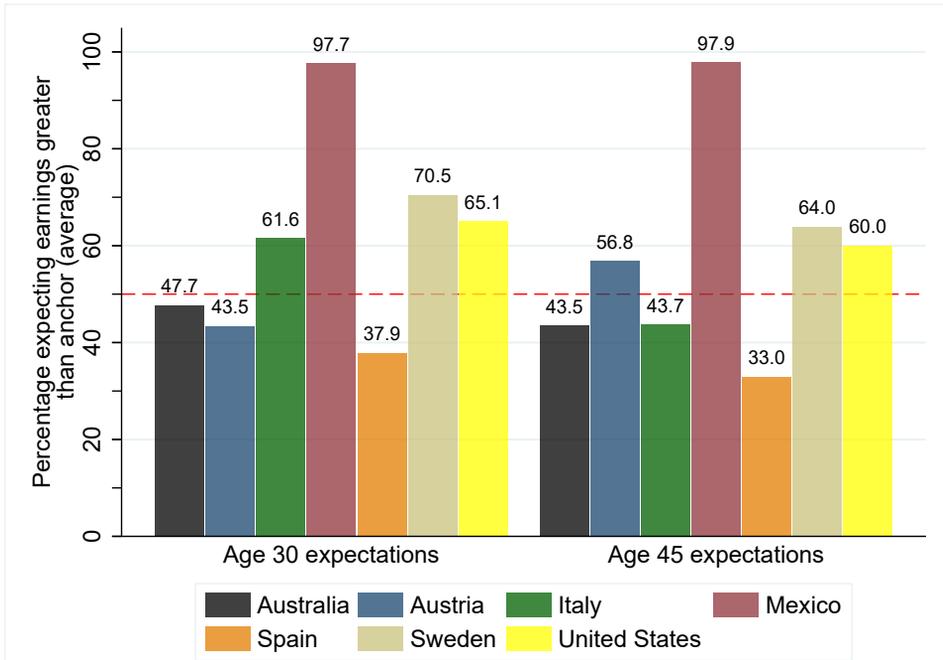
3.1.6 Discussion of labor market outcomes

Overall, our findings document that the labor market outcomes and future prospects of university students across the world have been adversely affected by the pandemic. In the seven countries in our sample, students have experienced own and family job loss at high rates, as well as reduced internship opportunities, and cancelled job offers. These events will likely hurt students in long-lasting ways (von Wachter, 2020).

While pervasive, the damaging effects of the pandemic have disproportionately affected students who already in normal times face greater disadvantage and barriers in the labor market: students from lower-income backgrounds, female students, and students belonging to racial minorities. Our results show that these groups of students were particularly more likely to experience job loss in their family, and, in most cases, also more likely to experience job loss themselves (both current jobs and canceled job offers).

Concurrently, the pandemic has increased the importance that low-income, female, and minority students place on positive future job characteristics, as well as the willingness to

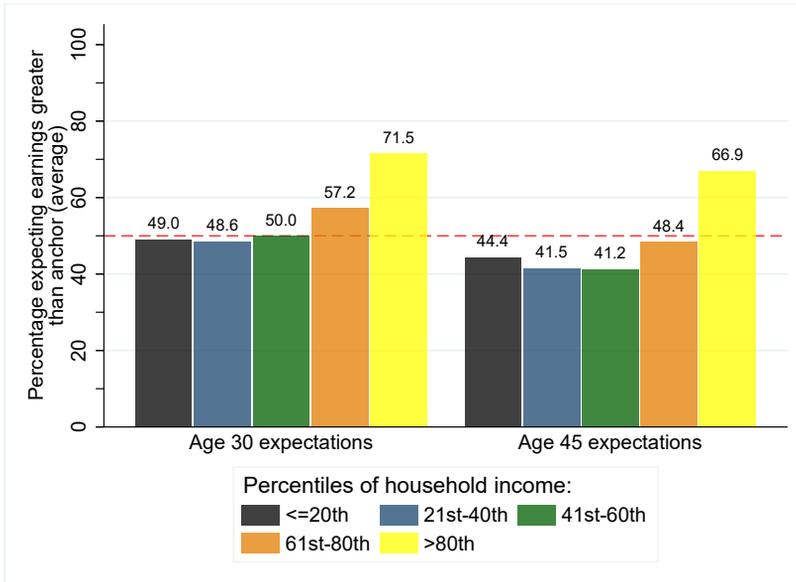
Figure 18: Earnings expectations at ages 30 and 45 (=1 if greater than current average), by country



Note: This figure presents mean values of responses to questions regarding earnings expectations. The exact questions were worded as "In 2019, the average annual earnings of a working 30 [45] year old with at least a Bachelor's degree was about \$60,000 [\$93,000]. What do you expect your earnings will be at age 30 [45]. Assume that there is no inflation between now and when you are 30 [45] and take into account any additional education you may obtain." and we discretize by generating an indicator variable that takes value of 1 if the ratio of individual answer and referenced average values is greater than 1. We multiply the indicator by 100. Left-hand side set of bars presents these values for age 30 while right-hand side set of bars presents these values for age 45. Black bars are for Australia, navy for Austria, green for Italy, maroon for Mexico, orange for Spain, khaki for Sweden, and yellow for the United States. Sample sizes are 3584 for Australia, 359 for Austria, 3121 for Italy, 524 for Mexico, 5093 for Spain, 430 for Sweden, and 15169 for the United States. Average values at age 30 [45] are USD 60,000 [USD 93,000] for the United States, AUD 85,000 [AUD 132,000] for Australia, EUR 44,181 [EUR 68,736] for Austria, SEK 370,000 [SEK 543 000] for Sweden, EUR 1,300 [EUR 2,200] for Italy (monthly reference), EUR 2,200 [EUR 3,400] for Spain (monthly reference), and MXN 13,000 [MXN 18,000] for Mexico (monthly reference). All reference values are for 2019 except for Mexico, Italy, and Spain where they are from 2018, 2016 and 2014, respectively, the last years for which publicly available data is accessible.

Covid Economics 79, 27 May 2021: 152-217

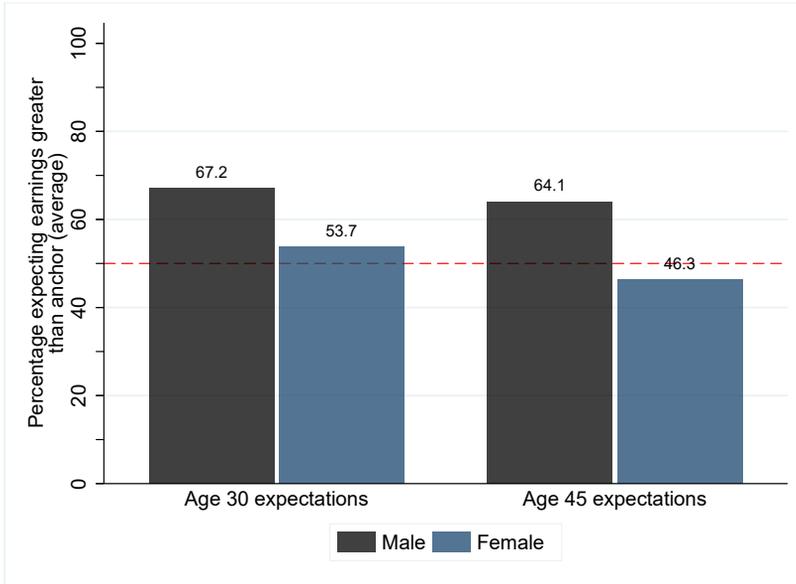
Figure 19: Earnings expectations at ages 30 and 45 (=1 if greater than current average), by parental income



Note: This figure presents mean values of responses to questions regarding earnings expectations. The exact questions were worded as “In 2019, the average annual earnings of a working 30 [45] year old with at least a Bachelor’s degree was about \$60,000 [\$93,000]. What do you expect your earnings will be at age 30 [45]. Assume that there is no inflation between now and when you are 30 [45] and take into account any additional education you may obtain.” and we discretize by generating an indicator variable that takes value of 1 if the ratio of individual answer and referenced average values is greater than 1. We multiply the indicator by 100. Left-hand side set of bars presents these values for age 30 while right-hand side set of bars presents these values for age 45. The responses are stratified by student’s household (parents) quintile which is country-specific based on national income distribution. Black bars are for bottom quintile, navy for 21st to 40th percentile, green for 41st to 60th percentile, orange for 61st to 80th percentile, and yellow for top quintile. Respective sample sizes are 1593, 2973, 4109, 5667, and 8155. Average values at age 30 [45] are USD 60,000 [USD 93,000] for the United States, AUD 85,000 [AUD 132,000] for Australia, EUR 44,181 [EUR 68,736] for Austria, SEK 370,000 [SEK 543 000] for Sweden, EUR 1,300 [EUR 2,200] for Italy (monthly reference), EUR 2,200 [EUR 3,400] for Spain (monthly reference), and MXN 13,000 [MXN 18,000] for Mexico (monthly reference). All reference values are for 2019 except for Mexico, Italy, and Spain where they are from 2018, 2016 and 2014, respectively, the last years for which publicly available data is accessible. Equivalent regression analyses with and without controls are presented in panel A of Table A5.

Covid Economics 79, 27 May 2021: 152-217

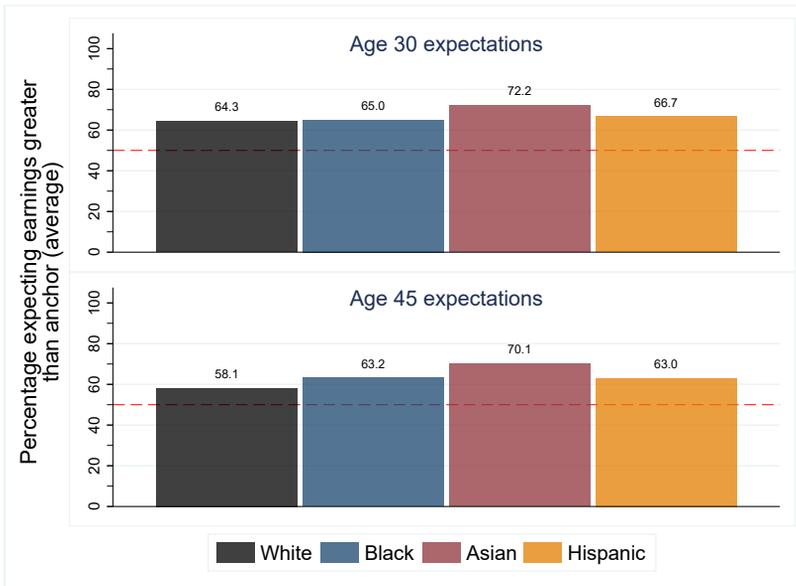
Figure 20: Earnings expectations at ages 30 and 45 (=1 if greater than current average), by gender



Note: This figure presents mean values of responses to questions regarding earnings expectations. The exact questions were worded as “In 2019, the average annual earnings of a working 30 [45] year old with at least a Bachelor’s degree was about \$60,000 [\$93,000]. What do you expect your earnings will be at age 30 [45]. Assume that there is no inflation between now and when you are 30 [45] and take into account any additional education you may obtain.” and we discretize by generating an indicator variable that takes value of 1 if the ratio of individual answer and referenced average values is greater than 1. We multiply the indicator by 100. Left-hand side set of bars presents these values for age 30 while right-hand side set of bars presents these values for age 45. Black bars are for males while navy are for females. Respective sample sizes are 8695 and 18001. Average values at age 30 [45] are USD 60,000 [USD 93,000] for the United States, AUD 85,000 [AUD 132,000] for Australia, EUR 44,181 [EUR 68,736] for Austria, SEK 370,000 [SEK 543 000] for Sweden, EUR 1,300 [EUR 2,200] for Italy (monthly reference), EUR 2,200 [EUR 3,400] for Spain (monthly reference), and MXN 13,000 [MXN 18,000] for Mexico (monthly reference). All reference values are for 2019 except for Mexico, Italy, and Spain where they are from 2018, 2016 and 2014, respectively, the last years for which publicly available data is accessible. Equivalent regression analyses with and without controls are presented in panel B of Table A5.

Covid Economics 79, 27 May 2021: 152-217

Figure 21: Earnings expectations at ages 30 and 45 (=1 if greater than current average), by race/ethnicity (US only)



Note: This figure presents mean values of responses to questions regarding earnings expectations. The exact questions were worded as “In 2019, the average annual earnings of a working 30 [45] year old with at least a Bachelor’s degree was about \$60,000 [\$93,000]. What do you expect your earnings will be at age 30 [45]. Assume that there is no inflation between now and when you are 30 [45] and take into account any additional education you may obtain.” and we discretize by generating an indicator variable that takes value of 1 if the ratio of individual answer and referenced average values is greater than 1. We multiply the indicator by 100. The responses are stratified by race/ethnicity and gender for the United States only. Top panel presents age 30 while bottom panel age 45 expectations. Left-hand side set of bars presents these values for males while right-hand side set of bars presents values for females. Black bars are for Whites, navy for Blacks, green for Asians, and orange for Hispanics. Sample sizes are 9368, 869, 1289, 1750 for Whites, Blacks, Asians, and Hispanics, respectively. Equivalent regression analyses with and without controls are presented in panel C of Table A5.

accept negative ones. Two hypotheses help explain why low-income, female, and minority students place more importance on good job characteristics as a result of the pandemic. First, their households generally being hit harder by the pandemic (Figures 3, 4, and 5) might lead these students to have more pessimistic expectations about their future labor market.⁸ More pessimistic expectations could in turn lead to lower faith in future on-the-job search outcomes and a greater value placed on landing a good job right away. Such higher importance could be aspirational, even if the actual chances of finding such a good job are diminished by the pandemic. A second potential explanation could arise even if all students have similar expectations about the future labor market, through students from wealthier and more advantaged backgrounds having stronger family safety nets and professional networks. Such insurance against income loss, job loss, or health shocks, might make students from more advantaged backgrounds less reliant on good job attributes. Note that the results in Figures 15, 16, and 17—showing that poorer, female, and minority students are more willing to accept negative job characteristics—are not consistent, however, with explanations in which students from less advantaged backgrounds have become more demanding of positive work conditions. While aspirations to land good jobs are stronger for these students, they are also more willing to work under negative conditions if necessary.

3.2 Educational outcomes

Educational prospects of students have also been affected as a result of the COVID-19 pandemic. In this section, we analyze the particular effects across several dimensions including: educational consequences and challenges, future schooling plans, and changes in studying characteristics due to the pandemic.

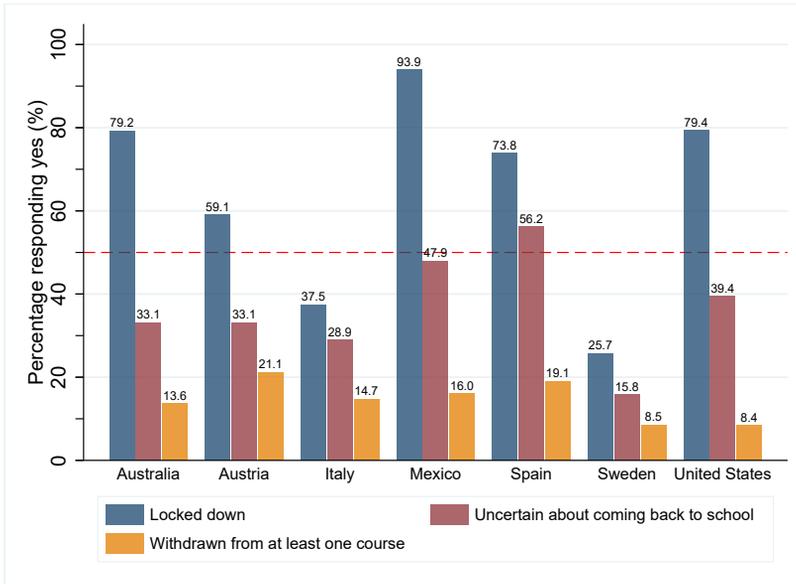
3.2.1 Educational consequences

Results by country. Figure 22 shows differences in educational consequences of the pandemic across countries. Most students in all countries, except in Sweden, were under lockdown measures with certain degree of variation in the percentage of affected students. For example, in Mexico, almost a 94% of students were affected while in Austria only 59% of students. In Sweden, only a 25.7% of students were under lockdown and a 37.5% of students in Italy.⁹ When it comes to the possibility of not returning to the current university in Fall 2020, 56% of students in Spain and a and a 48% of students in Mexico respond affirmatively, whereas only a 15.8% of students in Sweden considered this possibility. In the remaining countries (Australia, Austria and United States), about a third of the students

⁸Kuchler and Zafar (2019) show that personally experiencing unemployment affects individuals' expectations about the aggregate unemployment rate, arguing that this is consistent with "naive extrapolation." Roth and Wohlfart (2020) show that when expectations are manipulated, people extrapolate recession expectations to personal economic expectations, and those who do more so are people who are more exposed to macroeconomic risk.

⁹This relatively low number for Italy is due to the fact that the survey in the one Italian university in our sample launched in late Summer 2020 and, as opposed to other questions that specifically ask about the Spring semester, the question on lockdown referred to "right now."

Figure 22: Education disruptions, by country



Note: This figure presents mean values of responses to the following three questions/statements: (1) I am “locked down”, “quarantined”, “staying home”, or “sheltering in place” (navy bars); (2) Is it possible that the COVID-19 pandemic might lead to your not returning to your current university in Fall 2020 (maroon bars); and (3) Have you withdrawn from any of your courses since the COVID-19 pandemic? (orange bars). The responses are stratified by country. Sample sizes differ by question and country. These are, respectively for questions (1) to (3): for Australia 3645, 3986, 4799; for Austria 320, 405, 506; for Italy 2435, 2919, 3868; for Mexico 525, 501, 593; for Spain 5311, 4657, 6454; for Sweden 377, 469, 565; for the United States 15650, 16750, 19630.

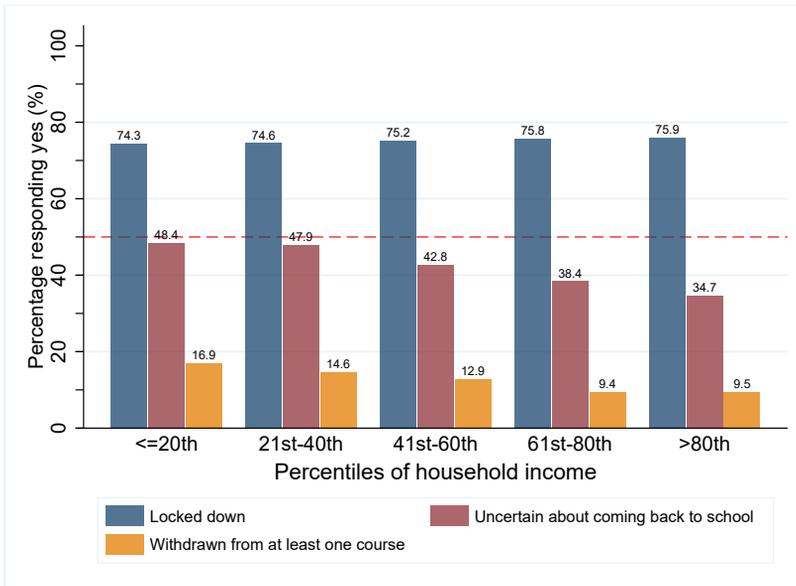
thought about not returning to the universities in which they were enrolled at the onset of the pandemic. Finally, a relatively smaller fraction of students, between 21% in Austria and a 8% in United States, have withdrawn from at least one course since the start of the pandemic.

Parental income differences. Figure 23 shows differences in educational consequences by parental income. There are practically no differences in the percentage of students under lockdown measures. On the other hand, there are large differences when it comes to the uncertainty of coming back to school in Fall 2020. Around 50% of students in bottom quintiles are uncertain, whereas only a 35% of students in the top quintile report uncertainty regarding return to their pre-pandemic university. Similarly, students from the top two quintiles are less affected (around a 9%) by having withdrawn from any course, whereas in the bottom two quintiles the percentage of students is higher (more than 14%).

Gender differences. Figure 24 shows differences in educational consequences by gender. In short we do not find any striking differences in these measures by student’s gender. As expected, females and males seem to be equally affected by lockdown measures (74.4% of

Covid Economics 79, 27 May 2021: 152-217

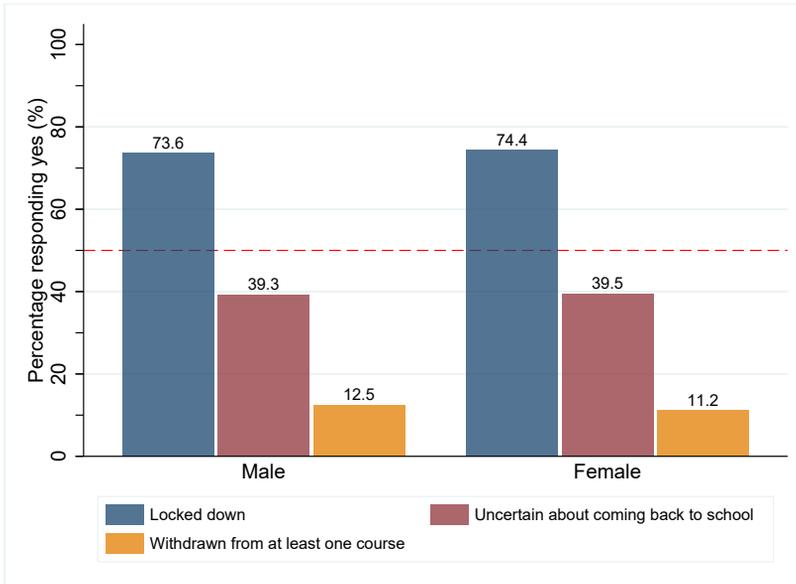
Figure 23: Education disruptions, by parental income



Note: This figure presents mean values of responses to the following three questions/statements: (1) I am "locked down", "quarantined", "staying home", or "sheltering in place" (navy bars); (2) Is it possible that the COVID-19 pandemic might lead to your not returning to your current university in Fall 2020 (maroon bars); and (3) Have you withdrawn from any of your courses since the COVID-19 pandemic? (orange bars). The responses are stratified by student's household (parents) quintile which is country-specific based on national income distribution. Sample sizes differ by question and quintile. These are, respectively for questions (1) to (3): for bottom quintile 1637, 1472, 1802; for 21st-40th percentile 2897, 2584, 3223; for 41st-60th percentile 3953, 3687, 4479; for 61st-80th percentile 5327, 5025, 6055; for top quintile 7808, 7487, 8861. Equivalent regression analyses with and without controls are presented in panel A of Table A6.

Covid Economics 79, 27 May 2021: 152-217

Figure 24: Education disruptions, by gender

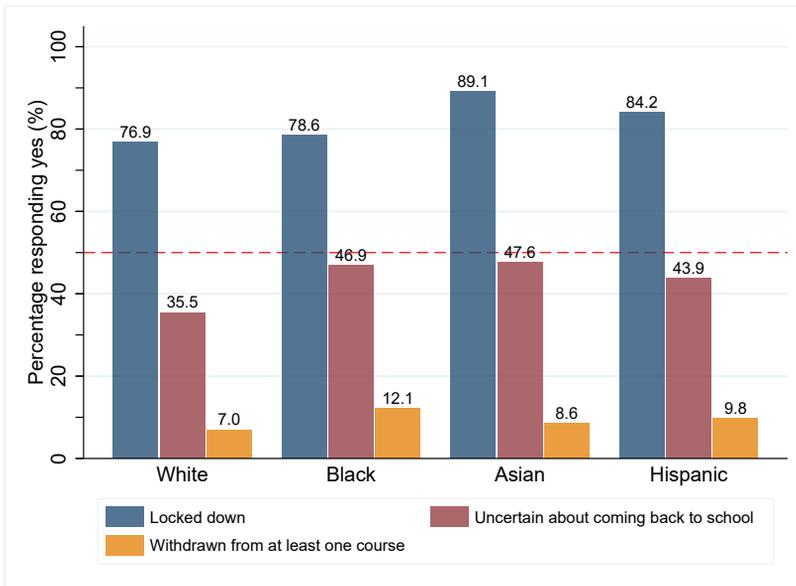


Note: This figure presents mean values of responses to the following three questions/statements: (1) I am “locked down”, “quarantined”, “staying home”, or “sheltering in place” (navy bars); (2) Is it possible that the COVID-19 pandemic might lead to your not returning to your current university in Fall 2020 (maroon bars); and (3) Have you withdrawn from any of your courses since the COVID-19 pandemic? (orange bars). The responses are stratified by gender. Sample sizes differ by question and gender. These are, respectively for questions (1) to (3): for males 8123, 7985, 9582; and for females 18707, 17156, 20946. Equivalent regression analyses with and without controls are presented in panel B of Table A6.

females and 73.6% of males), but they also have similar probabilities of having withdrawn from courses (12.5% of males and 11.2% of females) and being uncertain about coming back to school (39.5% of females and 39.3% of males).

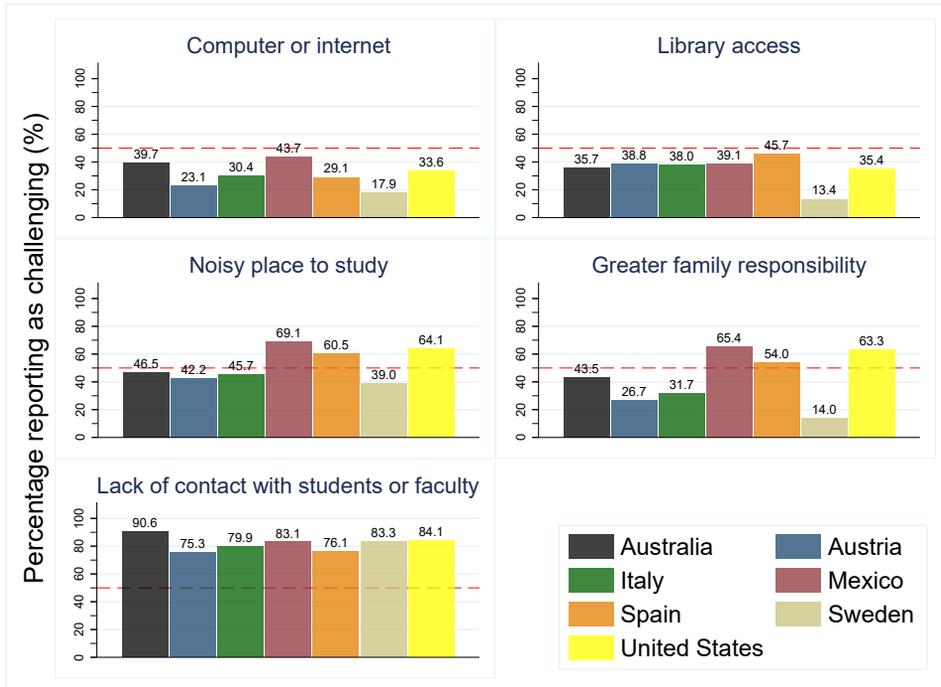
Racial differences in the US. Figure 25 shows racial differences in educational consequences in the United States. Asian students have been affected in a higher proportion by lockdown measures (89%) and uncertainty about coming back in Fall 2020 (47.6%). This could be related to the fact that states with larger Asian populations, e.g. California, imposed stricter social distancing measures at the beginning of the pandemic. White students are the least likely to report uncertainty about returning to school (around 36%). Black students had the highest propensity to have withdrawn from at least one course (12.1%). Around 9-10% of Asian and Hispanic students had withdrawn from at least one course, and White were the least likely to drop a course (7%).

Figure 25: Education disruptions by race/ethnicity (US only)



This figure presents mean values of responses to the following three questions/statements: (1) I am “locked down”, “quarantined”, “staying home”, or “sheltering in place” (navy bars); (2) Is it possible that the COVID-19 pandemic might lead to your not returning to your current university in Fall 2020 (maroon bars); and (3) Have you withdrawn from any of your courses since the COVID-19 pandemic? (orange bars). The responses are stratified by race/ethnicity and gender for the United States only. Sample sizes differ by question as well as race/ethnicity and gender. These are, respectively for questions (1) to (3): for Whites 9395, 9078, 10530; for Blacks 959, 844, 1000; for Asians 1436, 1316, 1552; for Hispanics 1933, 1754, 2007. Equivalent regression analyses with and without controls are presented in panel C of Table A6.

Figure 26: Education challenges, by country

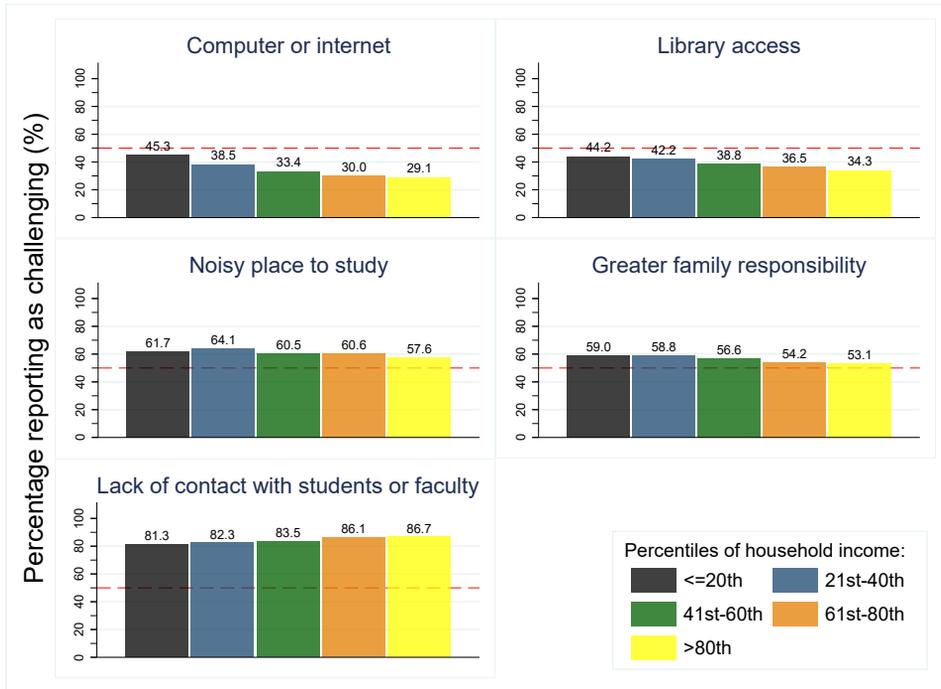


Note: This figure presents mean values of responses to question regarding challenges to completing coursework. The exact question was worded as “Did/does your situation since the COVID-19 pandemic present any challenges to completing your courses successfully? (check all that apply)” with the following options: (1) insufficient computer resources or internet problems, (2) library closed or insufficient library resources, (3) lack of a quiet place to study, (4) increased family responsibilities, and (5) (a) lack of contact with other students or (b) lack of contact with faculty. Each option is depicted as a separate panel and the responses are stratified by country. Black bars are for Australia, navy for Austria, green for Italy, maroon for Mexico, orange for Spain, khaki for Sweden, and yellow for the United States. Sample sizes are 4643 for Australia, 490 for Austria, 3651 for Italy, 586 for Mexico, 5997 for Spain, 485 for Sweden, and 18700 for the United States.

3.2.2 Educational challenges

Results by country. Figure 26 shows country differences in educational challenges to completing coursework faced by students due to the pandemic. For students in all countries, the most significant issue is lack of contact with other students or faculty (varying from a 75% in Austria up to a 91% in Australia) followed by noisy place to study (varying from a 39% in Sweden up to a 69% in Mexico). There are large differences across countries in the proportion of students reporting greater family responsibilities, from a 14% in Sweden up to a 65% in Mexico. Finally, students also report as challenges insufficient library access (especially in Spain, with 45.7% of students) and computer or internet problems (especially in Mexico and Australia, with 43% and 40% of students, respectively).

Figure 27: Education challenges, by parental income



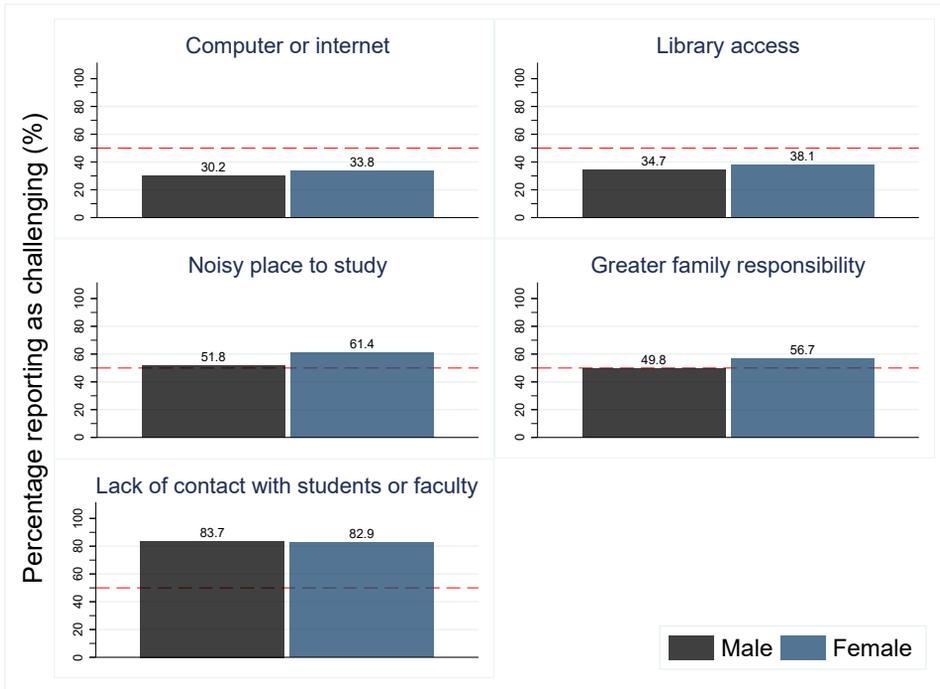
Note: This figure presents mean values of responses to question regarding challenges to completing coursework. The exact question was worded as “Did/does your situation since the COVID-19 pandemic present any challenges to completing your courses successfully? (check all that apply)” with the following options: (1) insufficient computer resources or internet problems, (2) library closed or insufficient library resources, (3) lack of a quiet place to study, (4) increased family responsibilities, and (5) (a) lack of contact with other students or (b) lack of contact with faculty. Each option is depicted as a separate panel and the responses are stratified by student’s household (parents) quintile which is country-specific based on national income distribution. Black bars are for bottom quintile, navy for 21st to 40th percentile, green for 41st to 60th percentile, orange for 61st to 80th percentile, and yellow for top quintile. Sample sizes across quintiles are 1755, 3115, 4298, 5735, and 8373. Equivalent regression analyses with and without controls are presented in panel A of Table A7.

Parental income differences. Figure 27 shows differences in educational challenges by parental income. In general, a higher proportion of students from the top quintile report that the lack of contact with other students and faculty is the most important challenge (86.7%), whereas the rest of the challenges are more important for the two bottom quintiles, including computer/internet issues and greater family responsibilities. For example, only 29% of students in the top income quintile reports problems with computer or internet access while this proportion raises to 45%, or by more than 50 percent, for those in the bottom income quintile.

Gender differences. Figure 28 shows differences in educational challenges by gender. Here we do not find any striking gender differences in demand for contact with other students and faculty. An 84% of males and a 83% of females report this as a challenge to complet-

Covid Economics 79, 27 May 2021: 152-217

Figure 28: Education challenges, by gender

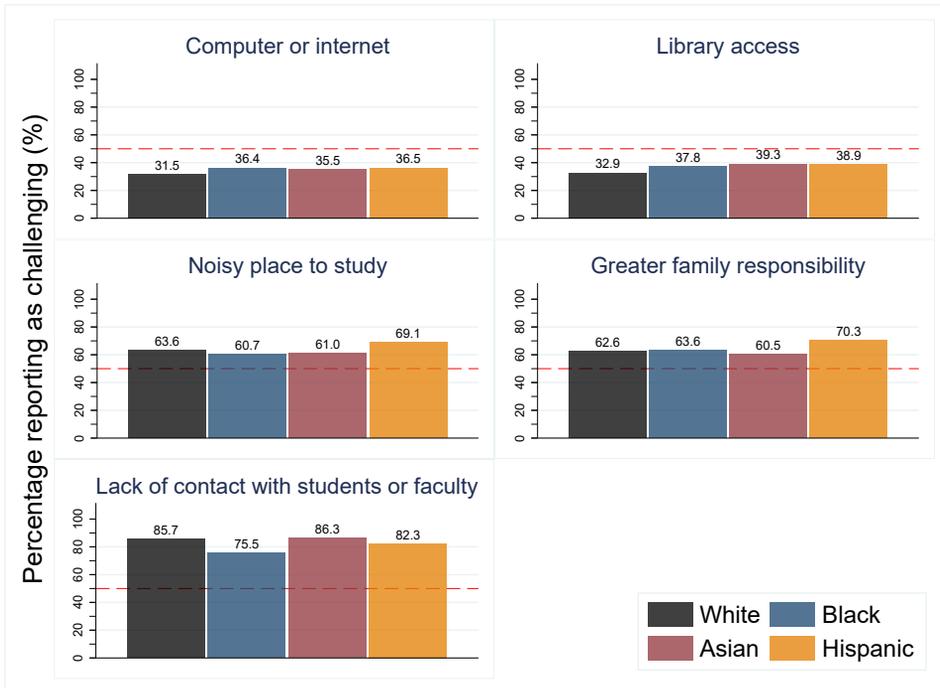


Note: This figure presents mean values of responses to question regarding challenges to completing coursework. The exact question was worded as “Did/does your situation since the COVID-19 pandemic present any challenges to completing your courses successfully? (check all that apply)” with the following options: (1) insufficient computer resources or internet problems, (2) library closed or insufficient library resources, (3) lack of a quiet place to study, (4) increased family responsibilities, and (5) (a) lack of contact with other students or (b) lack of contact with faculty. Each option is depicted as a separate panel and the responses are stratified by gender. Black bars are for males while navy bars are for females. Sample sizes for males and females respectively are 8920 and 20012. Equivalent regression analyses with and without controls are presented in panel B of Table A7.

ing their coursework. Regarding the rest of the challenges, we find a higher proportion of females than males reporting greater family responsibilities (56.7%), noisy place to study (61.4%), library access (38%) and computer or internet problems (34%).

Racial differences in the US. Figure 29 shows racial differences in educational challenges in the United States. In general, Hispanics are most likely to report computer or internet problems (36.5%), library access (38.9%), noisy place to study (69%) and greater family responsibilities (70%). In the case of lack of contact with other students and faculty, however, the proportion of students reporting this challenge is highest among Asians (86%) and Whites (85.7%).

Figure 29: Education challenges, by race/ethnicity (US only)



Note: This figure presents mean values of responses to question regarding challenges to completing coursework. The exact question was worded as “Did/does your situation since the COVID-19 pandemic present any challenges to completing your courses successfully? (check all that apply)” with the following options: (1) insufficient computer resources or internet problems, (2) library closed or insufficient library resources, (3) lack of a quiet place to study, (4) increased family responsibilities, and (5) (a) lack of contact with other students or (b) lack of contact with faculty. Each option is depicted as a separate panel and the responses are stratified by race/ethnicity for the United States only. Black bars are for Whites, navy bars are for Blacks, maroon bars are for Asians, and orange bars are for Hispanics. Sample sizes for Whites, Blacks, Asians, and Hispanics respectively are 9963, 952, 1491 and 1951. Equivalent regression analyses with and without controls are presented in panel C of Table A7.

Covid Economics 79, 27 May 2021: 152-217

3.2.3 Uncertainty about returning to school

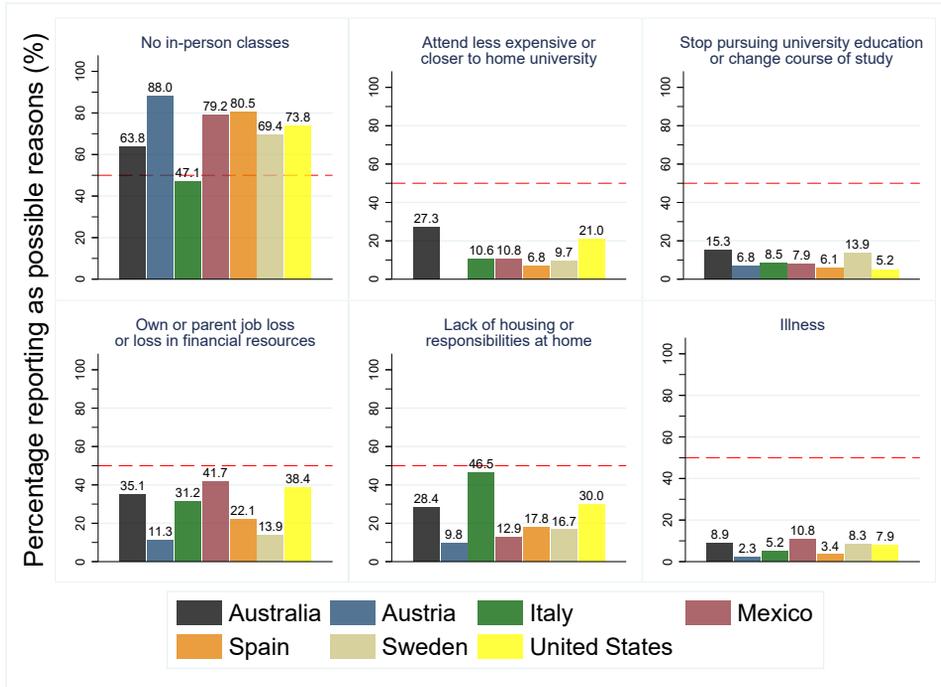
Results by country. Figure 30 shows differences among countries when it comes to the reasons why there is uncertainty about returning to school. In all countries, a significant percentage of students think that no in-person classes is the most important reason behind this uncertainty regarding Fall 2020 semester. Some differences between countries are also visible: in Austria and Spain a 88% and 80.5% of students, respectively, reported this reason whereas in Australia it is a 63.8% of students and in Italy only a 47.1%. On the other hand, Mexico, US and Australia, compared to other places, are the countries with a higher proportion of students reporting other reasons such as own or parent job loss or loss of financial resources (41.7% in Mexico, 38.4% in the US, and 35.1% in Australia). Italy shows a higher percentage of students reporting lack of housing or responsibilities at home (46.5%), and Australia is also the country reporting attending less expensive or closer to home university (27%) and stop pursuing college education or change the field of study (15%) as important reasons. A significant percentage of students in US also reported familiar job losses and financial issues (38.4%). In the case of US, two other reasons stand out: lack of housing and familiar responsibilities (30%) and attend less expensive or closer universities (21%).

Parental income differences. Figure 31 shows differences across the parental income distribution. There is a positive correlation between parental income and the percentage of students that reported no in-person classes as a factor behind the uncertainty to returning to classes in Fall 2020. The correlation between the importance of other reasons (mainly related to financial resources, familiar responsibilities or labour market consequences in the household) and parental income, however, is just the opposite. In the case of parental job loss or experiencing financial problems, the percentage of students in bottom quintiles that reported this reason (48.9%) is higher than in top quintiles (23%). In the case of lack of housing or responsibilities at home, this reason was reported by a 37.5% of students in the bottom quintile whereas in the top quintile only a 19.6% of students reported it. These results make sense given the labor market findings that we reported above.

Gender differences. Differences by gender are shown in Figure 32. In short, we do not find any meaningful differences across males and females. A similar percentage of males and females reported all the main reasons.

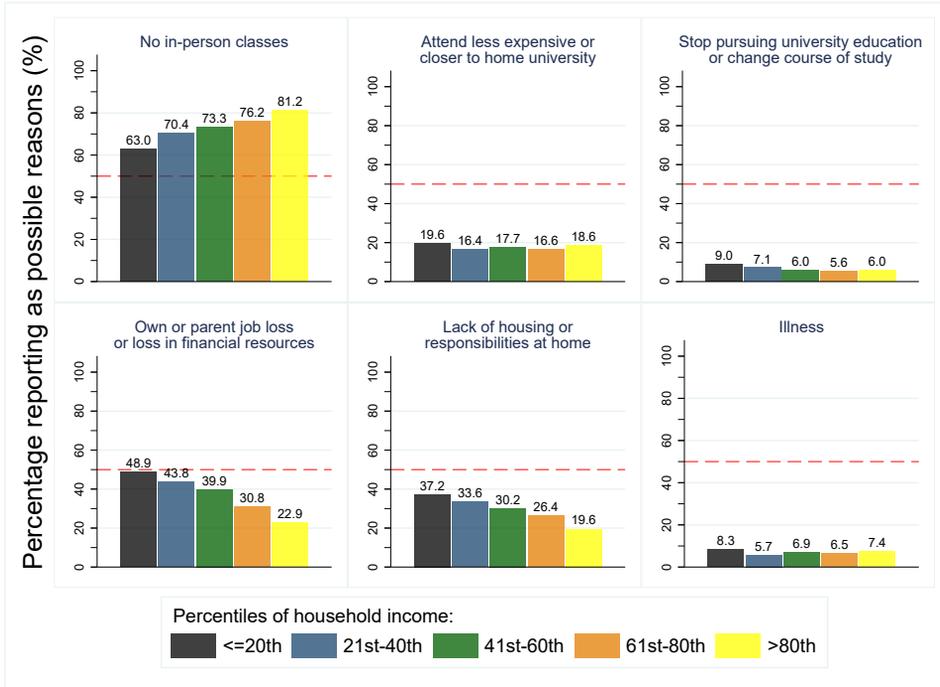
Racial differences in the US. Figure 33 shows the same reasons but by race and ethnicity for the US sample. White and Hispanic students are most likely to list no in-person classes as a source of uncertainty. At the same time, Hispanics along with Black students are most concerned about financial resources and family responsibilities which makes sense given that these two groups were disproportionately affected when it comes to the labor market. On the other hand, we do not find any striking differences when it comes to attending less

Figure 30: Reasons behind uncertainty about returning to school in Fall 2020, by country



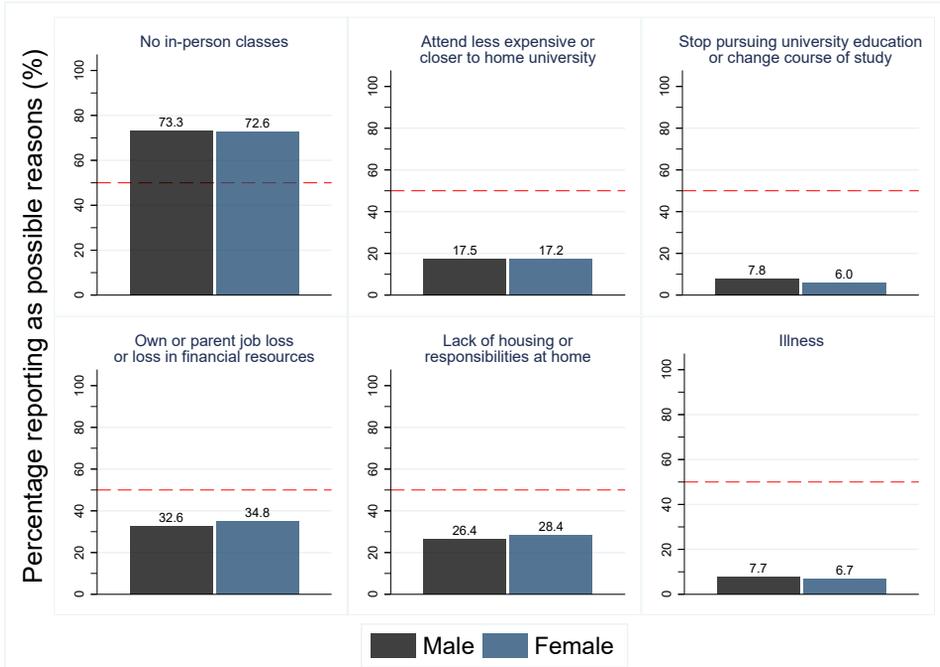
Note: This figure presents mean values of responses to question regarding reasons behind the possibility of not returning to university. The exact question was worded as “What factors might lead to your not returning to your university in Fall 2020? (check all that apply)” with the following options: (1) in-person classes do not resume, (2) want to go to university closer to home, (3) want to go to less expensive university (not asked in Austria), (4) want to change course of study, (5) want to stop going to university, (6) one or more parents laid off, (7) lost own job, (8) other loss of financial resources, (9) lack of housing, (10) responsibilities at home, and (11) illness. These questions were only presented to students who responded “yes” to the following question: “Is it possible that the COVID-19 pandemic might lead to your not returning to your current university in Fall 2020?”. For succinctness, the figure combines the following questions (2) and (3) as second panel, questions (4) and (5) as third panel, questions (6), (7), and (8) as fourth panel, questions (9) and (10) as fifth panel. Each option is depicted as a separate panel and the responses are stratified by country. Black bars are for Australia, navy for Austria, green for Italy, maroon for Mexico, orange for Spain, khaki for Sweden, and yellow for the United States. Sample sizes are 1298 for Australia, 133 for Austria, 792 for Italy, 240 for Mexico, 2552 for Spain, 72 for Sweden, and 6526 for the United States.

Figure 31: Reasons behind uncertainty about returning to school in Fall 2020, by parental income



Note: This figure presents mean values of responses to question regarding reasons behind the possibility of not returning to university. The exact question was worded as “What factors might lead to your not returning to your university in Fall 2020? (check all that apply)” with the following options: (1) in-person classes do not resume, (2) want to go to university closer to home, (3) want to go to less expensive university (not asked in Austria), (4) want to change course of study, (5) want to stop going to university, (6) one or more parents laid off, (7) lost own job, (8) other loss of financial resources, (9) lack of housing, (10) responsibilities at home, and (11) illness. These questions were only presented to students who responded “yes” to the following question: “Is it possible that the COVID-19 pandemic might lead to your not returning to your current university in Fall 2020?”. For succinctness, the figure combines the following questions (2) and (3) as second panel, questions (4) and (5) as third panel, questions (6), (7), and (8) as fourth panel, questions (9) and (10) as fifth panel. Each option is depicted as a separate panel and the responses are stratified by student’s household (parents) quintile which is country-specific based on national income distribution. Black bars are for bottom quintile, navy for 21st to 40th percentile, green for 41st to 60th percentile, orange for 61st to 80th percentile, and yellow for top quintile. Sample sizes across quintiles are 698, 1217, 1560, 1900 and 2570 for all but second panel. Sample sizes across quintiles for the second panel (“Attend less expensive or closer to home university”) are 693, 1202, 1549, 1885, and 2538. Equivalent regression analyses with and without controls are presented in panel A of Tables A8 and A9.

Figure 32: Reasons behind uncertainty about returning to school in Fall 2020, by gender



Note: This figure presents mean values of responses to question regarding reasons behind the possibility of not returning to university. The exact question was worded as “What factors might lead to your not returning to your university in Fall 2020? (check all that apply)” with the following options: (1) in-person classes do not resume, (2) want to go to university closer to home, (3) want to go to less expensive university (not asked in Austria), (4) want to change course of study, (5) want to stop going to university, (6) one or more parents laid off, (7) lost own job, (8) other loss of financial resources, (9) lack of housing, (10) responsibilities at home, and (11) illness. These questions were only presented to students who responded “yes” to the following question: “Is it possible that the COVID-19 pandemic might lead to your not returning to your current university in Fall 2020?”. For succinctness, the figure combines the following questions (2) and (3) as second panel, questions (4) and (5) as third panel, questions (6), (7), and (8) as fourth panel, questions (9) and (10) as fifth panel. Each option is depicted as a separate panel and the responses are stratified by gender. Black bars are for males while navy bars are for females. Sample sizes for males and females respectively are 3092 and 6668 for all but second panel. Sample sizes for males and females respectively for the second panel (“Attend less expensive or closer to home university”) are 3049 and 6591. Equivalent regression analyses with and without controls are presented in panel B of Tables A8 and A9.

Covid Economics 79, 27 May 2021: 152-217

expensive university or one that is closer to home. Black students were the most likely, however, than any other race/ethnic group to report they would discontinue their university education.

3.2.4 Changes in studying time due to the pandemic

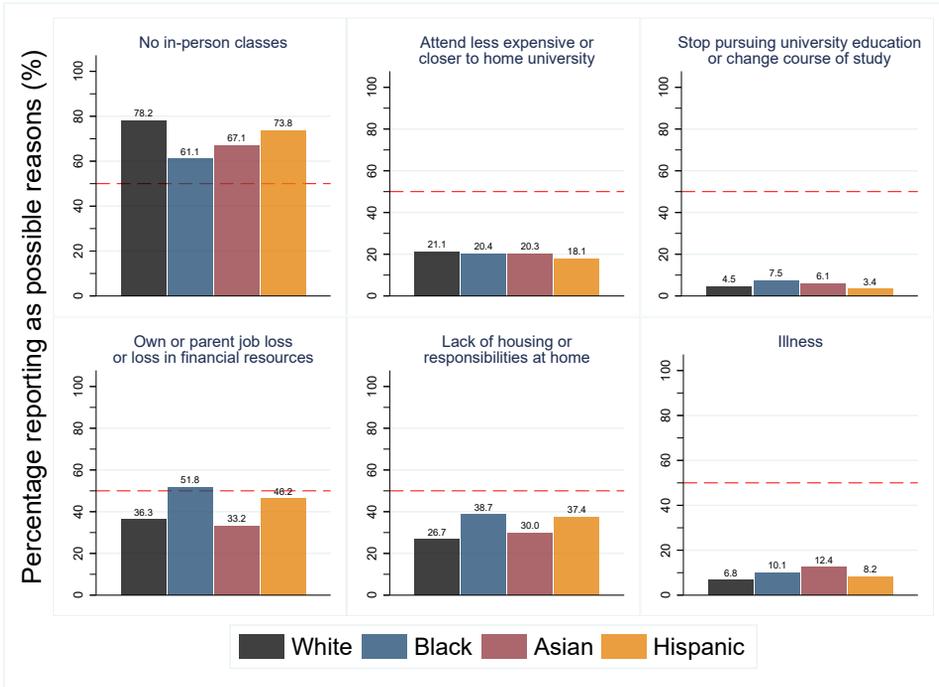
Results by country. Figure 34 shows differences across countries in the changes in the studying time due to the COVID-19 pandemic. In general, students devoting fewer hours (less than 15 hours per week) have increased their studying time as a consequence of pandemic in all the countries. In some cases, as in Sweden, the percentage of students in this category has changed from a 17% to a 26%. The percentage of students devoting a higher amount of hours per week before the pandemic (16-30 hours and over 30 hours per week), however, has been reduced in most countries. In some cases like the US, the percentage of students devoting between 16 and 30 hours has fallen from 47.3% to 32.6% and the percentage of students devoting more than 30 hours has fallen from a 23% to a 13%. There is one exception in this latter case; in Spain, the percentage of students devoting over 30 hours per week has raised from 38.5% up to a 44%. Thus, it appears that on average students in all countries shifted from more to less hours of studying time.

Parental income differences. Figure 35 shows the percentage of students by time devoted to study before and after the pandemic. In this case, the general pattern explored by country is consistent across different income levels. We observe a shift from studying 16 or more hours a week to between 1 and 15 hours a week. Interestingly, this shift appears larger for more compared to less affluent households. For example, rate of studying between 1 and 15 hours increases by 16 percentage points for those in the bottom income quintile and by 20 percentage points for those in the top.

Gender differences. Figure 36 shows documents differences by gender. Here we do not find any striking differences in studying time either before or during the pandemic.

Racial differences in the US. Figure 37 shows studying patterns by students race/ethnicity in the US. First, we observe differences in studying times across racial-ethnic categories. For example, before the pandemic, Asian students were most likely to study over 30 hours per week at 30.9% compared with only 17.9% for Black students. We observe reductions in study time across all groups considered, however, these are not uniform. The rate of studying only 1 to 15 hours increased by 26 percentage points for Whites but only by 21 percentage points for Black students who were most likely to study less in the pre-pandemic period. Conversely, declines in studying over 30 hours a week range from 13 percentage points for Asian students to 8 percentage points for Black students.

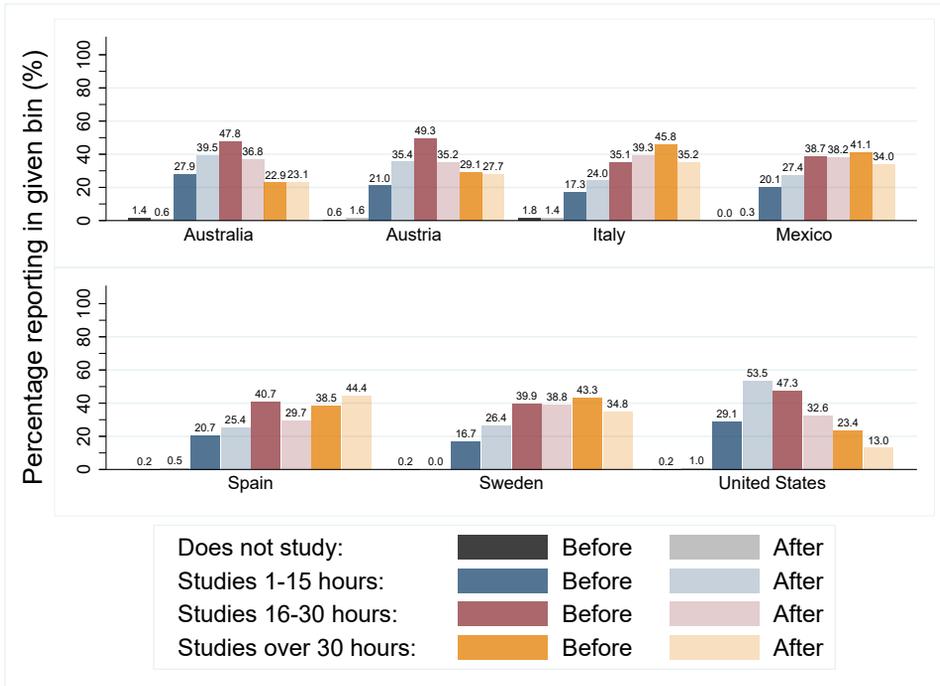
Figure 33: Reasons behind uncertainty about returning to school in Fall 2020, by race/ethnicity (US only)



Note: This figure presents mean values of responses to question regarding reasons behind the possibility of not returning to university. The exact question was worded as “What factors might lead to your not returning to your university in Fall 2020? (check all that apply)” with the following options: (1) in-person classes do not resume, (2) want to go to university closer to home, (3) want to go to less expensive university (not asked in Austria), (4) want to change course of study, (5) want to stop going to university, (6) one or more parents laid off, (7) lost own job, (8) other loss of financial resources, (9) lack of housing, (10) responsibilities at home, and (11) illness. These questions were only presented to students who responded “yes” to the following question: “Is it possible that the COVID-19 pandemic might lead to your not returning to your current university in Fall 2020?”. For succinctness, the figure combines the following questions (2) and (3) as second panel, questions (4) and (5) as third panel, questions (6), (7), and (8) as fourth panel, questions (9) and (10) as fifth panel. Each option is depicted as a separate panel and the responses are stratified by race/ethnicity for the United States only. Black bars are for Whites, navy bars are for Blacks, maroon bars are for Asians, and orange bars are for Hispanics. Sample sizes for Whites, Blacks, Asians, and Hispanics respectively are 3195, 388, 621 and 764. Equivalent regression analyses with and without controls are presented in panel C of Tables A8 and A9.

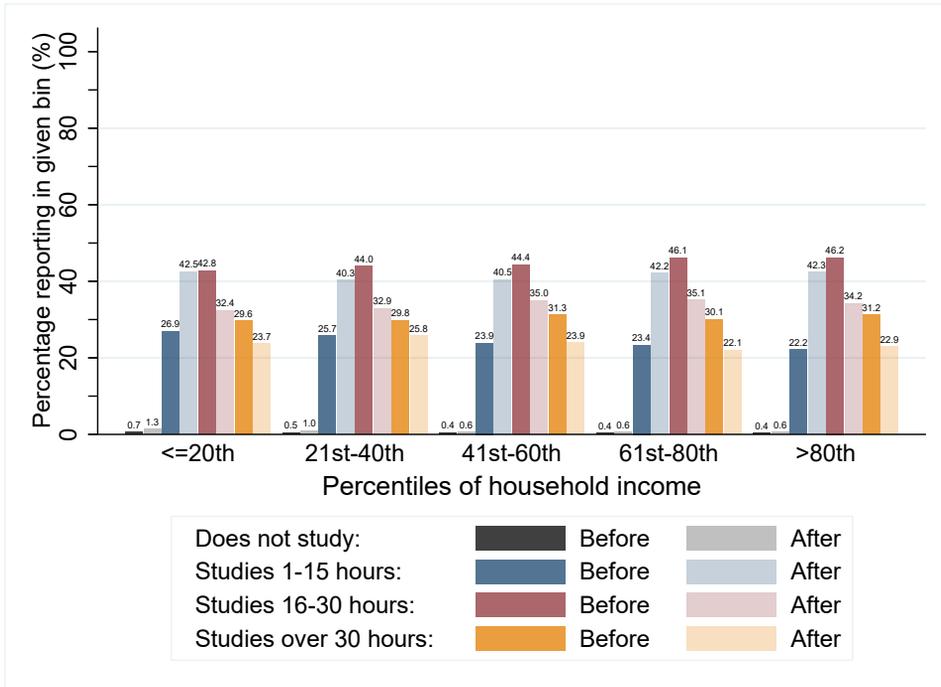
Covid Economics 79, 27 May 2021: 152-217

Figure 34: Changes in studying time, by country



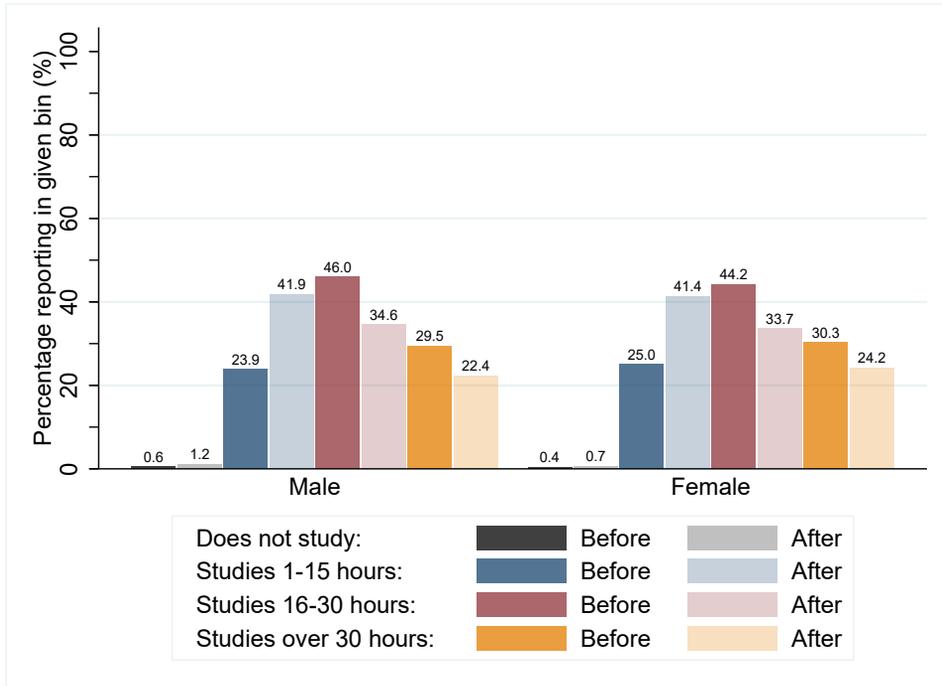
Note: This figure presents mean values of responses to questions regarding studying time of students prior to and since COVID-19 pandemic. The exact questions were "Before [Since] the COVID-19 pandemic, about how many hours per week did [do] you devote to academic work? (for example: attending class, reading class materials, attending labs, doing problem sets, writing papers, etc.)". Respondent had multiple options including: "None", 5-hour intervals above zero (e.g., "About 1-5 hours per week"), up to "More than 40 hours per week". We aggregated these responses to dichotomous scale of four variables depicted in this figure. Variables are multiplied by 100 and sum to 100 within a question. Black bars represent no studying, navy bars represent studying between 1 and 15 hours per week, maroon bars represent studying 16 to 30 hours per week, and orange bars represent studying more than 30 hours per week. Solid bars are for studying situation before while faded bars are for studying situation after the start of COVID-19 pandemic. Sample is divided by country. Top panel presents results for Australia, Austria, Italy, and Mexico while bottom panel presents results for Spain, Sweden, and the United States. Sample sizes are 4747 for Australia, 505 for Austria, 3851 for Italy, 591 for Mexico, 6434 for Spain, 564 for Sweden, and 19505 for the United States.

Figure 35: Changes in studying time, by parental income



Note: This figure presents mean values of responses to questions regarding studying time of students prior to and since COVID-19 pandemic. The exact questions were "Before [Since] the COVID-19 pandemic, about how many hours per week did [do] you devote to academic work? (for example: attending class, reading class materials, attending labs, doing problem sets, writing papers, etc.)". Respondent had multiple options including: "None", 5-hour intervals above zero (e.g., "About 1-5 hours per week"), up to "More than 40 hours per week". We aggregated these responses to dichotomous scale of four variables depicted in this figure. Variables are multiplied by 100 and sum to 100 within a question. Black bars represent no studying, navy bars represent studying between 1 and 15 hours per week, maroon bars represent studying 16 to 30 hours per week, and orange bars represent studying more than 30 hours per week. Solid bars are for studying situation before while faded bars are for studying situation after the start of COVID-19 pandemic. The responses are stratified by student's household (parents) quintile which is country-specific based on national income distribution. Sample sizes are 1796 for bottom quintile, 3211 for 21st to 40th percentile, 4473 for 41st to 60th percentile, 6041 for 61st to 80th percentile, and 8835 for top quintile.

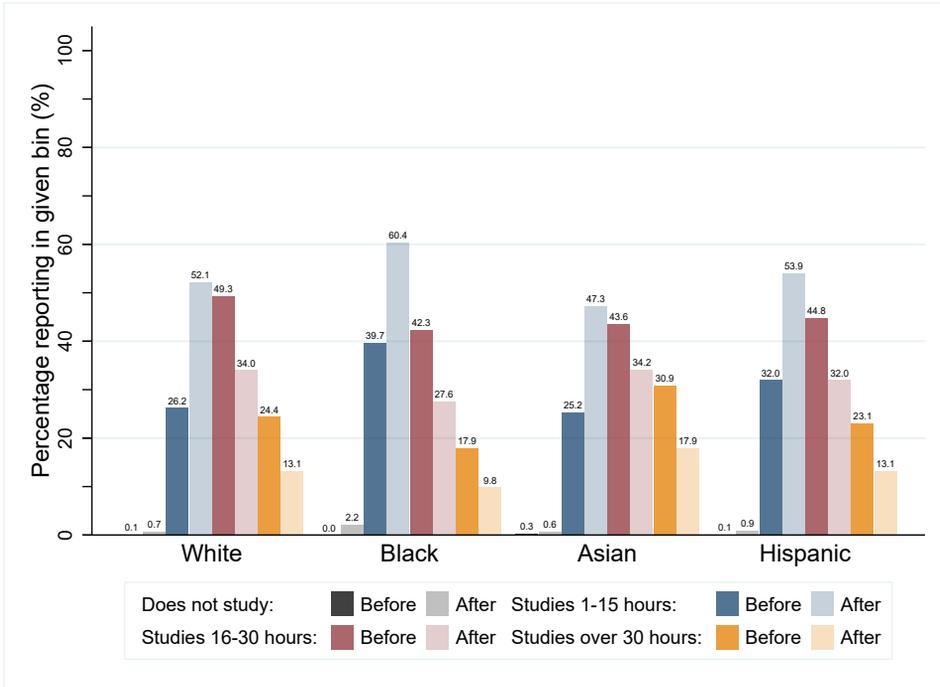
Figure 36: Changes in studying time, by gender



Note: This figure presents mean values of responses to questions regarding studying time of students prior to and since COVID-19 pandemic. The exact questions were "Before [Since] the COVID-19 pandemic, about how many hours per week did [do] you devote to academic work? (for example: attending class, reading class materials, attending labs, doing problem sets, writing papers, etc.)". Respondent had multiple options including: "None", 5-hour intervals above zero (e.g., "About 1-5 hours per week"), up to "More than 40 hours per week". We aggregated these responses to dichotomous scale of four variables depicted in this figure. Variables are multiplied by 100 and sum to 100 within a question. Black bars represent no studying, navy bars represent studying between 1 and 15 hours per week, maroon bars represent studying 16 to 30 hours per week, and orange bars represent studying more than 30 hours per week. Solid bars are for studying situation before while faded bars are for studying situation after the start of COVID-19 pandemic. The responses are stratified by student's gender. Sample sizes are 9537 for males and 20881 for females.

Covid Economics 79, 27 May 2021: 152-217

Figure 37: Changes in studying time, by race/ethnicity (US only)



Note: This figure presents mean values of responses to questions regarding studying time of students prior to and since COVID-19 pandemic. The exact questions were "Before [Since] the COVID-19 pandemic, about how many hours per week did [do] you devote to academic work? (for example: attending class, reading class materials, attending labs, doing problem sets, writing papers, etc.)". Respondent had multiple options including: "None", 5-hour intervals above zero (e.g., "About 1-5 hours per week"), up to "More than 40 hours per week". We aggregated these responses to dichotomous scale of four variables depicted in this figure. Variables are multiplied by 100 and sum to 100 within a question. Black bars represent no studying, navy bars represent studying between 1 and 15 hours per week, maroon bars represent studying 16 to 30 hours per week, and orange bars represent studying more than 30 hours per week. Solid bars are for studying situation before while faded bars are for studying situation after the start of COVID-19 pandemic. The responses are stratified by race/ethnicity and gender for the United States only. Sample sizes are 10504, 999, 1550, and 2002, for Whites, Blacks, Asians, and Hispanics, respectively.

Covid Economics 79, 27 May 2021: 152-217

3.2.5 Discussion on educational outcomes

The GC19SS survey shows that COVID-19 has affected students' experiences at university across all analyzed countries. In this section, we have studied results on educational consequences, challenges, reasons behind the uncertainty to return to classes and changes in studying time.

Across all samples, the main reason for considering not returning to university is the lack of in-person classes, which suggests important implications from the university perspective. On the other, hand other factors such as the role of financial resources or educational challenges exhibit more heterogeneous patterns.

The most important differences by country are based on family responsibilities (practically non-existent in Sweden to over 60% in the US) and in-person classes (especially important in Austria and Spain). Family responsibilities also exhibit large differences, with Swedish students being less affected opposite to what happened in countries like Mexico and Spain, where students reported a higher worrying. Physical barriers (as library and internet access) are also important in some countries like Spain, Australia and Mexico, which is possibly associated to rural-urban segregation.

We discover relatively large differences by household income when it comes to educational challenges related to infrastructure such as computer or internet as well as library access. This could be due to the fact that lower SES students disproportionately rely on university resources and infrastructure for their educational success. Interestingly, these students were least concern with lack of in-person classes and most troubled by financial and family concerns. This makes sense if poorer students treat university education as investment rather than consumption good, however, they were also most likely to stop pursuing tertiary education altogether.

Although, we did not find many striking differences by gender, except for perhaps women disproportionately reporting lack of quiet place to study and elevated family responsibilities, we did observed gaps by racial-ethnic groups in the US. In the US it appears that Hispanic students were particularly affected when it comes to educational challenges. They were the most likely to report being limited by noisy study place and greater family responsibilities. Furthermore, Black and Hispanic students are more likely not to return to school in Fall 2020, due to reasons like lack of housing, family responsibilities, loss of own job, and other financial losses. Lack of contact with other students of lack of in-person classes seems to be more important for Asians and Whites. Finally, Black students are most likely to stop pursuing university education.

3.3 Health outcomes

3.3.1 COVID-19 incidence

We describe how the pandemic affected the health outcomes of students and their families, both in terms of direct COVID-19 incidence as well as mental health issues. We first document to what degree have students experienced COVID-19 symptoms, tested positive for COVID-19, or had a family member or acquaintance die from COVID-19.

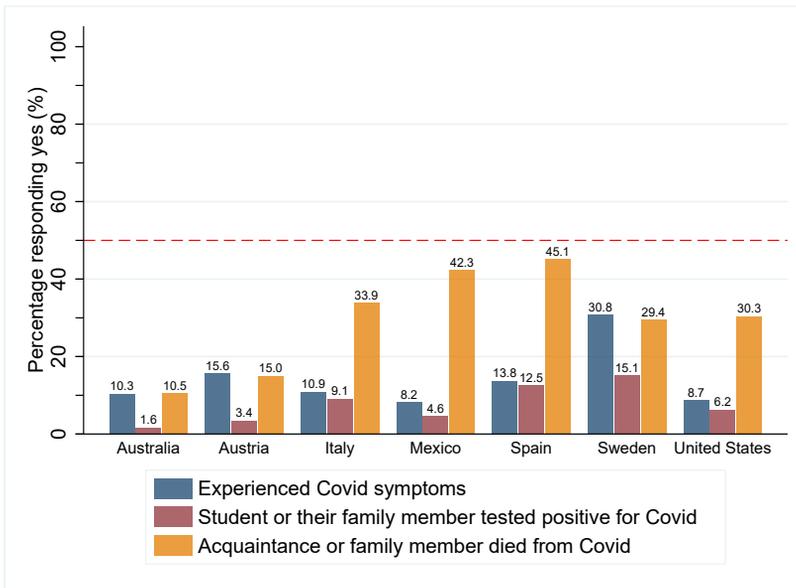
Results by country. Figure 38 shows how COVID-19 incidence has differentially affected undergraduates across countries in the survey. Students in Sweden, which did not close down its economy although universities moved to remote instruction for the most part, were by far the most likely to experience COVID-19 symptoms (31%). They were followed by Austria and Spain (16% and 14%, respectively), and Italy, Australia, the US, and Mexico (11%, 10%, 9%, and 8%, respectively). Testing positively for COVID-19, either the student themselves or their family member, also varied across countries. Sweden and Spain had the highest rates of positive testing (15% and 13%, respectively), followed by Italy (9%), the US (6%), Mexico (5%), Austria (3%), and Australia (2%). Students also reported relatively high rates of having lost an acquaintance or family member to COVID-19: 45% did so in Spain, 42% in Mexico, 34% in Italy, 30% in the US, 29% in Sweden, 15% in Austria, and 11% in Australia.

Parental income differences. Figure 39 shows COVID incidence by parental income quintiles. Incidence across students of different socioeconomic backgrounds was fairly similar, without clear parental income gradients. Across groups, between 9%–11% experienced COVID symptoms, 6%–7% tested positive for COVID (either themselves or their family member), and 30%–33% had an acquaintance or family member die from COVID-19.

Gender differences. COVID-19 incidence was also quite similar for male and female students. Figure 40 shows that 10% of both men and women experienced symptoms, while 7% of both genders tested positive (either themselves or their family member). A slight difference arises in the fraction reporting having an acquaintance or family member die from COVID-19, with 28% of men doing so compared to 32% of women.

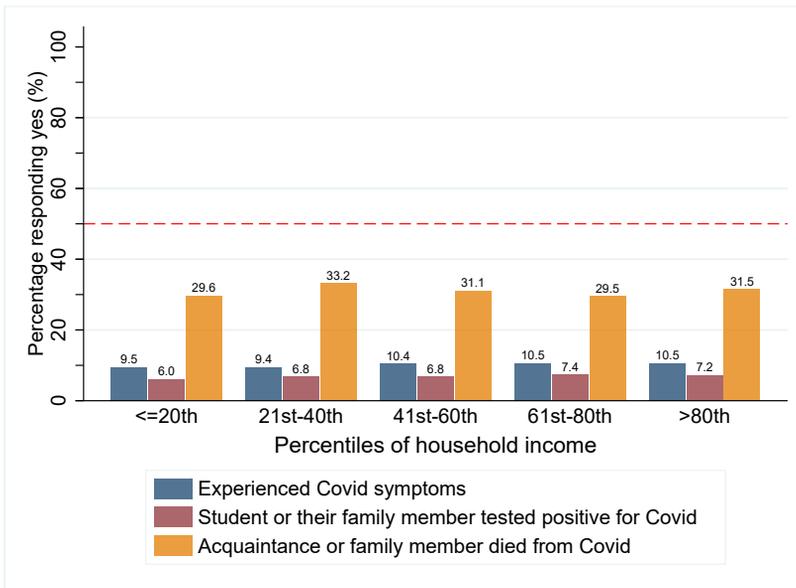
Racial differences in the US. Figure 41 shows that COVID-19 incidence in the US was quite different for students of different races/ethnicities. White students were the more likely to report having experienced symptoms (9%), compared to Hispanics (8%), Blacks (7%), and Asians (6%). In terms of testing positive and deaths, however, Blacks and Hispanics were the hardest hit. Among Blacks and Hispanics, 8% had a positive test (either themselves or their family member), compared to 6% of Whites and 3% of Asians. Black students experienced by far the most deaths among acquaintances and family members (42% of them did), followed by Hispanics (32%), and Asians and Whites (28%).

Figure 38: COVID incidence, by country



Note: This figure presents mean values of responses to the following five questions/statements: (1) I have experienced symptoms (dry cough, fever, aches) that are consistent with COVID-19 (navy bars); (2) (a) I have been positively diagnosed with COVID-19 or (b) One of my immediate family members (parents, siblings, partner) has been positively diagnosed with COVID-19 (maroon bars); (3) (a) One of my immediate family members (parents, siblings, partner) has died from COVID-19 or (b) I know someone outside of my immediate family who has died from COVID-19 (orange bars). Sample is divided by country. Sample sizes are 3645 for Australia, 320 for Austria, 2435 (questions 1 and 2) and 1480 (question 3) for Italy, 525 for Mexico, 5311 for Spain, 377 for Sweden, and 15650 for the United States.

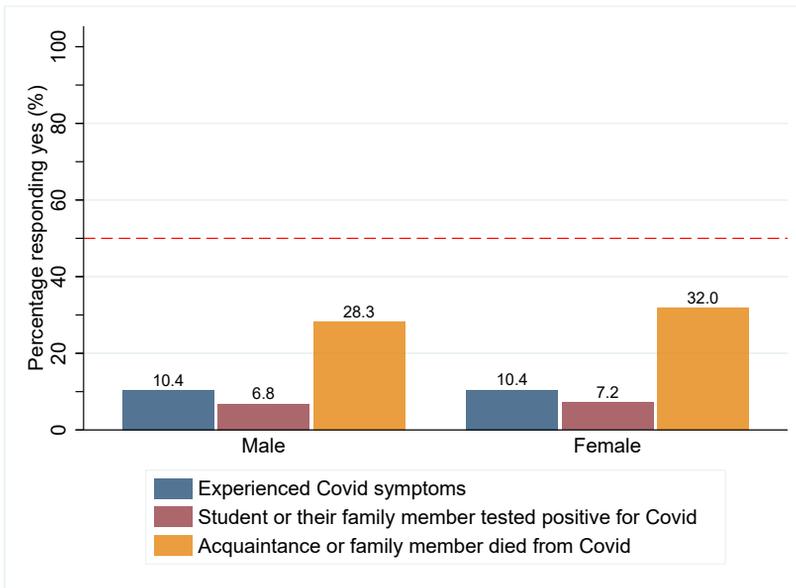
Figure 39: COVID incidence, by parental income



Note: This figure presents mean values of responses to the following five questions/statements: (1) I have experienced symptoms (dry cough, fever, aches) that are consistent with COVID-19 (navy bars); (2) (a) I have been positively diagnosed with COVID-19 or (b) One of my immediate family members (parents, siblings, partner) has been positively diagnosed with COVID-19 (maroon bars); (3) (a) One of my immediate family members (parents, siblings, partner) has died from COVID-19 or (b) I know someone outside of my immediate family who has died from COVID-19 (orange bars). The responses are stratified by student’s household (parents) quintile which is country-specific based on national income distribution. Sample sizes for questions (1) and (2) are 1637 for bottom quintile, 2897 for 21st to 40th percentile, 3953 for 41st to 60th percentile, 5327 for 61st to 80th percentile, and 7808 for top quintile. Equivalent numbers for question (3) are 1556, 2765, 3814, 5124, and 7566. Equivalent regression analyses with and without controls are presented in panel A of Table A10.

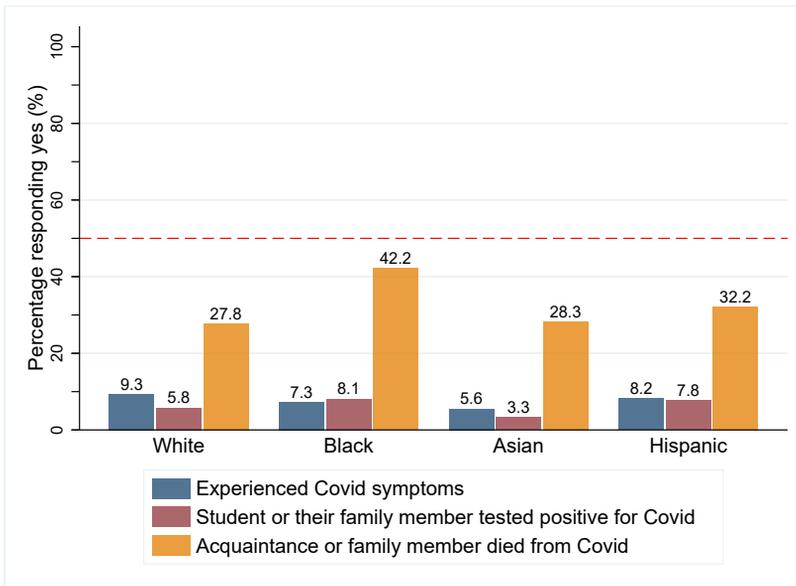
Covid Economics 79, 27 May 2021: 152-217

Figure 40: COVID incidence, by gender



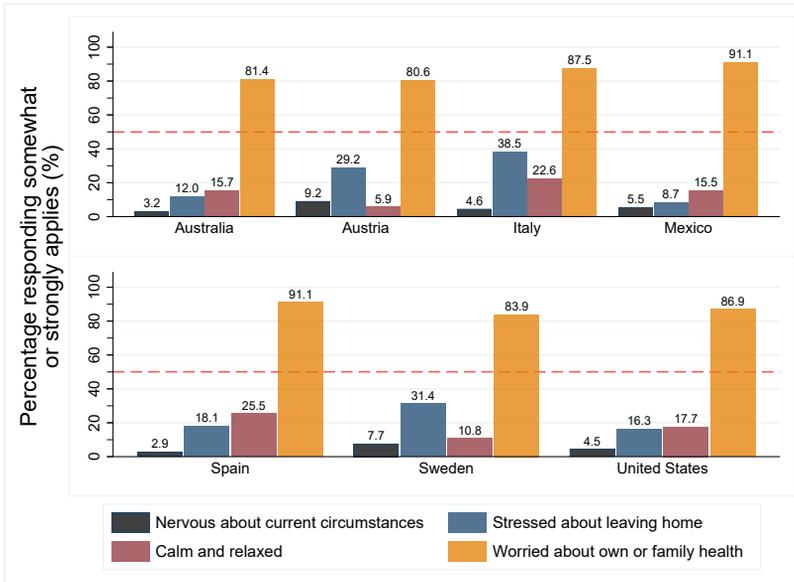
Note: This figure presents mean values of responses to the following five questions/statements: (1) I have experienced symptoms (dry cough, fever, aches) that are consistent with COVID-19 (navy bars); (2) (a) I have been positively diagnosed with COVID-19 or (b) One of my immediate family members (parents, siblings, partner) has been positively diagnosed with COVID-19 (maroon bars); (3) (a) One of my immediate family members (parents, siblings, partner) has died from COVID-19 or (b) I know someone outside of my immediate family who has died from COVID-19 (orange bars). The responses are stratified by student's gender. Sample sizes are 8123 for males and 18707 for females for questions (1) and (2) while they are 7850 and 18039 for question (3). Equivalent regression analyses with and without controls are presented in panel B of Table A10.

Figure 41: COVID incidence, by race/ethnicity (US only)



Note: This figure presents mean values of responses to the following five questions/statements: (1) I have experienced symptoms (dry cough, fever, aches) that are consistent with COVID-19 (navy bars); (2) (a) I have been positively diagnosed with COVID-19 or (b) One of my immediate family members (parents, siblings, partner) has been positively diagnosed with COVID-19 (maroon bars); (3) (a) One of my immediate family members (parents, siblings, partner) has died from COVID-19 or (b) I know someone outside of my immediate family who has died from COVID-19 (orange bars). The responses are stratified by race/ethnicity and gender for the United States only. Sample sizes are 9395, 959, 1436, and 1933, for Whites, Blacks, Asians, and Hispanics, respectively. Equivalent regression analyses with and without controls are presented in panel C of Table A10.

Figure 42: Mental health issues related to the pandemic, by country



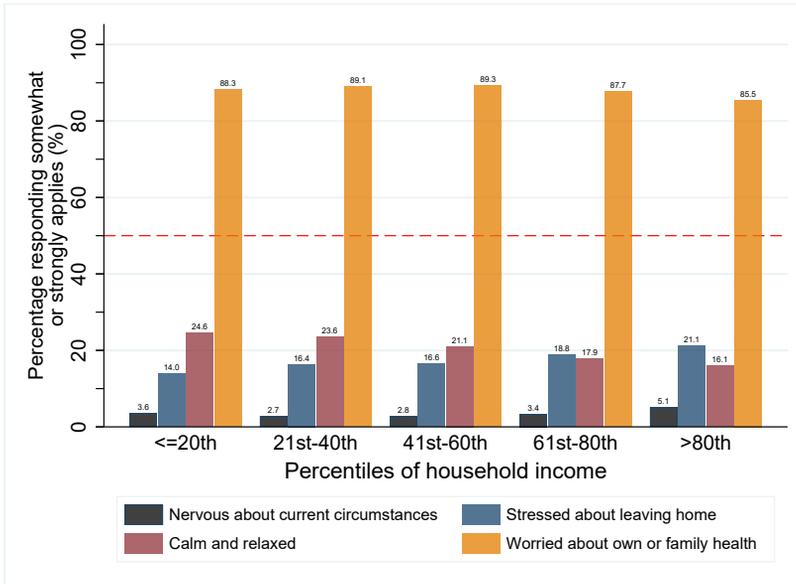
Note: This figure presents mean values of responses to the following five questions/statements: (1) I am nervous when I think about current circumstance (black bars); (2) I feel stressed about leaving my house (navy bars); (3) I am calm and relaxed (maroon bars); (4) (a) I am worried about my health or (b) I am worried about the health of my family members (orange bars). Sample is divided by country. Top panel presents results for Australia, Austria, Italy, and Mexico while bottom panel presents results for Spain, Sweden, and the United States. Sample sizes are 4165 for Australia, 459 for Austria, 3655 for Italy, 541 for Mexico, 5594 for Spain, 509 for Sweden, and 17053 for the United States.

3.3.2 Mental health

We also investigated the extent to which the pandemic has affected undergraduates' mental health. We focus on the level of nervousness and stress that they felt regarding the pandemic, as well as how worried they were about their health and that of their families.

Results by country. Figure 42 shows that the vast majority of students across countries was worried about their health or that of their family members, from 81% in Austria to 91% in Mexico and Spain. A fair amount of students—which varies across countries—reported being stressed about leaving home: 39% in Italy, 31% in Sweden, 29% in Austria, 18% in Spain, 16% in the US, 12% in Australia, and 9% in Mexico. The fraction explicitly reporting being nervous about current circumstances also varied across countries, ranging from 3% in Spain and Australia up to 9% in Austria. Lastly, it was not uncommon for some students to report being calm and relaxed. In Spain, 26% reported so, followed by Italy (23%), the US (18%), Australia and Mexico (16%), Sweden (11%), and Austria (6%).

Figure 43: Mental health issues related to the pandemic, by household income



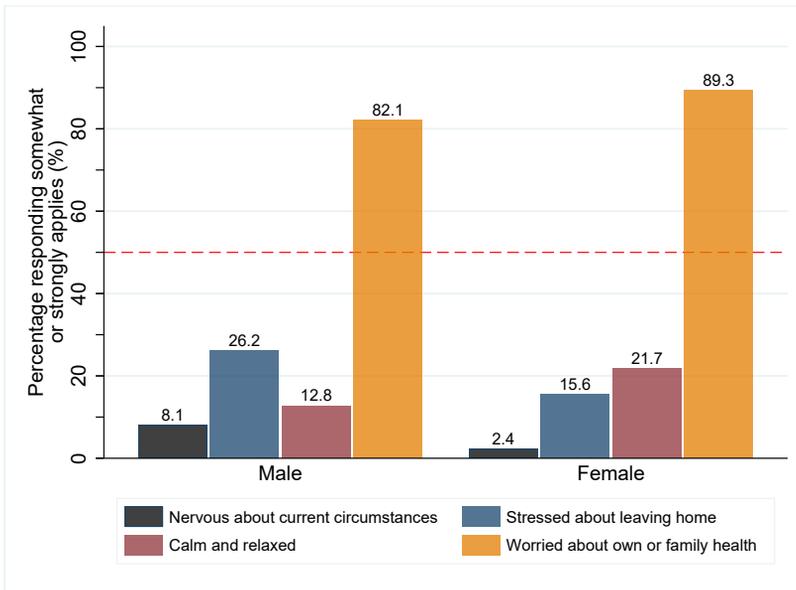
Note: This figure presents mean values of responses to the following five questions/statements: (1) I am nervous when I think about current circumstance (black bars); (2) I feel stressed about leaving my house (navy bars); (3) I am calm and relaxed (maroon bars); (4) (a) I am worried about my health or (b) I am worried about the health of my family members (orange bars). The responses are stratified by student’s household (parents) quintile which is country-specific based on national income distribution. Sample sizes are 1793 for bottom quintile, 3212 for 21st to 40th percentile, 4456 for 41st to 60th percentile, 6046 for 61st to 80th percentile, and 8824 for top quintile. Equivalent regression analyses with and without controls are presented in panel A of Table A11.

Parental income differences. Figure 43 shows mental health outcomes by parental income. Worrying about own or family health was quite similar across groups, with a small difference between those in the top quintile (86%) compared to the other four quintiles (88%–89%). At the same time, students from wealthier backgrounds were more stressed about leaving their home (21% of top-quintile vs. 14% of bottom-quintile students), somewhat more likely to report being nervous about current circumstances (5.1% vs. 3.6%), and less likely to feel calmed and relaxed (16% vs. 25%).

Gender differences. Figure 44 shows that women were more likely than men to be worried about their own or their family’s health (89% of women vs. 82% of men). They were less likely to be nervous about current circumstances (2.4% vs. 8.1%), less stressed about leaving home (16% vs. 26%), and more likely to feel calm and relaxed (22% vs. 13%).

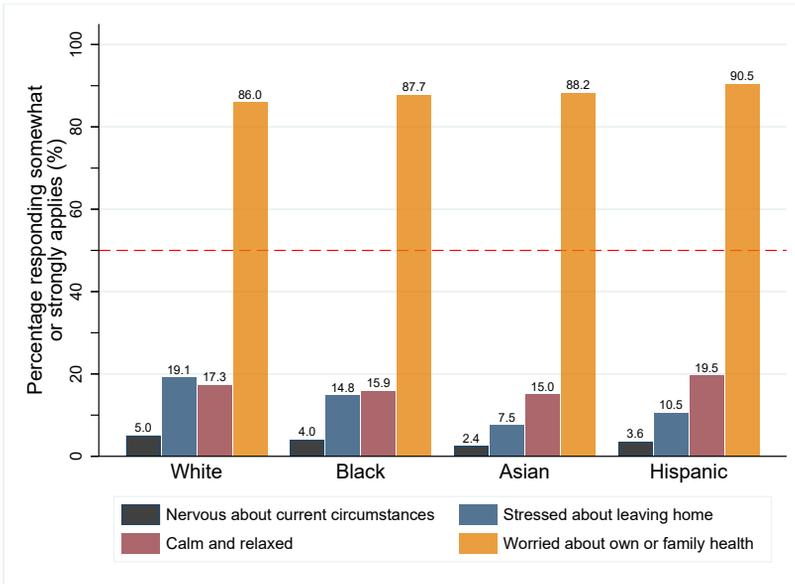
Racial differences in the US. Figure 45 shows that, across all races/ethnicities, the fraction of students being worried about their own or their family’s health was very high, but particularly so among Hispanics (91%), followed by Asians and Blacks (88%), and then Whites (86%). Whites, however, were the most likely to report being stressed about leaving home

Figure 44: Mental health issues related to the pandemic, by gender



Note: This figure presents mean values of responses to the following five questions/statements: (1) I am nervous when I think about current circumstance (black bars); (2) I feel stressed about leaving my house (navy bars); (3) I am calm and relaxed (maroon bars); (4) (a) I am worried about my health or (b) I am worried about the health of my family members (orange bars). The responses are stratified by student’s gender. Sample sizes are 9515 for males and 20812 for females. Equivalent regression analyses with and without controls are presented in panel B of Table A11.

Figure 45: Mental health issues related to the pandemic, by race/ethnicity (US only)



Note: This figure presents mean values of responses to the following five questions/statements: (1) I am nervous when I think about current circumstance (black bars); (2) I feel stressed about leaving my house (navy bars); (3) I am calm and relaxed (maroon bars); (4) (a) I am worried about my health or (b) I am worried about the health of my family members (orange bars). The responses are stratified by race/ethnicity and gender for the United States only. Sample sizes are 10457, 994, 1527, and 1995, for Whites, Blacks, Asians, and Hispanics, respectively. Equivalent regression analyses with and without controls are presented in panel C of Table A11.

(19%, compared to 15% Blacks, 11% Hispanics, and 8% Asians) or being nervous about current circumstances (5% of Whites, compared to 4% Blacks, 3.6% Hispanics, and 2.4% Asians). Hispanics were the most likely to report being calm and relaxed (20%), followed by Whites (17%), Blacks (16%), and Asians (15%).

3.3.3 Discussion of health outcomes

The GC19SS survey shows that the dramatic health consequences of COVID-19 felt around the world also acutely affected undergraduate students. Across the countries in our sample, a substantial number of students experienced COVID-19 symptoms and large fractions of them had an acquaintance or family member die from COVID-19. Their mental health also took a toll, with substantial fractions feeling nervous about the pandemic or stressed about leaving home. Almost all of them were worried about their own health or that of their family.

While we find no large differences in COVID-19 incidence by parental income or gender, substantial disparities by race/ethnicity arise in the US. According to the US Center for Disease Control, African Americans have been hardest hit by COVID-19, with roughly twice the documented infection rate, five times the hospitalization rate and twice the death rate,

Covid Economics 79, 27 May 2021: 152-217

compared with Whites.¹⁰ This is also to certain degree reflected in our data. Prior to the pandemic, Black-White differences in mortality were staggering, with age-adjusted mortality rates for Blacks equal to the same levels for Whites from thirty years ago (Wrigley-Field, 2020).¹¹ Thus, the pandemic, if anything, likely exacerbated these differences.

In the GC19SS, Whites were more likely to report experiencing symptoms, but Blacks and Hispanics experienced positive tests and the death of someone in their social or family network at significantly higher rates. Compared to racial minorities and women, Whites and men were more likely to be nervous and/or stressed about the pandemic and its consequences.

4 Conclusions

The global COVID-19 pandemic has affected educational experiences of university students in most countries, harmed their employment status as well as that of their family members, and created concerns about physical and mental health. College students' situations and perspectives have been transformed as a result of both the health crisis and the economic impact on the labor market and household conditions.

To learn about how the pandemic affected college students' education experiences, labor market prospects, and physical and mental health, a group of researchers designed the Global COVID-19 Student Survey (GC19SS). The survey asks students about their and their families' employment situations, changes in career considerations, earnings expectations, education experiences and challenges, uncertainty about returning to school, changes in study habits, mental health, and incidence of COVID-19. We analyze the data from the GC19SS by stratifying by country, parental income, gender, and race/ethnicity (US only).

College students' and their parents experienced high rates of job loss during the pandemic, particularly in the US, Spain and Australia and for students from lower-income households. Many graduating seniors had accepted job offers only to have them rescinded. The cancellation of job offers was particularly extensive in Spain (58%), for students from lower-income households (56%), and for Hispanics in the US (36%). In addition, many students had internships planned for the summer, and a large share were cancelled due to the pandemic. The percentages of students with cancelled internships varies across countries, with over half of the internships cancelled for students in Spain, the US, and Sweden. Internship cancellations tended to be more common for women (55%) than for men (50%), but the percentages with internships cancelled does not vary across household income or race/ethnicity (in the US).

The pandemic triggered changes in career considerations as well as expectations about future earnings. Over half of college students across all countries, household income groups, genders, and races/ethnicities (in the US) consider job security, paid sick leave, and flexible

¹⁰Hispanics have similar infection and hospitalization rates to Blacks, however, the COVID19 death rate is similar to Whites. See <https://www.cdc.gov/coronavirus/2019-ncov/COVID-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html>

¹¹White excess mortality, associated with COVID-19, would need to increase by a factor of six to reach the best Black mortality rates outside of the pandemic (Wrigley-Field, 2020).

work arrangements to be more important given the pandemic. We note, however, there are sizable differences across countries, for household income percentiles, between men and women, and for different races/ethnicities. In terms of earnings expectations at the ages of 30 and 45, students in some of the countries expect to be earning above average incomes at either or both 30 and 45 years of age. Exceptions include students from Spain and Australia. Around 70% of students with household incomes above the 80th percentile expect to earn more than the average at both ages. Slightly less than half of students with parental incomes below the 20th percentile expect to be earning less than the average at both ages. Over 60% of men expect to be earning above average incomes at 30 and 45 years old, but only 53% of women expect to be earning above average income at age 30 and 46% expect to earn above average incomes at age 45.

The pandemic induced changes in how students allocate their time. The share of students who reported not working before the pandemic compared with the analogous share during the pandemic increased dramatically across countries, parental income groups, genders, and races/ethnicities (US only). For those who worked before, the vast majority of them report working less after the onset of the pandemic. Students who tended to study more before, however, reduced time allocated to studying. The patterns also hold, with only a few exceptions, across the stratifying variables.

College students in most the countries in the survey report high degree of uncertainty about attending college in Fall 2020. However, that uncertainty is much smaller in Sweden than it is other countries, such as Spain, Mexico, and the US. The percentage of students who are uncertain about returning to school is highest for students with parental incomes at or below the 40th percentile. Likewise, the uncertainty associated with returning to school is also higher among members of minority groups in the US (Black, Asian, Hispanic). The primary reason behind the uncertainty of returning to school is the prospect of the university offering no in-person classes. Loss of financial resources or job losses either for the student or their parents also impacts the uncertainty of returning to school, but this varies by parental income and race/ethnicity (in the US). The main challenge students faced was lack of contact with other students and faculty, but having a noisy place to study and greater family responsibilities are noteworthy challenges in some countries (e.g., Mexico, Spain, and the US), for students from households with lower parental incomes, for women, and for Hispanic students in the US.

In terms of health, the percentage of students having tested positive or knowing someone who tested positive for COVID-19 varies widely across countries, with students in Sweden with the highest rates followed by Spain. The incidence of COVID-19 appears fairly invariant to parental income and gender, but Black and Hispanic students were more likely to have tested positive or know someone who did than White and Asian students. The percentage of students who have an acquaintance or family member that died from COVID-19 is highest in Spain (45%) and Mexico (42%), but the percentage does not vary sizeably across parental income. The rate is slightly higher for women over men (32% vs. 28%) and, in the US, the percentage of Black students who know someone who died from

COVID-19 is 42% versus much lower rates for White (28%), Asian (28%), and Hispanic (32%) students. A large share of students across all countries, parental income groups, genders, and races/ethnicities (in the US) report being worried about their own health or the health of family members. Male students tend to be more nervous and stressed about the pandemic than women, while a larger share of women than men report being calm and relaxed.

As of May 2021 the COVID-19 pandemic caused at least 150 million infections and over 3 million deaths but these consequences have not been uniform across countries, socioeconomic groups or races and ethnicities. In this paper we documented the consequences of the pandemic for university students across seven countries and 29 institutions. Despite the varying penetration of the virus it appears that students across settings suffered from the pandemic in terms of their labor market, educational and health outcomes. On the one hand, the degree of heterogeneity in these consequences was relatively small. On the other hand, however, we detected patterns that if anything will likely deepen the inequalities that existed prior to the pandemic with lower-SES students, females, and minorities bearing the disproportionate burden. We hope that gaps or lack of thereof in some cases, which we identified in this paper will guide university administrators and policy makers in more effectively overcoming the consequences of COVID-19 pandemic for the tertiary education. Our subsequent data collection will also allow us to verify to what degree the uncertainties and worries of students in our sample materialized providing one of the first international panel evidence on student's experiences during this unprecedented health shock.

References

- Albanesi, S. and J. Kim (2021). The gendered impact of the COVID-19 recession on the US labor market. Working Paper 28505, National Bureau of Economic Research.
- Alon, T., S. Coskun, M. Doepke, D. Koll, and M. Tertilt (2021). From mancession to shecession: Women's employment in regular and pandemic recessions. Working Paper 28632, National Bureau of Economic Research.
- Alon, T., M. Doepke, J. Olmstead-Rumsey, and M. Tertilt (2020). This Time It's Different: The Role of Women's Employment in a Pandemic Recession. Working Paper 27660, National Bureau of Economic Research.
- Aucejo, E. M., J. French, M. P. U. Araya, and B. Zafar (2020). The impact of covid-19 on student experiences and expectations: Evidence from a survey. *Journal of Public Economics* 191, 104271.
- Aucejo, E. M., J. F. French, and B. Zafar (2021). Estimating students' valuation for college experiences. Working Paper 28511, National Bureau of Economic Research.
- Barrero, J. M., N. Bloom, and S. J. Davis (2021). Why working from home will stick. Working Paper 28731, National Bureau of Economic Research.
- Blaskó, Z., P. da Costa, and S. V. Schnepf (2021). Learning loss and educational inequalities in europe: Mapping the potential consequences of the COVID-19 crisis. Discussion Paper 14298, IZA.
- Browning, M. H. E. M., L. R. Larson, I. Sharaievska, A. Rigolon, O. McAnirlin, L. Mullenbach, S. Cloutier, T. M. Vu, J. Thomsen, N. Reigner, E. C. Metcalf, A. D'Antonio, M. Helbich, G. N. Bratman, and H. O. Alvarez (2021, 01). Psychological impacts from covid-19 among university students: Risk factors across seven states in the united states. *PLOS ONE* 16, 1–27.
- Chetty, R., J. Friedman, N. Hendren, M. Stepner, and T. O. I. Team (2020). How Did COVID-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data. Working Paper 27431, National Bureau of Economic Research.
- Couch, K. A., R. W. Fairlie, and H. Xu (2020). Early evidence of the impacts of COVID-19 on minority unemployment. *Journal of Public Economics* 192, 104287.
- Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics* 133(4), 1645–1692.
- Fetzer, T., M. Witte, L. Hensel, J. Jachimowicz, J. Haushofer, A. Ivchenko, S. Caria, E. Reutskaja, C. Roth, S. Fiorin, and et al. (2020). Perceptions of an insufficient government response at the onset of the COVID-19 pandemic are associated with lower mental well-being. Technical report, PsyArXiv.
- Hardy, B. L. and T. D. Logan (2020). Racial economic inequality amid the COVID-19 crisis. Technical report, The Hamilton Project.
- Kuchler, T. and B. Zafar (2019). Personal experiences and expectations about aggregate outcomes. *The Journal of Finance* 74(5), 2491–2542.

- Logel, C., P. Oreopoulos, and U. Petronijevic (2021, May). Experiences and coping strategies of college students during the covid-19 pandemic. Working Paper 28803, National Bureau of Economic Research.
- Polyakova, M., V. Udalova, G. Kocks, K. Genadek, K. Finlay, and A. N. Finkelstein (2021). Racial disparities in excess all-cause mortality during the early COVID-19 pandemic varied substantially across states. *Health Affairs* 40(2), 307–316.
- Razai, M. S., H. K. Kankam, A. Majeed, A. Esmail, and D. R. Williams (2021). Mitigating ethnic disparities in covid-19 and beyond. *BMJ* 372.
- Rodríguez-Planas, N. (2020). Hitting Where It Hurts Most: COVID-19 and Low-Income Urban College Students. pp. 25.
- Roth, C. and J. Wohlfart (2020). How do expectations about the macroeconomy affect personal expectations and behavior? *Review of Economics and Statistics* 102(4), 731–748.
- Shaaban, A. N., B. Peleteiro, and M. R. O. Martins (2020). The writing's on the wall: On health inequalities, migrants, and coronavirus. *Frontiers in Public Health* 8.
- UNESCO (2012). World atlas of gender equality in education. Technical report.
- von Wachter, T. (2020). The persistent effects of initial labor market conditions for young adults and their sources. *Journal of Economic Perspectives* 34(4), 168–94.
- Wrigley-Field, E. (2020). US racial inequality may be as deadly as COVID-19. *Proceedings of the National Academy of Sciences* 117(36), 21854–21856.

Federal Reserve communication and the COVID-19 pandemic¹

Jonathan Benchimol,² Sophia Kazinnik³ and Yossi Saadon⁴

Date submitted: 19 May 2021; Date accepted: 25 May 2021

Have the content, sentiment, and timing of the Federal Reserve (Fed) communications changed across communication types during the COVID-19 pandemic? Did similar changes occur during the global financial and dot-com crises? We compile dictionaries specific to COVID-19 and unconventional monetary policy (UMP) and utilize sentiment analysis and topic modeling to study the Fed's communications and answer the above questions. We show that the Fed's communications regarding the COVID-19 pandemic concern matters of financial volatility, contextual uncertainty, and financial stability, and that they emphasize health, social welfare, and UMP. We also show that the Fed's communication policy changes drastically during the COVID-19 pandemic compared to the GFC and dot-com crisis in terms of content, sentiment, and timing. Specifically, we find that during the past two decades, a decrease in the financial stability sentiment conveyed by the Fed's interest rate announcements and minutes precedes a decrease in the Fed's interest rate.

1 This paper does not necessarily reflect the views of the Bank of Israel, the Federal Reserve Bank of Richmond, or the Federal Reserve System. We thank Itamar Caspi (discussant) and participants at the 133rd American Economic Association (ASSA) annual meeting and the Bank of Israel research seminars for their useful comments.

2 Bank of Israel, Research Department.

3 Federal Reserve Bank of Richmond, Quantitative Supervision & Research (QSR).

4 Bank of Israel, Research Department.

Copyright: Jonathan Benchimol, Sophia Kazinnik and Yossi Saadon

1. Introduction

At the outbreak of COVID-19, most central banks declared that they would take “all necessary steps” to mitigate the impact of the pandemic on their economies and decreased their interest rates to the zero-lower bound. Although conventional monetary policy tools have proven almost ineffective in the ensuing unprecedented economic crisis, unconventional monetary policy (UMP) tools have drastically changed both within and between central banks. At the same time, the content, sentiment, and timing of the central banks’ communications regarding their UMP measures and sentiments have changed as well.

However, there are no studies on how the content, sentiment, and timing of the communications of the Federal Reserve (Fed) change. In this paper, we study such changes across three different communication types (namely, Fed fund rate announcements, Federal Open Market Committee minutes, and Fed chairman speeches) during the past two decades and three different economic crises (namely, the GFC, dot-com, and COVID-19 crises). We pay particular attention to the Fed’s communications regarding financial stability and conventional and unconventional monetary policies.

Using state-of-the-art text-mining methodologies, we first analyze the Fed’s communications throughout 2020 to show they were uncertain¹ and heterogeneous during the COVID-19 crisis over time and across communications types. We then analyze how the Fed’s communications relate to the outbreak of the pandemic and subsequently derive potential policy implications from this analysis. Finally, we show that the Fed’s communications and actions have been more reactive to the COVID-19 crisis than to the global financial crisis (GFC) and the dot-com crisis. Indeed, some of the Fed’s communications and UMP-related actions appear to have anticipated the spread of COVID-19. Taken together, our findings show that the Fed’s communication policy has been drastically different during the COVID-19 pandemic than during the GFC and dot-com crises.

Central banks communicate on a variety of topics, through different channels, and with well-defined objectives (Hansen et al., 2019; Benchimol et al., 2020a). Central bank communication aims to inform (e.g., current and future policy objectives and decisions), explain (e.g., past, current, and future economic outlook and decisions), and influence (e.g., current and future uncertainty and financial decisions). These communications are usually published and stored as text (Haldane and McMahon, 2018).

The COVID-19 pandemic has affected all sectors of the global economy (Chetty et al., 2020). In particular, the effect of the pandemic on financial markets and social welfare has led to changes in monetary policy and threatened financial stability (Daly, 2020; Craig et al., 2021). Naturally, central banks have played an influential role during the COVID-19 crisis, and have adapted their communication policies to the current global economy. As in the GFC, central banks have managed the COVID-19 crisis using UMP tools (e.g., forward guidance, quantitative easing, funding and lending

¹ Communications are “uncertain” in the sense that they use nonspecific word such as “approximate,” “contingent,” “indefinite,” and “uncertain.”

facilities, adjustments to market operations, negative or dual interest rates, etc.). Their objective has been to decrease uncertainty and increase financial stability.

In this paper we examine the Fed's communications on conventional and unconventional monetary policies during the COVID-19 crisis by focusing on the most significant types of communications. Specifically, we look at Fed fund rate (FFR) announcements, Federal Open Market Committee (FOMC) minutes (policy decision discussions and deliberations), and speeches given by the Fed chairman. Like most central banks whose primary policy instrument (the nominal interest rate) has a zero or negative lower bound, the Fed's policy stance and subsequent expectation shaping is implemented through other channels, including communications, quantitative easing (QE), balance sheet policies, lending facilities, fiscal and money drops, forward guidance, and other market operations.² Although these decisions are undeniably related to global economic stability, which may involve monetary policy, credibility, and independence risks, the COVID-19 pandemic has led to unprecedented central bank monetary policy decisions. We examine the Fed's communications related to these decisions over the past two decades. This analysis allows us to study whether the Fed successfully implemented clear and transparent communications to support UMP measures addressing the economic challenges caused by the COVID-19 pandemic. Our descriptive analysis shows that Fed's communication was used in a timed and targeted way, showcasing Fed's increasing experience in crisis-specific communication management. The methodologies and technical tools used in this paper, such as R functions and applications to central bank texts, are available in Benchimol et al. (2020b).

In a highly uncertain economic environment such as that of the COVID-19 crisis, the standard for effective central bank communications would normally consist of straightforward and timely updates about current and near-term policy actions. Accordingly, financial stability updates have prevailed in the Fed's monetary policy and financial market-related communications. To proxy for the degree of financial stability conveyed in a central bank communication, we calculate a financial stability score for each relevant communication based on a word count of the terms that can also be found in the financial stability dictionary (Correa et al., 2021).

We study how monetary policy and financial stability are considered in central bank communications over time, from the dot-com crisis to the COVID-19 pandemic. Specifically, we analyze how the Fed's communication policy during the COVID-19 pandemic compares to that of the dot-com and GFC crises. The correlation of the sentiment and uncertainty of the Fed's communications to economic and financial variables, unconventional monetary policies, and the COVID-19 pandemic is also investigated. Finally, we focus on how the Fed's UMP-related communications have evolved during the past two decades.

We find that the content, timing, and sentiment of the Fed's communications exhibit noteworthy differences conditional on the crisis. Since the GFC, communications regarding UMP have become the "new normal," as reflected in all three communication types, namely, FFR announcements, FOMC minutes, and Fed chairman speeches. COVID-19 appears to have caused structural changes in the Fed's

² See, e.g., Bianchi et al. (2020) and Guerrieri et al. (2020).

communication content. We also find evidence for a link between conventional monetary policy and financial stability sentiment.

The remainder of the paper is organized as follows. Section 2 describes the data. Section 3 presents the results of the text analysis. Section 4 presents the results of the sentiment analysis. Section 5 presents the results of the topic modeling of the past and current economic situations. Section 6 examines the Fed's communications on unconventional monetary policy. Section 7 discusses the Fed's early communications on the pandemic. Section 8 compares the Fed's conventional monetary policy to its financial stability sentiment over the past two decades. Section 9 derives some policy implications, and Section 10 concludes. Section 12 presents the dictionaries used to analyze UMP and COVID-19 and additional results. The methodologies used in this paper are presented in Benchimol et al. (2020b).

2. Data

2.1 Text Data

Our study focuses on the most significant types of Fed's communications intended for public consumption. We gathered 776 of these communications for the period 2000–2020. The sample contains formal communications detailing monetary policy discussions (FFR announcements and FOMC minutes) as well as less formal communications (Fed chairman speeches).

Our dataset is summarized in Table 1. In addition to the texts, we use dictionaries for our text analysis purposes: namely, a well-known finance dictionary (Loughran and McDonald, 2011), a financial stability dictionary (Correa et al., 2021), an UMP dictionary (Christensen and Rising, 2017; Henry, 2008), and our own COVID-19 dictionary.³

Loughran and McDonald (2011) is a dictionary developed to measure the sentiment of financial texts better than general dictionaries. This dictionary is widely used in text analyses in the finance and economics literatures (Loughran and McDonald, 2016; Benchimol et al., 2020a). Loughran and McDonald (2011) also developed a dictionary to measure the uncertainty conveyed in financial texts.

Table 1. Descriptive Statistics: Federal Reserve Texts

	No. Texts	No. Words (average)	Sample
FFR Announcements	181	400	2000–2020
FOMC Minutes	170	6809	2000–2020
Chairman Speeches	425	2931	2000–2020
Total	776	3213	2000–2020

Notes: "Total" refers to the sum of all three communication types.

Sources: The Federal Reserve Board of Governors and FederalReserve.gov archives.

³ Our UMP and COVID-19 dictionaries are presented in the Appendix.

Correa et al. (2021) construct a dictionary explicitly tailored to financial stability contexts. This dictionary classifies words as positive or negative based on the sentiment they convey in financial stability reports.

For the UMP dictionary, we rely on a dictionary that translates central bank communications about future monetary policy into groups of positive and negative words (Christensen and Rising, 2017), and merge this dictionary with a more market-related dictionary (Henry, 2008). These two dictionaries are analyzed and compared in Erasmus and Hollander (2020).

Finally, we construct a COVID-19 dictionary by compiling relevant keywords that relate to the pandemic. We use this dictionary to capture the frequency (or “intensity”) of words associated with the COVID-19 pandemic in the Fed’s communications in order to identify virus-related content in those communications.

We describe how we apply these dictionaries to our sample in Section 3. See the Appendix for more details about the UMP and COVID-19 dictionaries.

2.2 COVID-19 Data

The database for the COVID-19 statistics is the COVID-19 Data Repository⁴ maintained by the Johns Hopkins University Center for Systems Science and Engineering (CSSE), with the support of the ESRI Living Atlas Team and the Johns Hopkins University Applied Physics Lab (APL). In our analysis, we essentially utilize the daily number of new COVID-19 cases.

2.3 Financial Data

The daily financial dataset used in this paper is collected through Bloomberg. It includes the SP500 equity index, the CBOE VIX, the nominal effective exchange rate (broad), and the nominal interest rate (FFR).

3. Methodology

This study aims to capture the change and impact of the Fed’s communications during the COVID-19 pandemic.⁵ To this end, we build text-based measures of uncertainty and sentiment in the Fed’s communications by utilizing an array of custom dictionaries, as described in the previous section.

We use three text-mining techniques: word counting, sentiment scoring, and topic modeling. First, we use simple word-counting procedures. Specifically, we count the terms related to UMP and COVID-19 that appear in the Fed’s communications. Second, we use sentiment scoring. This supervised machine-learning method allows us to measure sentiments conveyed by the Fed’s communications. Specifically, we use the Loughran and McDonald’s (2011) dictionary to proxy for sentiment and uncertainty in the Fed’s communications and build several sentiment scores and polarity indicators based on general (NRC, SentiWords, Hu&Liu, Jockers) and specialized (financial stability, UMP) dictionaries. Third, we use topic modeling. This unsupervised machine-learning method allows us to extract and examine the thematic

⁴ The full dataset can be downloaded from <https://github.com/CSSEGISandData/COVID-19>

⁵ Central bank communications are defined as significant communications such as FFR announcements, FOMC minutes, and Fed chairman speeches.

content of the Fed's communications. Specifically, we use the latent Dirichlet allocation (LDA) algorithm to compare the content of Fed's communications to economic and financial developments.

The results presented below are obtained with the statistical software R. The R functions and packages used in this paper are described in Benchimol et al. (2020b).

3.1 Word Counting

To estimate the amount of coverage of the COVID-19 pandemic in each of the Fed's communications, we construct a COVID-19-specific dictionary, presented in the Appendix. By counting the number of COVID-19-related words in each Fed communication, we can estimate how the Fed perceived the severity of the pandemic at that time.

In addition, we construct another dictionary that captures communications regarding the Fed's UMP measures by merging two existing dictionaries, those of Erasmus and Hollander (2020) and Henry (2008). We then calculate the overall sentiment related to UMP measures using this merged dictionary.

Finally, we explore whether the onset of the COVID-19 pandemic impacted communication clarity.

3.2 Sentiment Scoring

To measure the sentiment of the Fed's communications, we implement several methods of capturing text sentiment. The general methodology is based on counting positive and negative words according to a specific dictionary. The sentiment score is calculated by dividing the difference between the number of positive and negative words by the total amount of words.

To this end, we use Correa et al.'s (2021) dictionary, which is specifically tailored to capture financial stability sentiment. Correa et al. (2021) explain movements in financial cycle indicators related to credit, asset prices, systemic risk, and monetary policy rates and classify words (positive/negative) based on the sentiment conveyed in financial stability reports. The second dictionary focuses on forward-guidance and quantitative measures (Christensen and Rising, 2017) that we merge with another dictionary that focuses more on the regulatory context, structural attributes, and dual informational-promotional role of earnings press releases (Henry, 2008).

We use another set of dictionaries to capture different dimensions of the sentiment expressed in the text, such as the Loughran and McDonald (2011) sentiment dictionary, as well as a set of commonly used sentiment dictionaries in the text-mining literature, such as the Jockers, NRC, and Hu&Liu dictionaries. We use these dictionaries in conjunction with the so-called valence shifters (i.e., negators, amplifiers/intensifiers, de-amplifiers/downtoners) to capture nuances in the sentiment of the relevant text.

Last but not least, we construct two types of sentiment indicators based on the Loughran and McDonald (2011) dictionary. One is the standard score measure described above. The other is a polarity measure that includes the possibility of neutral, positive, negative, very positive, or very negative sentiment, according to the sentiment of the words immediately preceding and following the considered word.

Section 4 presents sentiment scores produced from these dictionaries for FFR announcements, FOMC minutes, and Fed chairman speeches. For these three communication types, we note a sharp decrease in sentiments in the first quarter of 2020, as well as a spike in uncertainty-related words during that same period. This finding suggests that the Fed's communications reflect its willingness to address the ongoing developments during the COVID-19 pandemic proactively.

3.3 Topic Modeling

Our goal in this section is to identify underlying themes that drive the Fed's communications and to capture theme prevalence over time. For this purpose, we rely on the so-called topic modeling approach. This algorithm allows us to identify a small number of verbal themes that best explain thematic variation over time, and we use it to capture the Fed's assessments of economic and financial risks in real-time.

Topic modeling is an unsupervised machine-learning technique that does not require any training or dictionary-based analysis. Here, we use LDA, which works in the following way. The LDA algorithm views each document as a mixture of topics that are present in the body of the text. The algorithm scans a set of relevant documents (FFR announcements, FOMC minutes, and Fed chairman speeches), detects words and phrases within them, and automatically clusters word groups (i.e., topics) that best characterize a set of documents. In essence, the algorithm identifies the different topics represented in the document, and calculates the prevalence of each of them (Blei et al., 2003).

In the following Section 4, we present the results of the word-counting and sentiment-scoring analysis and interpret them by providing a descriptive picture of the sentiments the Fed's communications conveyed. Section 5 presents the results of the topic modeling analysis. Overall, we find that at the onset of the COVID-19 pandemic, the topics related to policy intervention gained prominence, at the expense of other topics. We relate UMP to our various indicators and the Fed's communication types in Section 6 and control for the severity of the outbreak by including the number of COVID-19 cases at the time of the communication in Section 7. We also describe and relate our text-mining indicators to financial stability and interest rate decisions over the past two decades in Section 8.

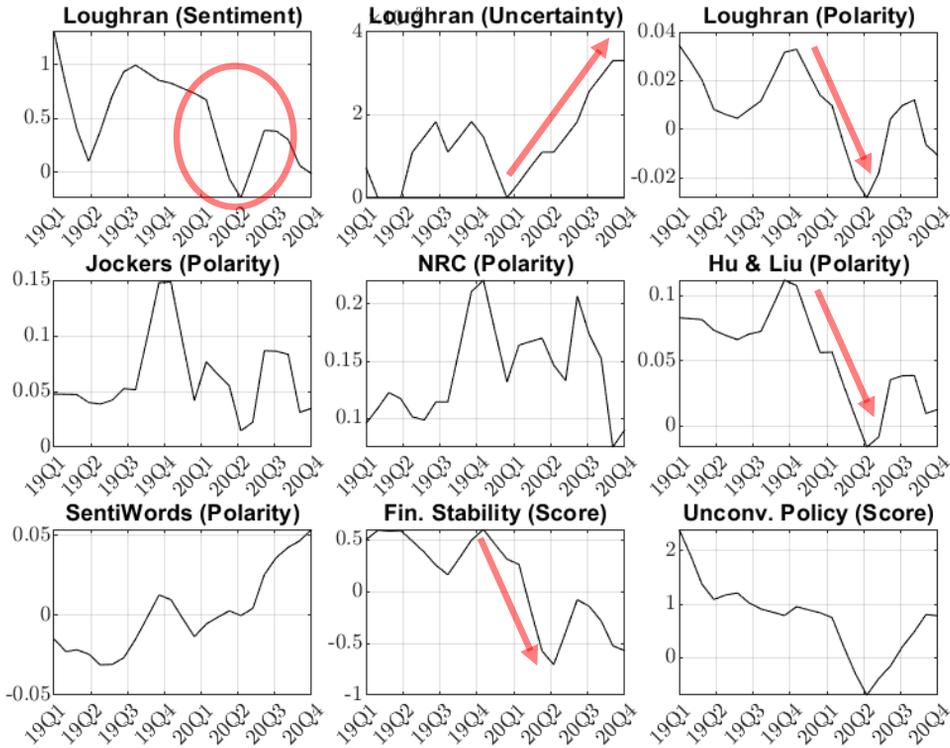
4. Sentiment Analysis

Figure 1 presents several sentiment indicators extracted from the text analysis of FFR announcements. The sentiments sharply degraded following the COVID-19 outbreak in China in January 2020 and the US in March 2020. This date range corresponds to an increase in the contextual uncertainty indicator (the number of words reflecting uncertainty, scaled by text length).

Based on the Loughran and McDonald (2011) dictionary, our new polarity indicator decreases but displays a more optimistic sentiment for 2020Q3. FFR announcements summarize the current state of the economy and monetary policy decisions. Figure 1 shows that the shock that occurred in January 2020 lasts up until April 2020.

The sharp decrease in sentiments in 2020Q1-Q2, and the increase in the contextual uncertainty in 2020Q1-Q3, correspond to the beginning of the COVID-19 crisis.

Figure 1. Sentiment Scores in FFR Announcements



Notes: Solid black lines represent sentiment score values. Red arrows represent trends related to the COVID-19 outbreak.

Figure 1 also shows that the SentiWords, Jockers, and NRC polarity indexes are (apparently) less informative, while the Hu&Liu polarity index displays similar dynamics to our Loughran and McDonald-based polarity index. SentiWords, a high-coverage polarity index, captures an interesting increase in the sentiment of FFR announcements from the beginning of the COVID-19 crisis. A potential explanation for this increase is that the Fed used a different communication strategy in this crisis than in the GFC and dot-com crisis.⁶ The financial stability sentiment has been decreasing into negative territory since 2019Q4, which means that more negative financial stability-related words were present in the FFR announcements than positive ones.

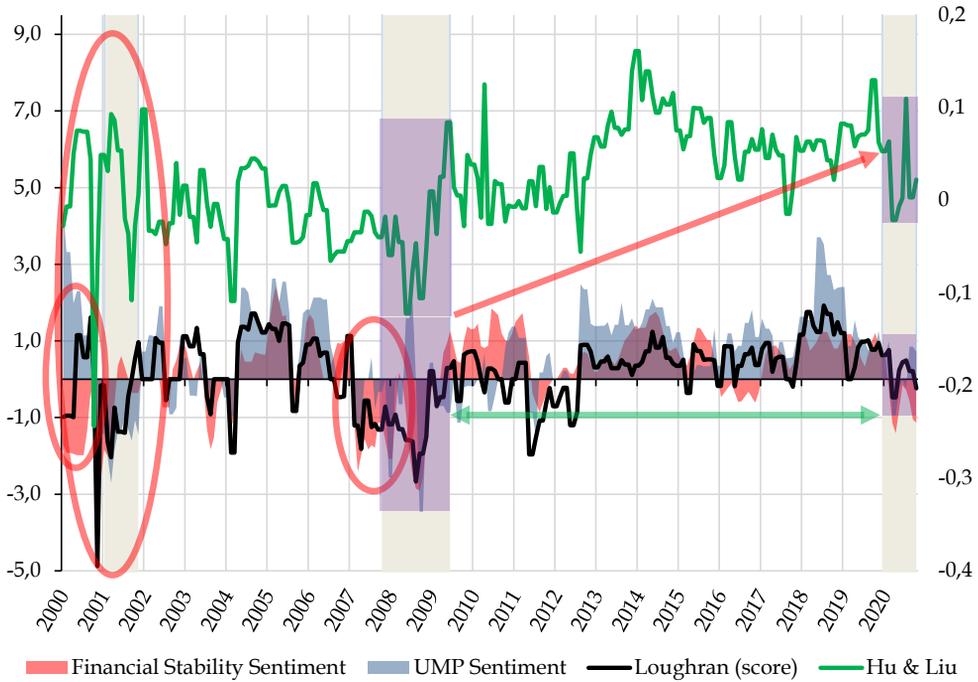
Interestingly, the UMP sentiment decreased mainly because this indicator includes forward-guidance sentiment but also words found in standard dictionaries.

⁶ SentiWords polarity index sharply decreases for the GFC and dot-com crisis. The results of the full sample are available in the Appendix (Figure A1).

Note that these results are based on FFR announcements, which are the most supervised and controlled⁷ communication type.

Figure 2 presents selected sentiment indicators over the full sample. As proxied by our sentiment measures, the Fed’s communications differ significantly in sentiment during the COVID-19 crisis, compared to the GFC and the dot-com crisis.

Figure 2. A Tale of Three Crises: Sentiment.



Notes: The gray shaded area represents the NBER recessions periods. Solid lines represent sentiment scores computed from FFR announcements.

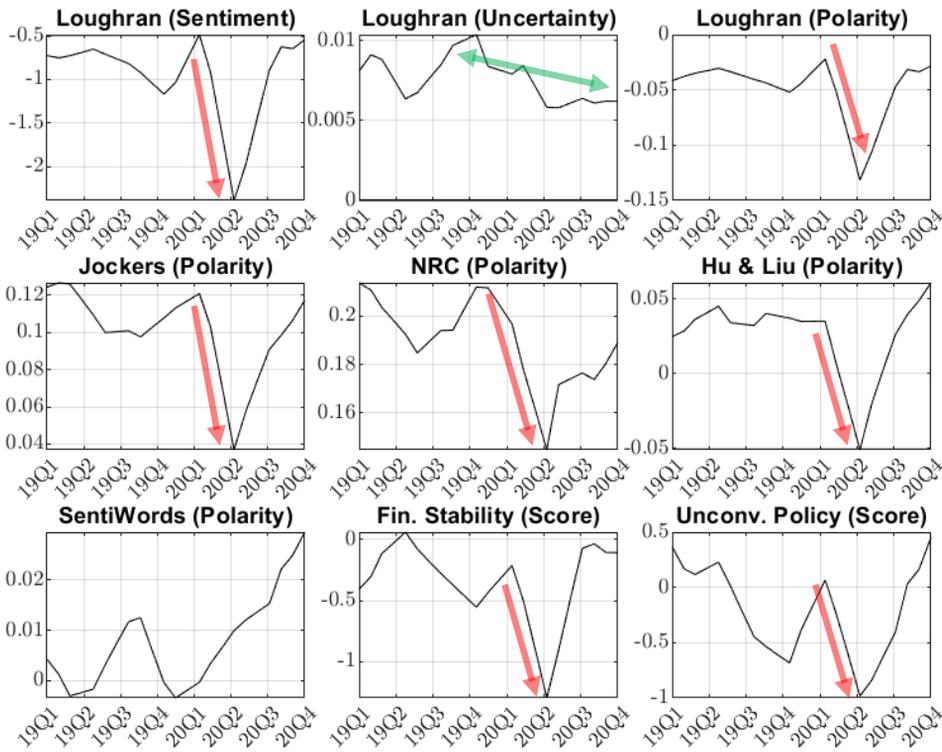
The financial stability sentiment sharply deteriorated before the GFC and dot-com crisis, whereas it is significantly positive before the COVID-19 pandemic. The FFR announcements appear to be a significant predictor of future conventional monetary policy (see Section 8).

The Hu and Liu sentiment polarity index improved from the GFC until the COVID-19 crisis. It decreased less in the COVID-19 crisis than in the other crises. While the Loughran and McDonald sentiment index did not improve between the crises, it decreased less in the COVID-19 crisis than in the other crises. Also, the volatility of these indicators was less pronounced in the COVID-19 crisis than in the GFC and dot-com crisis. Although Figure 2 presents the tale of three crises, it also clearly shows a tale of three communication policies.

⁷ By the spokesperson, and other Fed’s departments or officials.

Figure 3 presents the same indicators for the FOMC minutes. The dynamics for almost all sentiment indicators display a sharp deterioration in 2020Q2, which is much more pronounced for the FOMC minutes than for the FFR announcements.

Figure 3. Sentiment Scores for the Fed’s FOMC Minutes



Notes: Solid black lines represent sentiment score values. Red arrows represent trends related to the COVID-19 outbreak.

This probably reflects the gap between the description of the current economic situation (FFR announcements) and the discussions and tentative solutions to the COVID-19 crisis (discussed in the minutes). The Loughran and McDonald score and polarity indexes showed a sharp decrease in sentiment related to financial uncertainty from January to April 2020.

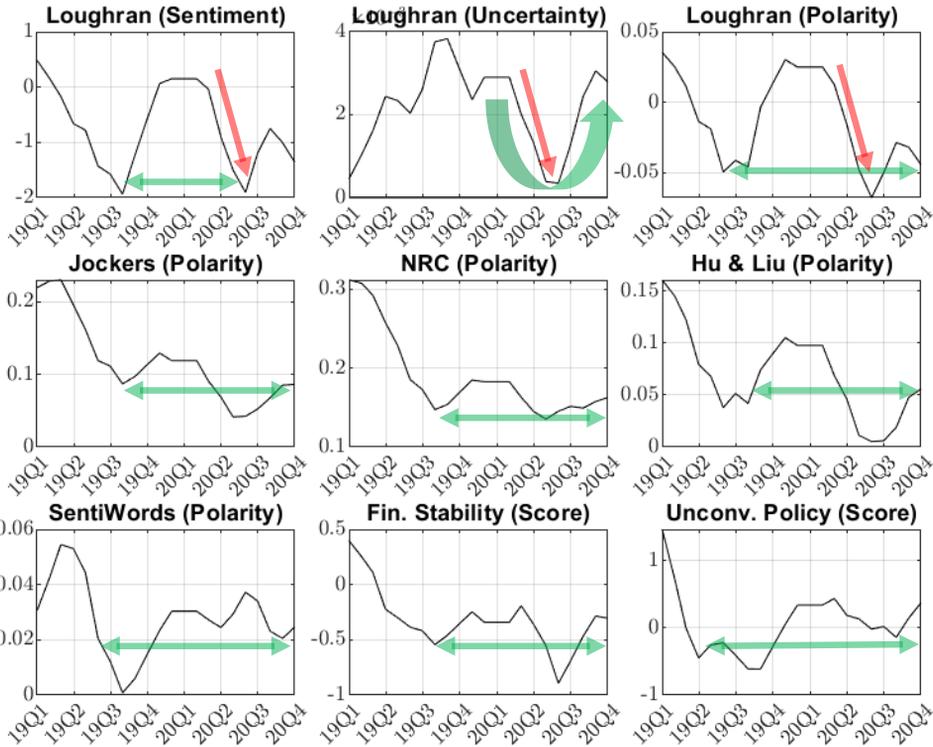
The UMP sentiment score decreases until 2020Q2 and then sharply increases until it becomes positive. This phenomenon corresponds to the more positive language adopted in FOMC minutes regarding UMP steps taken during the COVID-19 crisis. Figure 3 also shows that according to FOMC minutes (i.e., according to policymakers during monetary policy committee discussions), financial stability was perceived to be at risk in 2020Q2. This was effectively the case in reality but less so than during the GFC, as explained in the Appendix (Figure A5).

The high coverage of the SentiWords index captures an interesting pattern of increasing sentiment from the beginning of the COVID-19 crisis. This may result from the Fed’s communication strategy to calm and reassure economic agents with the use

of more positive words. This was not the case during the GFC, where the SentiWords indicator sharply declined to historically low levels.⁸

Figure 4 presents the sentiment indicators for the official speeches of the chairman of the Federal Reserve.

Figure 4. Sentiment Scores for Fed Chairman Speeches



Notes: Solid black lines represent sentiment score values. Red arrows represent trends related to the COVID-19 outbreak. Green arrows represent trends that remain constant before and after the COVID-19 outbreak.

The sentiment conveyed by Fed chairman speeches decreases less in comparison to the minutes and announcements. A potential explanation is that the speeches might be aimed at managing expectations more than the minutes and announcements are. Although the economic situation worsened and sentiments degraded from February 2020 onward, contextual uncertainty decreased.

The sentiment conveyed by the Fed chairman speeches is generally more volatile than the sentiment conveyed by FFR announcements and minutes (Benchimol et al., 2020a). However, the small sample makes the sentiment indicators for the COVID-19 crisis presented in Figure 4 less volatile than those for the GFC.⁹

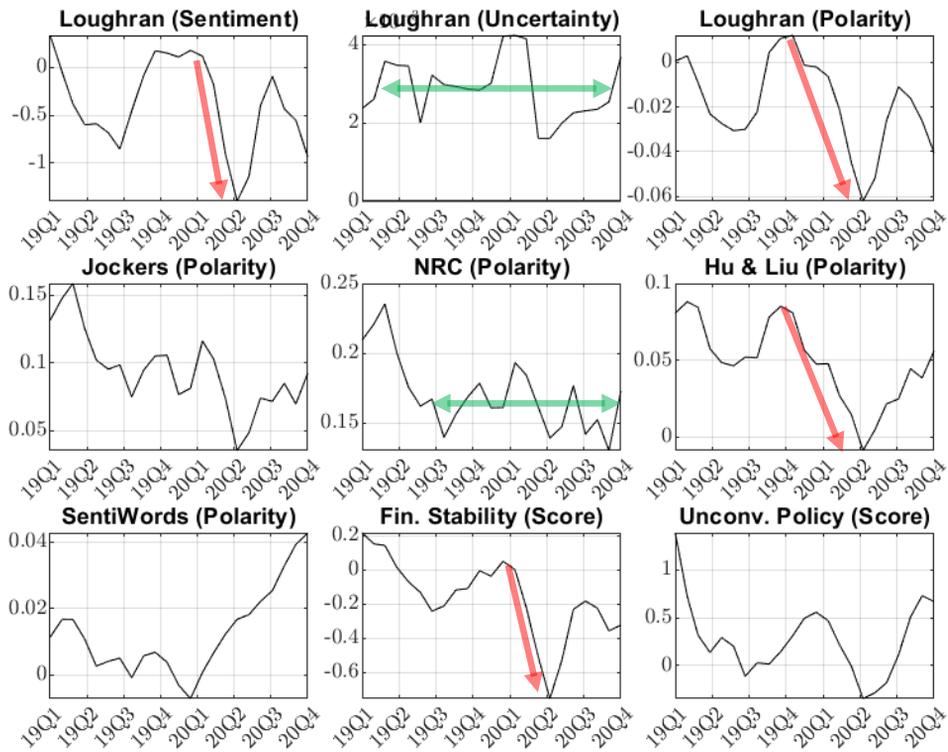
⁸ The results of the full sample are available in the Appendix (Figure A2).

⁹ The results of the full sample are available in the Appendix (Figure A3).

As for the FFR announcements and minutes, the large scope of the SentiWords dictionary captures this specific Fed’s communication policy held during the COVID-19, which seems to have been in force at least up until 2020Q4. The Loughran and McDonald dictionary also shows a decrease in the use of uncertainty-related words in the Fed chairman speeches since the COVID-19 outbreak in China, which may also result from a specific communication policy.

Figure 5 presents aggregated sentiment indicators based on FFR announcements, FOMC minutes, and Fed chairman speeches.

Figure 5. Sentiment Scoring of Main Fed Communications



Notes: Solid black lines represent sentiment score values. Red arrows represent trends related to the COVID-19 outbreak. Green arrows represent trends that remain constant before and after the COVID-19 outbreak.

We provide a global picture of the Fed’s communications by aggregating all communication types in our dataset. Figure 5 shows a sharp decrease in the uncertainty sentiment, which is mainly driven by minutes and speeches. The financial stability sentiment sharply decreases from January to April 2020, which may be correlated with volatility measures such as the CBOE VIX (see Section 7).

The results presented in Figures 1–5 point to the same conclusion. Namely, it is highly likely that the Fed has globally implemented the same communication policy across all communication types (FFR announcements, FOMC minutes, Fed chairman

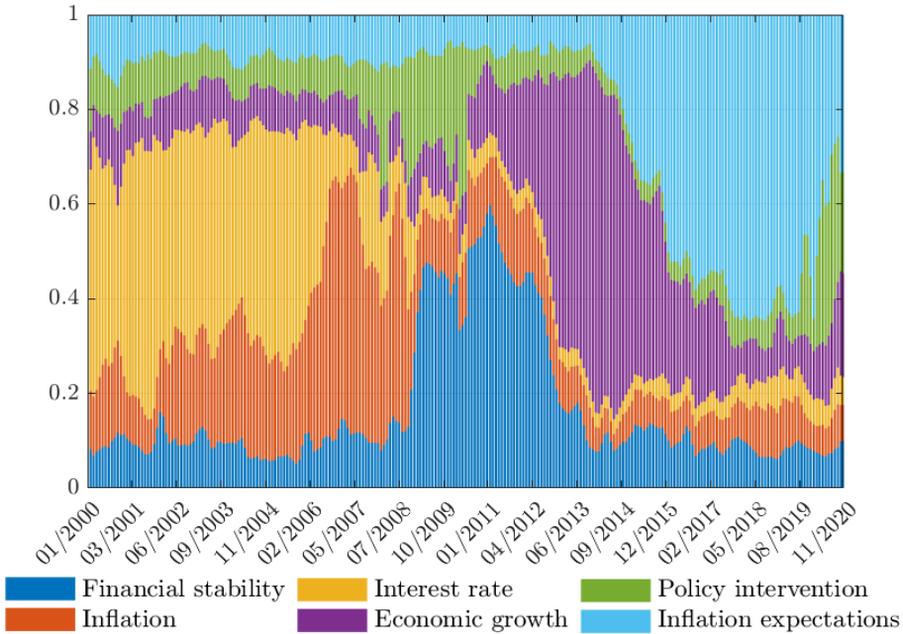
speeches) during the COVID-19 crisis. These results also demonstrate the different sentiments involved in the COVID-19 crisis compared to the GFC.¹⁰

This section covers the special sentiment deterioration that occurred from January to April of 2020. The recovery of sentiment into positive territory took place following this deterioration, with some sentiment measures rising above pre-crisis levels. Looking at the pronounced differences in sentiment over time, we find evidence that the Fed’s communications were used to shape the narrative and manage expectations.

5. Topic Modeling

In this section, we focus on the topics extracted from our sample of texts. It is important to note that the topic modeling methodology does not use any predetermined dictionaries. In contrast to sentiment analysis, it is a more structural and unsupervised approach to interpreting word-topic linkages in texts.

Figure 6. A Tale of Three Crises: Topics.



Note: Bars represent the topic probability computed from FFR announcements. For clarity and robustness, we restrict attention to the six most frequently discussed topics.

Source: Benchimol et al. (2020a).

Figure 6 presents the six topics extracted from FFR announcements over the past two decades. It shows that discussion of policy interventions was more pronounced during the COVID-19 than during the other crises. Interestingly, Figure 6 shows that

¹⁰ The results of the full sample are available in the Appendix (Figure A4).

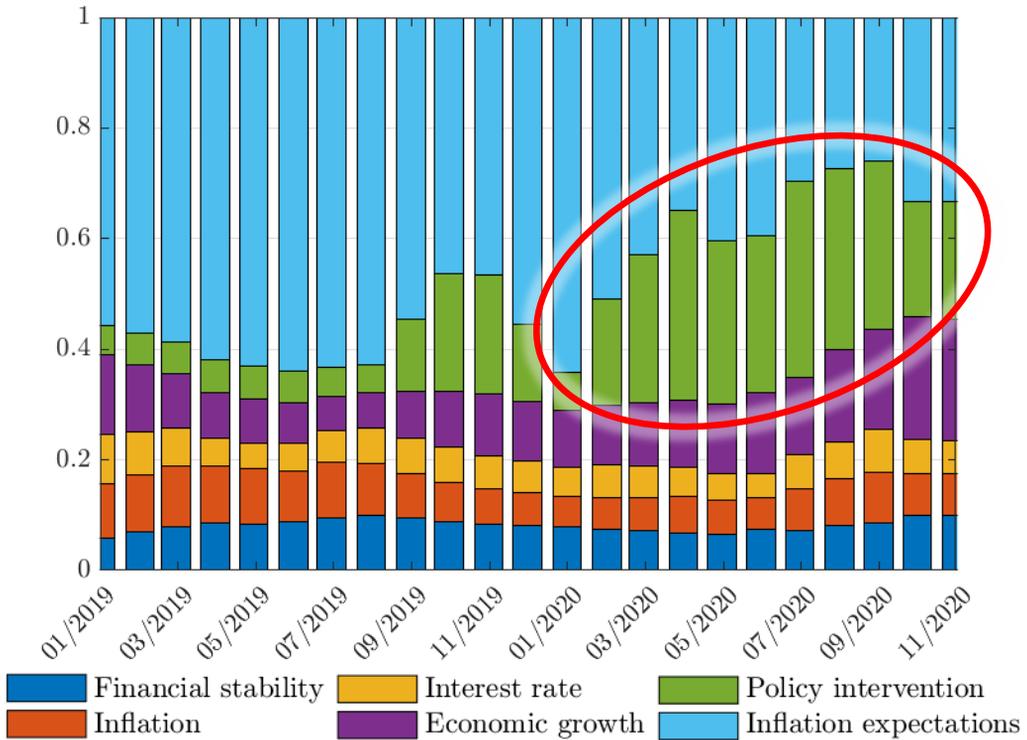
Covid Economics 79, 27 May 2021: 218-256

the topic of inflation expectations decreased in importance while the topic of economic growth, which includes economic growth considerations and concerns, increased.

Another interesting observation from Figure 6 is related to each crisis’s relative influence on the Fed’s communications. While the dot-com crisis had almost no effect on the topics conveyed to the public in the FFR announcements, the GFC and COVID-19 crises strongly shaped the topics conveyed in these announcements.

Figure 7 shows that the probability that policy intervention was discussed in FFR announcements significantly increased after the COVID-19 outbreak in China. However, this topic had begun to increase earlier, which indicates that it may be at least partly related to previous concerns unrelated to COVID-19.

Figure 7. Topic Analysis of FFR Announcements



Notes: For clarity and robustness, we restrict attention to the six most frequently discussed topics.

A sharp decrease in the topic probability of inflation expectations coincides with the COVID-19 outbreak, which is in line with monetary policy considerations at that time, when the focus shifted to policy intervention.

To a lesser extent, inflation is less discussed, and economic growth more discussed, in FFR announcements during the COVID-19 crisis.

The increase in the topic probability of policy intervention in FFR announcements decreases the topic probability of inflation expectations and, to a lesser extent, inflation. Although the Fed’s main objective is to stabilize prices, this

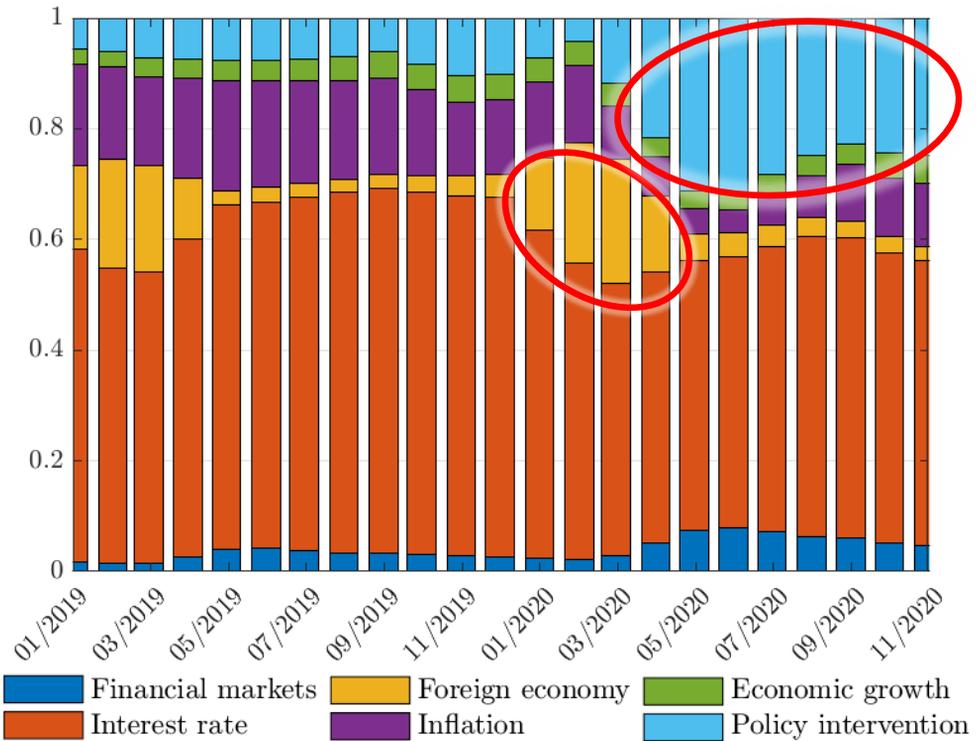
Covid Economics 79, 27 May 2021: 218-256

finding demonstrates that its FFR announcements were less related to inflation concerns after the COVID-19 outbreak.

The topic of economic growth slightly increased after the COVID-19 outbreak in China. This finding indicates the Fed’s concern that a pandemic would pose a threat to economic growth.

Figure 8 presents the topic analysis of FOMC minutes. It shows that the COVID-19 outbreak significantly shaped the discussions of the Federal Open Market Committee.

Figure 8. Topic Analysis of FOMC Minutes



Notes: For clarity and robustness, we restrict attention to the six most frequently discussed topics.

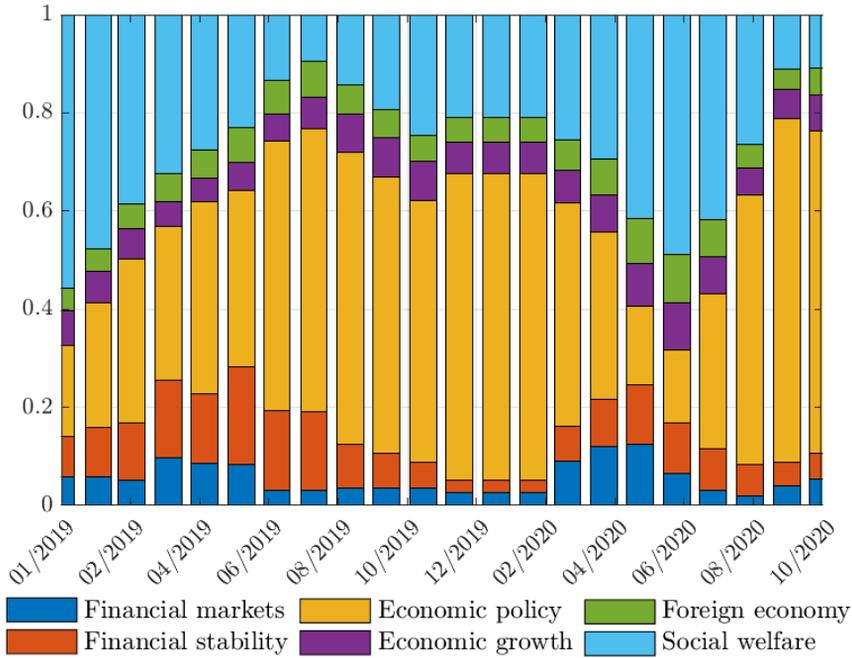
Interestingly, like the FFR announcements, the FOMC minutes are also influenced by the topic of policy intervention, even though the interest rate is the most prominent topic. The probability of discussion of inflation also decreased for the FOMC minutes due to the COVID-19 outbreak, while discussions of policy intervention and financial markets increased.

The topics conveyed by the Fed’s FOMC minutes reflect a sharp increase in coverage of policy intervention and foreign economy. However, the coverage of foreign economy had begun to increase even before the COVID-19 outbreak while inflation and interest rate topics had begun to decrease.

Figure 9 presents the topic analysis of Fed chairman speeches. These speeches focused on social welfare concerns after the COVID-19 outbreak, similar to the pre-

crisis concerns about education and inequality in the US. This finding shows that Fed chairman speeches are often devoted to issues unrelated to its primary objective of stabilizing prices, such as education, health, and development economics, including family and labor markets.¹¹

Figure 9. Topic Analysis of Fed Chairman Speeches



Notes: For clarity and robustness, we restrict attention to the six most frequently discussed topics.

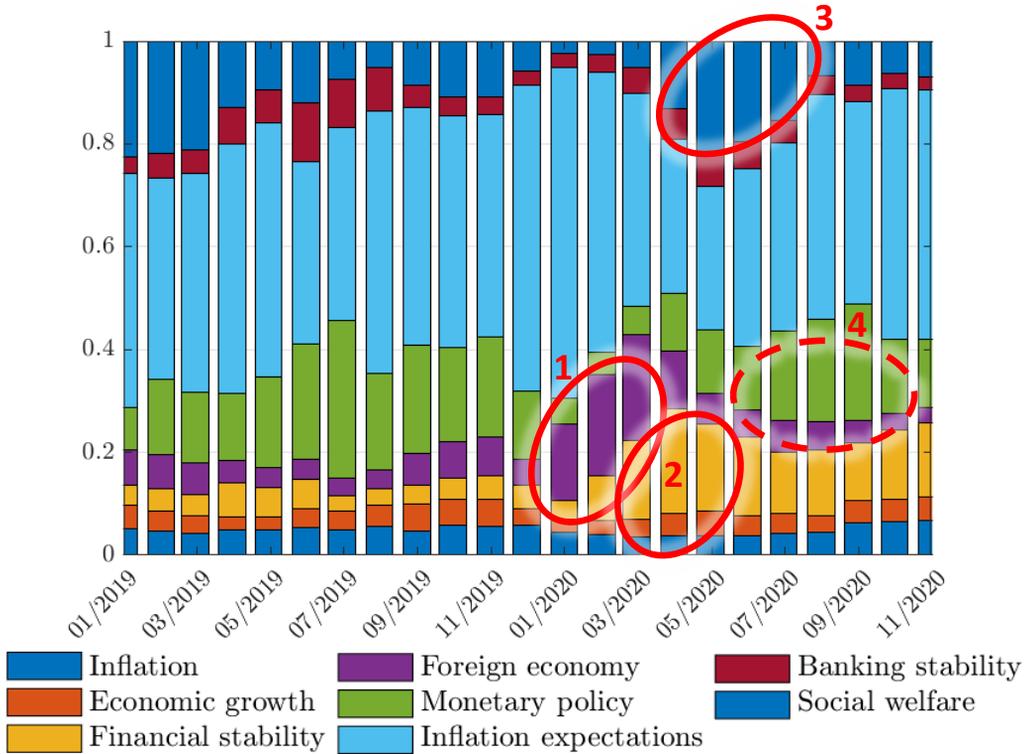
The fact that the speeches are less supervised and held to a broader audience than traditional FFR announcements and FOMC minutes, which are more focused on inflation and output growth, may explain the increase in discussion of social welfare issues. We also observe an increase in discussions of economic policy after June 2020. This may be due to COVID-19 spillovers, but we cannot reject the US election effect. Before the COVID-19 outbreak, economic policy considerations occupied the attention of most Fed chairman speeches.

Figure 10 presents the topics discussed in the Fed’s communications in the aggregate. We provide a global picture of the topic modeling of the Fed’s communications by aggregating all communication types in our dataset, namely, FFR announcements, FOMC minutes, and Fed chairman speeches. Because of the considerable number of texts analyzed, and their respective, often different, characteristics, we were constrained to include a larger number of topics for the global

¹¹ The most frequently used words and word fragments (root words) in the context of the topic of social welfare are *communiti, economi, educ, work, develop, research, busi, job, peopl, mani, help, opportun, can, import, family.*

topic modeling. Taken together, these topics include most of the topics described in Figures 5 to 7.

Figure 10. Topic Analysis of Main Fed Communications



Notes: For clarity and robustness, we restrict attention to the six most frequently discussed topics.

Note that the monetary policy topic contains some references to UMP and unemployment, which may be attributed to the Fed’s dual mandate.

The topic of inflation expectations continues to attract most of the Fed’s attention overall, which the COVID-19 crisis and long-term interest rate concerns may have reinforced.

Figure 10 shows three prevalent topics that emerged in quick succession. First, the Fed’s communications regarding the foreign economy increased as the COVID-19 pandemic spread from China to the rest of the world (1). Second, the Fed’s communications regarding financial stability increased as fears about the financial system due to the impact of the COVID-19 crisis increased (2). Third, the Fed’s communications regarding social welfare increased as the potential need of Americans for additional relief plans from the government and the Fed increased (3). The Fed’s communications regarding conventional monetary policy (especially average inflation targeting) and UMP decreased at the beginning of the COVID-19 pandemic but increased thereafter (4).

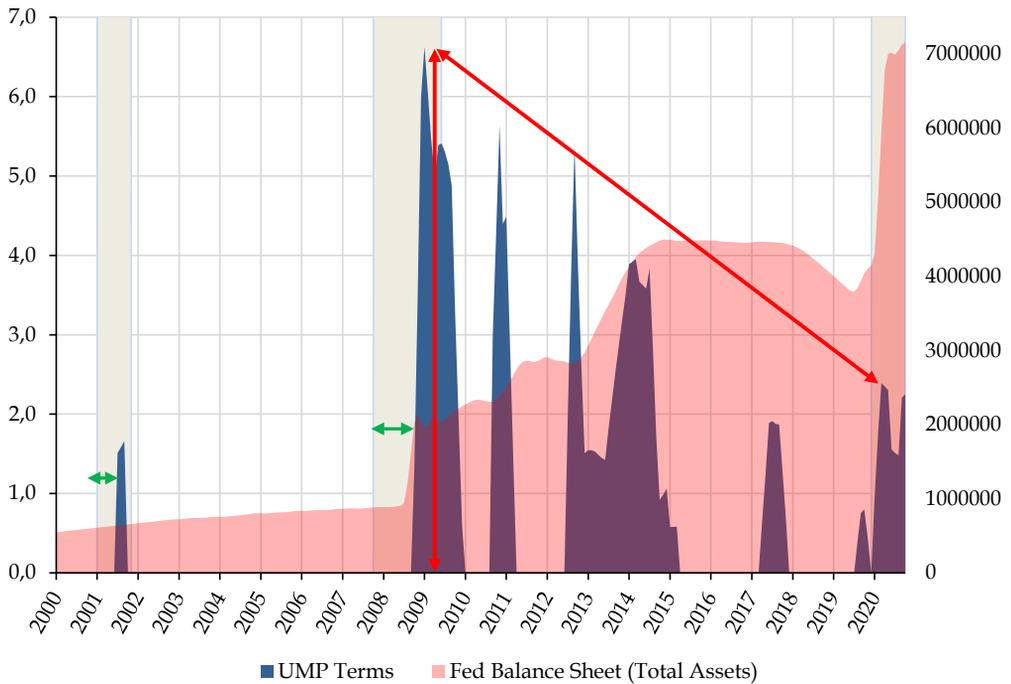
Overall, the content and timing of the Fed’s communications exhibit differences across the three crises discussed above. Unlike the GFC and dot-com crises, the COVID-19 crisis has caused the focus of the Fed’s communications to shift away from discussions of inflation expectations to discussions of policy intervention.

Interestingly, the topic of policy intervention is much more prevalent in the Fed’s communications during the Covid-19 crisis, compared to GFC and dot-com crises. It seems that the policymakers not only implement policy interventions but discuss these interventions differently across the crises, and that the content, sentiment, and timing of these communications are conditioned on the crisis.

6. Unconventional Monetary Policy

This section analyzes the link between the Fed’s communications, its actions, UMP, and COVID-19. Figure 11 compares UMP terms with UMP measures as reflected by the Fed’s balance sheet.

Figure 11. Unconventional Monetary Policy in FFR Announcements



Notes: The gray shaded area represents NBER recession periods. The red shaded area represents total assets (minus eliminations from consolidation) in the Fed’s balance sheet in millions of US dollars. The blue shaded area represents the word-counting indicator based on our UMP dictionary presented in the Appendix. The red arrows indicate UMP terms’ intensity from the onset of the GFC to the outbreak of COVID-19. The green arrows indicate the timing of the Fed’s UMP communications.

Source: Board of Governors of the Federal Reserve System (US).

Covid Economics 79, 27 May 2021: 218-256

Figure 11 shows that the Fed communicated more extensively about UMP during the GFC than during the COVID-19 crisis. However, it is important to note the timing. The Fed communicated about and acted against the GFC after a delay of nine months, whereas it hastened to do so during the COVID-19 crisis and the communications and actions were more clearly coordinated.

Figure 11 shows that the Fed's communications regarding UMP during the COVID-19 crisis (according to a word count of the terms listed in our UMP dictionary in the Appendix) correspond to effective UMP measures that led to Fed balance sheet changes with several lags.

It is notable that whereas actions were implemented *before* they were communicated during the GFC, they were implemented *after* they were communicated after the GFC and during the COVID-19 crisis.

Another feature of our dictionary is that it can be used to identify other periods beyond the above-noted crises when substantial UMP measures were taken to support the US economy. Indeed, as can be seen in Figure 11, each communication peak related to a UMP measure influenced the Fed's balance sheet shortly after the communication shock. The communications about UMP in 2013–2014 were devoted to conveying the message that these expansionary policies would cease (and effectively did so according to the Fed's balance sheet), thus proving that our UMP dictionary captures tapering communication policies.

The Fed implemented this gradual reversal of QE easing policies to mitigate economic growth expectations. The "tapering" effectively started in 2013 when Ben Bernanke, the Fed chairman at the time, commented that the Fed would lower the amount of purchased assets each month if economic conditions, such as inflation and unemployment, continued to be favorable.¹²

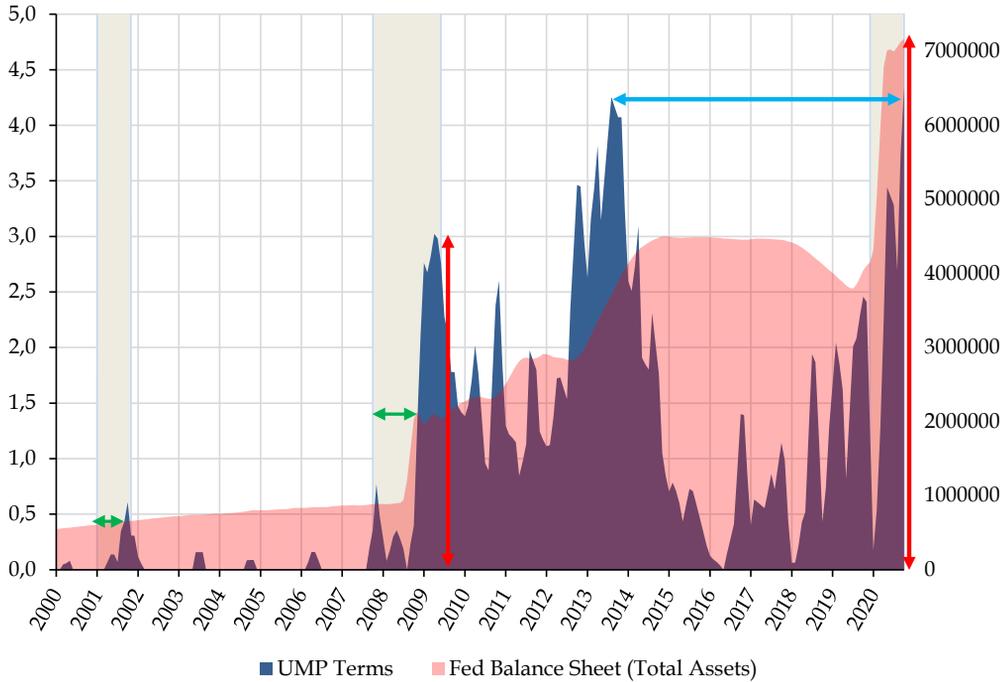
Figure 12 shows that FOMC minutes discuss UMP actions relatively earlier for the COVID-19 crisis than for the GFC. The quantity of these UMP discussions is similar to tapering discussions held by the FOMC in 2013, and the overall level is higher in late 2020 compared to the GFC period.

It is worth noting the difference between the FFR announcements and the FOMC minutes with respect to UMP. Although FOMC minutes are less supervised and longer than FFR announcements, pre-COVID-19 crisis UMP FOMC discussions were more intense than pre-GFC or even during the first half of the GFC, a behavior confirmed by comparing Figures 11 and 12.

In summarizing the discussions held between monetary policy committee members, FOMC minutes typically contain more UMP terms (discussions or controversies about potential solutions or policy implementations) than FFR announcements do. It is interesting to note that such terms have remained in frequent use since the GFC.

¹² "Tapering" refers to gradually reducing the Fed's asset purchases, not altogether eliminating them.

Figure 12. Unconventional Monetary Policy and Minutes



Notes: The gray shaded area represents NBER recession periods. The red shaded area represents total assets (minus eliminations from consolidation) in the Fed's balance sheet in millions of US dollars. The blue shaded area represents the word-counting indicator based on our UMP dictionary presented in the Appendix. The red arrows indicate UMP terms' intensity from the onset of the GFC to the outbreak of COVID-19. The green arrows indicate the timing of the Fed's UMP communications. The blue arrow indicates that UMP discussions in the Fed's minutes during the pre-tapering period were almost as high as during the COVID-19 crisis.

Source: Board of Governors of the Federal Reserve System (US).

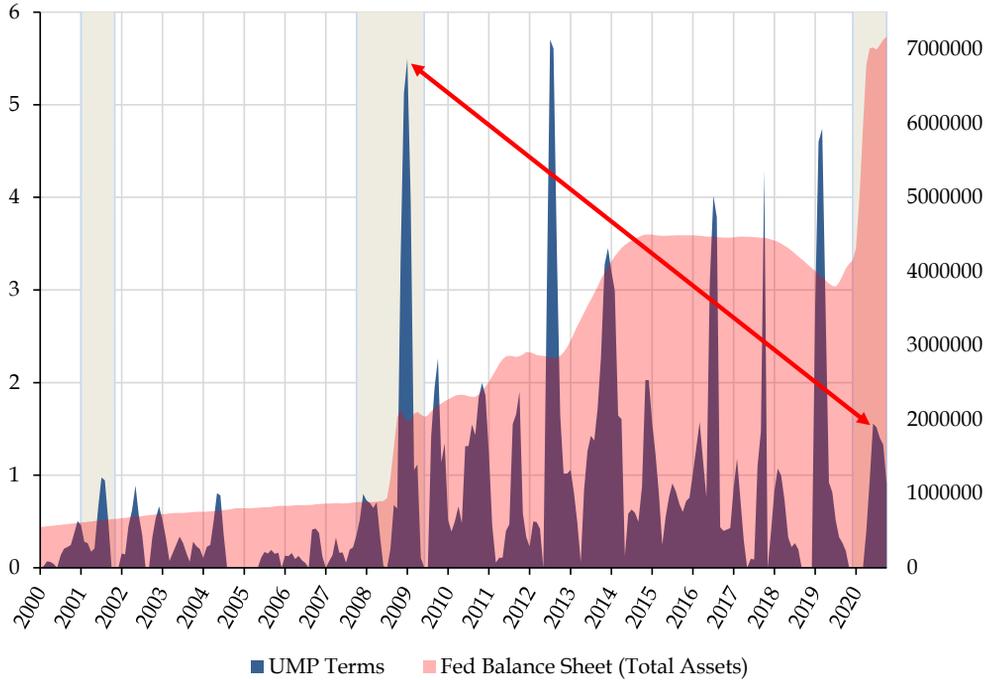
Figure 13 presents the word counts related to UMP terms in Fed chairman speeches.

Figure 13 shows that relevant communication shocks, like during the GFC or the third round of monthly purchases of Treasury securities and mortgage-backed securities (MBS) in September 2012 (third QE), are strongly related to the UMP content of the Fed chairman speeches. Indeed, during or after each communication peak, the dynamics of the Fed's balance sheet changed. Most of the peaks that did not influence the Fed's balance sheet were related to forward-guidance communications.

The Fed chairman speeches appear to be the privileged platform for mentioning UMP terms. It is interesting to compare the frequency of UMP terms in the speeches delivered before and after each crisis. As the Figure 13 shows, the frequency of the UMP terms in the speeches increases in the wake of the GFC, dot-com, and COVID-19 crises. However, the frequency is higher, and earlier, in the COVID-19 crisis, compared to the other crises.

Covid Economics 79, 27 May 2021: 218-256

Figure 13. Unconventional Monetary Policy and Fed Chairman Speeches



Notes: The gray shaded area represents NBER recession periods. The red shaded area represents total assets (minus eliminations from consolidation) in the Fed’s balance sheet in millions of US dollars. The blue shaded area represents the word-counting indicator based on the dictionary presented in the Appendix.

Source: Board of Governors of the Federal Reserve System (US).

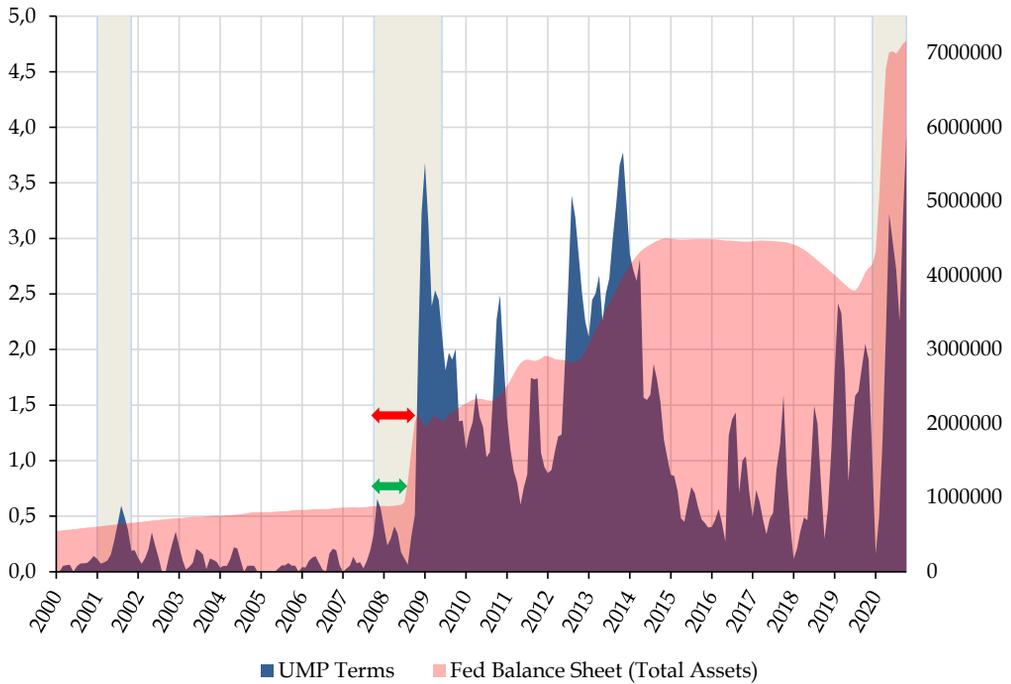
Figure 14 presents an aggregated UMP indicator for main Fed communications. The figure shows differences in the timing of UMP communications and actions for the COVID-19 and GFC crises

Notable is the post-GFC “new normal,” where UMP communications and actions are more frequent than in the pre-GFC period. This continuous need for UMP tools may eventually transform their unconventional character into a more conventional or regular one.

Overall, communicating about QE and forward-guidance (UMP) actions became the “new normal” for the Fed since the GFC (Bernanke, 2020), while the frequency of UMP terms remains higher for the COVID-19 crisis than for previous crises.

Covid Economics 79, 27 May 2021: 218-256

Figure 14. Unconventional Monetary Policy in Main Fed Communications



Notes: The gray shaded area represents NBER recession periods. The red shaded area represents total assets (minus eliminations from consolidation) in the Fed’s balance sheet in millions of US dollars. The blue shaded area represents the word-counting indicator based on the dictionary presented in the Appendix. The arrows indicate the time periods between the beginning of the GFC crisis and the first UMP communications (red) and the first UMP measures (green) influencing the Fed’s balance sheet.
Source: Board of Governors of the Federal Reserve System (US).

7. COVID-19

In this section we examine the use of COVID-19 terms with our dictionary presented in the Appendix. We compare these terms with UMP and contextual uncertainty terms, and financial volatility and new COVID-19 cases.

Figure 15 presents the repartition of COVID-19-related terms used in the main Fed’s communications in 2020.

The figure shows that the Fed chairman speeches anticipated the waves of new COVID-19 cases. One has to consider this result cautiously since the first tests started later in the US compared to other countries. Nevertheless, the speeches anticipated the spillovers of the virus from China, focusing on the US economy.

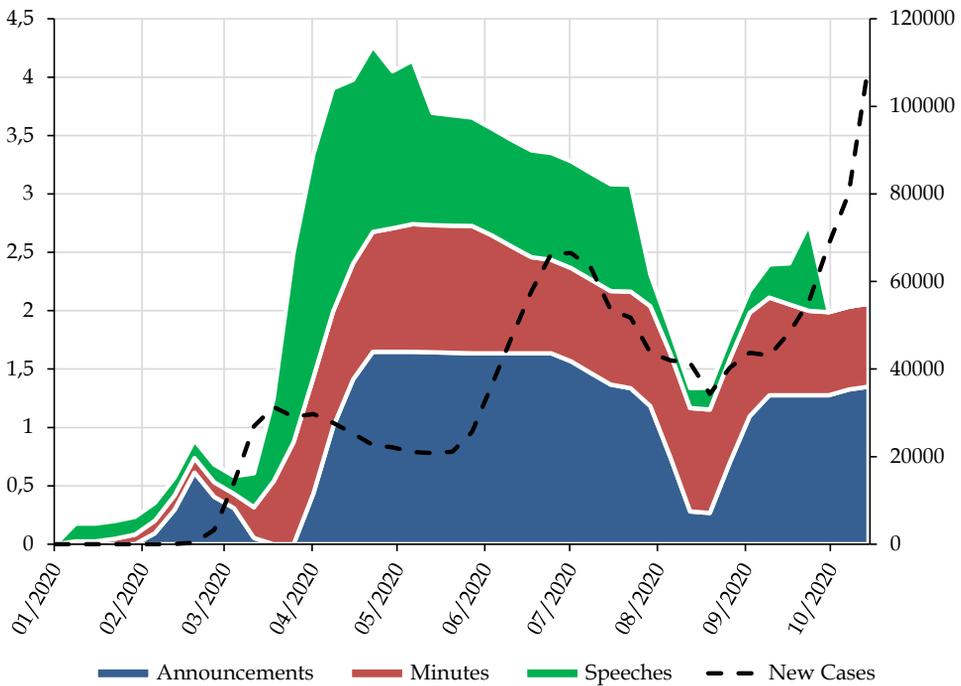
It is worth noting that the Fed chairman speeches provide a more timely and flexible communication vehicle than FOMC minutes and FFR announcements. They are disseminated quickly and informally compared to the other communication types, and allow health, political or foreign considerations that are less discussed in the other communication types.

Covid Economics 79, 27 May 2021: 218-256

Interestingly, Figure 15 presents two types of COVID-19 waves: the first type of wave plots new cases of COVID-19 based on medical statistics from the COVID-19 Data Repository, and the second type of wave plots the intensity of COVID-19-related terms in the Fed’s communications based on our COVID-19 dictionary presented in the Appendix. It is visually apparent that the Fed communication waves precede the virus waves.¹³

The magnitude and severity of the COVID-19 virus were rapidly understood and communicated to the public by the Fed via its FFR announcements and chairman speeches. The FFR announcements used more COVID-19-related terms than the other communication types and contributed better to the first communication wave than the speeches, but they lagged a few weeks behind the first Fed chairman speeches mentioning COVID-19-related terms.

Figure 15. COVID-19 and Main Fed Communications



Notes: The shaded areas represent the word-counting indicator for each communication type based on our COVID-19 dictionary presented in the Appendix. The dashed line represents the number of new COVID-19 cases in the US (right axis).

Source: Johns Hopkins University Center for Systems Science and Engineering (CSSE).

The decrease in the intensity of COVID-19 terms in the Fed chairman speeches in the second quarter of 2020 is directly correlated to the decrease in positive sentiment reported in Figure 5 for the same period. It is also correlated to the increase in the

¹³ Granger causality tests also confirm this finding, but given the few observations available, the results are not reported.

topicality of social welfare in the Fed chairman speeches during this period, as reported in Figure 9. Consequently, both the topics and the sentiments of the Fed chairman speeches were affected by the COVID-19 outbreak.

Interestingly, the increase in the SentiWords sentiment exposed in Section 4 seems to be confirmed by the decrease in the frequency of COVID-19-related terms after 2020Q3 despite the increase in the number of new cases during that quarter. In other words, the Fed's communications conveyed a more positive message than the reality of the pandemic and its economic spillovers would warrant. This seems to be the result of a crisis-specific communication strategy.

Overall, the waves of Fed communications about the COVID-19 crisis anticipated the waves of new COVID-19 cases. The Fed chairman speeches communicated about the first wave of COVID-19 earlier than the other communication types (FFR announcements and FOMC minutes). This result again confirms that speeches are less supervised than announcements and minutes and thus allow the Fed to communicate in a timelier manner. It also indicates that the Fed had an early understanding of the severity and magnitude of the COVID-19 pandemic and its economic spillovers.

Figure 16 compares new COVID-19 cases in the US and word-counting indicators based on the dictionaries of UMP and COVID-19 terms. This figure also includes the Chicago Board Options Exchange (CBOE) Volatility Index (VIX) to compare market volatility and potential financial uncertainty underlying the virus outbreak with UMP and COVID-19 terms.

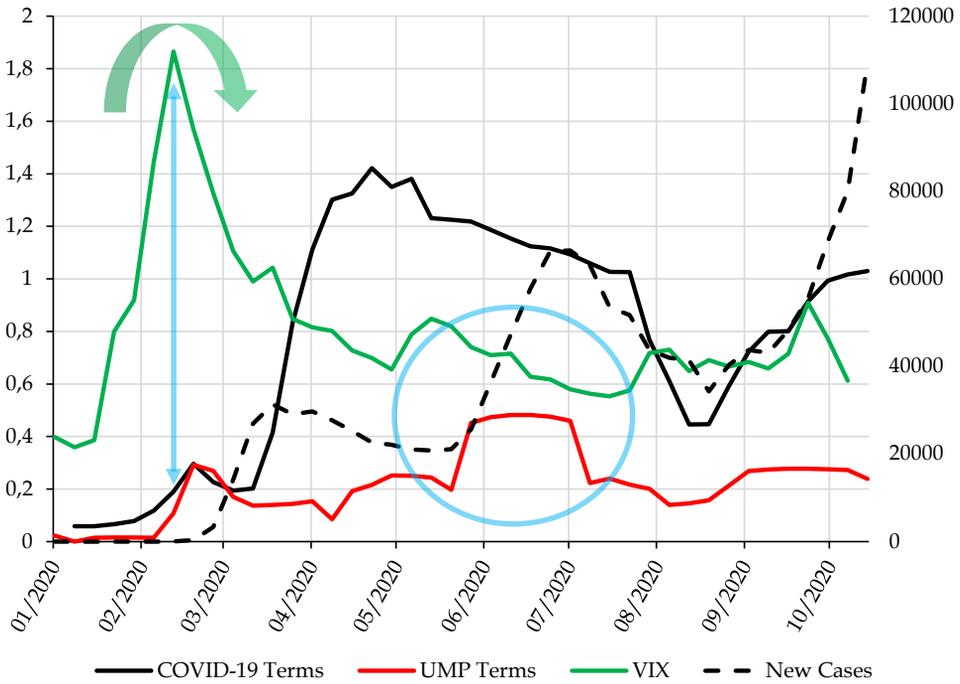
Figure 16 shows that the COVID-19-related terms in the Fed's communications in January 2020 about the virus outbreak in China considerably upset financial markets in the US. However, the UMP-related terms in the Fed's subsequent communications, about the UMP actions taken by the Fed in response to the pandemic, helped to decrease this financial volatility.

The CBOE VIX dramatically increased with the COVID-19 outbreak in China and several other countries, including the US. COVID-19 mentions in the Fed's communications preceded UMP considerations and waves of new cases in the US.

Between May and July of 2020, the Fed extensively communicated about unconventional monetary policies. During this period, although new cases of COVID-19 significantly increased, the Fed's communications and actions slightly decreased financial volatility.

Following this period, the increase in the frequency of UMP-related terms in the Fed's communications as the pandemic continued to worsen may have stabilized the volatility of financial markets.

Figure 16. COVID-19 and UMP Terms in Main Fed Communications



Notes: The dark blue line represents the word-counting indicator based on our dictionary of COVID-19 terms presented in the Appendix. The dashed line represents the number of new COVID-19 cases in the US (right axis). The green arrow indicates the CBOE VIX reversal due to UMP communications (and actions) highlighted with the blue arrow. The blue circle highlights the other relations between increases in UMP communications and decreases in the VIX.

Sources: Bloomberg and Johns Hopkins University Center for Systems Science and Engineering (CSSE).

Figure 17 presents our COVID-19 and UMP word-counting indicators based on Loughran and McDonald’s (2011) dictionary of contextual uncertainty terms.

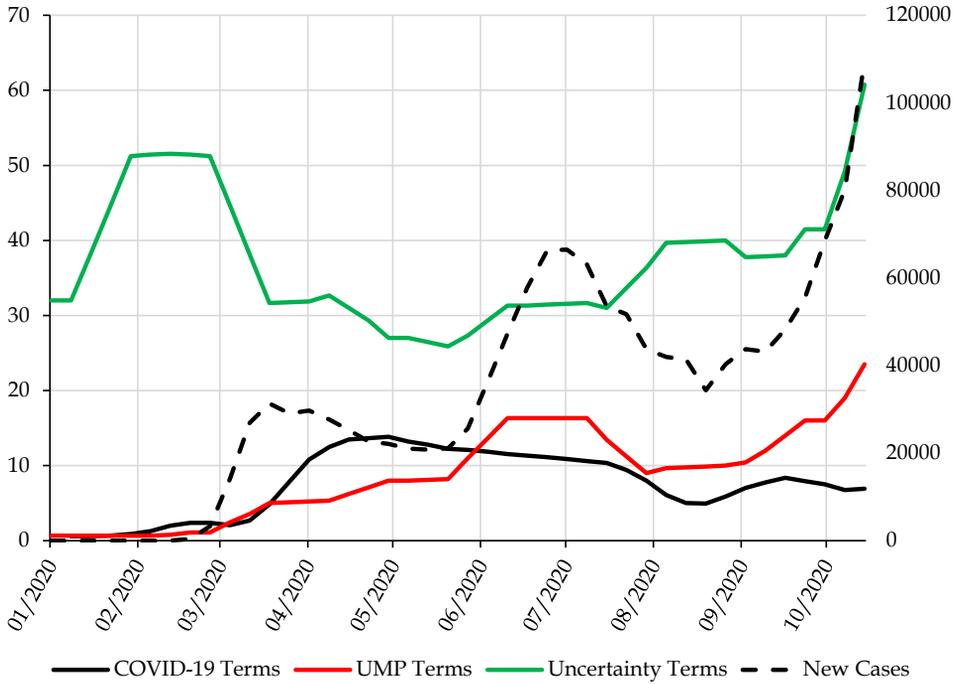
Uncertainty in the Fed’s communication is related to the number of UMP terms found in those communications, what we interpret as the “uncertainty effect”. COVID-19, UMP, and uncertainty comove in the Fed’s communications, especially during the second half of 2020. This is not necessarily the case at the beginning of the COVID-19 sample period, mainly because the suddenness of the outbreak of the virus took everyone by surprise and increased the frequency of the uncertainty-related terms before the others. The “uncertainty effect” appears during crisis periods necessitating UMP to mitigate market and economic uncertainty.

Figure 17 also demonstrates the anticipatory effects of uncertainty- and UMP-related terms in the Fed’s communication regarding COVID-19. The increase in the use of uncertainty terms appears to precede increases in new cases of the virus.

The correlation between the contextual uncertainty from Loughran and McDonald’s (2011) dictionary and the UMP-related terms from the UMP dictionary

presented in the Appendix is significantly positive at 0.44 for weekly average communications between 2000 and 2020 (i.e., 1090 observations).

Figure 17. COVID-19, UMP, and Uncertainty in Main Fed Communications



Notes: The dashed line represents the number of new COVID-19 cases in the US (right axis). The contextual uncertainty indicator is the number of uncertainty terms according to Loughran and McDonald (2011).

Source: Johns Hopkins University Center for Systems Science and Engineering (CSSE).

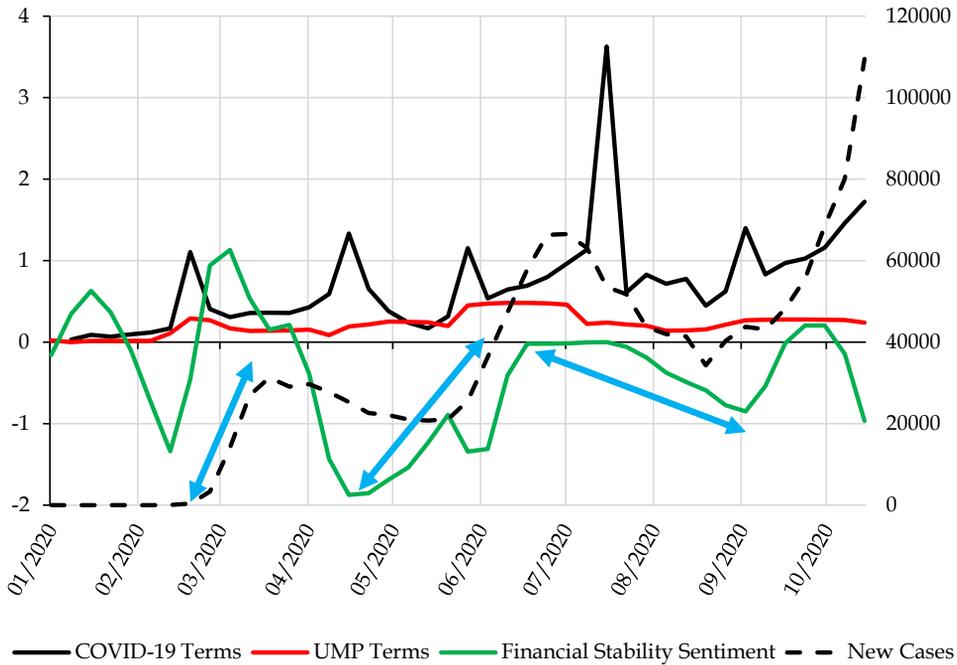
Figure 18 presents our COVID-19 and UMP word-counting indicators together with the financial stability sentiment (Correa et al., 2021) and the number of new COVID-19 cases in the US.

Except at the beginning of the COVID-19 crisis, increases in the sentiment associated with financial stability are correlated to increases in the number of new virus cases. The end-of-sample decrease in the financial stability sentiment is partly driven by the lack of chairman’s speeches during this period.

Figure 18 shows that the decrease in sentiment associated with financial stability lags a few weeks behind the increases in both COVID-19- and UMP-related terms in the Fed’s communications. This result is not surprising given that discussions and decisions related to financial stability generally occur after shocks to financial stability. The several deteriorations in financial stability sentiment that precedes the increase in the number of new COVID-19 cases may confirm the anticipatory effect of the Fed’s discussions of their stabilization policies.

Covid Economics 79, 27 May 2021: 218-256

Figure 18. COVID-19, UMP, and Financial Stability in Main Fed Communications



Notes: The blue line represents the word-counting indicator based on our COVID-19 dictionary presented in the Appendix. The dashed line represents the number of new COVID-19 cases in the US (right axis). The financial stability index is rescaled to match scale constraints. The blue arrows indicate similar trends between the number of new COVID-19 cases and the financial stability index.

Sources: Johns Hopkins University Center for Systems Science and Engineering (CSSE).

Following the GFC, the Fed’s communications anticipated effective UMP implementations (actions). The timing and magnitude of these implementations were dramatically different between crises. It was shown in Figure 18 that the Fed’s communications about the COVID-19 crisis also anticipated waves of new COVID-19 cases. The UMP implementations were indeed aimed at reducing market volatility together with COVID-19 spillovers.

We have also shown that the contextual uncertainty in the Fed’s communications well anticipated COVID-19 waves. Finally, it appears that the decrease in sentiment associated with the Fed’s communications about financial stability generally anticipated increases in the number of new COVID-19 cases. The anticipatory effects of contextual uncertainty in the Fed’s communications seem to confirm its early understanding of the COVID-19 spreading and its economic spillovers.

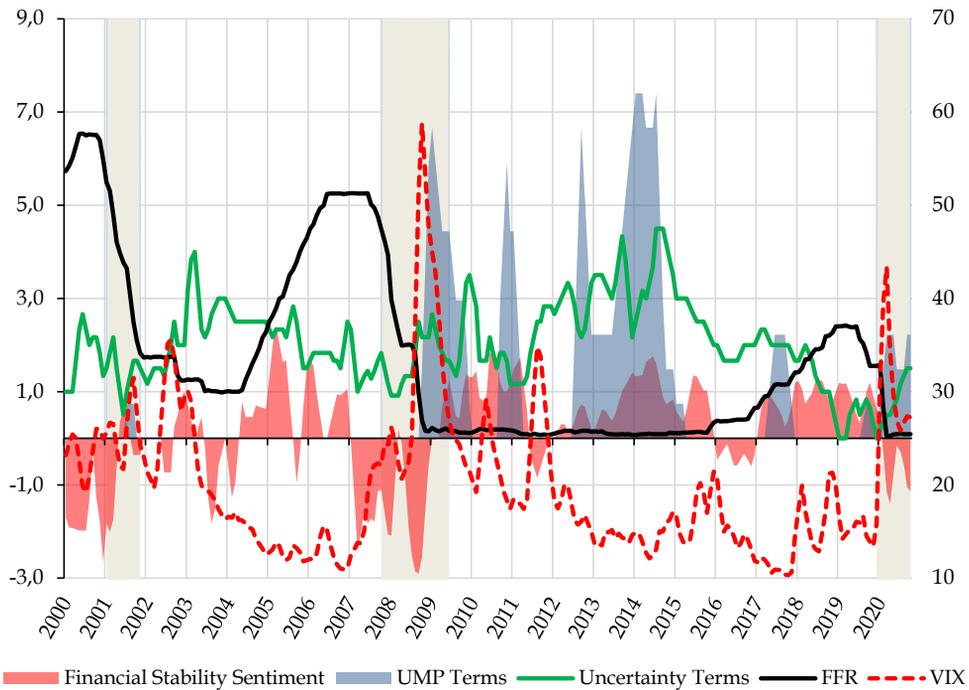
8. Financial Stability

UMP- and uncertainty-related terms are associated with financial stability and volatility. This section focuses on the financial stability sentiment and contextual

uncertainty and their respective dynamics relative to the conventional monetary policy instrument (the FFR) and financial market volatility (CBOE VIX).

Figure 19 compares the financial stability sentiment, UMP- and uncertainty-related terms from FFR announcements and the CBOE VIX.

Figure 19. Financial Stability and Monetary Policy in FFR Announcements



Notes: The gray shaded area represents NBER recession periods. The right axis indicates the VIX level.

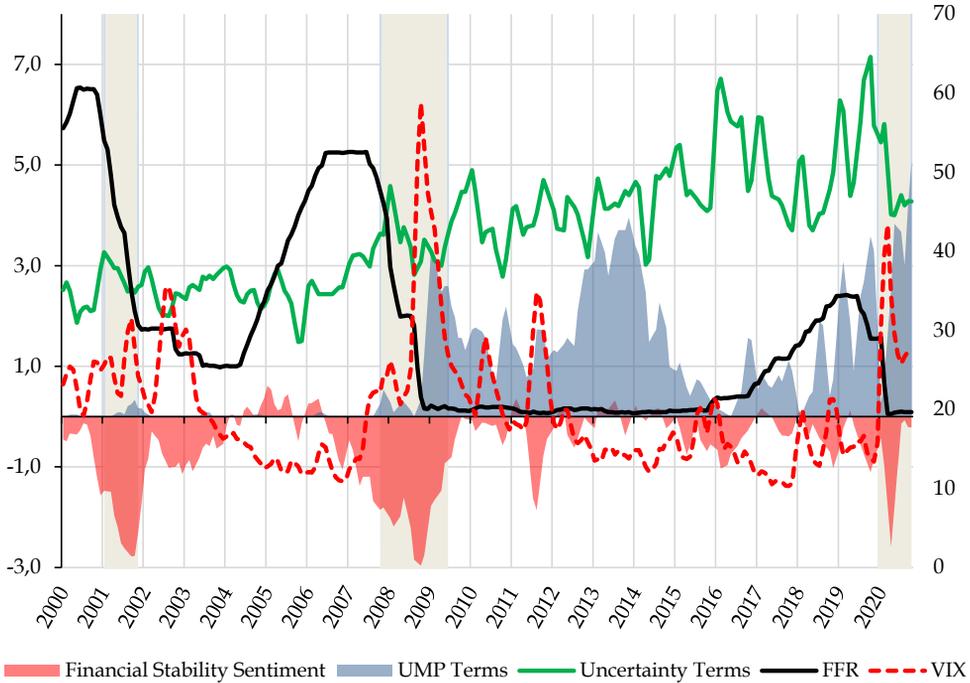
Figure 19 shows the Fed used fewer uncertainty-related words in its FFR announcements, and so voluntarily limited contextual uncertainty, during the COVID-19 crisis than in previous years and previous crises. This suggests a communication strategy that aimed to decrease market uncertainty (volatility) when the number of new virus cases sharply increased. Before the dot-com and GFC crises, a strong decrease in the positive sentiment associated with financial stability sentiment occurred, but this was not the case in the COVID-19 crisis. This is mainly due to the unpredictable characteristics of this crisis as well as its rapid spillovers on market uncertainty rather than the banking system.

The UMP-related terms played a significant role in decreasing market volatility during the dot-com and GFC crises but also during the COVID-19 crisis. In all of these crises, FFR announcements about the implementation of UMP measures reduced the VIX.

The financial stability sentiment is closely related to the FFR level. A decrease in the financial stability sentiment generally corresponds to a decrease in the FFR.

Figure 20 is the same as Figure 19 except that FOMC minutes are considered instead of FFR announcements.

Figure 20. Financial Stability and Monetary Policy in FOMC Minutes



Notes: The gray shaded area represents NBER recession periods. The right axis indicates the VIX level.

Interestingly, although each crisis was preceded by an increase in the use of uncertainty-related terms in FOMC minutes, a clear tendency to reduce uncertainty-related words during crises is observed in Figure 20, similar to Figure 19. The VIX is negatively correlated to financial stability sentiment, and both precede UMP terms, which generally leads to UMP actions being taken to stabilize the markets and financial stability fears and financial volatility (VIX).

As FOMC minutes provide detailed information on the monetary policy committee views about the suitable and near-term policy stance and the US economic outlook, they convey financial stability sentiments and UMP terms earlier than FFR announcements.

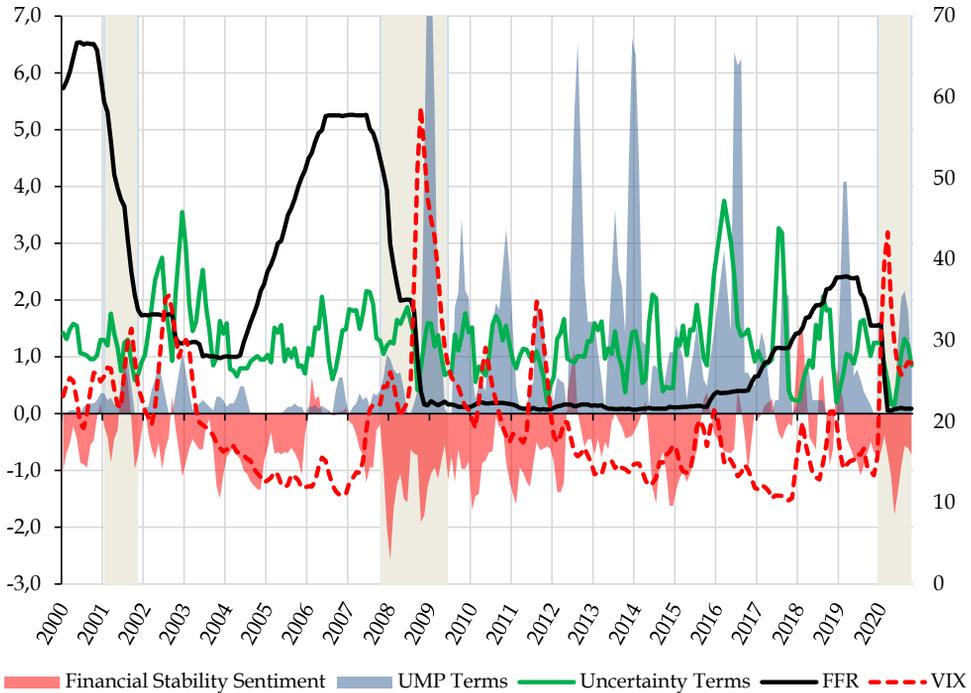
The FFR increases correspond to high financial stability sentiment levels, except during 2012–2015 where the financial stability sentiment was driven up by UMP communications and actions.

Figure 21 presents Fed chairman speeches together with the FFR announcements and financial volatility.

The financial stability sentiment present in the speeches is less indicative of the future FFR compared to the announcements and minutes. Chairman speeches may convey a negative financial stability sentiment even as the FFR increases, which on

average does not happen in FFR announcements and FOMC minutes (see Figures 19 and 20).

Figure 21. Financial Stability and Monetary Policy in Fed Chairman Speeches



Notes: The gray shaded area represents NBER recession periods. The right axis indicates the VIX level.

Comparing the period between the dot-com and GFC crises (P1) with the period between the GFC and the COVID-19 crises (P2) is informative. While few Fed chairman speeches contained UMP-related terms during P1, the “new normal” is on its way to being established during P2. More interestingly, the UMP- and uncertainty-related terms are relatively correlated during P2, whereas this correlation is nonexistent during P1. This comparison is, to a lesser extent, also valid for the VIX- and uncertainty-related terms, which are less correlated during P1 compared to P2.

Figure 21 shows that uncertainty-related terms in Fed chairman speeches were fewer during the COVID-19 crisis than in the previous years and previous crises. However, the contrast is less stark for the announcements and the minutes. The strong instability in the speeches is due to the wide-ranging fields and objectives they cover, coupled with the fact that speeches are generally less supervised than announcements and minutes.

The communication related to UMP occurred after the volatility peaks during the GFC and COVID-19 crises. Figures 20 and 21 highlight the Fed interventionism policy, which most central banks of developed countries employ: after each FFR decreases, UMP communication, usually followed by actions, compensates for the

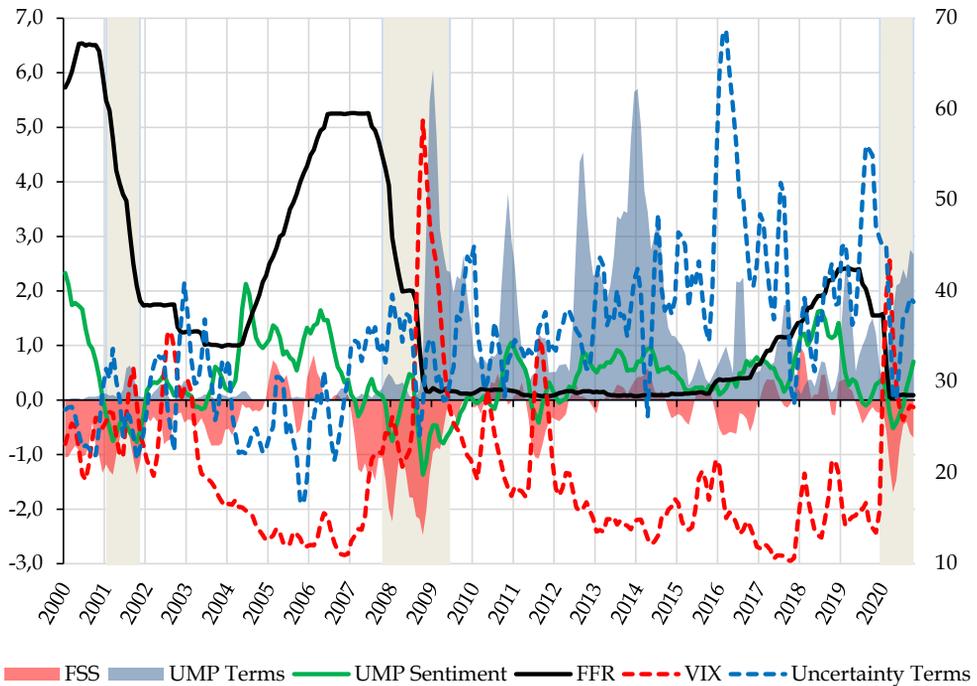
inability of the central bank to use the nominal interest rate, their main policy instrument, stuck at the ZLB.

Figure 22 aggregates the Fed’s three communication types to present a global picture of the Fed’s communications.

Our previous finding that the Fed had a crisis-specific communication strategy is confirmed. Indeed, Figure 22 supports the finding that, during crises, the Fed decreased the sentiment associated with UMP measures around the same time that it decreased the frequency of uncertainty-related terms in its communications.

Decreases in financial stability sentiment generally precede VIX increases, except in the COVID-19 crisis. A potential explanation for this finding is that the COVID-19 crisis was less predictable than the dot-com and GFC crises.

Figure 22. Financial Stability and Monetary Policy in Main Fed Communications



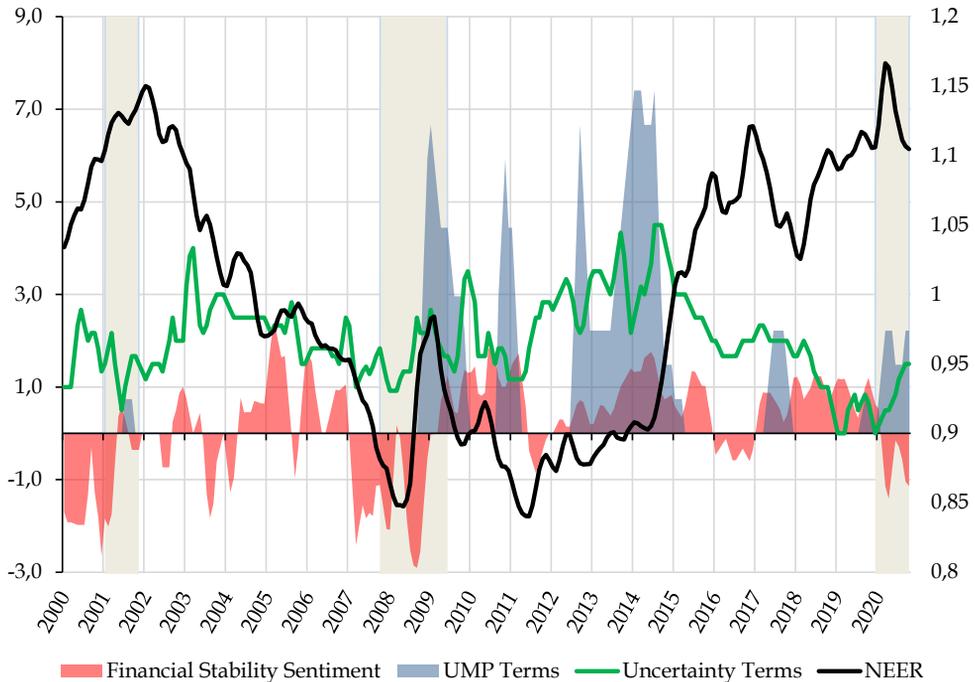
Notes: The gray shaded area represents NBER recession periods. FSS stands for “financial stability sentiment.” The right axis indicates the VIX and uncertainty terms levels.

Following the GFC, a new normal was established in which the Fed’s communications came to be increasingly fed by discussions of UMP tools, including forward-guidance measures. Typically, these UMP discussions were characterized by a high-level contextual uncertainty. This new normal was partly upset by the COVID-19 crisis, where the Fed adopted a communication strategy to use fewer uncertainty-related terms in their UMP communications.

Covid Economics 79, 27 May 2021: 218-256

The previous figures were focused on financial stability, but the Fed's communications also addressed economic stability. Figure 23 relates the word-counting indicators discussed above to the nominal effective exchange rate (NEER).

Figure 23. Financial Stability, NEER, and FFR Announcements



Notes: The gray shaded area represents NBER recession periods. The NEER corresponds to the amount of US dollars needed to purchase foreign currency (right axis). The financial stability sentiment, UMP, and uncertainty terms are related to FFR announcements.

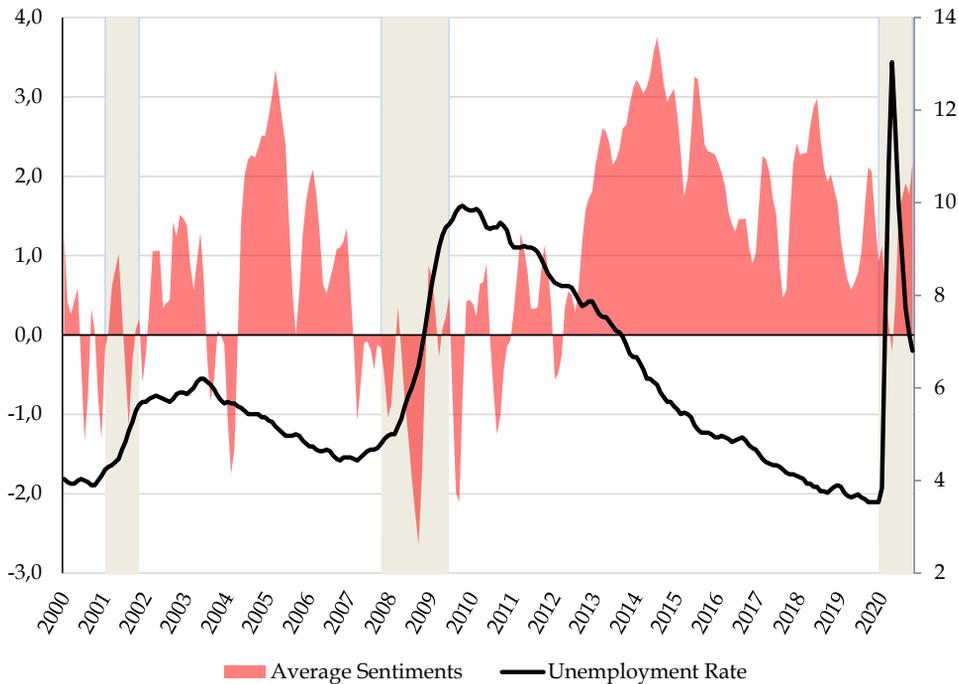
Figure 23 shows that increases in the NEER are correlated to decreases in the frequency of uncertainty-related words in the Fed's FFR announcements. It also shows UMP communications (and actions) generally decrease the NEER except during the tapering period where the NEER increased. High NEER levels also correspond to low levels of uncertainty-related terms in the Fed's communications.

Figure 24 relates the sentiment indicators discussed above to another measure of economic stability: the unemployment rate.

Figure 24 shows that sentiment and the unemployment rate are almost always inversely related. When the aggregated sentiment is positive, the unemployment rate tends to decrease. A switch from a positive to a negative aggregated sentiment usually coincides with an increase in the unemployment rate. The unpredictable nature of the COVID-19 crisis makes this statement debatable but not necessarily wrong. Unlike previous crises, the COVID-19 crisis increased the unemployment rate in the US from 3.5% to 13% in a short period of time, between January and May 2020. The unemployment rate decreased to lower levels after May 2020, around 6%. The sharp

and short-time shock on the aggregate sentiment was even shorter than the one on the unemployment rate, which may indicate a crisis-specific communication strategy toward communication optimism during crises.

Figure 24. Sentiment and Unemployment



Notes: The gray shaded area represents NBER recession periods. The right axis indicates the unemployment rate levels. The average sentiment aggregate contains an equally weighted average of sentiments according to the Loughran and McDonald (score and polarity), Hu&Liu (polarity), Jockers (polarity), NRC (polarity), SentiWords (polarity), UMP (score), and financial stability (score) sentiments. To achieve a balanced aggregated indicator for each communication type, we weight this average sentiment aggregate for FFR announcements more than for FOMC minutes, which in turn are weighted more than for chairman speeches.

Figure 24 confirms this intuition again that the Fed adopted a communication strategy to convey positive sentiments during the COVID-19 crisis.

To summarize, we show that the financial stability sentiment relates to FFR decisions in FFR announcements and FOMC minutes but not significantly enough in speeches. Increases in conventional monetary policy are often preceded by increases in financial stability sentiment, except during the period where UMP terms are used by Fed's communications, which generally involve actions improving financial stability.

We have also shown that the positive aggregated sentiment in the main Fed's communications correlates with decreasing unemployment. Except in times of

significant UMP steps, the NEER correlates with the level of uncertainty in the Fed's communications.

9. Policy Implications

The Fed implemented more unconventional monetary policies during the COVID-19 pandemic than during the dot-com and GFC crises. Moreover, it did so in a concise time window due to the abrupt upward slope of adverse shocks to the economy that the COVID-19 restrictions generated. The Fed's experience in crisis-specific communication and UMP tools acquired during the dot-com and GFC also contributed to understand better and address the COVID-19 crisis. To be successful, the Fed's UMP steps needed to be supported by clear and transparent communications and engagement with both the financial markets and the public.

We show that both the supervised and unsupervised learning methods we employ demonstrated that the Fed's communications during the COVID-19 crisis sharply differ from those of previous crises. Comparing the terms, sentiments, and topics conveyed by the Fed's communications with COVID-19 and financial data confirms that the Fed adopted a specific communication strategy during the COVID-19 crisis that also differs from the one adopted during the GFC and dot-com crises. We conclude that the Fed is getting better at using its communications to manage crises.

Our analysis determines that this communication policy consists of conveying optimism to the public during the worst periods of the pandemic while discussing (and implementing) earlier than in the previous crisis (Figure 14) unconventional monetary policies by justifying their importance in mitigating risks and uncertainties (Figures 16 and 17).

Another critical finding corroborates the Fed's forward-looking ability and its appropriate use of communication to convey a determined sentiment and justify UMP before each wave of the virus or each worsening of the financial conditions due to the virus's spillovers.

While we are still far from recovery, our results show that communications regarding the adopted policies and emergency programs allowed them to be perceived as a useful tool supporting economic recovery.

The Fed may have conducted a specific communication strategy for the COVID-19 crisis. This potential strategy conveyed less uncertainty and more optimism to the public while promoting UMP measures for managing the crisis situation. We interpret this behavior as conveying optimism without affecting transparency. The Fed's timely communications, together with its actions, succeeded in stabilizing financial markets.

10. Conclusion

This paper provides a comprehensive analysis of central bank communication during the past two decades, emphasizing the COVID-19 pandemic.

We show that both the supervised and unsupervised learning methods we employ determined that the Fed's communications during the COVID-19 crisis sharply differ from those of previous crises. Comparing the terms, sentiments, and topics conveyed by the Fed's communications with COVID-19 and financial data

confirms that the Fed adopted a specific communication strategy during the COVID-19 crisis that also differs from the one adopted during the GFC and dot-com crises.

During the COVID-19 pandemic, the Fed's communications emphasized topics of health, social welfare, and UMP interventions, which appear to be related to the conveyed sentiments. The Fed's communications regarding COVID-19 and UMP typically touch on the topics of financial volatility, uncertainty, and stability.

The content, sentiment, and timing of the Fed's communications changed in the COVID-19 crisis compared to previous crises. In particular, the sentiments of the Fed's communications significantly changed during the COVID-19 crisis compared to the GFC. Following the GFC, communicating about UMP became a "new normal" in the Fed's minutes and chairman speeches. Interestingly, we have shown that a negative financial stability sentiment usually precedes conventional monetary policy accommodation, except under the ZLB.

COVID-19 caused structural changes in the Fed's communication content. The Fed may have implemented a specific communication policy for the COVID-19 crisis that contrasted with its communication policy during the dot-com and GFC crises.

11. References

- Benchimol, J., Kazinnik, S., and Saadon, Y., 2020a. Communication and transparency through central bank texts. Paper presented at the 132nd Annual Meeting of the American Economic Association, January 3–5, 2020, San Diego, CA, United States.
- Benchimol, J., Kazinnik, S., and Saadon, Y., 2020b. Text mining methodologies with R: An application to central bank texts. Occasional Paper #2020.01, Bank of Israel.
- Bernanke, B.S., 2020. The new tools of monetary policy. *American Economic Review* 110(4), 943–983.
- Bianchi, F., Faccini, R., Melosi, L., 2020. Monetary and fiscal policies in times of large debt: Unity is strength. CEPR Discussion Paper #14720.
- Blei, D. M., Ng, A. Y., and Jordan, M. I., 2003. Latent Dirichlet allocation. *Journal of Machine Learning Research* 3, 993–1022.
- Craig, B. R., Phelan, T., Siedlarek, J.-P., 2021. Modeling behavioral responses to COVID-19. Economic Commentary, Federal Reserve Bank of Cleveland 2021(05), 1–6.
- Correa, R., Garud, K., Londono, J. M., and Mislav, N., 2021. Sentiment in central banks' financial stability reports. *Review of Finance* 25(1), 85–120.
- Daly, M. C., 2020. Is the Federal Reserve contributing to economic inequality? *FRBSF Economic Letter* 2020(32), 1–7.
- Erasmus, R., and Hollander, H., 2020. A forward guidance indicator for the South African Reserve Bank: Implementing a text analysis algorithm. Working Paper WP04/2020, Stellenbosch University, Department of Economics.
- Chetty, R., Friedman, J. N., Hendren, N., and Stepner, M., 2020. The economic impacts of COVID-19: Evidence from a new public database built using private sector data. NBER Working Paper #27431.
- Giacomini, R., and White, H., 2006. Tests of conditional predictive ability. *Econometrica* 74(6), 1545–1578.

- Guerrieri, V., Lorenzoni, G., Straub, L., Werning, I., 2020. Macroeconomic implications of COVID-19: Can negative supply shocks cause demand shortages? NBER Working Paper #26918.
- Haldane, A., McMahon, M., 2018. Central bank communications and the general public. *AEA Papers and Proceedings* 108, 578–583.
- Hansen, S., McMahon, M., Tong, M., 2019. The long-run information effect of central bank communication. *Journal of Monetary Economics* 108(C), 185–202.
- Henry, E., 2008. Are investors influenced by how earnings press releases are written? *Journal of Business Communication* 45(4), 363–407.
- Loughran, T., and McDonald, B., 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance* 66(1), 35–65.
- Loughran, T., and McDonald, B., 2016. Textual analysis in accounting and finance: A survey. *Journal of Accounting Research* 54(4), 1187–1230.
- Nyman, R., Kapadia, S., Tuckett, D., Gregory, D., Ormerod, P., and Smith, R., 2018. News and narratives in financial systems: Exploiting big data for systemic risk assessment. Bank of England Working Paper #704, Bank of England.

12. Appendix

12.1 Unconventional Monetary Policy Dictionary

Our targeted lexicon (presented in Table 2) was constructed by collecting words related to unconventional monetary policies from Fed communications using topic modeling and Bag-of-Words (BoW); see Benchimol et al. (2020b).

Table 2. Unconventional Monetary Policy Lexicon

asset purchases	depreciation pressure	market disrupt	risk premium
helicopter	direct lending	market functioning	securities purchases
QE	ELB	monetary base	stagflation
securities purchases	foreign exchange reserve	monetary stimulus	support
balance sheet	forward guidance	money supply	support liquidity
business support	funding	negative policy	supporting corporat
credit facilit	insolvency	negative rate	swap line
credit impair	intervention	NIRP	unconventional
deferral	lending facilit	quantitative easing	ZLB
deflation	lower bound	relaxing regulatory	

Source: Words and root words were extracted mainly from Fed communications.

12.2 COVID-19 Dictionary

Table 3 was constructed essentially from terms related to COVID-19 that appeared in both media (e.g., Google Trends search queries) and recent Fed communications (BoW) using the same methodology as in Table 2.

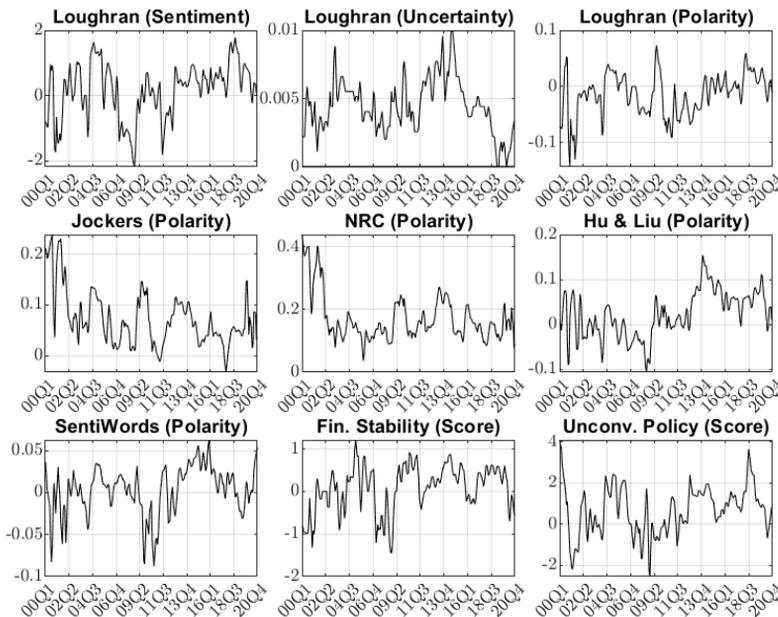
Table 3. COVID-19 Lexicon

acute	elderly	infect	pandemic	severe acute
cases	emergency	infection	pneumonia	sickness
confin	epidem	infection rate	quarantine	spreading
contagio	epidemic	lockdown	relief	syndrom
corona	hcov	mask	reproduction rate	testing
coronavirus	health	medical	respirator	vaccin
covid	hospital	morbid	respiratory	virus
death	hubei	morbidity rate	sars	wave
disabilit	human	mortal	sars cov	wuhan
disease	illness	ncov	sarscov	
disorder	inception rate	outbreak	sars-cov	

Source: Words and root words were extracted mainly from media and Fed communications.

12.3 Full Sample Figures

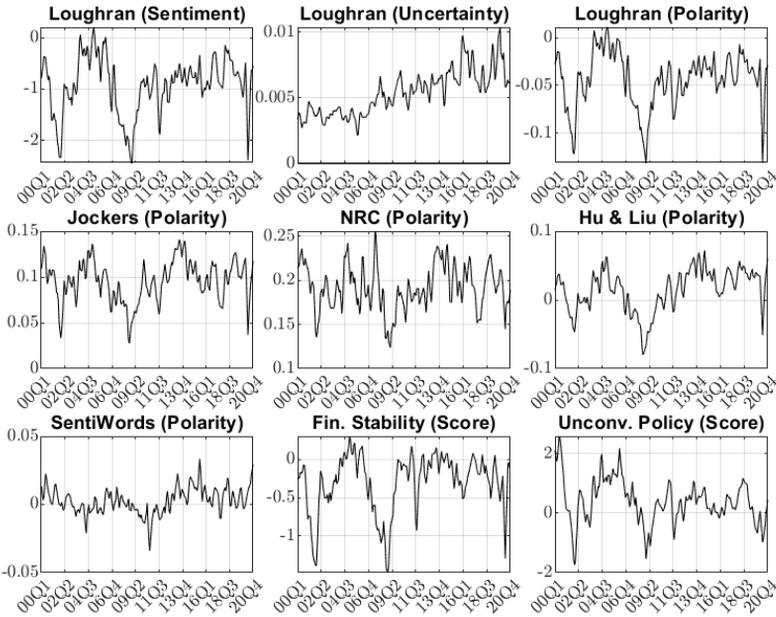
Figure A1. Sentiment Scores in FFR Announcements – Full Sample



Notes: Solid black lines represent sentiment score values. Source: Benchimol et al. (2020a).

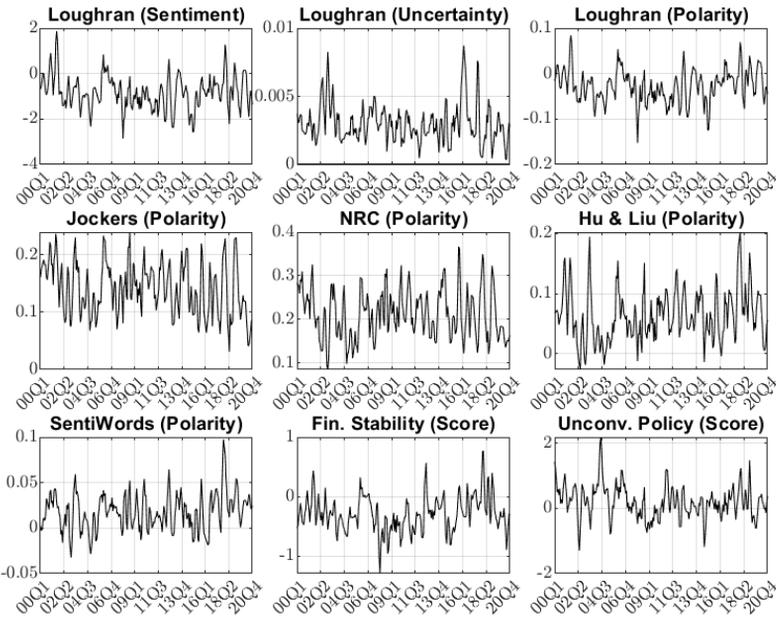
Covid Economics 79, 27 May 2021: 218-256

Figure A2. Sentiment Scores for the Fed’s FOMC Minutes - Full Sample



Notes: Solid black lines represent sentiment score values. Source: Benchimol et al. (2020a).

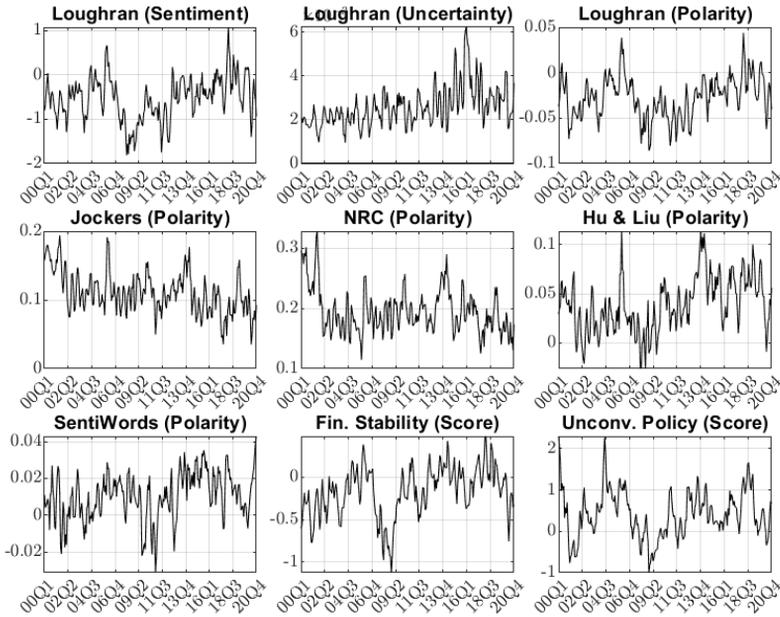
Figure A3. Sentiment Scores for Fed Chairman Speeches - Full Sample



Notes: Solid black lines represent sentiment score values. Source: Benchimol et al. (2020a).

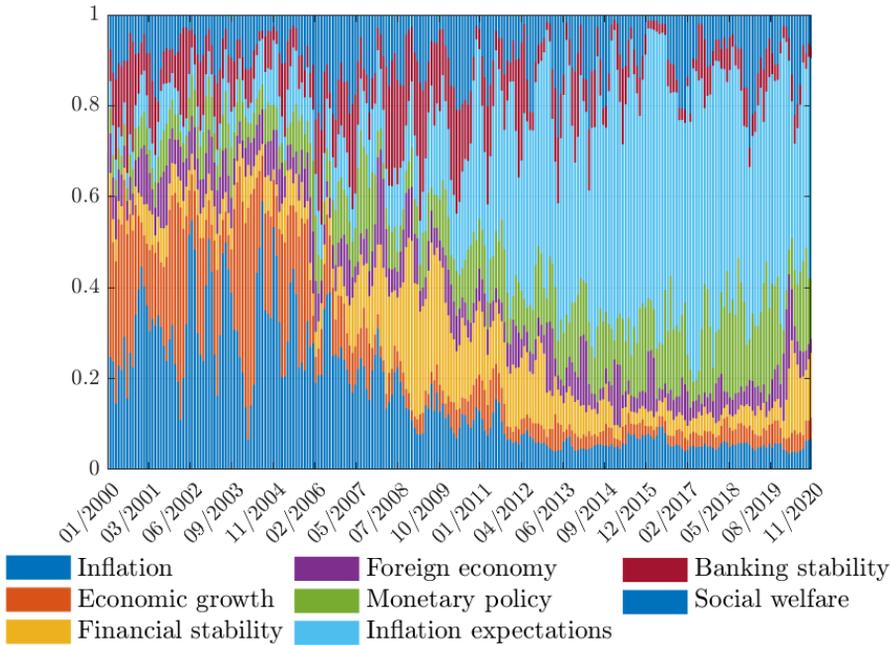
Covid Economics 79, 27 May 2021: 218-256

Figure A4. Sentiment Scoring of Main Fed Communications – Full Sample



Notes: Solid black lines represent sentiment score values. Source: Benchimol et al. (2020a).

Figure A5. Topic Analysis of Main Fed Communications



Notes: For clarity and robustness, we restrict attention to the six most frequently discussed topics.