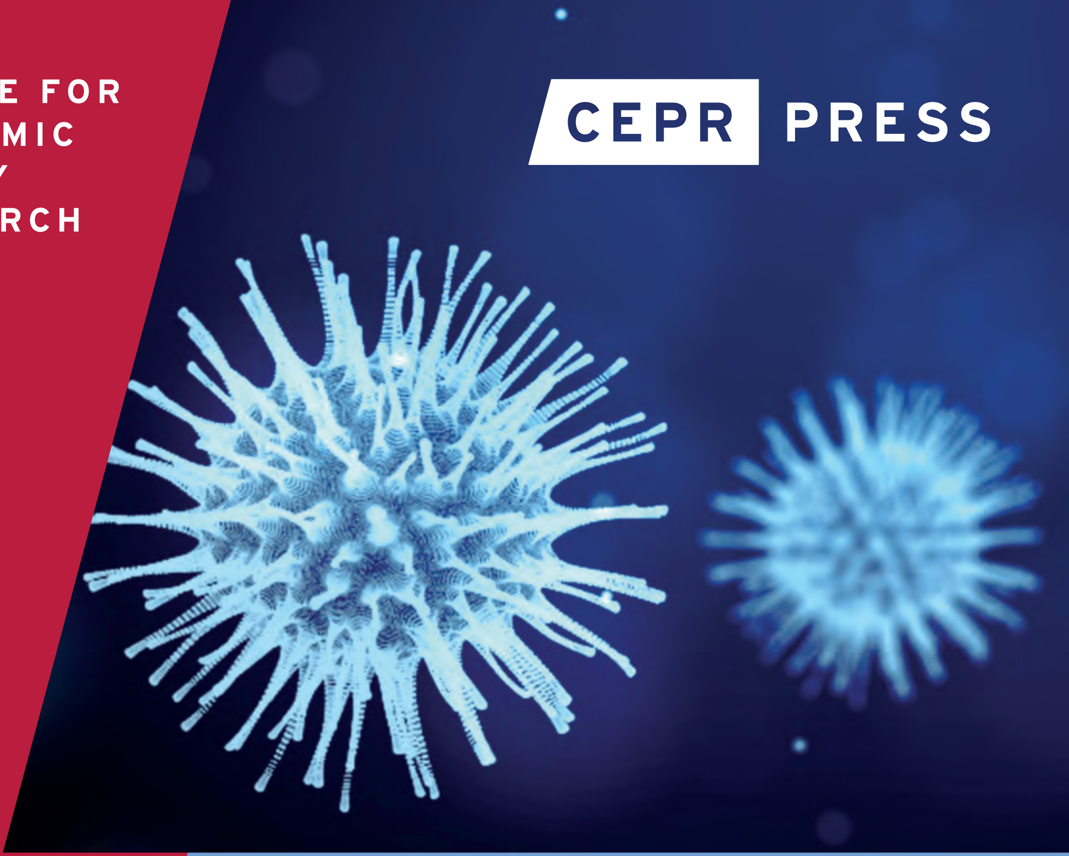
CENTRE FOR
ECONOMIC
POLICY
RESEARCH

CEPR PRESS



COVID ECONOMICS

VETTED AND REAL-TIME PAPERS

ISSUE 38
16 JULY 2020

Covid Economics
Vetted and Real-Time Papers

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

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

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

The Centre for Economic Policy Research (CEPR)

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

Editorial Board

Ethics

Covid Economics will feature high quality analyses of economic aspects of the health crisis. However, the pandemic also raises a number of complex ethical issues. Economists tend to think about trade-offs, in this case lives vs. costs, patient selection at a time of scarcity, and more. In the spirit of academic freedom, neither the Editors of *Covid Economics* nor CEPR take a stand on these issues and therefore do not bear any responsibility for views expressed in the articles.

Submission to professional journals

The following journals have indicated that they will accept submissions of papers featured in *Covid Economics* because they are working papers. Most expect revised versions. This list will be updated regularly.

American Economic Review

American Economic Review, Applied Economics

American Economic Review, Insights

American Economic Review, Economic Policy

American Economic Review, Macroeconomics

American Economic Review, Microeconomics

American Journal of Health Economics

Canadian Journal of Economics

*Econometrica**

Economic Journal

Economics of Disasters and Climate Change

International Economic Review

Journal of Development Economics

*Journal of Econometrics**

Journal of Economic Growth

Journal of Economic Theory

*Journal of the European Economic Association**

Journal of Finance

Journal of Financial Economics

Journal of International Economics

*Journal of Labor Economics**

Journal of Monetary Economics

Journal of Public Economics

Journal of Political Economy

Journal of Population Economics

*Quarterly Journal of Economics**

Review of Economics and Statistics

*Review of Economic Studies**

Review of Financial Studies

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

Covid Economics
Vetted and Real-Time Papers

Issue 38, 16 July 2020

Contents

Mutual fund performance and flows during the COVID-19 crisis[1]

Luboš Pástor[2] and M. Blair Vorsatz[3]

Date submitted: 10 July 2020; Date accepted: 10 July 2020

We present a comprehensive analysis of the performance and flows of U.S. actively managed equity mutual funds during the COVID-19 crisis of 2020. We find that most active funds underperform passive benchmarks during the crisis, contradicting a popular hypothesis. Funds with high sustainability ratings perform well, as do funds with high star ratings. Fund outflows largely extend pre-crisis trends. Investors favor funds that apply exclusion criteria and funds with high sustainability ratings, especially environmental ones. Our finding that investors remain focused on sustainability during this major crisis suggests they view sustainability as a necessity rather than a luxury good.

Covid Economics 38, 16 July 2020: 1-36

1 We are grateful to the University of Chicago Booth School of Business for research support. The views in this paper are the responsibility of the authors, not the institutions they are aliated with.

2 Charles P. McQuaid Professor of Finance, University of Chicago Booth School of Business. CEPR Research Fellow. NBER Research Associate and Board member, National Bank of Slovakia.

3 Ph.D. student in Finance, University of Chicago Booth School of Business.

1. Introduction

Active equity mutual funds are well known to have underperformed passive benchmarks, net of fees.[1] Despite its long-lasting underperformance, the active management industry remains large, managing tens of trillions of dollars. The existence of a large underperforming industry appears puzzling because an alternative—passive funds—is easily available to investors.

One popular hypothesis is that investors are willing to tolerate this underperformance because active funds outperform in periods that are particularly important to investors. This hypothesis is first formulated by Moskowitz (2000) who asks whether mutual funds provide a hedge against recessions. Glode (2011) formalizes this hypothesis by building a model in which a fund manager generates active returns that depend on the state of the economy. In equilibrium, the manager chooses to work harder in periods when investors' marginal utility of consumption is higher because investors are willing to pay for this insurance. If active funds deliver high returns in periods when investors need them the most then these funds' unconditional performance understates the funds' true abilities.

We test this hypothesis by analyzing the performance of active mutual funds during the COVID-19 crisis of 2020. This crisis is particularly suitable for the task at hand for two reasons. First, it has led to an unprecedented output contraction and the fastest increase in unemployment on record. Investors surely want to hedge against such a severe crisis. Second, active managers have an opportunity to perform well during this crisis because the crisis has created unusually large price dislocations in financial markets. In the equity market, the S&P 500 index experienced its steepest descent in living memory, losing 34% of its value in the five-week period between February 19 and March 23, 2020 before bouncing back by over 30% by the end of April. The sharp response of equity markets to COVID-19 is analyzed in a growing number of studies.[2] In the bond market, liquidity evaporated in March 2020, not only for corporate bonds (e.g., Kargar et al., 2020, and O'Hara and Zhou, 2020) but also for the usually-liquid Treasuries (e.g., Schrimpf, Shin, and Sushko, 2020). Until liquidity improved following the interventions from the Federal Reserve, its temporary shortage created massive market disruptions. For example, in the corporate bond market, bonds traded at large discounts to credit default swaps, and ETFs traded at large discounts to net asset values (Haddad, Moreira, and Muir, 2020). In addition, the Treasury market witnessed significant

[1]See Jensen (1968), Elton, Gruber, Das, and Hlavka (1993), Malkiel (1995), Gruber (1996), Carhart (1997), Wermers (2000), Pástor and Stambaugh (2002), and Fama and French (2010), and others.

[2]For evidence at the aggregate stock market level, see, for example, Alfaro et al. (2020) and Gormsen and Koijen (2020). For cross-sectional evidence, see Bretscher, Hsu, and Tamoni (2020), Ding et al. (2020), Fahlenbrach, Rageth, and Stulz (2020), Gerding, Martin, and Nagler (2020), Pagano, Wagner, and Zechner (2020), Ramelli and Wagner (2020), and others.

Covid Economics 38, 16 July 2020: 1-36

CEPR PRESS

mispricing between bonds and bond futures (Schrimpf, Shin, and Sushko, 2020). These price dislocations are due to a combination of factors including record-high volatility and traders working from home. Under the hypothesis that active funds outperform during recessions, they should find it particularly easy to outperform when markets are rife with mispricing.

Contrary to this hypothesis, we find that active funds underperform their passive benchmarks during the COVID-19 crisis. We define the crisis period as the ten-week period between February 20 and April 30, 2020. We choose February 20 as the starting date because the stock market peaked on February 19 before its rapid descent. We choose April 30 as the ending date because it is a month-end by which the market largely rebounded, and also because it puts the market bottom on March 23 roughly in the middle of the crisis period. The ten-week crisis period is thus roughly evenly split between the crash and the recovery. Our evidence is based on daily returns of all U.S. active equity mutual funds.

The underperformance of active funds is particularly strong when measured relative to the S&P 500 benchmark. We find that 74.2% of active funds—about three quarters!—underperform the S&P 500 during the COVID-19 crisis. The average fund underperformance is −5.6% ($t = -5.37$) during the ten-week period, or −29.1% on an annualized basis.

While the S&P 500 is the most popular benchmark among U.S. equity funds, it is not appropriate for all funds given its large-cap focus. We consider three types of benchmarks that are tailored to each fund's investment style: Morningstar-designated FTSE/Russell benchmarks, fund-designated prospectus benchmarks, and factor-model benchmarks. We find that active funds also underperform these fund-specific benchmarks, although by lower margins. For example, 57.6% of funds underperform their FTSE/Russell benchmarks and 54.2% of funds underperform their prospectus benchmarks. The average fund underperformance relative to the FTSE/Russell benchmark is −2.1% ($t = -3.90$) during the crisis period, or −11% on an annualized basis. Relative to the prospectus benchmark, the average underperformance is −1.5% ($t = -2.49$) during the crisis, or −7.7% annualized.

Besides benchmark-adjusted fund returns, we also examine factor-adjusted returns by computing fund alphas relative to five different factor models. All five alphas are significantly negative on average during the crisis period, ranging from −7.6% annualized ($t = -3.25$) for the six-factor model that includes the five factors of Fama and French (2015) plus momentum, to −29.1% annualized ($t = -7.02$) for the CAPM. The fraction of funds with negative alphas ranges from 60.4% for the four-factor Carhart (1997) model to a stunning 80.2% for the CAPM. In short, active funds perform poorly during the COVID-19 crisis.

Prior tests of the same hypothesis arrive at a different conclusion. Moskowitz (2000)

Covid Economics 38, 16 July 2020: 1-36

shows that active funds' returns from 1975 to 1994 are higher during recessions by 6% per year, on average. Kosowski (2011) analyzes the period from 1962 to 2005 and finds that mutual fund alphas in recessions exceed those in expansions by 3% to 5% per year, on average. Glode (2011) reports that funds with poor unconditional performance generate countercyclical risk-adjusted returns in 1980 through 2005. Kacperczyk, van Nieuwerburgh, and Veldkamp (2016) find that fund alphas are 1.6% to 4.6% per year higher in recessions over the 1980–2005 period. Unlike our study, all of these studies examine periods in which recessions are substantially milder than the COVID-19 crisis.

While active funds as a whole underperform, their performance during the COVID-19 crisis exhibits substantial heterogeneity. One of the strongest predictors of performance is the sustainability rating from Morningstar. Morningstar assigns between one and five sustainability "globes" to each fund, with more globes denoting higher sustainability. We find that funds with more globes as of January 31, 2020 have higher benchmark-adjusted returns between February 20 and April 30, 2020. Remarkably, the relation is monotonic across the globe categories: five-globe funds outperform four-globe funds, which in turn outperform three-globe funds, etc. High-globe funds (those with four or five globes) significantly outperform the remaining funds within the same investment style by 14.2% per year ($t = 4.85$) in terms of FTSE/Russell benchmark-adjusted returns. This result is driven largely by environmental sustainability—funds with higher environmental ratings outperform those with lower ratings.

Our findings linking fund performance to sustainability resemble those of Nofsinger and Varma (2014) who find that socially responsible mutual funds tend to outperform during periods of market crises. Their findings are based on a sample of 240 U.S. domestic equity mutual funds in the period of 2000 through 2011, which includes two recessions (2001 and 2007–2009). We examine only one recession (2020) but many more funds. Another related study, Albuquerque et al. (2020), finds that U.S. firms with high environmental and social ratings earn comparatively high stock returns in the first quarter of 2020. Ding et al. (2020) report a similar finding based on corporate social responsibility ratings of firms in 56 countries.[3] Our fund-level evidence complements their stock-level evidence in highlighting the role of sustainability during the COVID-19 crisis. The high returns of sustainable funds and stocks suggest that market participants' tastes continue to shift toward green assets and green products during this crisis (Pástor, Stambaugh, and Taylor, 2020).

Besides sustainability ratings, another strong predictor of fund performance during the

[3]These findings echo those based on the 2008–2009 recession. Lins, Servaes, and Tamayo (2017) show that U.S. firms with higher environmental and social ratings perform better during that recession.

Covid Economics 38, 16 July 2020: 1-36

Covid Economics 38, 16 July 2020: 1-36

COVID-19 crisis is the fund's star rating from Morningstar. Star ratings assigned as of January 31, 2020 predict performance between February 20 and April 30, 2020 positively and significantly. Similar to Morningstar globes, the relation is monotonic: five-star funds outperform four-star funds, which outperform three-star funds, etc. Five-star funds significantly outperform one-star funds in terms of cumulative benchmark-adjusted returns. One additional star is associated with an increase in performance of 5.78% per year ($t = 2.84$) in terms of FTSE/Russell benchmark-adjusted returns. That is, a five-star fund outperforms a one-star fund of the same style by about 23% per year, on average.

Finally, we find that growth funds outperform value funds. This finding is only partly driven by the well-known fact that the growth style outperforms the value style during the crisis because we measure fund performance net of the fund's style. In other words, we find that growth funds beat value funds on a style-adjusted basis. This result is strong when the style adjustment is performed through a factor model but it is insignificant when the adjustment is based on the style benchmark. The mixed nature of this evidence suggests roles for both active management and the superior performance of the growth style in explaining the different performance of growth and value funds during the crisis.

In addition to fund performance, we analyze capital flows in and out of active mutual funds. During the COVID-19 crisis, active funds experience steady outflows that largely continue long-term trends. The outflows are rapid during the market crash but they continue, albeit at a slower pace, during the market rebound after March 23, 2020.

Fund flows vary substantially across funds. Similar to performance, crisis-period flows are predictable by funds' pre-crisis sustainability ratings. Flows are near-monotonic across the five globe categories, with five-globe funds having the largest net flows and one-globe funds having the lowest flows between February 20 and April 30, 2020. In particular, one-globe funds suffer outflows of 2.6% of assets under management over the ten-week period, whereas five-globe funds' net flows are roughly zero. This difference, which is statistically significant, is driven especially by environmental concerns. Furthermore, funds that apply exclusion criteria in their investment process receive net inflows during the crisis, whereas funds that do not apply exclusions experience outflows. It is well known that mutual fund investors have come to favor sustainability-oriented funds in the 2010s (e.g., Bialkowski and Starks, 2016, and Hartzmark and Sussman, 2019). We find that this pre-crisis trend toward sustainability continues during the COVID-19 crisis.

A popular perspective in traditional neoclassical economics is that sustainability issues, such as environmental quality, are "luxury goods" that are likely to be of concern only

Covid Economics 38, 16 July 2020: 1-36

to those whose more basic needs for food, housing, and survival are adequately met (e.g., Baumol and Oates, 1979).[4] This perspective predicts that interest in sustainability should subside during a major economic and health crisis. In contrast, we find that investors retain their commitment to sustainability during the COVID-19 crisis. This finding suggests that investors have come to view sustainability as a necessity rather than a luxury good.

The performance hypothesis rejected by our evidence—that active funds outperform in recessions—is not the only possible explanation for why active management remains popular despite its poor track record. Gruber (1996) suggests that some investors suboptimally rely on active management because they are influenced by advertising, brokers, institutional arrangements, or tax considerations. Pástor and Stambaugh (2012) argue that a large active management industry can be rationalized if investors believe that active managers face decreasing returns to scale. In their model, rational investors respond to past underperformance of active funds by withdrawing money, which improves those funds' future performance to the point where investors are indifferent between investing actively or passively.

Our focus on crisis-period fund performance is also related to the literature on time-varying fund manager skill. An important early contribution is Ferson and Schadt (1996). More recently, Kacperczyk, van Nieuwerburgh, and Veldkamp (2016) develop a model of optimal attention allocation over the business cycle. In their model, fund managers allocate more attention to idiosyncratic shocks in expansions and to aggregate shocks in recessions. Similarly, Kacperczyk, van Nieuwerburgh, and Veldkamp (2014) find that fund managers exhibit better stock picking in expansions and better market timing in recessions. We do not attempt to separate stock selection from market timing during the COVID-19 crisis because such an exercise would require time series of fund holdings, which are widely available only on a quarterly basis. Data availability also limits our ability to test the hypothesis that the profit opportunities created by COVID-19 lead active funds to trade more, improving their future performance (Pástor, Stambaugh, and Taylor, 2017). Whether funds increase their turnover in 2020, and whether this turnover causes better future fund performance, remains to be seen because the turnover data from the SEC are only annual.

The paper is organized as follows. In Section 2, we describe our data. In Section 3, we analyze fund performance and its determinants. In Section 4, we discuss fund flows and their determinants. Section 5 concludes. Additional empirical results and database construction details are located in the Appendix, which is available on the authors' websites.

[4]An example of this common view is the controversial "Summers memo" from 1991, in which the World Bank's Chief Economist suggests that the Bank should be encouraging more migration of dirty industries to the least-developed countries. One of the reasons given in the memo is that "the demand for a clean environment for aesthetic and health reasons is likely to have very high income elasticity."

2. Data

We use daily data from Morningstar Direct covering the period from January 1, 2017 to April 30, 2020. Our original sample covers 4,292 actively managed U.S. equity mutual funds, although we primarily focus on the 3,626 funds with at least one non-missing net return between February 20 and April 30, 2020. The latter sample represents $4.9 trillion of total net assets as of January 31, 2020.

Our fund universe is constructed largely following Pástor, Stambaugh, and Taylor (2015), with two main differences. First, we also include international and sector equity funds domiciled in the U.S. Second, we do not require funds to appear in both CRSP and Morningstar; we use only Morningstar data. As in Pástor, Stambaugh, and Taylor (2015), we use the Morningstar FundID variable to aggregate share classes to the fund level.[5]

We use keywords in the Morningstar Category variable and the prospectus benchmark to exclude bond funds, money market funds, real estate funds, target retirement funds, and other non-equity funds. We also exclude funds identified by Morningstar as passive index funds and funds whose name contains the word "index." In our baseline results, we also use a fund size filter to include only funds with at least $15 million of net assets on January 31, 2020. Excluding the smallest funds has been advocated by Elton, Gruber, and Blake (2001), among many others. This screen is particularly relevant for fund flows because modest dollar flows can translate into extreme percentage flows for the smallest funds. This subsample covers 2,764 funds and $4.891 trillion of total net assets.

Throughout our analysis, we use funds' returns net of the expense ratio because our goal is to measure the return delivered to clients after fees. Despite being very fresh (we downloaded data through April 30, 2020 in May 2020), the data appear to be free of salient errors. For example, none of our sample's 2,692,799 fund-level daily net return observations are below -90% and only one is larger than 100%. We adjust fund returns for daily benchmark returns, also obtained from Morningstar, and for daily factor returns, which we obtain from Ken French's data library along with the risk-free rate.

We rely on three main Morningstar categorization variables throughout our analysis: the

[5]Many funds have multiple share classes, which are tied to the same pool of assets but have different fee structures. Since different share classes of the same fund have the same Morningstar FundID value, we can use the FundID variable to aggregate the share classes up to the fund level. Specifically, we compute a fund's total net assets by summing total net assets across the fund's share classes, setting the fund-level variable to missing if total net assets are missing for any of the share classes on that date. The fund's net returns, net expense ratio, and turnover ratio are averaged (lag-asset-weighted) across all share classes with non-missing values.

Covid Economics 38, 16 July 2020: 1-36

Covid Economics 38, 16 July 2020: 1-36

Morningstar Category, the Morningstar Institutional Category, and the Global Category. In our full sample of 3,626 funds, there are 39 Global Categories, 52 Morningstar Categories, and 93 Morningstar Institutional Categories. Each of these variables classifies a fund based on its investment style, sector, and geographical orientation. The Global Categories are the coarsest classification system and are used by Morningstar as groupings within which sustainability can be ranked. These categories include U.S. equity large-cap blend, U.S. equity small-cap, energy sector equity, and Latin America equity. We use the Global Categories for style fixed effects. The Morningstar Category variable is built on the 3-by-3 style box of size tilts (large-cap vs. small-cap) and growth vs. value style tilts. We use this variable to follow Pástor, Stambaugh, and Taylor (2015) in their classification of funds into equity and non-equity categories. Morningstar uses these groupings to rank performance in terms of star ratings. Last, the Morningstar Institutional Category variable is built on an extended version of the 3-by-3 style box with size tilts including micro-cap and giant and style tilts including deep value and high growth. We use this finest classification system for clustering our standard errors. While this is conservative relative to the more standard treatment of clustering at the fund level, we believe this appropriately accounts for how the health crisis shock may generate residual correlation among funds with similar strategies. For further details of our data construction, see the Appendix.

3. Fund Performance

Figure 1 provides a preliminary look at the performance of active funds during the COVID-19 crisis, along with the performance of the most popular passive benchmark: the S&P 500 index. We normalize the levels of both the S&P 500 and each fund's net asset value to 100 as of February 19, 2020. For each day t after February 19, we compute the price indices for each fund as well as the S&P 500 by compounding the corresponding daily returns:

$$F_t = 100(1+r_1^F)(1+r_2^F)...(1+r_t^F) \tag{1}$$

$$B_t = 100(1+r_1^B)(1+r_2^B)...(1+r_t^B) , \tag{2}$$

where F_t is the fund price index, B_t is the price index for the passive benchmark, r_t^F is the fund's net return on day t, and r_t^B is the benchmark return. Figure 1 plots both B_t and the average value of F_t across all funds. The figure also plots a 95% confidence interval around average F_t. Standard errors are clustered on the Morningstar Institutional Category, both here and in all subsequent figures reporting confidence intervals.

Covid Economics 38, 16 July 2020: 1-36

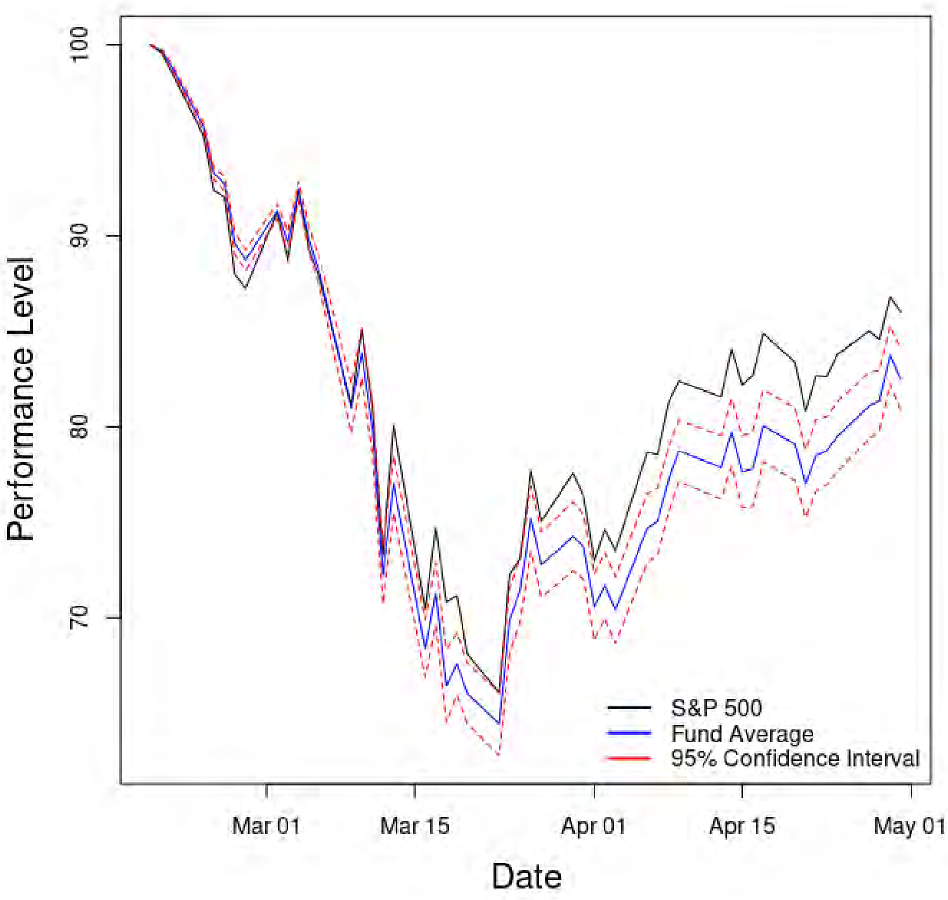


Figure 1. Average Fund Performance vs. the S&P 500 During the Crisis. This figure plots the performance of the average active equity mutual fund against the S&P 500 in February 20 through April 30, 2020. Both price indices are initialized at 100 on February 19, 2020 and computed by compounding daily returns. The fund average is computed by adding the average difference between the fund price index and the S&P 500 price index to the S&P 500 price index. Standard errors are estimated for this difference and are clustered on the Morningstar Institutional Category. 95% confidence intervals are plotted in red.

Computing F_t in equation (1) requires all of the fund's daily returns starting February 20 through day t. Any gap in the fund's return series, however short, would necessitate the fund's deletion from the average calculation. To avoid deleting too many funds, we replace any missing returns by the average return across all funds with the same FTSE/Russell

Covid Economics 38, 16 July 2020: 1-36

benchmark on the same day, thus preserving the average level of performance across funds. We only replace missing returns for which there exists a non-missing return later in the fund's history by April 30. That is, we do not replace any missing returns at the end of a fund's history because funds that stop reporting returns to Morningstar may no longer be alive. Altogether, we replace 19,124 missing returns, which account for 13.9% of our sample. We apply this "patch" not only in Figure 1 but also in Figures 2 through 8.[6] We do not replace missing returns in Tables 1 through 3 because the analysis behind those tables does not require funds to have continuous return series.

Figure 1 shows that the S&P 500 loses 34% of its value between February 19 and March 23, before gaining 30% by April 30, 2020. The average active fund performs similarly, but it significantly underperforms the S&P 500 during the crisis. The April 30 price index levels are 86.01 for the S&P 500 but only 82.45 for the average active fund.

Given its focus on large-cap stocks, the S&P 500 is not the most appropriate benchmark for every fund. Several large-cap technology stocks performed well during the crisis, making the S&P 500 hard to beat. We thus compare each fund's return also to the returns of two benchmarks tailored to the fund's investment style: the prospectus benchmark and the FTSE/Russell benchmark. The prospectus benchmark is chosen by the fund itself (with some potential for strategic choice, as discussed by Sensoy, 2009), whereas the FTSE/Russell benchmark is assigned to each fund by Morningstar based on the fund's holdings.

Figure 2 compares fund performance to the FTSE/Russell benchmark (Panel A), the prospectus benchmark (Panel B), and the S&P 500 (Panel C). Unlike Figure 1, which plots index levels, Figure 2 plots the cumulative performance of the average active fund relative to the benchmark. Specifically, at each date t after February 19, 2020, the figure plots the average value of $\log(F_t) - \log(B_t)$, where F_t and B_t are defined in equations (1) and (2).

Figure 2 shows that active funds significantly underperform their benchmarks, on average. As of April 30, 2020, the average underperformance over the ten-week period is 1.53% relative to the FTSE/Russell benchmark, 0.94% relative to the prospectus benchmark, and 4.77% relative to the S&P 500. This underperformance is highly statistically significant in Panel C, and marginally significant in Panels A and B. Moreover, Figure 2 underestimates the actual underperformance due to a mild survivorship bias because average F_t can only be computed across funds that have survived through time t. During the ten-week period, 22 funds drop out of our sample, so their returns are excluded from the plot as of April 30.

[6] All figures look virtually identical if we replace missing returns not by average fund returns but rather by the returns on the fund's FTSE/Russell benchmark on the same day.

Covid Economics 38, 16 July 2020: 1-36

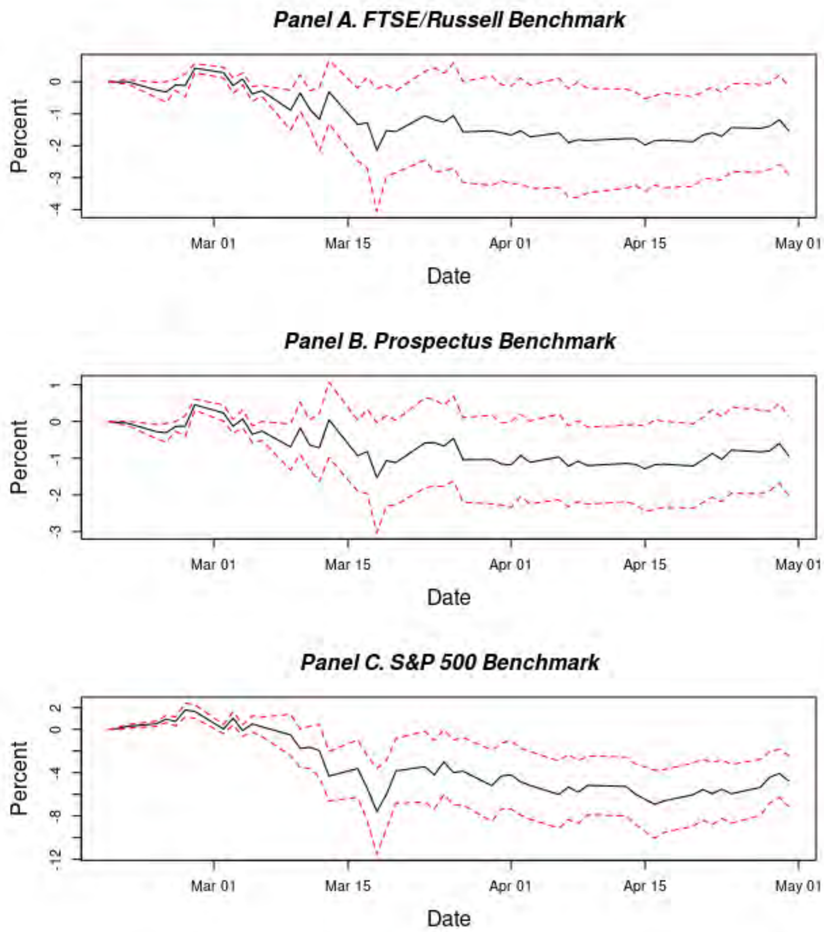


Figure 2. Average Benchmark-Adjusted Fund Performance. This figure plots the cumulative compound performance of the average active equity mutual fund in February 20 through April 30, 2020 relative to three benchmarks: the Morningstar-designated FTSE/Russell benchmark (Panel A), the prospectus benchmark (Panel B), and the S&P 500 (Panel C). Relative performance is measured by $\log(F_t) - \log(B_t)$, where F_t and B_t are the cumulative compounded daily returns of the average fund and the benchmark, respectively. Standard errors are estimated for this difference and are clustered on the Morningstar Institutional Category. 95% confidence intervals are plotted in red.

Table 1 reports average benchmark-adjusted fund performance in a way that is immune to the survivorship bias. For each fund, live or dead, we take all of the fund's available daily returns in the given time period and subtract the same days' returns on the corresponding benchmark. We report annualized averages of those benchmark-adjusted returns in the first three columns of Panel A of Table 1. The average fund underperforms its FTSE/Russell benchmark by 11.02% per year, with a t-statistic of −3.90. Average underperformance relative to the prospectus benchmark is slightly smaller, 7.70% per year, but still significant (t = −2.49). The average fund underperforms the S&P 500 by a whopping 29.12% per year (t = −5.37). We observe underperformance not only during the full ten-week crisis period but also during both subperiods, the first of which captures the market crash and the second the recovery, and also in the pre-crisis period (October 1, 2019 to January 31, 2020).

Table 1
Fund Performance

This table describes active equity mutual funds' performance against both benchmarks and factor models. Panel A reports simple averages across funds of estimated deltas and alphas, all reported in annualized percentage terms. The deltas are average differences between the fund's net returns and its benchmark returns. The benchmarks are the FTSE/Russell benchmark, the prospectus benchmark, and the S&P 500. The alphas are estimated intercepts from the regressions of excess net fund returns on factor returns. The factor models are described in the text. Panel B reports the fraction of funds that underperform (i.e., have a negative delta or alpha). The time periods are: crisis (February 20 to April 30, 2020); crash (February 20 to March 23, 2020); recovery (March 24 to April 30, 2020); and pre-crisis (October 1, 2019 to January 31, 2020). Standard errors are clustered on the Morningstar Institutional Category. t-statistics are in brackets.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$\Delta^{FTSE/Rus}_{Bench}$	Δ^{Prosp}_{Bench}	$\Delta^{S\&P500}_{Bench}$	α^{CAPM}	α^{FF3}	α^{Car4}	α^{FF5}	$\alpha^{FF5+Mom}$
Panel A. Average Fund Performance (%)								
Crisis	-11.02	-7.70	-29.12	-29.11	-11.30	-7.84	-9.90	-7.62
	[-3.90]	[-2.49]	[-5.37]	[-7.02]	[-5.16]	[-3.22]	[-4.51]	[-3.25]
Crash	-7.91	-8.19	-64.31	-80.94	-37.75	-38.98	-49.84	-51.11
	[-1.54]	[-1.33]	[-4.77]	[-8.33]	[-4.25]	[-3.90]	[-5.91]	[-5.77]
Recovery	-12.68	-7.55	-5.81	8.47	17.49	20.76	18.75	22.83
	[-3.94]	[-2.13]	[-1.14]	[1.46]	[2.83]	[4.12]	[3.00]	[4.42]
Pre-Crisis	-2.11	-1.28	-3.03	-5.14	-3.12	-3.01	-2.65	-2.59
	[-2.86]	[-2.09]	[-1.80]	[-3.87]	[-4.37]	[-3.64]	[-3.22]	[-2.88]
Panel B. Fraction of Funds Underperforming (%)								
Crisis	57.59	54.17	74.24	80.15	69.66	60.35	67.80	60.43
Crash	51.62	48.57	63.48	83.79	78.57	80.01	80.65	81.41
Recovery	55.73	55.64	55.77	53.09	39.53	34.79	40.08	34.28
Pre-Crisis	63.58	59.85	67.26	71.20	73.31	71.54	70.02	69.51

The remaining columns of Panel A of Table 1 report average fund alphas from five multifactor models: the capital asset pricing model (CAPM), the three-factor model of Fama and French (1993), the four-factor model of Carhart (1997), the five-factor model of Fama and French (2015), and a six-factor model that includes those five factors plus momentum. For a fund's alpha to be included in the average, the fund must have at least 15 non-missing net returns for the time period of interest. All five alphas are significantly negative during the crisis period, ranging from −7.62% per year ($t = -3.25$) for the six-factor model to −29.11% per year ($t = -7.02$) for the CAPM. The alphas are particularly negative during the crash period (February 20 to March 23, 2020), ranging from −37.75% to −80.94% per year across the five models.

Panel B of Table 1 shows that 57.6% of funds underperform their FTSE/Russell benchmarks during the crisis. Additionally, 54.2% of funds underperform prospectus benchmarks and 74.2% of funds—almost three quarters!—underperform the S&P 500. More than 80% of funds have negative CAPM alphas during the crisis period. The fraction of funds with negative alphas ranges from 60.4% to 80.2% across the five models. Regardless of how we look at the data, we see active funds underperforming during the crisis.

3.1. Sustainability

We find that funds with higher sustainability ratings perform better during the crisis. For each fund, Morningstar evaluates how well the fund's holdings perform on ESG issues relative to the fund's peer group (i.e., Morningstar Global Category). Morningstar uses company-level ESG scores from Sustainalytics to determine each fund's asset-weighted average unmanaged ESG risk exposure. Then, within each peer group, these scores are fitted to an approximate normal distribution to award 1, 2, 3, 4, or 5 sustainability globes to each fund.[7] Funds with 5 globes are the most sustainable and funds with 1 globe are the least sustainable. We find that funds with more globes perform better during the crisis.

Figure 3 shows the distributions of cumulative returns during the crisis across funds with different sustainability ratings, which are assigned as of January 2020. We collect funds in two groups: funds with 4 or 5 globes ("high sustainability") and funds with 1 or 2 globes ("low sustainability"). Panel A shows the distributions of cumulative total fund returns, whereas Panel B shows cumulative benchmark-adjusted returns, which are adjusted using FTSE/Russell benchmarks. Specifically, Panel A shows $\log(F_t)$, where F_t is in equation (1) and t corresponds to April 30, 2020, while Panel B shows $\log(F_t) - \log(B_t)$, where B_t is in equation (2) for the fund's FTSE/Russell benchmark. Both panels clearly show that more sustainable funds perform better in the crisis.

[7]Within each peer group, the top 10% of funds receive 5 globes, the next 22.5% receive 4 globes, the next 35% receive 3 globes, the next 22.5% receive 2 globes, and the bottom 10% receive 1 globe.

Covid Economics 38, 16 July 2020: 1-36

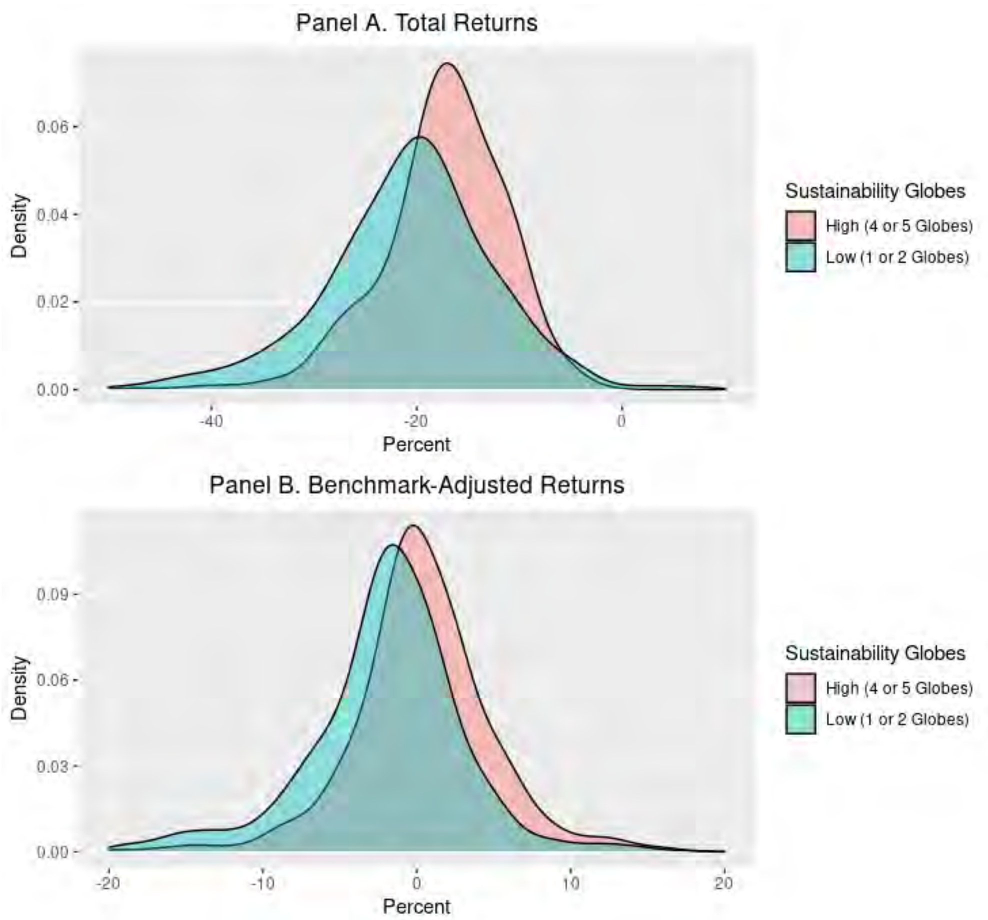


Figure 3. Cumulative Return Densities Across Sustainability Ratings. This figure plots densities of funds' cumulative returns from February 20 to April 30, 2020 for two categories of sustainability: high (four or five Morningstar globes) and low (one or two globes), both assigned as of January 2020. In Panel A, the cumulative returns are unadjusted, given by $\log(F_t)$ where $F_t = (1 + r_1^F)(1 + r_2^F) \ldots (1 + r_t^F)$ is the fund's cumulative total return. In Panel B, the cumulative returns are benchmark-adjusted, given by $\log(F_t) - \log(B_t)$, where B_t is the cumulative total return of the fund's FTSE/Russell benchmark.

Figure 4 presents the sustainability result from a perspective similar to Figure 2, plotting cumulative fund performance relative to the benchmark, or $\log(F_t) - \log(B_t)$. We consider the same three benchmarks as before: FTSE/Russell (Panels A and B), prospectus (Panels C and D), and the S&P 500 (Panels E and F). In the left panels (A, C, and E), we plot the average cumulative performance differences for each of the five globe groups. In the right panels (B, D, and F), we report 95% confidence intervals for high-sustainability funds (4 or 5 globes) and low-sustainability funds (1 or 2 globes).

Covid Economics 38, 16 July 2020: 1-36

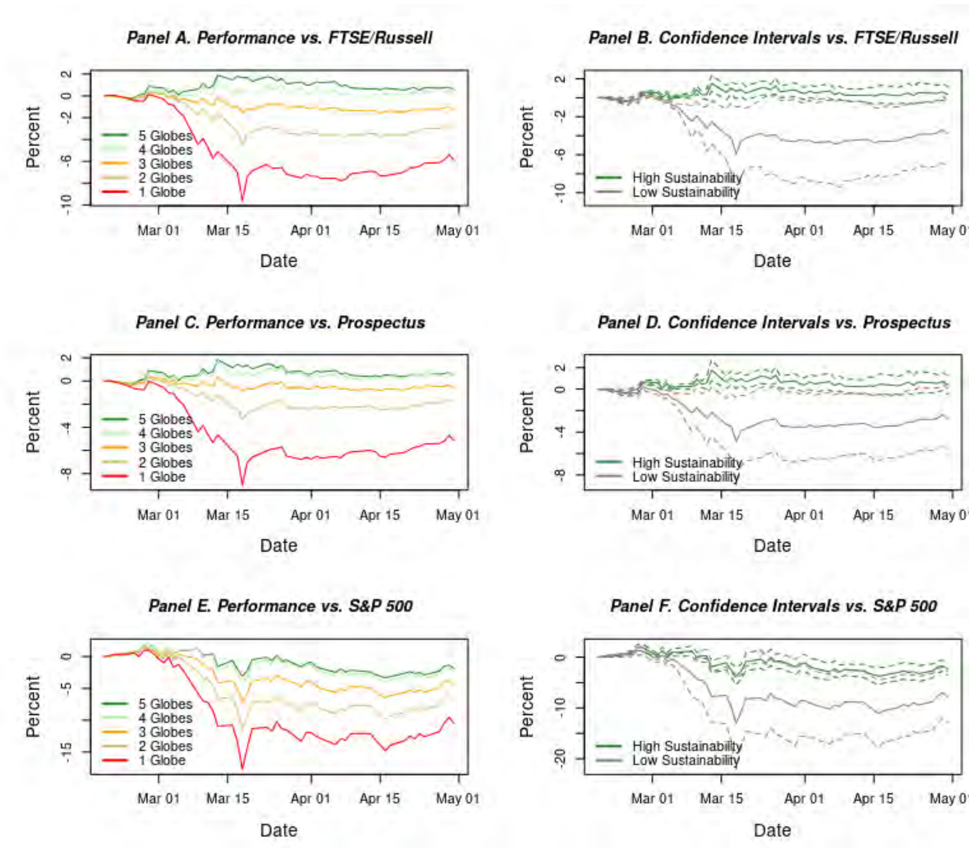


Figure 4. Benchmark-Adjusted Fund Performance: Sustainability Ratings. This figure plots the cumulative compound performance in February 20 through April 30, 2020 for fund categories with different numbers of Morningstar sustainability globes assigned as of January 2020. “High sustainability” denotes funds with 4 or 5 globes while “low sustainability” denotes funds with 1 or 2 globes. Performance is measured relative to the FTSE/Russell benchmark (Panels A and B), the prospectus benchmark (Panels C and D), and the S&P 500 (Panels E and F). Relative performance is measured by $\log(F_t) - \log(B_t)$, where F_t and B_t are the cumulative compounded daily returns of the average fund and the benchmark, respectively. Standard errors are estimated for this difference and are clustered on the Morningstar Institutional Category.

Remarkably, Figure 4 shows a monotonic relation between benchmark-adjusted fund performance and sustainability globes: five-globe funds outperform four-globe funds, which outperform three-globe funds, which in turn outperform two-globe funds, which beat one-globe funds. This monotonicity is present for all three benchmarks. The performance difference between high-sustainability funds and low-sustainability funds is marginally statistically significant.[8] The significance is stronger in the subsequent regression analysis in Table 2.

[8]This is a difference-in-difference type of calculation, where one difference is between the fund and its

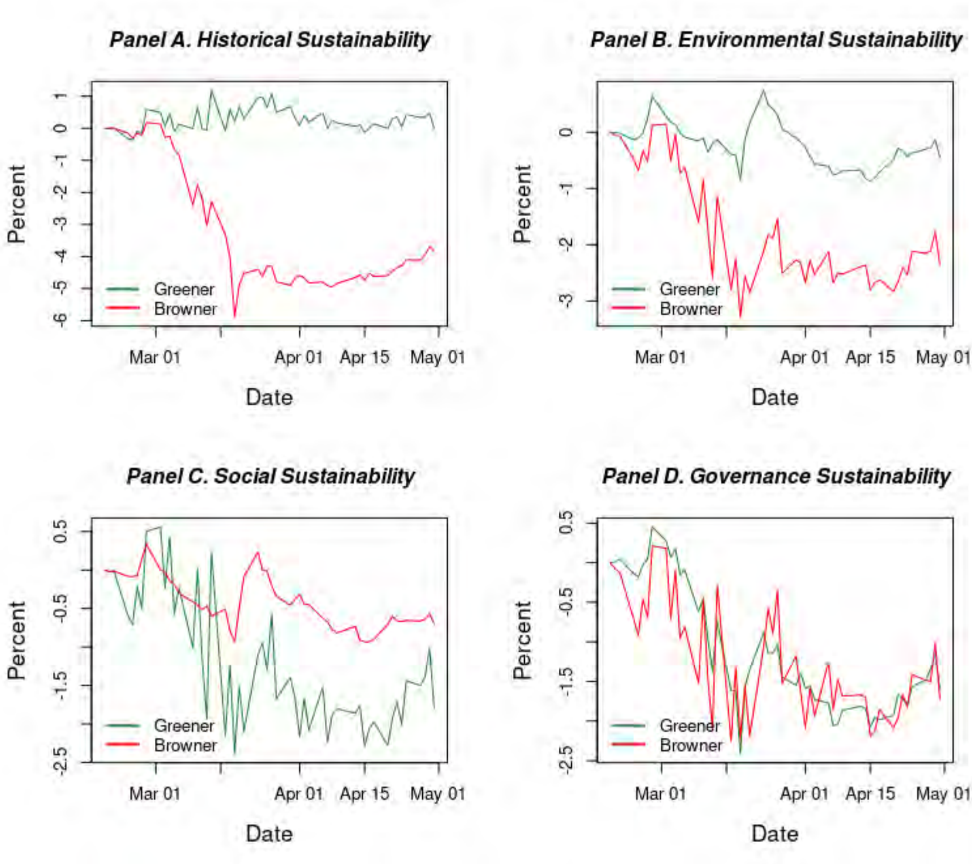


Figure 5. Benchmark-Adjusted Fund Performance: Sustainability Components. This figure plots the cumulative compound performance in February 20 through April 30, 2020 for fund categories with different Morningstar sustainability scores. These scores represent historical portfolio sustainability scores (Panel A), environmental scores (Panel B), social scores (Panel C), and governance scores (Panel D). For each of the four scores, the top 30% most sustainable funds are labeled as "greener" and the bottom 30% of funds are labeled "browner." Performance is measured relative to the FTSE/Russell benchmark by $\log(F_t) - \log(B_t)$, where F_t and B_t are the cumulative compounded daily returns of the average fund and the benchmark, respectively.

Given the important role of sustainability in determining fund performance, we investigate which dimensions of sustainability—E, S, or G—matter the most during the crisis. After sorting funds by their individual E, S, and G scores from Morningstar, we separate funds into two groups, "greener" (top 30%) and "browner" (bottom 30%), for each of the three scores. We do the same for the composite historical sustainability score, based on

benchmark and the other difference is between funds with different numbers of globes.

Covid Economics 38, 16 July 2020: 1-36

Covid Economics 38, 16 July 2020: 1-36

which Morningstar assigns globes to each fund.[9] We perform the greener-versus-browner comparisons in Figure 5, whose four panels are analogous to Panel B of Figure 4, except that sustainability globes are replaced by the four metrics described above. In all four panels, we benchmark funds against FTSE/Russell.

Figure 5 shows that funds with high sustainability scores outperform those with low scores. This result from Panel A is not surprising, given the prior results from Figure 4. More interesting, funds with high environmental (E) scores outperform those with low E scores (Panel B), whereas funds with high social (S) scores underperform those with low S scores (Panel C). According to Panel D, funds' governance (G) scores have no effect on performance. To make the figure easy to read, we do not show confidence intervals, but we do show them in the Appendix. Only the pattern in Panel B is statistically significant.

Figures 3 through 5 demonstrate that more sustainable funds perform better during the crisis. We further examine this result by conducting regression analysis, with two benefits. First, regressions allow us to see whether the result survives the inclusion of many control variables. Second, we remove the slight survivorship bias discussed earlier.

Table 2 analyzes the determinants of crisis-period fund performance in cross-sectional regressions with large numbers of controls. Panel A focuses on benchmark-adjusted performance, using FTSE/Russell benchmarks. Panel B considers factor-adjusted performance, using the four-factor Carhart (1997) model. The right-hand-side variables include indicators for sustainability, exclusions, and the growth investment style, as well as the Morningstar star rating. Fund-level controls include the log of fund age, the log of the fund's total net assets, fund turnover, expense ratio, cash position, the Morningstar medal rating, and market beta (in Panel B only because there is no beta estimation in Panel A). Industry controls are the fund's net investment position as a percentage of net assets in industries including energy, healthcare, and technology, among others. All regressions include style fixed effects, where style is measured at the level of the Morningstar Global Category. As a result, the relevant comparisons are across funds within the same investment style.

[9]A fund's individual E, S, and G scores do not simply add up to the fund's historical sustainability score. There does not appear to be a simple direct mapping between the two sets of scores.

Table 2
Determinants of Fund Performance During the Crisis

The table reports slope coefficients estimated from regressions of fund performance in February 20 to April 30, 2020 on fund characteristics and controls. In Panel A, the dependent variable is FTSE/Russell-benchmark-adjusted performance; in Panel B, it is the Carhart four-factor alpha. Both performance measures are estimated using simple returns and expressed in annualized percentage terms. Global category fixed effects are based on the Morningstar Global Category variable. Fund-level controls include the log of the fund's age in days, the log of the fund's January 31, 2020 total net assets (TNA), turnover ratio as of January 2020, net expense ratio as of January 2020, net cash position (as a percent of TNA) as of January 2020, Morningstar medal rating as of January 2020, and, in Panel B only, market beta estimated from the October 1, 2019 to January 31, 2020 period. Industry controls include the fund's net position as a percent of TNA in basic materials, communication services, consumer cyclical, consumer defensive, healthcare, industrials, real estate, technology, energy, financial services, and utilities. Standard errors are clustered on the Morningstar Institutional Category. *t*-statistics are in brackets.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Panel A. Benchmark-Adjusted Performance				
$\mathbb{I}$(4 or 5 Sustainability Globes)	14.21				11.51	8.61	9.76
	[4.85]				[3.22]	[2.26]	[2.60]
$\mathbb{I}$(Employs Exclusions)		8.61			5.47	2.03	2.79
		[3.26]			[2.44]	[1.05]	[1.24]
Star Rating			5.78		5.12	7.00	6.49
			[2.84]		[2.42]	[3.50]	[3.41]
$\mathbb{I}$(Growth Tilt)				12.43	7.24	9.39	5.15
				[2.35]	[1.16]	[1.70]	[0.75]
Global Category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund-Level Controls	No	No	No	No	No	Yes	Yes
Industry Controls	No	No	No	No	No	No	Yes
Observations	2,494	2,561	2,286	2,561	2,251	1,632	1,604
Adjusted R^2	0.06	0.05	0.06	0.06	0.06	0.12	0.15
			Panel B. Factor-Adjusted Performance				
$\mathbb{I}$(4 or 5 Sustainability Globes)	5.59				2.67	3.04	3.47
	[4.25]				[2.39]	[2.55]	[3.15]
$\mathbb{I}$(Employs Exclusions)		−0.89			−2.61	−3.46	−3.16
		[−0.50]			[−1.52]	[−2.12]	[−2.19]
Star Rating			3.15		2.51	3.25	3.13
			[3.35]		[2.79]	[5.92]	[5.42]
$\mathbb{I}$(Growth Tilt)				10.62	7.53	7.51	7.77
				[5.58]	[4.09]	[3.74]	[4.10]
Global Category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund-Level Controls	No	No	No	No	No	Yes	Yes
Industry Controls	No	No	No	No	No	No	Yes
Observations	2,233	2,363	2,104	2,363	2,020	1,522	1,494
Adjusted R^2	0.10	0.12	0.11	0.12	0.10	0.42	0.46

Covid Economics 38, 16 July 2020: 1-36

Covid Economics 38, 16 July 2020: 1-36

Table 2 confirms that funds with higher sustainability ratings perform better during the crisis. As before, we define high-sustainability funds as those with 4 or 5 globes. Column 1 of the table includes no controls other than style fixed effects. In column 1 of Panel A, the slope on the high-sustainability indicator is 14.21 (t = 4.85), indicating that high-sustainability funds outperform the remaining funds within the same style by 14.21% per year during the crisis. The high-sustainability indicator is also highly significant in column 1 of Panel B (t = 4.25), where fund returns are factor-adjusted rather than benchmark-adjusted. Sustainability thus remains a significant determinant of performance even after style fixed effects are included. This is not surprising, given our prior results, because both sustainability ratings and fund returns are style-adjusted, though in slightly different ways—sustainability ratings by Morningstar, with respect to the Morningstar Global Category, and returns by us, with respect to the fund's FTSE/Russell benchmark.

More interesting, sustainability remains significantly associated with fund performance after the inclusion of fund-level and industry controls. The slope on the high-sustainability indicator decreases as controls are added, but it remains both statistically and economically significant even when all controls are included: 9.76 (t = 2.60) in Panel A and 3.47 (t = 3.15) in Panel B. To summarize, we find that funds with high sustainability ratings perform better during the crisis.

A subset of funds employ exclusions in their investment process. These funds exclude from their portfolios stocks of firms such as tobacco producers or gun manufacturers whose business they deem unacceptable. Exclusions represent one possible approach to sustainability, one that discards the opportunities to engage with the firm as well as to benefit from the potential mispricing of the firm's stock. 107 of our funds employ exclusions, representing 3.9% of our TNA-screened sample.

Table 2 shows that funds that employ exclusions outperform same-style funds that do not employ exclusions by 8.61% per year (t = 3.26) in terms of benchmark-adjusted returns. However, this result weakens, and eventually loses significance, after adding enough control variables. The result does not obtain on a factor-adjusted basis; if anything, it goes the other way (Panel B). The exclusion aspect of sustainability therefore does not have a robust association with fund performance during the crisis.

CEPR PRESS

Covid Economics 38, 16 July 2020: 1-36

3.2. Star Ratings

Besides sustainability, the most important determinant of active fund performance during the crisis is the fund's star rating as of January 31, 2020. To calculate star ratings, Morningstar computes each fund's risk-adjusted performance over the prior three, five, and ten years relative to the fund's peer group. Averaging across the three periods, Morningstar awards 1, 2, 3, 4, or 5 stars to each fund, with 5 stars going to the best-performing funds.[10] We find that funds with higher star ratings perform better during the crisis.

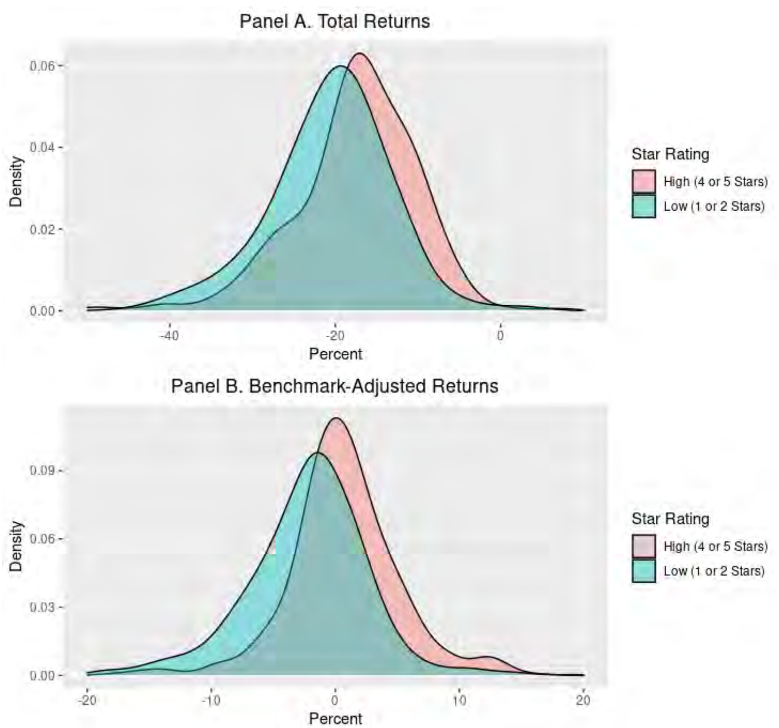


Figure 6. Cumulative Return Densities Across Star Ratings. This figure plots densities of funds' cumulative returns from February 20 to April 30, 2020 for two categories of star ratings: high (four or five Morningstar stars) and low (one or two stars), both assigned as of January 2020. In Panel A, the cumulative returns are unadjusted, given by $\log(F_t)$ where $F_t = (1 + r_1^F)(1 + r_2^F)\ldots(1 + r_t^F)$ is the fund's cumulative total return. In Panel B, the cumulative returns are benchmark-adjusted, given by $\log(F_t) - \log(B_t)$, where B_t is the cumulative total return of the fund's FTSE/Russell benchmark.

[10]As with the sustainability globes, within each peer group, the top 10% of funds receive 5 stars, the next 22.5% receive 4 stars, the next 35% receive 3 stars, the next 22.5% receive 2 stars, and the bottom 10% receive 1 star. A fund must have at least three years of performance to be considered for a rating, and depending on its age, a combination of three-year, five-year, and ten-year performance measures are averaged to construct the fund's raw performance score.

Figure 6 shows the distributions of cumulative returns during the crisis across funds with different star ratings, similar to Figure 3. We collect funds in two groups: funds with 4 or 5 stars ("high") and funds with 1 or 2 stars ("low"). Panel A shows the distributions of cumulative total fund returns whereas Panel B shows cumulative returns adjusted for FTSE/Russell benchmark returns. Both panels clearly show that funds with more stars perform better during the crisis.

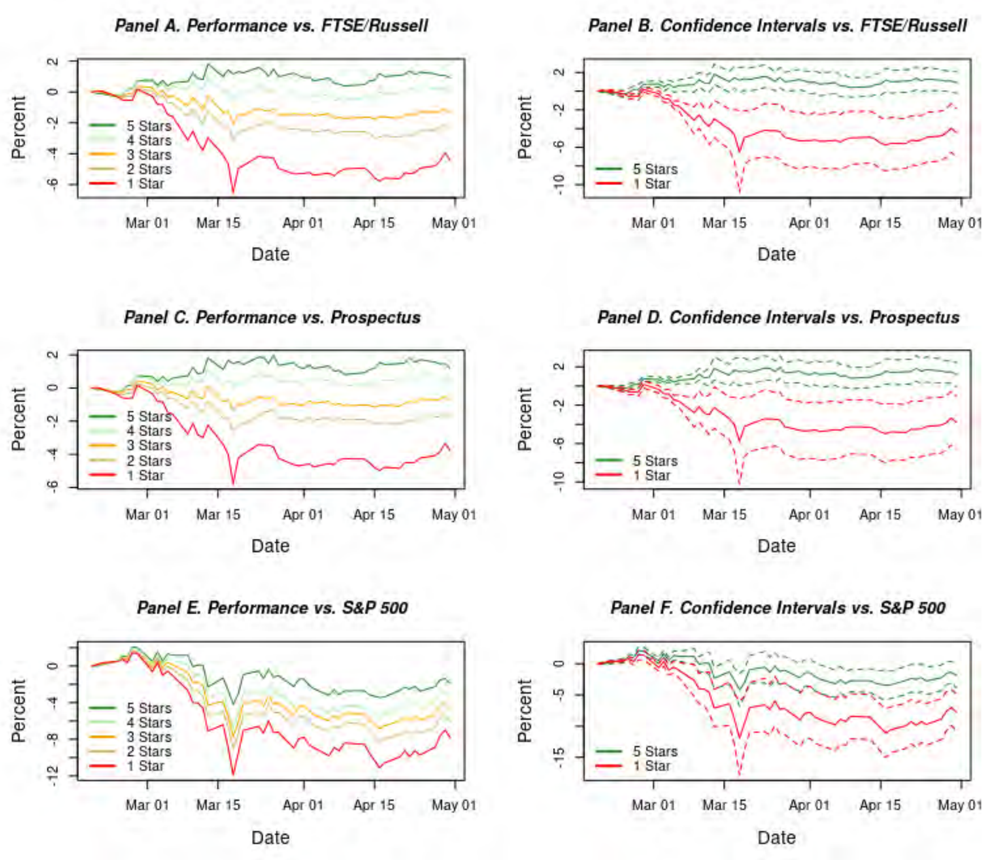


Figure 7. Benchmark-Adjusted Fund Performance: Star Ratings. This figure plots the cumulative compound performance in February 20 through April 30, 2020 for fund categories with different numbers of Morningstar stars assigned as of January 2020. Performance is measured relative to the FTSE/Russell benchmark (Panels A and B), the prospectus benchmark (Panels C and D), and the S&P 500 (Panels E and F). Relative performance is measured by $\log(F_t) - \log(B_t)$, where F_t and B_t are the cumulative compounded daily returns of the average fund and the benchmark, respectively. Standard errors are estimated for this difference and are clustered on the Morningstar Institutional Category.

CEPR PRESS

Figure 7 shows the same result from a different perspective. Similar to the layouts from Figures 2 and 4, Figure 7 plots cumulative benchmark-adjusted fund performance for groups of funds with different star ratings. The relation between benchmark-adjusted performance and star ratings is monotonic across the five star groups, with five-star funds performing the best and one-star funds performing the worst. This striking monotonicity is observed for all three benchmarks. The figure also shows that five-star funds outperform one-star funds by a significant margin for all three benchmarks.

Table 2 confirms the important role of the star rating in our regression setting with controls and style fixed effects. The star rating significantly predicts both benchmark-adjusted and factor-adjusted returns, with t-statistics ranging from 2.42 to 3.50 in Panel A and from 2.79 to 5.92 in Panel B. This is a surprising result—it is not clear a priori why Morningstar star ratings, which are computed before the crisis from historical risk-adjusted returns, should have such strong predictive power for fund performance during the crisis. The result is significant not only statistically but also economically. For example, the slope coefficient of 5.78 in column 3 of Panel A indicates that one extra star is associated with a higher crisis-period benchmark-adjusted return of 5.78% per year. Therefore, a five-star fund outperforms a one-star fund of the same style by four times that amount, 23.1% per year, on average.

Covid Economics 38, 16 July 2020: 1-36

3.3. Value versus Growth

Sustainability and stars are the most robust predictors of active fund performance during the crisis. Next in line, though less robust, is the value/growth investment style. We find that growth funds tend to outperform value funds. Importantly, we are not saying that the growth style outperforms the value style during the crisis—that is well known (e.g., HML's crisis-period return is −18%). What we are saying is that growth funds deliver higher returns than value funds on a style-adjusted basis.

To decide which funds follow the value and growth investment styles, we use the equity style box variable from Morningstar. We define value funds as funds classified as large-cap value, mid-cap value, or small-cap value. We define growth funds as funds classified as large-cap growth, mid-cap growth, or small-cap growth.

Figure 8 shows that growth funds outperform value funds, on average, for all three benchmarks. The outperformance is statistically significant when measured against the S&P 500 and prospectus benchmarks, but not against the FTSE/Russell benchmarks.

Covid Economics 38, 16 July 2020: 1-36

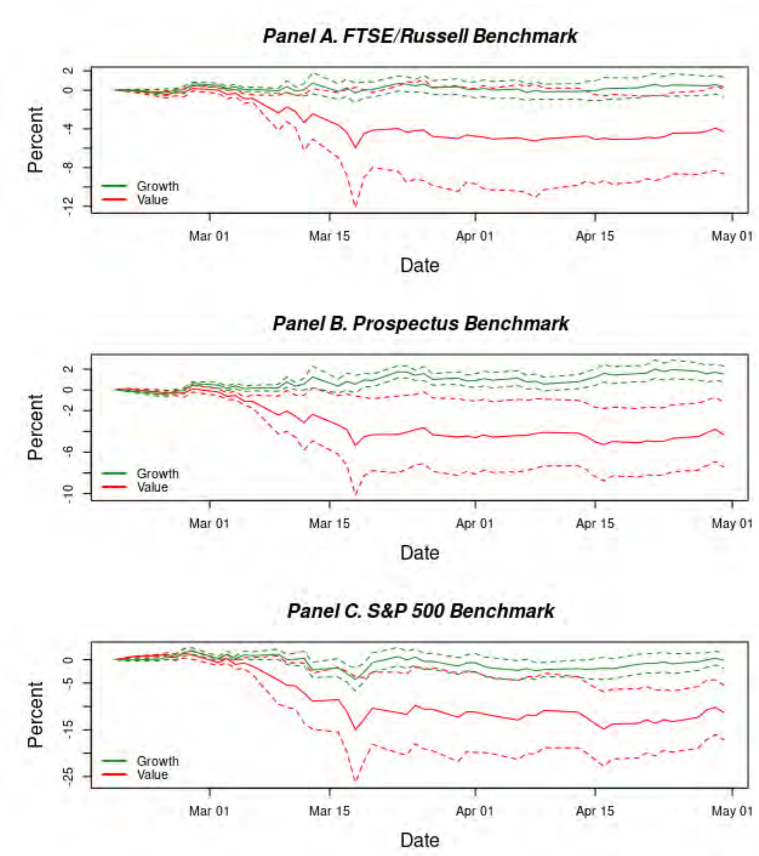


Figure 8. Benchmark-Adjusted Fund Performance: Growth vs. Value Funds. This figure plots the cumulative compound performance in February 20 through April 30, 2020 for growth vs value funds, as determined by the Morningstar equity style box. Performance is measured relative to the FTSE/Russell benchmark (Panel A), the prospectus benchmark (Panel B), and the S&P 500 (Panel C). Relative performance is measured by $\log(F_t) - \log(B_t)$, where F_t and B_t are the cumulative compounded daily returns of the average fund and the benchmark, respectively. Standard errors are estimated for this difference and are clustered on the Morningstar Institutional Category. 95% confidence intervals are shown.

Table 2 finds the same outperformance in our regression setting with controls and style fixed effects. In Panel B, where fund performance is factor-adjusted, the indicator variable for the growth tilt is always positive and significant, with t-statistics exceeding 3.7 in all specifications. For example, using the estimate from column 4, growth funds outperform non-growth funds by 10.62% per year (t = 5.58) during the crisis. In Panel A, the estimated slopes on the growth indicator are similar in magnitude to those in Panel B but their statistical significance is substantially weaker, with t-statistics ranging from 0.75 to 2.35 across the four specifications.

CEPR PRESS

3.4. Robustness

Our main results are robust to a variety of methodological modifications. As noted earlier, our sample is restricted to active equity funds with at least $15 million in total net assets as of January 31, 2020. However, we show in the Appendix that the results from Table 1 are extremely similar even if we do not impose this size screen.

Another screen that is commonly imposed on mutual fund samples is an age screen. Researchers often exclude young funds because of a concern about the incubation bias (Evans, 2010). This bias can appear if researchers analyze historical fund data with a delay that would allow the bias to creep in. The bias is not a concern in our study because there is no such delay—we constructed our fund sample in May 2020, shortly after the end of our sample period. Nonetheless, we show in the Appendix that our main results are extremely similar also when we exclude funds less than two years old from the sample.

Our tables report evidence based on simple returns. Our plots of cumulative performance are based on log (i.e., continuously compounded) returns because those returns cumulate over time in a tractable manner. This distinction is immaterial—our main table results are very similar if we replace simple returns by log returns, as we show in the Appendix.

To remove the effects of investment style, we include style fixed effects in our regressions. Nonetheless, we show in the Appendix that the regression results from Table 2 are similar if style fixed effects are excluded. The Appendix also reports subperiod results for Table 2 and its variations, dividing the full crisis period into the crash period (February 20 to March 23, 2020) and the recovery period (March 24 to April 30, 2020).

Recall that Figure 3 shows the distributions of crisis-period returns across two groups of funds, those with high sustainability ratings (4 or 5 globes) and low sustainability ratings (1 or 2 globes). In the Appendix, we present analogous plots showing three distributions corresponding to funds with 1, 3, and 5 globes, and also five distributions, one for each possible number of globes. Those plots are more cluttered but they convey the same message as Figure 3—that more sustainable funds perform better in the crisis.

Similarly, Figure 6 shows the distributions of returns across funds with high star ratings (4 or 5 stars) and low star ratings (1 or 2 stars). In the Appendix, we present analogous plots showing three distributions corresponding to funds with 1, 3, and 5 stars, and also five distributions, one for each possible number of stars. Again, those plots convey the same message as Figure 6: funds with more stars perform better during the crisis.

Covid Economics 38, 16 July 2020: 1-36

4. Fund Flows

Our key measure for assessing fund flows is the cumulative net fund flow percentage. Daily net fund flows, in dollars, are computed following Barber, Huang, and Odean (2016) as

$$FD_{i,t} = TNA_{i,t} - (1 + R_{i,t})TNA_{i,t-1} \ , \tag{3}$$

where $TNA_{i,t}$ is the total net assets of fund i on date t and $R_{i,t}$ is the net return of fund i on date t. To convert the dollar values of net fund flows into a cumulative percentage, the values of $FD_{i,t}$ are accumulated across the time period of interest and divided by the total net assets of fund i on the day before the period of interest begins. Given the sensitivity of the cumulative net flow percentage to missing values, we restrict consideration to funds with entirely non-missing daily net fund flows.[11]

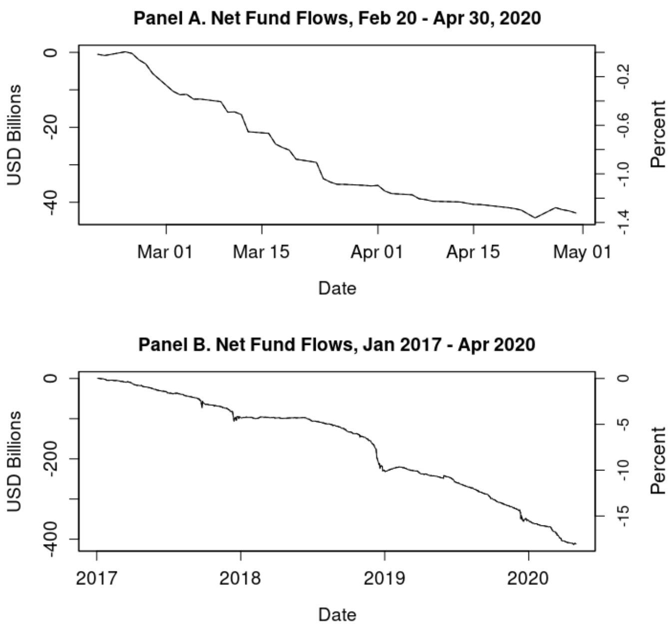


Figure 9. Aggregate Net Fund Flows. This figure plots aggregate net flows into active equity mutual funds during the crisis period (Panel A) and over the past three years (Panel B). Specifically, Panel A plots total cumulative net fund flows (in both USD billions and as a percent of February 19, 2020 aggregate total net assets) over the February 20 to April 30, 2020 period. Panel B covers the January 4, 2017 to April 30, 2020 period, and it expresses flows as a percent of January 3, 2017 total net assets.

[11] In our baseline fund size filtered sample, we retain 2,082 funds (75.3%) over the full February 20 to April 30, 2020 time period, 2,137 funds (77.3%) over the February 20 to March 23, 2020 time period, and 2,219 funds (80.3%) over the March 24 to April 30, 2020 time period.

Covid Economics 38, 16 July 2020: 1-36

Figure 9 shows the time series of cumulative net fund flows into active equity mutual funds, both in dollar terms and in percentage terms. Panel A shows that active funds experience steady outflows during the COVID-19 crisis of about 43 billion dollars, or 1.3% of assets under management. The pace of outflows is fairly rapid during the market crash between February 20 and March 23, 2020. Outflows continue, albeit at a slower pace, after the market rebound. Panel B shows that between January 2017 and April 2020, active funds experience outflows amounting to about 5% per year as a fraction of their initial assets. These steady outflows reflect the well-known ongoing trend toward passive investment management. Year 2020 does not stand out relative to prior years, indicating that crisis-period outflows largely extend their long-term pre-crisis trend.

4.1. Sustainability

We find that funds with higher sustainability ratings (i.e., more Morningstar globes) receive larger net flows during the crisis. Figure 10 plots cumulative net fund flows across funds with different globe ratings over the February 20 to April 30, 2020 time period. Panel A plots aggregate cumulative flows for each globe category, which we compute as total cumulative net flows into that category divided by that category's total net assets on February 19, 2020. The panel shows a near-monotonic relation between those ratings and net fund flows, with five-globe funds having the largest flows and one-globe funds the lowest flows. Panel B focuses on the five-globe and one-globe categories. For both of them, the panel plots average fund-level cumulative flows across all funds in the given category, scaled as a percent of the fund's February 19, 2020 total net assets, along with the 95% confidence intervals. The panel shows that five-globe funds receive significantly larger net flows than one-globe funds. In short, investors favor sustainable funds when moving their money during the crisis.

Figure 11 shows a similar pattern based on a different measure of sustainability: an indicator of whether the fund employs exclusions in its investment process. Funds that do not employ exclusions, which account for the vast majority of funds, experience net outflows during the crisis. However, funds that do employ exclusions experience net inflows, and the difference between the two groups is statistically significant.

Covid Economics 38, 16 July 2020: 1-36

CEPR PRESS

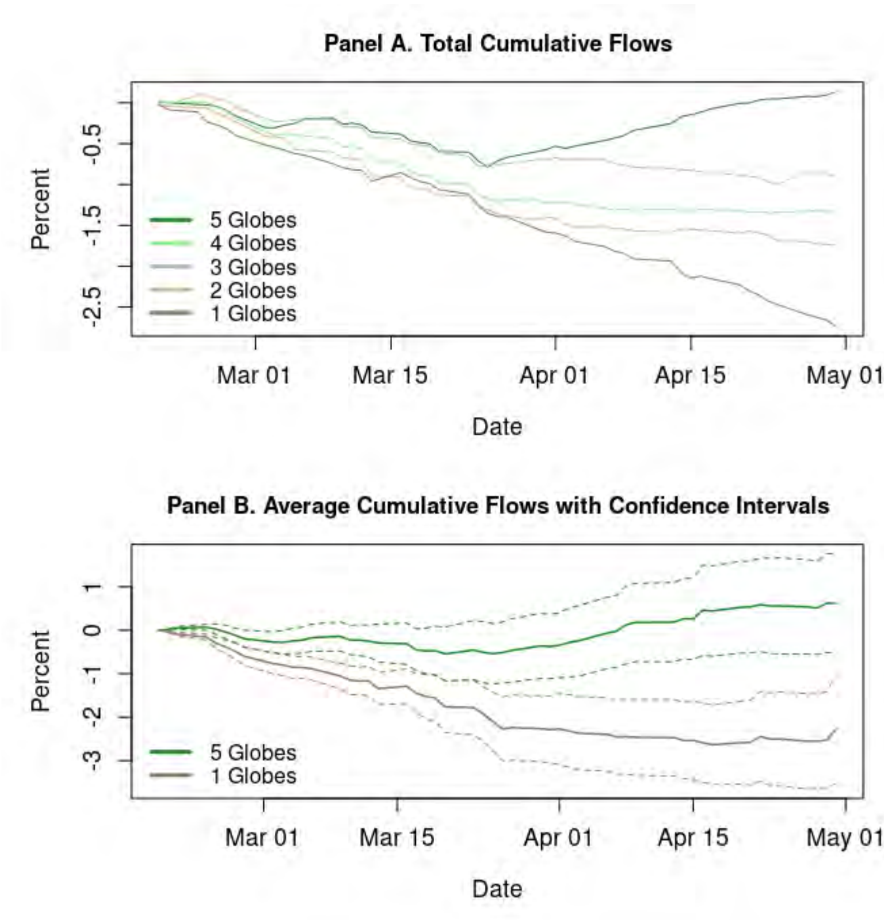


Figure 10. Fund Flows and Sustainability Ratings. This figure plots net fund flows over the February 20 to April 30, 2020 period for categories of funds sorted by Morningstar sustainability ratings. Panel A plots aggregate cumulative net flows for each of the five globe categories. Flows are aggregated within each category and accumulated over time, then scaled by the category's total net assets on February 19, 2020. Panel B plots the average across funds of cumulative net flows as a percent of the fund's February 19, 2020 total net assets, for the five- and one-globe categories only. Unlike in Panel A, the sample in Panel B is restricted to funds with at least \$15 million of total net assets as of January 31, 2020 and the net fund flow percentage is winsorized at the 2.5% and 97.5% levels. Panel B also plots 95% confidence intervals, with standard errors clustered on the Morningstar Institutional Category.

Covid Economics 38, 16 July 2020: 1-36

Covid Economics 38, 16 July 2020: 1-36

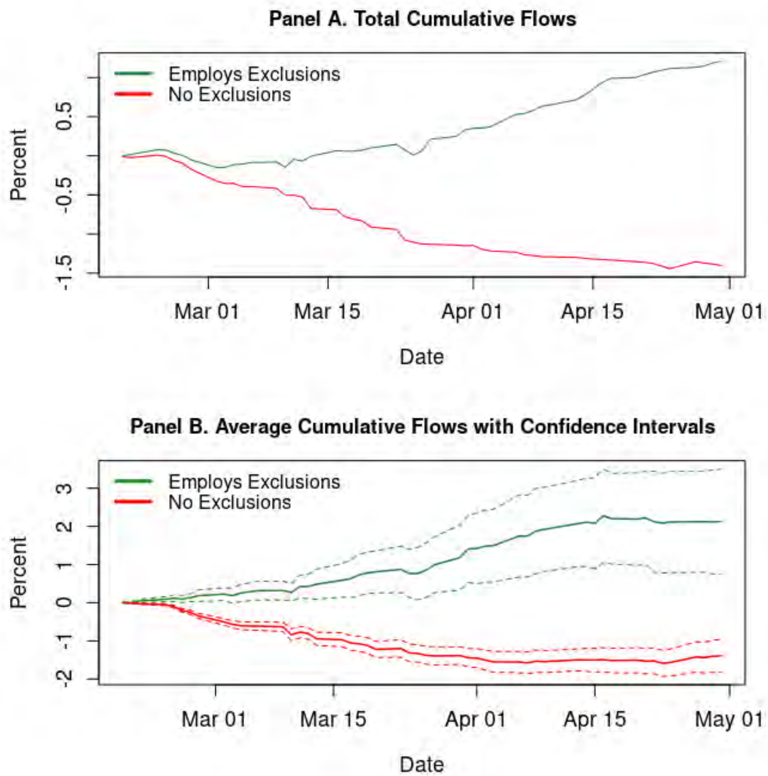


Figure 11. Fund Flows and Exclusions. This figure plots net fund flows over the February 20 to April 30, 2020 period for two categories of funds: those that do and do not employ exclusions in their investment process. Panel A plots aggregate cumulative net flows for both categories. Flows are aggregated within each category and accumulated over time, then scaled by the category's total net assets on February 19, 2020. Panel B plots the average across funds of cumulative net flows as a percent of the fund's February 19, 2020 total net assets, for both categories. Unlike in Panel A, the sample in Panel B is restricted to funds with at least $15 million of total net assets as of January 31, 2020 and the net fund flow percentage is winsorized at the 2.5% and 97.5% levels. Panel B also plots 95% confidence intervals, with standard errors clustered on the Morningstar Institutional Category.

Figure 12 unpacks sustainability into its E, S, and G dimensions. As before, we separate funds into high-E, low-E, high-S, low-S, high-G and low-G, where the high (low) group always denotes the top (bottom) 30% of funds. Figure 12 shows that cumulative aggregate net flows during the crisis are larger for high-E funds than for low-E funds. Low-E funds experience substantial outflows of 2.7% of assets, whereas the outflows from high-E funds are only 0.3%. Net flows are also larger for high-G funds than for low-G funds: low-G funds have outflows of 2.0% whereas high-G funds' outflows are only 1.0%. High-S funds actually experience larger outflows than low-S funds, but the difference between the two categories'

Covid Economics 38, 16 July 2020: 1-36

total flows is relatively small, only 0.7%, and the average outflow is in fact slightly larger for the low-S category. The effect of sustainability on fund flows thus seems driven by E and, to a lesser extent, G. Investors seem to have retained their focus on environmental issues even during the health crisis of 2020.

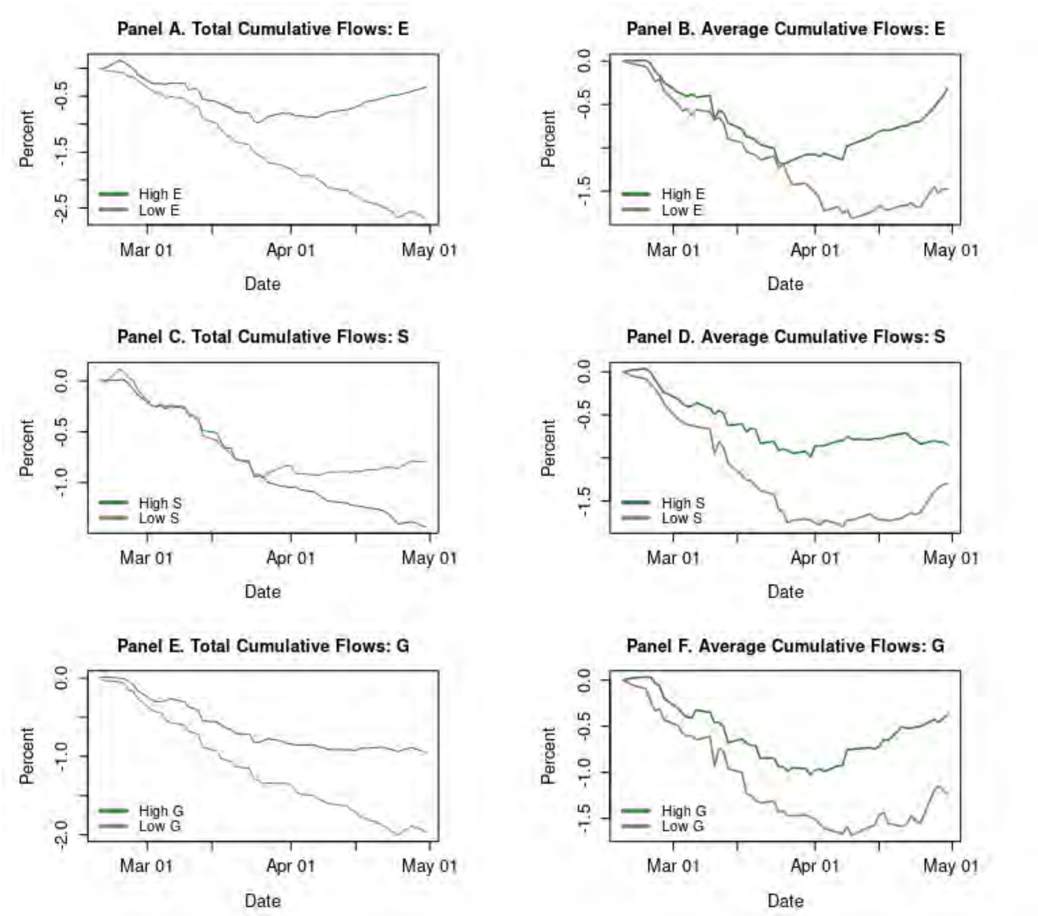


Figure 12. Fund Flows and ESG Scores. This figure plots net fund flows over the February 20 to April 30, 2020 period for funds in the top 30% ("high") and bottom 30% ("low") of environmental, social, and governance sustainability scores. The left panels plot aggregate cumulative net flows for the high and low E, S, and G categories. Flows are aggregated within each category and accumulated over time, then scaled by the category's total net assets on February 19, 2020. The right panels plot the average across funds of cumulative net flows as a percent of the fund's February 19, 2020 total net assets, for both the high and low categories. Unlike in the left panels, the samples in the right panels are restricted to funds with at least $15 million of total net assets as of January 31, 2020 and the net fund flow percentage is winsorized at the 2.5% and 97.5% levels.

Covid Economics 38, 16 July 2020: 1-36

Table 3 reports results from cross-sectional regressions of crisis-period net fund flows on the sustainability variables as well as style fixed effects and a large number of fund and industry controls. The sustainability variables remain significant even with these controls. The exclusions variable is the most robust, with t-statistics ranging from 2.50 to 4.02 across the seven different specifications. The indicator for high-E (i.e., green) funds is also generally significant, with t-statistics ranging from 1.89 to 3.09. The five-globe indicator is significant when included on its own as well as in several other specifications, but it loses significance when the exclusions indicator is included. Overall, these results indicate that investors favor sustainable funds while reallocating money during the pandemic of 2020.

Table 3
Determinants of Net Fund Flows During the Crisis

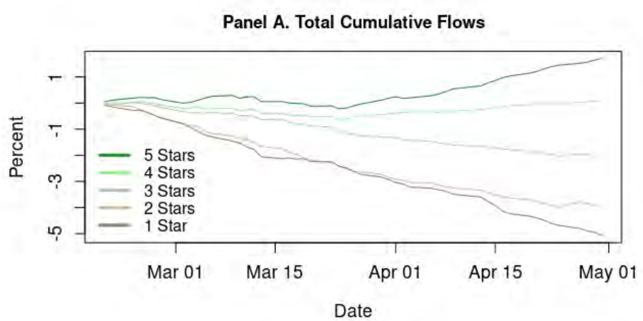
The table reports slope coefficients estimated from regressions of net fund flows in February 20 to April 30, 2020 on fund characteristics and controls. A fund's net flow is expressed as a percent of the fund's February 19, 2020 total net assets. Flows are winsorized at the 2.5% and 97.5% levels. Global category fixed effects are based on the Morningstar Global Category variable. Fund-level controls include an indicator for a growth tilt, the log of the fund's age in days, the log of the fund's January 31, 2020 total net assets (TNA), turnover ratio as of January 2020, net expense ratio as of January 2020, net cash position (as a percent of TNA) as of January 2020, Morningstar medal rating as of January 2020, and market beta estimated from the October 1, 2019 to January 31, 2020 period. Industry controls include the fund's net position as a percent of TNA in basic materials, communication services, consumer cyclical, consumer defensive, healthcare, industrials, real estate, technology, energy, financial services, and utilities. Standard errors are clustered on the Morningstar Institutional Category. t-statistics are in brackets.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\mathbb{I}$(5 Sustainability Globes)	1.76	0.92	1.36	1.44		0.70	1.16	1.22				
	[3.19]	[1.82]	[2.23]	[2.32]		[1.23]	[1.77]	[1.82]				
$\mathbb{I}$(Employs Exclusions)					2.84	2.61	2.70	2.75		2.69	2.95	3.11
					[3.37]	[3.09]	[2.50]	[2.53]		[4.02]	[3.46]	[3.43]
$\mathbb{I}$(Greener E)									1.67	1.02	1.04	1.91
									[3.09]	[2.08]	[1.89]	[2.72]
$\mathbb{I}$(Greener S)									−0.48	−0.43	−0.39	−0.88
									[−0.93]	[−0.83]	[−0.64]	[−1.22]
$\mathbb{I}$(Greener G)									0.76	0.63	1.00	1.26
									[1.39]	[1.09]	[1.31]	[1.69]
Star Rating		1.83	1.78	1.90		1.82	1.76	1.89		1.70	1.38	1.49
		[7.49]	[6.04]	[6.00]		[7.51]	[6.04]	[5.96]		[5.77]	[3.91]	[4.01]
Global Category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund-Level Controls	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Industry Controls	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Observations	2,082	1,863	1,434	1,390	2,082	1,863	1,434	1,390	1,503	1,348	1,037	1,020
Adjusted R^2	0.02	0.08	0.11	0.11	0.02	0.09	0.11	0.11	0.01	0.08	0.10	0.11

CEPR PRESS

4.2. Other Determinants of Fund Flows

Figure 13 plots cumulative net fund flows for funds with different star ratings as of January 31, 2020. Panel A shows a monotonic flow-star relation, with higher-star funds receiving higher net flows. Panel B shows that the differences in average cumulative net flows are statistically significant: five-star funds receive significantly larger average net flows than three-star funds, whose average flows are significantly larger than those of one-star funds. The positive flow-star relation is highly significant also after controlling for style fixed effects and the other controls (see Table 3). This relation is not surprising, however, because star ratings are constructed based on past returns, which are well known to have predictive power for fund flows. Star ratings essentially perform the role of catch-all controls for past returns in our regressions.



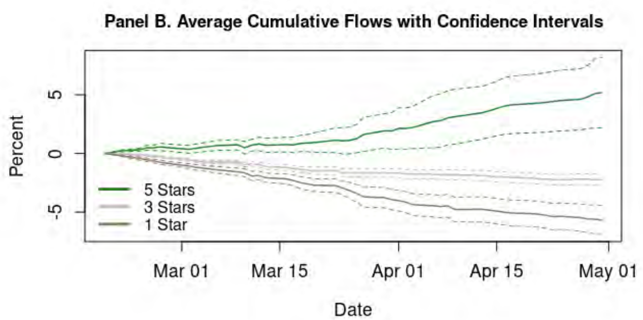


Figure 13. Fund Flows and Star Ratings. This figure plots net fund flows over the February 20 to April 30, 2020 period for categories of funds sorted by Morningstar star ratings. Panel A plots aggregate cumulative net flows for each of the five star categories. Flows are aggregated within each category and accumulated over time, then scaled by the category's total net assets on February 19, 2020. Panel B plots the average across funds of cumulative net flows as a percent of the fund's February 19, 2020 total net assets, for the five-, three-, and one-star categories only. Unlike in Panel A, the sample in Panel B is restricted to funds with at least $15 million of total net assets as of January 31, 2020 and the net fund flow percentage is winsorized at the 2.5% and 97.5% levels. Panel B also plots 95% confidence intervals, with standard errors clustered on the Morningstar Institutional Category.

Covid Economics 38, 16 July 2020: 1-36

Figure 14 shows that growth funds receive significantly larger net flows than value funds. This relation holds largely at the style level because it vanishes when we run regressions that include style fixed effects. We include the growth indicator variable among the fund-level controls in Table 3 but we suppress it because it is never statistically significant.

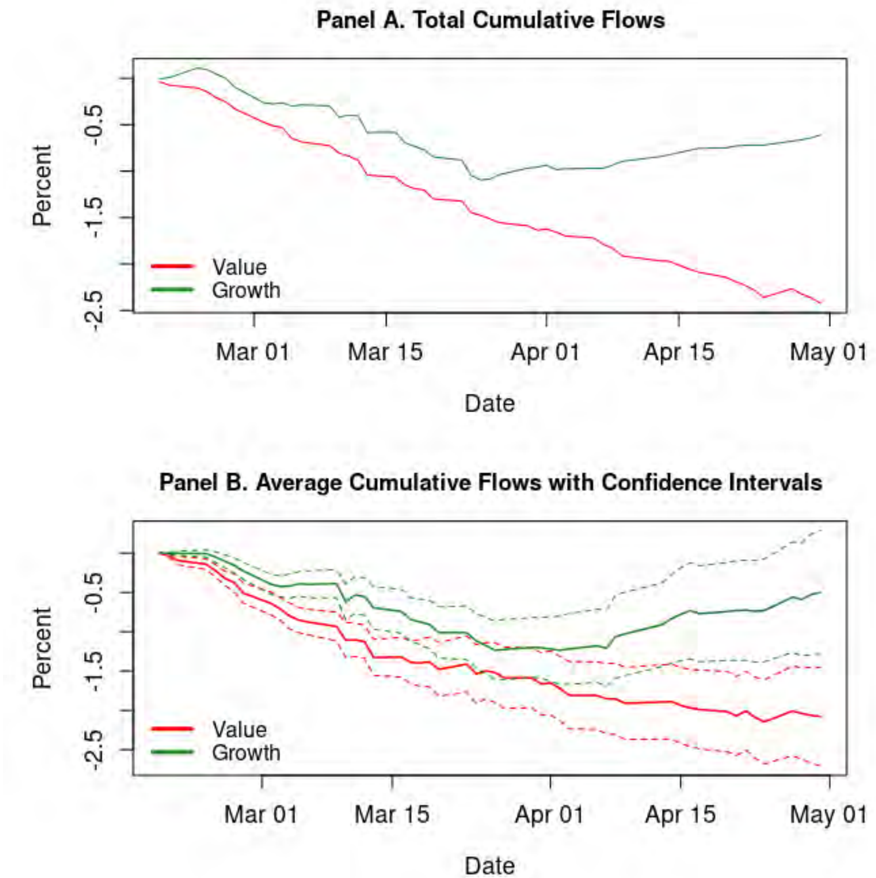


Figure 14. Fund Flows and Growth vs. Value Funds. This figure plots net fund flows over the February 20 to April 30, 2020 period for growth vs value funds, as determined by the Morningstar equity style box. Panel A plots aggregate cumulative net flows for both categories. Flows are aggregated within each category and accumulated over time, then scaled by the category's total net assets on February 19, 2020. Panel B plots the average across funds of cumulative net flows as a percent of the fund's February 19, 2020 total net assets, for both categories. Unlike in Panel A, the sample in Panel B is restricted to funds with at least $15 million of total net assets as of January 31, 2020 and the net fund flow percentage is winsorized at the 2.5% and 97.5% levels. Panel B also plots 95% confidence intervals, with standard errors clustered on the Morningstar Institutional Category.

Covid Economics 38, 16 July 2020: 1-36

Covid Economics 38, 16 July 2020: 1-36

As in our analysis of fund performance, we also report our regression results in specifications without style fixed effects. The results are quite similar to those reported in Table 3, as we show in the Appendix. The Appendix also reports subperiod (crash and recovery) results for Table 3, as well as for its version with no style fixed effects.

5. Conclusions

We analyze the performance and flows of U.S. active equity mutual funds during the COVID-19 crisis. We find that most active funds underperform passive benchmarks during the crisis, contradicting the hypothesis that active funds outperform in recessions. This underperformance is particularly large when measured against the S&P 500 index, but it is observed also relative to fund-specific style benchmarks. Funds with high sustainability ratings and high star ratings outperform those with low sustainability ratings and low star ratings, respectively. When reallocating capital across funds, investors favor funds with high sustainability ratings and funds that apply exclusion criteria. That investors retain their focus on sustainability during a major crisis indicates that sustainability is not just a luxury good.

While this paper appears to be the first to analyze the performance and flows of active funds during the COVID-19 crisis, it leaves plenty of room for future research. For example, while we focus on equity funds, fixed income funds also deserve a careful investigation. So do the sources of active funds' underperformance during the crisis. It would also be useful to extend our work to deepen our understanding of the dependence of investors' demand for sustainability on economic conditions.

REFERENCES

Albuquerque, Rui A., Yrjo Koskinen, Shuai Yang, and Chendi Zhang, 2020, Love in the time of COVID-19: The resiliency of environmental and social stocks, Working paper, Boston College.

Alfaro, Laura, Anusha Chari, Andrew N. Greenland, and Peter K. Schott, 2020, Aggregate and firm-level stock returns during pandemics, in real time, *Covid Economics* 4, 2–24.

Barber, Brad M., Xing Huang, and Terrance Odean, 2016, Which factors matter to investors? Evidence from mutual fund flows, *The Review of Financial Studies* 29, 2600–2642.

Baumol, William J. and Wallace E. Oates, 1979, *Economics, Environmental Policy, and Quality of Life*, Inglewood Cliffs, NJ: Prentice-Hall.

Bialkowski, Jedrzej, and Laura T. Starks, 2016, SRI funds: Investor demand, exogenous shocks and ESG profiles, Working paper.

Bretscher, Lorenzo, Alex Hsu, and Andrea Tamoni, 2020, The supply channel of uncertainty shocks and the cross-section of returns: Evidence from the COVID-19 crisis, Working paper, LBS.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Ding, Wenzhi, Ross Levine, Chen Lin, and Wensi Xie, 2020, Corporate immunity to the COVID-19 pandemic, Working paper, University of Hong Kong.

Elton, Edwin J., Martin Jay Gruber, Sanjiv Das, and Matthew Hlavka, 1993, Efficiency with costly information: A reinterpretation of evidence from managed portfolios, *Review of Financial Studies* 6, 1–22.

Elton, Edwin J., Martin Jay Gruber, and Christopher R. Blake, 2001, A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases, *Journal of Finance* 56.

Evans, Richard, 2010, Mutual fund incubation, *Journal of Finance* 65, 1581–1611.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2010, Luck versus skill in the cross section of mutual fund returns, *Journal of Finance* 65, 1915–1947.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Covid Economics 38, 16 July 2020: 1-36

Fahlenbrach, Rudiger, Kevin Rageth, and René M. Stulz, 2020, How valuable is financial flexibility when revenue stops? Evidence from the Covid-19 crisis, Working paper, Swiss Finance Institute.

Ferson, Wayne E., and Rudi W. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, 425–461.

Gerding, Felix, Thorsten Martin, and Florian Nagler, 2020, The value of fiscal capacity in the face of a rare disaster, Working paper, Bocconi University.

Glode, Vincent, 2011, Why mutual funds "underperform", *Journal of Financial Economics* 99, 546–559.

Gormsen, Niels Joachim, and Ralph SJ Koijen, 2020, Coronavirus: Impact on stock prices and growth expectations, Working paper, University of Chicago.

Gruber, Martin J., 1996, Another puzzle: The growth in actively managed mutual funds, *Journal of Finance* 51, 783–810.

Haddad, Valentin, Alan Moreira, and Tyler Muir, 2020, When selling becomes viral: Disruptions in debt markets in the COVID-19 crisis and the Fed's response, Working paper, UCLA.

Hartzmark, Samuel M., and Abigail B. Sussman, 2019, Do investors value sustainability? A natural experiment examining ranking and fund flows, *Journal of Finance* 74, 2789–2837.

Jensen, Michael C., 1968, The performance of mutual funds in the period 1945–1964, *Journal of Finance* 23, 389–416.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2014, Time-varying fund manager skill, *Journal of Finance* 69, 1455–1484.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2016, A rational theory of mutual funds' attention allocation, *Econometrica* 84, 571–626.

Kargar, Mahyar, Benjamin T. Lester, David Lindsay, Shuo Liu, and Pierre-Oliver Weill, 2020, Corporate bond liquidity during the COVID-19 crisis, Working paper, University of Illinois.

Kosowski, Robert, 2011, Do mutual funds perform when it matters most to investors? US mutual fund performance and risk in recessions and expansions, *Quarterly Journal of Finance 1*, 607–664.

Lins, Karl V., Henri Servaes, and Ane Tamayo, 2017, Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis, *Journal of Finance* 72, 1785–1824.

Covid Economics 38, 16 July 2020: 1-36

CEPR PRESS

Malkiel, Burton G., 1995, Returns from investing in equity mutual funds 1971 to 1991, *Journal of Finance* 50, 549–572.

Moskowitz, Tobias J., 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses: Discussion, *Journal of Finance* 55, 1695–1703.

Nofsinger, John, and Abhishek Varma, 2014, Socially responsible funds and market crises, *Journal of Banking and Finance* 48, 180–193.

O'Hara, Maureen, and Xing Alex Zhou, 2020, Anatomy of a liquidity crisis: Corporate bonds in the COVID-19 crisis, Working paper, Cornell University.

Pagano, Marco, Christian Wagner, and Josef Zechner, 2020, Disaster resilience and asset prices, *Covid Economics* 21, 1–39.

Pástor, Ľuboš, and Robert F. Stambaugh, 2002, Mutual fund performance and seemingly unrelated assets, *Journal of Financial Economics* 63, 315–349.

Pástor, Ľuboš, and Robert F. Stambaugh, 2012, On the size of the active management industry, *Journal of Political Economy* 120, 740–781.

Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor, 2015, Scale and skill in active management, *Journal of Financial Economics* 116, 23–45.

Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor, 2017, Do funds make more when they trade more?, *Journal of Finance* 72, 1483–1528.

Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor, 2020, Sustainable investing in equilibrium, *Journal of Financial Economics*, forthcoming.

Ramelli, Stefano, and Alexander F. Wagner, 2020, Feverish stock price reactions to COVID-19, Working paper, University of Zurich.

Schrimpf, Andreas, Hyun Song Shin, and Vladyslav Sushko, 2020, Leverage and margin spirals in fixed income markets during the COVID-19 crisis, Working paper, BIS.

Sensoy, Berk, 2009, Performance evaluation and self-designated benchmark indexes in the mutual fund industry, *Journal of Financial Economics* 92, 25–39.

Wermers, Russ, 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses, *Journal of Finance* 55, 1655–1695.

Covid Economics 38, 16 July 2020: 1-36

COVID-19 in emerging markets: Firm-survey evidence[1]

Thorsten Beck,[2] Burton Flynn[3] and Mikael Homanen[4]

Date submitted: 10 July 2020; Date accepted: 10 July 2020

Using survey responses across nearly 500 listed firms in 10 emerging markets from early April, we find the vast majority of firms were negatively affected by COVID-19. Firms reacted by reducing investment rather than payroll. There is a surprising degree of support vis-à-vis employees, customers, other stakeholders and broader society. Although stock prices initially reacted to the impact of the crisis, delayed stock price reactions suggest evidence of inefficient markets. Furthermore, we find evidence that stakeholder-centric firms experienced lower stock price declines during the crisis drawdown.

Covid Economics 38, 16 July 2020: 37-67

1 We are grateful for excellent assistance by Rabi Uddin Nasir, Michael Genio Aguila, Ivan Nechunaev and Ara Reyes Tisalona, and comments from participants at the Cass PhD workshop and RCN Seminar series. The views expressed in this paper are the authors' only and do not necessarily represent those of the PRI.
2 City's Business School, University of London, and CEPR
3 Terra Nova Capital, Evli Group and City's Business School, University of London.
4 PRI and City's Business School, University of London.

Covid Economics 38, 16 July 2020: 37-67

1. Introduction

The economic fallout from the COVID-19 pandemic has been severe as reflected in recently released GDP numbers for the first quarter of 2020 and predictions for the rest of the year (IMF, 2020). Whole industries have been shut and firm-level evidence confirms the aggregate number in terms of lost jobs, output and investment. Most of the evidence so far, however, has been for advanced countries. This paper presents firm-level evidence on the impact of COVID-19 for 10 emerging markets, combining survey answers with financial statement information and stock price returns.

Many advanced countries have reacted to the pandemic with lockdown policies that brought a large part of the economy to a standstill, including industries that rely on direct customer contact but also manufacturing and construction in some countries. Among emerging markets, there has been variation in responses, with some countries going for strict lockdowns, such as Vietnam, while others keeping their economies relatively open, such as Turkey. These differences in timing and severity of lockdown policies have also implications for firms and their employees. More generally, higher informality, fewer jobs that can be done from home, and more limited state capacity make both public-health oriented containment and their enforcement less effective, while limited fiscal space and limited access to international financial markets make economic support policies more difficult to implement (Djankov and Panizza, 2020). The trade-off between public health and economic survival is therefore more biased towards the latter in many emerging markets (Alon et al., 2020; Hevia and Neumeyer, 2020).

We present survey evidence on 488 firms across 10 emerging markets in Europe, Africa, the Middle East, and Asia on the impact of COVID-19, firms' financial and investment reaction and their approach towards employees, other stakeholders and the broader society. Specifically, we explore cross-country and cross-sector variation in (i) the effect of the pandemic on firms, (ii) firms' reaction in investment and labour demand and (iii) firms' reactions in their relationships with employees, suppliers and other stakeholders. We also test whether the effect of the crisis on firms and their reaction is reflected in stock returns.

Covid Economics 38, 16 July 2020: 37-67

As the pandemic started in China and then spread to Europe and North America, many emerging markets experienced first an economic and then a public health shock. Specifically, disruptions in supply chains and drop in international demand were often the first negative signals before the public health crisis reached many of these countries. This gave both governments and firms (the focus of this study) time to react to the pandemic. However, firm-level analysis suggests that many firms in developing countries have limited liquidity to survive a longer economic lockdown that deprives them of their revenue stream.[1] Furthermore, integration into global supply chains makes many firms more vulnerable during global recessions (Accetturo and Giunta 2019), while procyclical bank behaviour might result in lower access to external funding when it is needed most.

Firms can react in different ways to such an unprecedented shock. On the one hand, investment plans can be cancelled rather quickly; such a reaction might reflect firms' assumption that there will not be a V-shaped quick recovery but rather a drawn-out crisis or loss of access to external funding. On the other hand, many emerging markets' labour markets are often considered more flexible than those of advanced countries, so that quick reductions of payrolls are another option to reduce costs. However, relationships between firms and their stakeholders, such as employees, customers, suppliers and society at large, are often based not only on contractual but also informal and personal links, so that firms can react quickly and reduce commitment but might also have to worry about long-term implications of undermining important relationships with business partners and stakeholders.

Our survey results for almost 500 listed firms across 10 emerging markets show at least three in four firms being negatively impacted by the pandemic. Surprisingly few firms, however, expect to breach their covenants or see a need for raising additional capital. About half of the firms have received or expect to receive government support. Firms have reacted primarily reducing investment spending and much less through layoffs. Meanwhile,

[1] Bosio et al. (2020a) find that in high- and middle-income countries, without revenues, the average firms have liquidity to survive between 12 and 38 weeks, while Bosio (2020b) find for firms in lower and lower-middle income countries, the average survival time across industries ranges from 6 to 28 week and 6 and 18 weeks once collapsed export demand is taken into account.

some firms cut back on executive compensation, and more firms expanded employee benefits than cut them. The large majority of firms have acted before their governments imposed measures, shown flexibility vis-à-vis customers other stakeholders, and provided donations to support society at large or have shifted business operations to fulfil pandemic needs. This shows a picture of firms focusing on short-term needs of stakeholders, protecting labour and long-term relationships, thus a picture not consistent with short-term focused shareholder maximisation value, but longer-term value maximisation for both share and other stakeholders.

A second major finding is that there is significant variation across countries, sectors and firms in both impact and reaction. While sector-variation matters most for variation in the impact of the crisis and government support, country-variation seems to matter most for variation in firm reaction to the crisis. Interestingly, variation in whether firms reacted to the pandemic before government-imposed actions varies more with firm-level characteristics than across countries or sectors.

Third, delayed share price reactions to the impact of the crisis and firms' reactions suggest inefficient markets in which astute investors were able to capture excess returns. While there was a reaction between late February and late April, firms that stated in early April that they were negatively affected by the pandemic saw a 6.1% lower cumulative return between late April and late June. Additionally, firms that claimed to be positively affected saw a 18.3% higher return in the following two months though this result is not statistically significant owing to the small sample size (only 13 such firms). Furthermore, firms that stated in April that they were reducing investment saw a lower return both between late April and late June. These delayed reactions to publicly available information are not consistent with the semi-strong form of efficient market hypothesis.

Fourth, we also document that more stakeholder-centric firms (e.g. those which did not reduce employee benefits, took measures to protect stakeholders or made donations to fight the pandemic) experienced lower stock price declines and did not significantly underperform during the late April to late June period, suggesting that the financial markets valued these stakeholder-centric corporations more than their counterparts during the crisis.

Covid Economics 38, 16 July 2020: 37-67

Our paper is part of a small but rapidly expanding literature exploring the impact of the pandemic. While several papers have used stock market data to explore the real sector implications of the COVID-19 recession (e.g., Alfaro et al., 2020 and Capelle-Blancard and Desrozier, 2020) and variation of stock market returns across firms (e.g. Ding et al., 2020, and Ramelli and Wagner, 2020), there are few papers using survey data capturing firms' reactions to the COVID-19 crisis and those few primarily focus on advanced rather than developing countries (e.g., Bennedsen et al., 2020 for Denmark and Li et al., 2020, for the US). While there are household surveys in developing countries (e.g., Gomez et al., 2020 and Rahman et al., 2020), our paper relies on (one of) the first firm-level surveys across developing countries.

Before proceeding, we would like to state a few caveats. First, the firms in our survey are medium sized to large, formal and listed. The survey does not provide any insights about the reaction of smaller and/or informal firms. It is notable, however, that together the operating income of the surveyed firms make up between 0.5% and 8% of the respective GDP.[2] More importantly, Alfaro et al. (2019) show that consistent with Gabaix (2011) large firms are systemically important in emerging markets, as idiosyncratic shocks to large firms significantly correlate with GDP growth. Second, survey responses might be biased, though their use has become increasingly popular in corporate finance, with evidence showing that self-reporting is significantly correlated with publicly and verifiable information. Finally, our survey provides insights at a specific point in time in early April, in what in most countries was the early stage of the pandemic and government responses to it. However, we did follow up with an additional survey in early June asking for an update on prior survey responses and found that little had changed.

The remainder of the paper is structured as follows. Section 2 discusses our data collection process. Section 3 presents some preliminary analysis, while section 4 provides more formal regression analysis. Section 5 shows the relationship between the impact of and reaction of

[2] As we do not have value-added data for individual firms, we estimate the share of the firms in the economy by dividing the sum of 2018 operating income of all sample firms in USD by 2018 GDP in USD of their respective countries. The ratio varies between 0.2% for Turkey to 7.6% for South Africa.

Covid Economics 38, 16 July 2020: 37-67

firms to the pandemic and buy-and-hold returns. Section 6 concludes.

2. Data

In early April 2020, we sent out a short 13-question survey to 630 firms across 10 emerging markets, including Bangladesh, Indonesia, Malaysia, Pakistan, Philippines, Saudi Arabia, South Africa, Thailand, Turkey and Vietnam.[3] All our firms are listed on stock exchanges and are active across all sectors of the economy.

Out of the 630 firms, 488 responded, 98% of them between April 2 and 24, thus a response rate of 78%.[4] In early June, we followed up with these 488 firms which answered the first set of questions with additional questions and inquiring whether their responses to the first survey questions had changed. 415 firms of the 488 firms responded, thus a response rate of 85%. Appendix A1 lists the survey questions and Appendix Table A2 lists the exact split of responses.

We augment the survey data with financial statement and market information on all firms, specifically, sales growth over the past year to capture the firm's growth path, Tobin's q (market value / book value of equity) to proxy for growth opportunities, log of total assets in million USD to capture the size of the firm, cash divided by total assets to gauge the liquidity position, capital expenditures divided by total assets as investment indicator and the leverage ratio as measured by debt-to-assets ratio. We report descriptive statistics in Table 1. In regression analysis, we winsorise sales growth, cash divided by total assets and capital expenditures divided by total assets at the 1st and 99th percentiles to avoid the impact of outliers.

[3] Between June 2019 and March 2020, the portfolio managers of a Finnish-based emerging markets fund requested one-on-one private meetings with the CEOs of 1,476 firms, comprised of all listed companies with greater than $250,000 of average daily trading volume across these ten countries. Various executives (mainly CEOs, CFOs, and IR personnel) at 630 of the firms accepted the meeting requests. In April 2020, the fund managers emailed the executives they met at all 630 firms with 13 questions regarding how the coronavirus pandemic is affecting their firm and how their firm is responding.

[4] 91% of responses were received by Apr 17; the last response was received on 16 June.

Covid Economics 38, 16 July 2020: 37-67

Table 1: Summary Statistics

Panel A: April Survey - operational and governance variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	N	mean	sd	min	max	p25	p50	p75
Have your operations been affected by the pandemic? (negatively)	439	0.697	0.460	0	1	0	1	1
Have you or might you breach covenants?	439	0.0524	0.223	0	1	0	0	0
Would you need to raise capital to weather the pandemic?	439	0.194	0.396	0	1	0	0	0
Do you expect to receive government support?	439	0.358	0.480	0	1	0	0	1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	N	mean	sd	min	max	p25	p50	p75
Have you reduced investment spending?	439	0.579	0.494	0	1	0	1	1
Have you made any layoffs?	439	0.0820	0.275	0	1	0	0	0
Made any changes to employee benefits? (negatively)	416	0.0721	0.259	0	1	0	0	0
Have you made changes to the dividend or share buyback plan?	439	0.187	0.390	0	1	0	0	0
Have you made changes to executive compensation plans?	439	0.175	0.381	0	1	0	0	0

Panel B: April survey – stakeholder variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	N	mean	sd	min	max	p25	p50	p75
Have you taken any measures to protect your stakeholders?	439	0.911	0.285	0	1	1	1	1
Continued to pay employees or services providers for disrupted services?	439	0.465	0.499	0	1	0	0	1
Provided financial flexibility to any customers or business partners?	439	0.392	0.489	0	1	0	0	1
Made any donations to help fight the pandemic?	439	0.781	0.414	0	1	1	1	1

Panel C: June survey – additional variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	N	mean	sd	min	max	p25	p50	p75
Did you take action before the government?	267	0.700	0.459	0	1	0	1	1
Did your revenue decline in the last two months?	208	0.856	0.352	0	1	1	1	1
How long before you expect to recover to the pre-pandemic revenue?	137	2.620	0.979	1	5	2	2	3

Panel D: Firm controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	N	mean	sd	min	max	p25	p50	p75
Sales Growth	439	1.351	3.495	0.0181	57.20	0.326	0.692	1.345
Leverage	439	25.73	18.76	0	93.02	8.845	25.53	38.69
Total Assets	439	5,701	14,923	4.536	87,503	285.7	875.0	3,539
Cash / Total Assets	439	8.308	8.908	0.0553	61.60	2.705	5.707	10.71
Capex / Total Assets	439	4.488	4.866	0.0029	25.84	0.647	3.150	6.706
Sales Growth	439	0.102	0.335	-0.661	3.777	-0.027	0.0673	0.170

Notes: The variable "How long before you expect to recover to the pre-pandemic revenue?" takes a value between 1 and 5. The variable is equal to one if companies expect recovery in the first half of 2020, two for second half of 2020, three for first half of 2021 and so forth. In total 4.5% of the firms responded with 1, 56% with 2, 25% with 3, 5% with 4 and 8% with 5, out of a total of 156 responses that provided full details of recovery expectations (not adjusting for whether firm controls are available). The variable "Made any changes to employee benefits? (negatively)" has less than 439 observations, because whether benefits were negatively or positively changed was extrapolated from qualitative answers provided by the corporate responses. In few cases, it was not possible to evaluate whether they were positive, negative or affected at all. Firm controls were collected from the S&P Capital IQ Platform. Total Assets, Cash / Total Assets, Capex / Total Assets and Sales Growth are all winsorized at the 1st and 99th percentile.

Covid Economics 38, 16 July 2020: 37-67

Further, we use share prices of these firms to compute buy-and-hold returns between February 24 and April 24, and between April 24 and June 24. The sample using stock prices is smaller than our overall sample, as we do not have stock price data for Bangladeshi firms – the Dhaka Stock Exchange was closed for two months beginning on 26 March. All financial statement and stock price data were retrieved from S&P Capital IQ.

We formally test whether there is a selection bias in survey responses, i.e., whether the survey respondents differ significantly from the non-respondents. Specifically, in Appendix Table A3 we test for differences across the firm-level variables discussed above between these two samples and do not find any significant difference across firm characteristics except that non-respondents had higher sales growth in the previous years. This assures us that the survey responses provide an adequate picture of firms in this segment. We also test for the differences across the firm-level variables between firms that responded to both April and June surveys and those that responded only to the April survey and find again mostly insignificant differences, except that pre-crisis capital expenditures are higher for firms responding only to the survey in April.

The survey questions are divided into three blocks. In the first block, we asked firms whether their operations were affected by the pandemic, they might breach covenants, they need to raise capital and they expect to receive government support. A second block considered adjustment in firms' operation and plans; whether they have reduced investment, laid off staff, made changes to employee benefits, made changes to executive compensation and made changes to the dividends or share buyback plans. Finally, we asked questions on how firms dealt with business partners and society more generally; whether they voluntarily have taken any measures to protect their employees and other stakeholders, continued to pay employees or service providers for disrupted services, provided financial flexibility to any customers or business partners and made any donations to help fight the pandemic or shifted business operations to fulfil pandemic needs. In addition to yes/no answers, most firms gave more details on some of these questions, which we exploit below.

In June, we asked firms whether there had been any updates on their responses to the April survey, plus three additional questions. First, by how much approximately their sales had

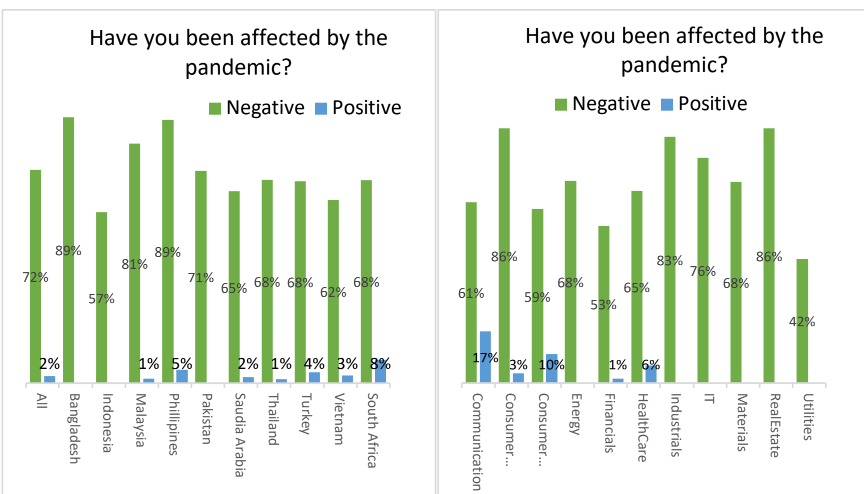
Covid Economics 38, 16 July 2020: 37-67

fallen over the past two months (April and May). Second, by when they expected a recovery to the pre-crisis level of revenues. Third, whether they adjusted their operations to the circumstances of COVID-19 before government measures came into place or following the issue of such measures.

3. COVID-19 in emerging markets

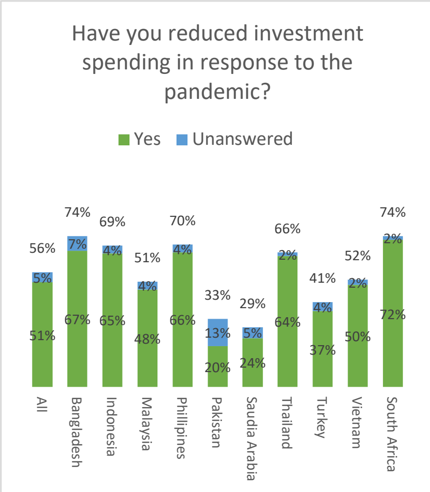
The survey responses from April and June show clearly an overwhelmingly negative impact of the pandemic on firms, a quick reaction of firms to the pandemic and quite some flexibility in terms of their relationships with stakeholders. However, we also find significant variation across countries and across sectors. Table 1 lists the descriptive statistics of the different variables. Figure 1 shows the variation for selected survey variables across countries and sectors, and Appendix Tables A4 and A5 show the variation for all the survey variables across countries and sectors.

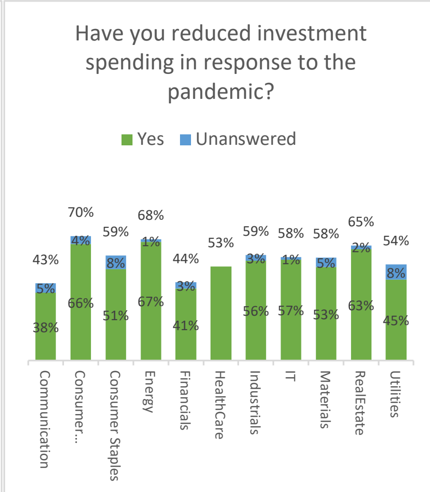
Figure 1: COVID-19 across countries and sectors

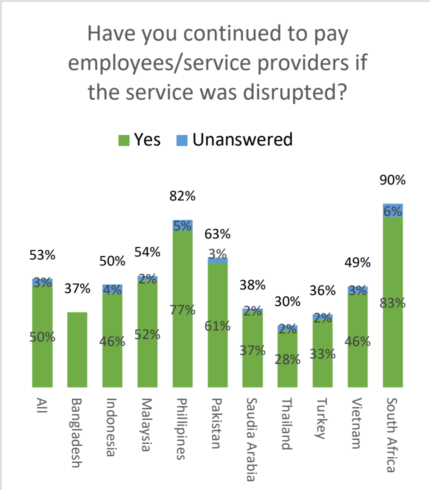


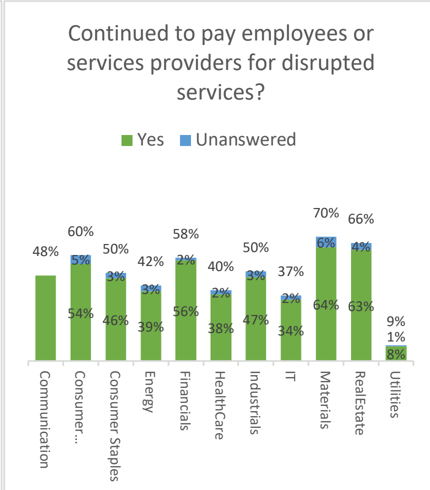
Covid Economics 38, 16 July 2020: 37-67

Covid Economics 38, 16 July 2020: 37-67









The vast majority of surveyed firms have been negatively affected by the COVID-19 crisis. Specifically, in April 2020, 69% of firms report a negative impact and less than 3% a positive impact, with 15% reporting no impact and 5% not answering the questions. The remaining 11% of firms that answered the question reported an impact but did not specify the type of impact. However, the effect varies a lot across countries and sectors; ranging from a negative impact to 87% of firms in the Philippines to 55% in Indonesia, and between 85% in real estate and 43% in utilities (Figure 1). The sectoral variation is not surprising, as firms in essential sectors (e.g., utilities) were hit less and firms in less essential sectors (e.g., consumer discretionary) and sectors relying heavily on physical customer interaction (e.g., real estate) were hit most.

Covid Economics 38, 16 July 2020: 37-67

In early June, we followed up with a question of how much revenues had fallen in April and May. While most firms did not give a specific number, 50% of firms indicated a fall and 28% declined to answer, thus confirming their response from April that they were negatively impacted by the pandemic. Meanwhile, 13% reported minimal impact and 9% said revenues actually increased.

In June, we also asked firms by when they expected a recovery to pre-crisis revenues. There is wide variation in responses to this question, ranging from Q2 2020 to sometime in 2023: 11% of firms did not see any significant deterioration, consistent with responses above. 24% of firms expect a recovery in 2020, 12% expect a recovery in 2021 and 3% expect a recover in 2022 or later, while 49% did not answer the question. The fact that half of firms did not answer the question points to the very high uncertainty that firms across the globe, including in emerging markets, currently operate under.

Few firms responded affirmatively to the question on whether they have breached or expect to breach any covenants. While only six firms (1.2%) responded affirmatively in April, 93% responded negatively. If we consider the share of firms responding affirmatively or not responding as the upper bound for the likelihood that firms breach their covenants, we find that up to 6% might be breaching their covenants. Six additional firms reported breach of covenants in the June survey round (a total of 2.4%).

Similarly, few firms expect to need to raise capital in the near future because of the pandemic. Specifically, 12% (15%) responded affirmatively in April (June), while 80% responded negatively. However, there is quite some variation across countries, with 45% of firms in Bangladesh stating that they need to raise capital, but only 1% in South Africa, as well as across sectors, from 28% in energy to 5% in finance.

On the other hand, more than one third of firms (35%) expected to receive government support, ranging from 49% in Saudi Arabia to 19% in South Africa, and from 54% in consumer discretionary goods to none in utilities. This share only increased to 37% in June at the time of our second survey. There is also quite some variation in the form of government support firms have received or expect to receive, with tax or social security support relief and wage and employee benefits being listed most often, followed by debt relief. Support with rental or utility payments or other fees and relaxation of regulations are named less.

In terms of firm reactions to the pandemic, half of surveyed firms (50%) have reduced their investment spending, ranging from 65% in Thailand to 16% in Pakistan, 67% in energy and real estate to 35% in communications and financials when asked in April (Figure 1). This only increased to 52% during our second survey in June. On the other hand, very few firms (3%) have laid off employees – even taking into account non-responses only gets us to 9%, suggesting that the cuts fell primarily on capital rather than wage spending. It is more, 12% of firms have expanded employee benefits and only 7% have reduced them. Meanwhile, 17% of firms have reduced executive compensation.

In addition to cutting expenditures, another way to strengthen firms' financial position is to reduce/suspend dividends or stop share buy-backs. Across the ten countries, 13% of all firms planned to do so, ranging from 17% in South Africa to none in the Pakistan, from 20% in energy and industrials to 2% in consumer staple goods. The proportion of firms reducing or suspending dividends increased to 18% in June.

Most of the surveyed firms have taken actions to support their stakeholders and the broader society. 91% of the firms have taken measures to protect employees and stakeholders beyond government-imposed regulations and 47% have continued to pay employees and vendors even if service was disrupted. 100% of firms in Bangladesh and Pakistan reported taking voluntary measures to protect employees and other stakeholders while only 79% in Turkey did so, and 100% of firms in the IT and utilities sectors compared to only 85% of consumer staples companies. 71% of firms in South Africa continued paying employees for disrupted services compared to only 28% of firms in Thailand, and 60% of firms in the real estate sector voluntarily paid employees while only 21% of utilities firms did the same. 41% of surveyed firms have provided financial flexibility to customers or business partners, ranging from 70% of firms in the Philippines to 15% in Bangladesh, and from 62% of financials to 18% of health care.

Nearly four in five surveyed firms (77%) have made donations to support the broader society in their community and country or shifted business operations to fulfil pandemic

Covid Economics 38, 16 July 2020: 37-67

needs. Donations have taken on the form of cash and PPE and other medical or hygienic products, as well as food. To provide some examples: 26 firms shifted business operations to manufacture personal protective equipment or sanitizing/cleaning items, while 51 firms provided various services free of charge, such as servicing of ambulances by an auto manufacturer, development of a donations platforms by tech companies, provision of hotel rooms to medical workers to stay during curfew period, quarantine facilities by several hotel groups as well as a software company campus which hosts over 2,000 people, care for COVID-19 patients by a hospital group seeking only to recover costs from the government, free storage of coronavirus-related supplies for NGOs fighting the virus by a self-storage company, potential construction/conversion for temporary hospitals by a construction firm, provision of bus fleet by a taxi company for drop off and pick up services for doctors and nurses, assistance to government of citizen repatriation by an airline and free food packs for medical workers, police force, military personnel and logistics drivers by a gas station operator.

In the June follow-up survey, we also asked firms whether they adjusted their operations to the circumstances of COVID-19 before government guidelines came into place or following the issue of such guidelines. The large majority of firms (64%) reported that they adjusted operations before government guidelines came into effect while 27% reported that they waited until government measures were implemented before reacting. This may be due to the fact that companies in emerging markets generally do not rely on their government to lead; as one broker in Mexico put it, “companies, business men and women, do not count at all in our government, so we need to organize ourselves for this new era.”[5] However, there is again variation across countries and sectors. For example, 72% of firms in Thailand took proactive action whereas only 33% of firms in Turkey adjusted operations before government measures, and 83% of firms in the utilities sector made adjustments prior to the government’s guidelines while only 45% of energy firms did so. It is interesting to note (and we will return to this below) that country-variation in such action does not vary with how quickly countries imposed lockdowns.

[5] Private message to one of the paper’s authors on May 5, 2020.

Covid Economics 38, 16 July 2020: 37-67

4. Explaining variation in COVID-19 impact and responses

While the previous section explored simple correlations, we now offer a more formal regression analysis to explore which (i) firm characteristics, (ii) country factors and (iii) sector factors can explain variation in the impact of COVID-19 and firm-reaction to the pandemic. While our dependent variables are binary, we use linear regression models to avoid the incidental parameter problem given that we saturate our models with country- and sector fixed effects. Specifically, we run the following regression:

$$Y_{ijk} = \alpha \Sigma I_k + \beta \Sigma C_j + \gamma X_i + \varepsilon_{ijk} \qquad (1)$$

Where the subscripts i, j, and k denote firm, country and sector respectively. As discussed above, we winsorise several of the firm variables at 1^{st} and 99^{th} percentile. The omitted country is Bangladesh and the omitted sector is communication, so that the other country (sector) fixed effects denote differences in the respective country (sector) to Bangladesh (communication). We report heteroskedasticity consistent error terms.

Before presenting regression results, we assess whether country, sector or firm characteristics are most important in explaining variation in the different survey responses. Specifically, Appendix Table A6 shows regressions with (i) firm-level variables, (ii) country dummies, (iii) sectoral dummies and (iv) all three groups, for three of our survey variables. Overall, country-variation seems to matter most for variation in adjustments in investment spending and donations, while variation in government support varies most across sectors. For other survey questions (not reported), firm-level variables seem to matter most for breaching (or not) covenants, whether firms took actions before the government issued guidance, and expected time to recovery, sector-variation matters most for the negative impact of the crisis and changes to executive compensation, while for other variables it is either country variation that matters or there is a mix of firm-, sector- and country-factors that matter most.

The results in Table 2 show that larger firms are less likely to raise capital to weather the pandemic, more likely to reduce investment spending, provide financial flexibility to customers and make donations or shift operations to fill pandemic needs. Larger firms also

Table 2: COVID-19 Corporate Impact and Responses

Panel A

VARIABLES	(1) Have you been affected by the pandemic? (negatively)	(2) Have you or might you breach covenants?	(3) Would you need to raise capital to weather the pandemic?	(4) Do you expect to receive government support?
Tobin's q	-0.008*	0.010	0.006	0.001
	(0.005)	(0.006)	(0.005)	(0.006)
Log of Total Assets	0.011	0.008	-0.029**	0.003
	(0.015)	(0.006)	(0.013)	(0.014)
Cash / Total Assets	-0.001	-0.000	-0.007***	-0.002
	(0.003)	(0.001)	(0.002)	(0.003)
Capex / Total Assets	0.005	-0.003	-0.004	0.005
	(0.005)	(0.002)	(0.005)	(0.006)
Leverage	0.000	0.002***	0.000	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
Sales Growth	0.099*	0.001	0.088	-0.117**
	(0.055)	(0.018)	(0.058)	(0.053)
Have you been affected by the pandemic? (negatively)		-0.051*	0.022	0.057
		(0.026)	(0.041)	(0.051)
Observations	439	439	439	439
R-squared	0.143	0.102	0.160	0.118

Panel B

VARIABLES	(1) Have you reduced investment spending?	(2) Have you made changes to the dividend or share buyback plan?	(3) Have you made changes to exec com plans?	(4) Have you made any layoffs?	(5) Made any negative changes to employee benefits?
Tobin's q	0.011***	0.017***	0.002	0.008	-0.005
	(0.003)	(0.004)	(0.007)	(0.006)	(0.004)
Log of Total Assets	0.046***	0.007	-0.001	-0.003	-0.004
	(0.014)	(0.012)	(0.011)	(0.008)	(0.007)
Cash / Total Assets	-0.002	-0.000	0.000	-0.000	-0.000
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)
Capex / Total Assets	-0.003	-0.003	-0.001	0.005	-0.003
	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)
Leverage	0.002	0.001	0.002**	0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Sales Growth	0.047	-0.014	-0.018	-0.008	0.012
	(0.056)	(0.054)	(0.043)	(0.031)	(0.030)
Have you been affected by the pandemic? (negatively)	0.143***	0.031	0.076**	-0.021	0.061***
	(0.054)	(0.040)	(0.036)	(0.031)	(0.023)
Observations	439	439	439	439	416
R-squared	0.18	0.121	0.122	0.090	0.101

Covid Economics 38, 16 July 2020: 37-67

Covid Economics 38, 16 July 2020: 37-67

Panel C				
	(1)	(2)	(3)	(4)
VARIABLES	Have you taken any measures to protect your stakeholders?	Continued to pay employees or services providers for disrupted services?	Provided financial flexibility to any customers or business partners?	Made any donations to help fight the pandemic?
Tobin's q	0.002	-0.007**	0.004	-0.013***
	(0.002)	(0.003)	(0.005)	(0.003)
Log of Total Assets	0.011	0.014	0.046***	0.050***
	(0.010)	(0.015)	(0.015)	(0.013)
Cash / Total Assets	0.003*	0.000	-0.005*	0.001
	(0.002)	(0.003)	(0.003)	(0.003)
Capex / Total Assets	0.002	-0.001	-0.005	0.001
	(0.003)	(0.006)	(0.005)	(0.005)
Leverage	-0.000	-0.003***	0.001	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
Sales Growth	0.002	0.128**	0.030	-0.059
	(0.033)	(0.056)	(0.067)	(0.063)
Have you been affected by the pandemic? (negatively)	0.064*	0.184***	0.083	-0.033
	(0.033)	(0.051)	(0.050)	(0.042)
Observations	439	439	439	439
R-squared	0.083	0.215	0.206	0.176

Panel D			
	(1)	(2)	(3)
VARIABLES	Did you take action before the government?	Did your revenue decline in the last two months?	How long before you expect to recover to the pre-pandemic revenue?
Tobin's q	-0.009	0.003	-0.023*
	(0.025)	(0.003)	(0.013)
Log of Total Assets	0.026	0.005	0.196***
	(0.019)	(0.018)	(0.056)
Cash / Total Assets	-0.003	0.001	0.002
	(0.003)	(0.004)	(0.013)
Capex / Total Assets	0.014**	0.003	0.013
	(0.006)	(0.004)	(0.014)
Leverage	0.001	-0.000	0.006
	(0.002)	(0.002)	(0.005)
Sales Growth	0.020	-0.022	-0.082
	(0.079)	(0.055)	(0.442)
Have you been affected negatively by the pandemic?	-0.092	0.142*	0.140
	(0.063)	(0.081)	(0.211)
Observations	267	208	137
R-squared	0.107	0.130	0.283
Country FE	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes

*Notes: Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1*

Covid Economics 38, 16 July 2020: 37-67

expect a longer time until recovery. Firms with a higher Tobin's q and thus higher market/book valuation are more likely to reduce investment and more likely to need to raise capital to weather the pandemic; further, they are less likely to voluntarily continue to pay employees or services providers for disrupted services and make donations or shift operations to fill pandemic needs, but expect a quicker recovery. Firms with higher sales growth over the past 12 months are less likely to receive government support and more likely to continue to pay employees or services providers for disrupted services. Firms with higher cash reserves (relative to their assets) are less likely to need to raise capital to weather the pandemic, while more leveraged firms are more likely to breach covenants on their loans and reduce executive compensation but less likely to continue to voluntarily pay employees or services providers for disrupted services. Firms with higher investment spending before the pandemic are more likely to have adjusted their operation before government guidance was issued.

In summary, financially more vulnerable firms see themselves more affected and reacted more strongly to the pandemic, while larger firms acted more pro-actively to the threat of the pandemic. We also find that firms more negatively affected by the pandemic are more likely to have reduced investment spending, executive compensation and benefits, but are also more likely to have continued to pay employees or services providers for disrupted service; however, this might not indicate any causality but rather common shocks.

We run several additional analyses using the same regression set-up. First, for a subset of firms we have information on the share of domestic and international revenues. While we expect a higher negative impact and a faster reaction of more export-oriented firms, we do not find any evidence for that, with the exception that firms with a higher share of domestic revenue are more likely to state that they need to raise capital to weather pandemic and to make negative changes to benefits, suggesting that the international trade channel was not the primary channel through which the pandemic negatively impacted growth in emerging markets; importantly, however, we regard these findings as tentative, given that we have this variable available only for half of our sample.[6] Second, we explore whether differences in the pandemic's impact and reaction of firms can be explained by variation in countries'

[6] Results are available on request.

decision whether or not and when to impose lockdowns. Specifically, we consider the difference between the date the firm returned the April survey and the date when the government in the respective country imposed a stringent lockdown, which included school closing, closure of non-essential businesses and mobility constraints. Only five of the ten countries imposed such a strict lock-down, though at different dates, starting with the Philippines on 17 March, followed by Malaysia on 22 March, South Africa on 26 March, Vietnam on 1 April and Saudi Arabia on 6 April. The other five countries imposed less stringent lockdowns between 22 March in Turkey and 10 April in Indonesia.[7] We code the length of time in lockdown as the date we received the survey response minus the date the country entered strict lockdown if the country had imposed a strict lockdown by the time we received the survey, and assign a zero to firms in countries which had not yet imposed a strict lockdown as of the date we received their response. As most of the variation in these regressions is explained at the country-level (with within-country variation determined by the day of survey return), we drop country fixed effects from these regressions.

The results in Table 3 show that firms in countries that implemented a lock-down earlier are more likely to be negatively affected by the pandemic, less likely to make negative changes to benefits, and more likely to continue to voluntarily pay employees or services providers for disrupted services and to provide financial flexibility to customers or business partners.

5. The delayed effect of firm responses on stock prices

We next gauge the effect of firms' reactions to the crisis on firms' stock prices. Specifically, we regress the buy-and-hold stock return between 24 February and 24 April and between 24 April and 24 June on a set of country and sectoral fixed effects, the firm-characteristics discussed above and the different survey questions from the early April round.[8] Doing so,

[7] The data on containment policies are from Olivier Lejeune's web-site: https://github.com/OlivierLej/Coronavirus_CounterMeasures. We denote with strict lock-down any containment policy that has a five in the database ("All-day lockdown. Government requires citizens to shelter in place all day long. Citizens are allowed to come out to buy essential items.")

[8] We choose 24 April, as by this date, 98% of all first-round surveys had been received. 24 February is also the first trading date after a first local lockdown in Italy, thus can be seen as the date when the pandemic clearly had arrived in Europe. We also performed the regressions on a range of other dates to ensure the results were

Table 3: COVID-19 Corporate Impact and Responses and Lockdown Dates

VARIABLES	(1) Have you reduced investment spending?	(2) Made any changes to employee benefits? (negative)	(3) Continued to pay employees or services providers for disrupted services?	(4) Provided financial flexibility to any customers or business partners?	(5) Have you been affected by the pandemic? (negatively)
Difference of days between receiving the survey and the lockdown being implement	0.001	-0.003**	0.009***	0.007***	0.006***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Have you been affected by the pandemic? (negatively)	0.133**	0.053**	0.186***	0.076	
	(0.054)	(0.022)	(0.051)	(0.050)	
Tobin's q	0.014***	-0.007**	-0.008**	0.003	-0.008
	(0.003)	(0.003)	(0.003)	(0.005)	(0.006)
Log of Total Assets	0.057***	0.002	0.022	0.051***	0.002
	(0.014)	(0.007)	(0.014)	(0.014)	(0.013)
Cash / Total Assets	-0.003	-0.001	-0.001	-0.006**	-0.000
	(0.003)	(0.001)	(0.003)	(0.002)	(0.003)
Capex / Total Assets	0.000	-0.001	0.001	-0.003	0.005
	(0.005)	(0.003)	(0.006)	(0.005)	(0.005)
Leverage	0.002	0.000	-0.004***	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Sales Growth	0.113**	0.006	0.129**	0.057	0.085*
	(0.056)	(0.029)	(0.058)	(0.068)	(0.048)
Observations	439	416	439	439	439
R-squared	0.117	0.059	0.152	0.164	0.114
Country FE	No	No	No	No	No
Sector FE	Yes	Yes	Yes	Yes	Yes

*Notes: Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1*

we would like to assess how investors reacted to the differential impact of the crisis on different firms and to their varied reactions. We also include each firm's market beta, which we estimate by regressing one year of weekly excess returns (from December 31, 2018 through December 31, 2019) on a constant and the weekly market factor from the respective country.[9]

not sensitive to the specific date range and found consistent results. We do not include the questions from the June survey round.

[9] We do not report the coefficient estimates for firm characteristics and market beta to save space, but results are available on request.

Covid Economics 38, 16 July 2020: 37-67

Gauging stock market reaction between late February and late April (the initial outbreak period of the global pandemic) allows us to test whether markets incorporated early indications of the crisis impact and firm reactions, while gauging stock market reaction between late April and late June (two months after 98% of survey responses had arrived) allows to test the timely market reaction to public information.

The results in Table 4 (buy-and hold return between 24 February and 24 April) show that markets priced in the need for firms to raise capital, reduction in investment, changes to the dividend or share buyback, and layoffs. Specifically, firms that stated in early April that they were reducing investment saw a 3.7% lower return, firms that have made changes to the dividend or share buyback plan saw a 7.8% lower return, firms that needed to raise capital to weather the pandemic saw a 5.8% lower return and firms that made layoffs saw a 11.1% lower return. However, we do not find a significant market reaction to a negative or positive impact of the pandemic.

Given that this information was available in early April, the efficient market hypothesis would predict no significant market reaction to the impact of the pandemic or firms' reactions. However, the results in Table 5 (buy-and hold return between 24 April and 24 June) show that firms that stated in early April that they were negatively affected by the pandemic saw a 6.1% lower return than firms that were not or were even positively affected by the pandemic, suggesting (together with the statistically insignificant results in Table 4) that stock markets were slow to react to the negative impact of the pandemic on firms. Firms which reported a positive impact averaged a return 18.3% higher the following two months though this coefficient estimate is not significant, partially due to the rather small sample of positively affected firms (only 13 such firms). Firms that stated in early April that they were reducing investment saw a 6.1% lower return than firms that did not reduce investment. Similarly, we find that firms that voluntarily provide financial flexibility to any customers have 4.4% lower cumulative returns. This suggests that the market had a delayed

Covid Economics 38, 16 July 2020: 37-67

Covid Economics 38, 16 July 2020: 37-67

Table 4: Stock Returns and COVID-19 Corporate Responses, February - April

Panel A

	(1)	(2)	(3)	(4)	(5)
VARIABLES			Cumulative Returns		
Have you been affected by the pandemic? (positively)	0.102				
	(0.062)				
Have you been affected by the pandemic? (negatively)		-0.015			
		(0.020)			
Have you or might you breach covenants?			-0.071		
			(0.050)		
Would you need to raise capital to weather the pandemic?				-0.058**	
				(0.023)	
Do you expect to receive government support?					0.000
					(0.019)
Observations	408	408	408	408	408
R-squared	0.144	0.142	0.127	0.127	0.133

Panel B

	(1)	(2)	(3)	(4)	(5)
Have you reduced investment spending?	-0.037**				
	(0.018)				
Have you made changes to the dividend or share buyback plan?		-0.078***			
		(0.023)			
Have you made changes to executive compensation plans?			-0.035		
			(0.024)		
Have you made any layoffs?				-0.111***	
				(0.028)	
Made any changes to employee benefits? (negative)					-0.005
					(0.033)
Observations	408	408	408	408	386
R-squared	0.144	0.127	0.130	0.127	0.142

Panel C

	(1)	(2)	(3)	(4)
Have you taken any measures to protect your stakeholders?	0.077***			
	(0.027)			
Continued to pay employees for disrupted services		-0.010		
		(0.019)		
Provided financial flexibility to customers?			-0.006	
			(0.018)	
Made any donations to help fight the pandemic?				0.068***
				(0.021)
Observations	408	408	408	408
R-squared	0.127	0.127	0.135	0.128
Country FE		Yes	Yes	Yes
Sector FE		Yes	Yes	Yes
Firm Controls		Yes	Yes	Yes

*Notes: The dependent variables are cumulative returns measured February 24th – April 24th. All regressions also include market beta as an additional firm control. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1*

CEPR PRESS

Covid Economics 38, 16 July 2020: 37-67

Table 5: Stock Returns and COVID-19 Corporate Responses, April - June

Panel A					
	(1)	(2)	(3)	(4)	(5)
VARIABLES	Cumulative Returns				
Have you been affected by the pandemic? (positively)	0.183				
	(0.125)				
Have you been affected by the pandemic? (negatively)		-0.061**			
		(0.026)			
Have you or might you breach covenants?			0.008		
			(0.047)		
Would you need to raise capital to weather the pandemic?				-0.016	
				(0.027)	
Do you expect to receive government support?					-0.038*
					(0.022)
Observations	408	408	408	408	408
R-squared	0.144	0.142	0.127	0.127	0.133

Panel B					
	(1)	(2)	(3)	(4)	(5)
Have you reduced investment spending?	-0.061***				
	(0.023)				
Have you made changes to the dividend or share buyback plan?		-0.008			
		(0.026)			
Have you made changes to executive compensation plans?			-0.031		
			(0.026)		
Have you made any layoffs?				0.004	
				(0.036)	
Made any negative changes to employee benefits?					-0.072**
					(0.032)
Observations	408	408	408	408	386
R-squared	0.144	0.127	0.130	0.127	0.142

Panel C				
	(1)	(2)	(3)	(4)
Have you taken any measures to protect your stakeholders?	0.017			
	(0.036)			
Continued to pay employees or services providers for disrupted services?		-0.010		
		(0.022)		
Provided financial flexibility to any customers or business partners?			-0.044**	
			(0.020)	
Made any donations to help fight the pandemic?				-0.017
				(0.025)
Observations	408	408	408	408
R-squared	0.127	0.127	0.135	0.128

*Notes: The dependent variables are cumulative returns measured April 24th – June 24th. All regressions also include market beta as an additional firm control. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1* reaction to the available information, not consistent with the efficient market hypothesis.

The results in Tables 4 and 5 also show a positive market reaction to a stakeholder-centric approach by firms. Between late February and late April, firms that provided donations and took voluntary measures to protect stakeholders had, on average, 6.8% and 7.7% higher returns, respectively. Furthermore, firms that reduced employee benefits in April subsequently saw 7.2% lower stock returns than other firms. We also do not find any negative and significant reaction between late April and late June to firms providing donation or voluntarily taking measures to protect stakeholders. This suggests that the financial markets valued these stakeholder-centric corporations more than their counterparts during the crisis and that the market has indeed priced in attempts of firms to maintain long-term relationships of firms with employees, business partners and other stakeholders.

6. Conclusions

This paper presented firm-level evidence on the effect of the COVID-19 pandemic in 10 emerging markets as well as firms' reactions. Survey evidence shows an overall negative effect, but also a quick reaction by firms. Adjustments are mostly on the investment rather than labour side, with significant expectations of government support. We find quite some variation in impact and firm reaction across not only countries and sectors, but also firms with different characteristics.

Our results point to quick reaction of firms in emerging markets (often even before governments), their focus on accommodating their labour force and maintaining long-term relationships with stakeholders. We also find evidence of markets being slow to incorporate publicly available information and valuing stakeholder centric activities. Our findings provide a snapshot of the current situation and another stock-taking later this year might provide different findings, but they show the picture of corporations trying to make the best of a very bad situation.

Covid Economics 38, 16 July 2020: 37-67

References

Accetturo, Antonio and Anna Giunta (2018): Value Chains and the Great Recession: Evidence from Italian and German firms, *International Economics* 153, 55-68.

Alfaro, Laura , Gonzalo Asis, Anusha Chari and Ugo Panizza (2019): Corporate debt, firm size and financial fragility in emerging markets, *Journal of International Economics* 118, 1-19.

Alfaro, Laura, Anusha Chari, Andrew Greenland and Peter K. Schott (2020): Aggregate and firm-level stock returns during pandemics, in real time, *Covid Economics, Vetted and Real-Time Papers 4,* April.

Alon, Titan, Minki Kim, David Lagakos , Mitchell VanVuren (2020): Lockdowns in developing countries should focus on shielding the elderly, VoxEU 25 June

Bennedsen, Morten, Birthe Larsen, Ian Schmutte and Daniela Scur (2020): Preserving job matches during the COVID-19 pandemic: Firm-level evidence on the role of government aid, *Covid Economics, Vetted and Real-Time Papers 27,* June

Bosio, Erica, Simeon Djankov, Filip Jolevski and Rita Ramalho (2020a), Survival of Firms during Economic Crisis", World Bank Policy Research Working Paper No. 9239, Washington, DC.

Bosio, Erica, Filip Jolevski, Joseph Lemoine and Rita Ramalho (2020b) Survival of firms in developing economies during economic crisis, in: Djankov, Simeon and Ugo Panizza (Eds): Covid-19 in Developing Economies. CEPR Press, London, UK

Capelle-Blancard, Gunther and Adrien Desroziers (2020): The stock market is not the economy? Insights from the Covid-19 crisis, *Covid Economics, Vetted and Real-Time Papers 28,* June

Ding, Wenzhi, Ross Levine, Chen Lin, and Wensi Xie (2020): Corporate Immunity to the COVID-19 Pandemic, NBER Working Paper 27055.

Djankov, Simeon and Ugo Panizza, Eds. (2020): Covid-19 in Developing Economies. CEPR Press, London, UK.

Gabaix, Xavier, 2011. The granular origins of aggregate fluctuations. *Econometrica* 79, 733–772.

Gómez, Margarita, Andriy Ivchenko, Elena Reutskaja and Pablo Soto-Mota (2020): Behaviours, perceptions and mental wellbeing in high-income and low/middle-income countries at the beginning of COVID-19 pandemic, in: Djankov, Simeon and Ugo Panizza (Eds): Covid-19 in Developing Economies. CEPR Press, London, UK

Covid Economics 38, 16 July 2020: 37-67

Hevia, Constantino and Andy Neumeyer (2020): A perfect storm: COVID-19 in emerging economies, in: Djankov, Simeon and Ugo Panizza (Eds): Covid-19 in Developing Economies. CEPR Press, London, UK

IMF (2020): World Economic Outlook Update, June 2020: A Crisis Like No Other, An Uncertain Recovery Washington, DC

Li, Kai, Xing Liu, Feng Mai and Tengfei Zhang (2020): The Role of Corporate Culture in Bad Times: Evidence from the COVID-19 Pandemic, *Covid Economics, Vetted and Real-Time Papers 32,* June.

Rahman, H Z, N Das, I Matin, M A Wazed, S Ahmed, J Nusrat and U Zillur (2020), Livelihoods, Coping and Support during the COVID-19 Crisis, BGID-PPRC Rapid Response Research.

Ramelli, Stefano and Alexander F. Wagner (2020): Feverish Stock Price Reactions to COVID-19. CEPR Discussion Paper 14511.

Covid Economics 38, 16 July 2020: 37-67

Covid Economics 38, 16 July 2020: 37-67

Appendix
A1: Survey questions

Early April survey

1. Have your operations been affected by the coronavirus pandemic; or do you expect them to be affected?
2. Have you breached, or might you breach, any covenants as a result?
3. Will raising additional capital likely be necessary to weather the crisis?
4. In response to the coronavirus pandemic, have you:
   a. reduced or delayed any investment spending plans?
   b. made any layoffs?
   c. made changes to any employee benefits (e.g. health care, sick leave, compensation, etc)?
   d. made any changes to executive compensation plans?
   e. made any changes to dividend or share buyback plans?
   f. taken any measures to protect employees, customers, or other stakeholders; or to prevent the spread of the virus, in excess of government guidelines?
   g. voluntarily continued to pay any employees or services providers for services which are temporarily disrupted?
   h. provided extraordinary non-contractual financial flexibility to any customers, suppliers, or business partners?
   i. made any additional coronavirus-related donations (cash, supplies, human resources), or shifted any business operations to fulfil crisis needs (e.g. manufacture of masks, hand sanitizer, etc) to help fight the pandemic?
   j. received any coronavirus-related government support, or do you expect to be a beneficiary of any coronavirus-related government support?

Early June survey

1. In response to the coronavirus pandemic, have you:
2. Approx how much have revenues declined over the past two months?
3. By when do you expect a recovery to pre-crisis revenues?
4. Are there any updates to the answers you sent us in April (in the email below)?

Covid Economics 38, 16 July 2020: 37-67

Table A2: Survey answers

	April survey			June update	
	Yes	**No**	**No answer**	**Yes**	**New**
Operations affected	82%	15%	3%	82%	0%
Breach covenants (or likely to breach)	1%	93%	6%	1%	0%
Raising additional capital likely necessary	12%	80%	8%	28%	16%
Received government support (or expect to receive)	35%	57%	8%	36%	1%
Reduced or delayed investment spending plans	50%	41%	10%	50%	0%
Layoffs	3%	91%	5%	7%	3%
Changes to employee benefits	24%	70%	6%	27%	3%
Changes to executive comp	17%	76%	7%	23%	5%
Changes to dividend or share buyback plans	13%	80%	8%	14%	1%
Measures to protect employees or other stakeholders	91%	5%	5%	91%	0%
Continued to pay employees for disrupted services	47%	45%	8%	50%	3%
Financial flexibility to customers	41%	51%	9%	42%	2%
Donations or shifted business operations	77%	18%	5%	77%	0%

	Positive	**Negative**	**Not clear**	**Positive**	**New**
Operations affected	3%	69%	11%	3%	0%
Changes to employee benefits	12%	7%	5%	15%	3%

Approx how much have revenues declined over the past two months?	**Total**
Decline	49%
Increase	9%
Minimal or no impact/change	13%
No answer / declined to answer	28%

By when do you expect a recovery to pre-crisis revenues?	**Total**
2020	24%
2021	12%
2022 or beyond	3%
Maintaned or better performance	11%
Not sure / unclear / no answer	50%

Did you take action before gov't pandemic measures were put in place, or following the issue of gov't measures?	**Total**
Before	63%
Following	27%
No clear answer	10%

Covid Economics 38, 16 July 2020: 37-67

Table A3: Two-sample Tests for Equal Means

(a) Two-sample test for equal means:
Firms that responded to April survey compared to those that did not

	Obs 1	Obs 2	Mean 1	Mean 2	Dif	SE	t-value	p-value
Tobin's q	99	439	1.179	1.351	-.172	.361	-.5	.634
Total Assets	99	439	6019.028	5700.712	318.317	1702.793	.2	.852
Cash / Total Assets	99	439	7.955	8.309	-.353	1.019	-.35	.729
Capex / Total Assets	95	439	4.806	4.489	.318	.579	.55	.584
Leverage	99	439	22.55	25.734	-3.184	2.126	-1.5	.135
Sales Growth	97	439	.197	.102	.095	.044	2.15	.032

Notes: Obs 1 is equal to one if companies did not respond to the April survey. Obs 2 is equal to 1 if companies responded to the April survey.

(b) Two-sample test for equal means:
Firms that responded to April and June survey compared to those that only responded to April survey

	Obs 1	Obs 2	Mean 1	Mean 2	Dif	SE	t-value	p-value
Tobin's q	52	387	.956	1.405	-.449	.516	-.85	.385
Total Assets	52	387	3301.613	6023.071	-2721.45	2202.827	-1.25	.218
Cash / Total Assets	52	387	6.687	8.526	-1.839	1.315	-1.4	.163
Capex / Total Assets	52	387	5.795	4.313	1.481	.716	2.05	.039
Leverage	52	387	26.619	25.615	1.004	2.774	.35	.718
Sales Growth	52	387	.12	.1	.02	.05	.4	.69

Notes: Obs 1 is equal to one if companies responded only to the April survey. Obs 2 is equal to 1 if companies responded to both April and June survey.

Covid Economics 38, 16 July 2020: 37-67

Table A4: Country variation in survey responses

	Bangladesh	Indonesia	Malaysia	Pakistan	Philippines	Saudi Arabia	South Africa	Thailand	Turkey	Vietnam	**Total**	**June updated**
Response rate	53%	73%	90%	56%	84%	67%	79%	91%	83%	80%	**77%**	**85%**
Sample size	20	55	69	25	46	49	72	82	29	41	**488**	**415**
Operations affected	90%	75%	87%	72%	98%	73%	89%	79%	72%	78%	**82%**	**82%**
Positively affected	*0%*	*0%*	*1%*	*4%*	*4%*	*2%*	*7%*	*1%*	*3%*	*2%*	***7%***	***3%***
Negatively affected	*85%*	*55%*	*80%*	*68%*	*87%*	*63%*	*61%*	*68%*	*69%*	*61%*	***69%***	***69%***
Breach covenants (or likely to breach)	0%	0%	0%	8%	0%	4%	3%	0%	0%	0%	**1%**	**28%**
Raising additional capital likely necessary	45%	18%	13%	8%	11%	8%	1%	11%	14%	15%	**12%**	**36%**
Received government support (or expect to receive)	40%	35%	39%	44%	22%	49%	19%	39%	34%	37%	**35%**	**50%**
Reduced or delayed investment spending plans	55%	56%	45%	16%	61%	20%	63%	65%	34%	49%	**50%**	**7%**
Layoffs	0%	5%	0%	0%	7%	6%	3%	4%	3%	5%	**3%**	**27%**
Changes to employee benefits	10%	24%	17%	16%	37%	10%	25%	40%	7%	29%	**24%**	**23%**
Positive changes	*5%*	*9%*	*12%*	*8%*	*20%*	*0%*	*10%*	*26%*	*3%*	*10%*	***12%***	***15%***
Negative changes	*0%*	*5%*	*3%*	*8%*	*0%*	*4%*	*10%*	*16%*	*3%*	*12%*	***7%***	***7%***
Changes to executive comp	10%	22%	12%	8%	11%	12%	25%	26%	14%	15%	**17%**	**23%**
Changes to dividend or share buyback plans	10%	33%	7%	0%	13%	2%	17%	7%	24%	10%	**13%**	**14%**
Measures to protect employees or other stakeholders	100%	95%	94%	100%	93%	88%	85%	93%	79%	83%	**91%**	**91%**
Continued to pay employees for disrupted services	40%	42%	48%	60%	74%	33%	71%	28%	31%	41%	**47%**	**50%**
Financial flexibility to customers	15%	29%	38%	36%	70%	20%	51%	48%	28%	44%	**41%**	**42%**
Donations or shifted any business operations	80%	87%	83%	76%	91%	47%	78%	83%	59%	76%	**77%**	**77%**
JUNE SURVEY												
Response rate	60%	87%	87%	84%	93%	80%	79%	94%	83%	83%	**85%**	
Sample size	12	48	60	21	43	39	57	77	24	34	**415**	
Approx. how much have revenues declined over the past two months?												
Decline	58%	56%	35%	57%	63%	38%	40%	52%	38%	38%	**49%**	
Increase	0%	6%	5%	5%	7%	10%	7%	9%	8%	21%	**9%**	
Minimal or no impact/change	0%	4%	20%	24%	2%	13%	4%	9%	25%	21%	**13%**	
No answer / declined to answer	42%	33%	40%	14%	28%	38%	49%	30%	29%	21%	**28%**	
By when do you expect a recovery to pre-crisis revenues?												
2020	8%	33%	15%	33%	19%	21%	5%	27%	25%	26%	**24%**	
2021	17%	10%	13%	5%	16%	13%	9%	14%	8%	9%	**12%**	
2022 or beyond	0%	4%	3%	0%	5%	3%	2%	8%	0%	0%	**3%**	
Maintained or better performance	8%	8%	15%	14%	0%	13%	2%	8%	21%	24%	**11%**	
Not sure / unclear / no answer	50%	42%	52%	43%	49%	36%	68%	38%	29%	29%	**50%**	
Did you take action before gov't pandemic measures were put in place, or following the issue of gov't measures?												
Before	50%	60%	62%	62%	56%	49%	58%	71%	33%	53%	**63%**	
Following	8%	35%	33%	14%	30%	33%	26%	19%	17%	21%	**27%**	
No clear answer	8%	4%	2%	5%	2%	3%	2%	5%	29%	21%	**10%**	

Covid Economics 38, 16 July 2020: 37-67

Table A5: Sector variation in survey responses

	Comms	Cons Discr	Cons Staples	Energy	Fin	Health Care	Ind	IT	Mat	Real Estate	Utilities	**Total**	**June updated**
Response rate	*82%*	*75%*	*87%*	*81%*	*71%*	*85%*	*84%*	*82%*	*76%*	*76%*	*61%*	***77%***	***85%***
Sample size	*23*	*65*	*47*	*25*	*77*	*17*	*70*	*31*	*66*	*53*	*14*	***488***	***415***
Operations affected	87%	92%	74%	80%	70%	76%	90%	84%	77%	96%	50%	**82%**	**82%**
Positively affected	*17%*	*3%*	*11%*	*0%*	*1%*	*6%*	*0%*	*0%*	*0%*	*0%*	*0%*	***7%***	***3%***
Negatively affected	*61%*	*83%*	*51%*	*68%*	*52%*	*65%*	*83%*	*74%*	*65%*	*85%*	*43%*	***69%***	***69%***
Breach covenants (or likely to breach)	4%	2%	0%	0%	0%	0%	1%	3%	2%	2%	0%	**1%**	**28%**
Raising additional capital likely necessary	13%	18%	9%	28%	5%	12%	20%	6%	6%	11%	7%	**12%**	**36%**
Received government support (or expect to receive)	39%	54%	23%	52%	17%	35%	49%	39%	32%	30%	0%	**35%**	**50%**
Reduced or delayed investment spending plans	35%	58%	45%	64%	35%	41%	54%	55%	47%	64%	43%	**50%**	**7%**
Layoffs	0%	9%	4%	0%	1%	6%	7%	0%	2%	2%	0%	**3%**	**27%**
Changes to employee benefits	39%	29%	19%	24%	22%	24%	27%	19%	17%	30%	14%	**24%**	**23%**
Positive changes	*39%*	*8%*	*11%*	*16%*	*16%*	*12%*	*10%*	*3%*	*8%*	*11%*	*14%*	***12%***	***15%***
Negative changes	*0%*	*14%*	*2%*	*4%*	*3%*	*6%*	*13%*	*10%*	*8%*	*8%*	*0%*	***7%***	***7%***
Changes to executive comp	4%	31%	9%	8%	12%	12%	26%	23%	12%	25%	0%	**17%**	**23%**
Changes to dividend or share buyback plans	13%	18%	2%	20%	14%	12%	20%	13%	6%	8%	7%	**13%**	**14%**
Measures to protect employees or other stakeholders	96%	89%	85%	92%	88%	82%	91%	100%	88%	94%	100%	**91%**	**91%**
Continued to pay employees for disrupted services	48%	46%	45%	36%	52%	29%	44%	32%	56%	60%	21%	**47%**	**50%**
Financial flexibility to customers	48%	35%	36%	24%	62%	18%	37%	45%	23%	58%	29%	**41%**	**42%**
Donations or shifted any business operations	96%	74%	77%	76%	79%	88%	70%	71%	76%	79%	93%	**77%**	**77%**
JUNE SURVEY													
Response rate	*91%*	*83%*	*85%*	*88%*	*84%*	*82%*	*90%*	*97%*	*74%*	*85%*	*86%*	*85%*	
Sample size	*21*	*54*	*40*	*22*	*65*	*14*	*63*	*30*	*49*	*45*	*12*	*415*	
Approx. how much have revenues declined over the past two months?													
Decline	43%	67%	43%	55%	29%	43%	46%	33%	57%	51%	42%	**49%**	
Increase	14%	2%	15%	5%	11%	14%	8%	7%	10%	2%	8%	**9%**	
Minimal or no impact/change	10%	2%	10%	14%	14%	14%	16%	17%	14%	2%	25%	**13%**	
No answer / declined to answer	33%	30%	33%	27%	46%	29%	30%	43%	18%	44%	25%	**28%**	
By when do you expect a recovery to pre-crisis revenues?													
2020	14%	26%	25%	14%	14%	14%	21%	20%	31%	24%	17%	**24%**	
2021	5%	11%	8%	18%	9%	0%	19%	13%	8%	18%	8%	**12%**	
2022 or beyond	0%	7%	3%	0%	2%	0%	8%	0%	0%	7%	0%	**3%**	
Maintained or better performance	10%	4%	20%	5%	6%	36%	10%	23%	8%	2%	17%	**11%**	
Not sure / unclear / no answer	62%	41%	38%	55%	54%	50%	37%	37%	51%	38%	50%	**50%**	
Did you take action before gov't pandemic measures were put in place, or following the issue of gov't measures?													
Before	81%	50%	60%	45%	57%	64%	54%	67%	55%	60%	83%	**63%**	
Following	10%	30%	23%	32%	25%	21%	30%	20%	39%	24%	0%	**27%**	
No clear answer	10%	2%	3%	9%	11%	7%	6%	10%	4%	7%	0%	**10%**	

Table A6: R-Squared Analyses

Panel A				
	(1)	(2)	(3)	(4)
VARIABLES	Have you reduced investment spending?			
Observations	439	439	439	439
Adjusted R-squared	0.054	0.098	0.040	0.133
Firm Controls	Yes	No	No	Yes
Country FE	No	Yes	No	Yes
Sector FE	No	No	Yes	Yes

Panel B				
	(5)	(6)	(7)	(8)
VARIABLES	Made any donations to help fight the pandemic?			
Observations	439	439	439	439
Adjusted R-squared	0.056	0.067	0.003	0.124
Firm Controls	Yes	No	No	Yes
Country FE	No	Yes	No	Yes
Sector FE	No	No	Yes	Yes

Panel C				
	(9)	(10)	(11)	(12)
VARIABLES	Do you expect to receive government support?			
Observations	439	439	439	439
Adjusted R-squared	0.026	0.018	0.051	0.062
Firm Controls	Yes	No	No	Yes
Country FE	No	Yes	No	Yes
Sector FE	No	No	Yes	Yes

*Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1*

CEPR PRESS

Covid Economics **Issue 38, 16 July 2020**

Financial markets and news about the coronavirus[1]

Harry Mamaysky[2]

Date submitted: 8 July 2020; Date accepted: 13 July 2020

I examine how financial markets interact with news about the COVID-19 pandemic. A twelve topic model optimizes the trade-off between number of topics and topic coherence. Using this model, I show that before mid-March 2020 markets react more to the same quantum of news when volatility is higher – a phenomenon I call hypersensitivity. Formal tests identify a structural break in mid-March, post which markets are no longer hypersensitive. In the hypersensitive stage, markets are overly volatile and overreact to news. Despite hypersensitivity, lagged prices better forecast future COVID-19 case counts than do lagged news.

Covid Economics 38, 16 July 2020: 68-128

1 I thank seminar participants from Columbia, QWAFAxNEW, UBS, and the Wolfe Research NLP and ML Conference, as well as Charles Calomiris and Paul Glasserman, for helpful suggestions.

2 Columbia Business School.

Covid Economics 38, 16 July 2020: 68-128

1 Introduction

In times of great economic and market stress, asset prices often respond dramatically to news about underlying economic or market conditions. This was true during the global financial crisis (GFC) of 2008–2009, when news about each successive multi-billion dollar loss would cause financial markets to wildly gyrate. Yet during the GFC, news data sets and the tools to analyze them were not yet widely available. Since the GFC, the use of natural language processing (NLP) in finance and economics has experienced sizable growth. The COVID-19 pandemic of 2020 has emerged as the first global crisis that can be analyzed in real-time using NLP tools developed in the last decade. Rather than draw anecdotal conclusions about how news and markets interact, researchers can systematically analyze the news-markets relationship. In this paper, I attempt such an analysis.

Over the time period from January to April of 2020, I analyze the 72,263 Reuters news articles that mention "coronavirus" or "COVID-19." I find there are twelve distinct topics into which these stories can be classified. The choice of twelve topics optimally balances the desire to have more topics to better represent news flow with the requirement that the topics are coherent – that is that these topics contain words that "sensibly" belong together, as in Newman et al. (2010). The prevalence of certain topics in news articles effectively tracks the narrative about the evolution of the pandemic. From early discussions of the health and market impact of the coronavirus, the narrative shifts to a discussion of the impact of the pandemic on corporations (2 topics), of its effect on credit markets (2 topics), and of the policy responses by central banks and governments. The other topics focus on currencies, European economies, oil and commodities, and sports.

As in Calomiris and Mamaysky (2019), I construct a measure of topical sentiment by interacting topical frequency with sentiment, the latter measured using the Loughran and McDonald (2011) dictionary. Topical sentiment is contemporaneously related to price changes in the SP500 stock index, in the VIX volatility index (which measures implied volatility on short-term SP500 options), in a high-yield corporate bond index (HY), and in yield changes in 2- and 10-year US Treasuries. Daily topical sentiment is able to explain a large portion of the variation in daily returns of these five asset classes. I argue that a contemporaneous relationship between asset returns and news is to be expected under the log-linear return approximation of Campbell and Shiller (1988) and Campbell (1991).

I check to see whether asset prices or the news-based measures are better forecasters of future COVID-19 case counts. Given investor focus on understanding potential coronavirus outcomes over my sample period, contemporaneous market responses to news would be justified if news were a good forecaster of future case incidence. I find that five of the top seven forecasters of future COVID-19 case counts are prices of the asset classes mentioned above, which optimally forecast future case counts at lags of two to four weeks. Lags of various news measures also forecast future case counts, but not as well as the market price series. In all cases, lags are chosen to maximize the absolute value of the correlation of the lagged series with future COVID-19 case counts. Lagged market prices are generally better forecasters of future COVID-19 case incidence than lagged news.

The contemporaneous relationship between news measures and asset returns becomes considerably larger when the VIX index has been elevated over the prior 10 trading days. The same piece of news has a considerably larger effect on asset prices when lagged implied volatility is higher. I refer to this phenomenon as *hypersensitivity*. Thus a portion of the increased volatility during the coronavirus crisis is caused not by progressively more extreme news flow, but by progressively larger market responses to similar news flow. In the hypersensitive phase, market prices respond to dimensions of news flow that are ultimately uninformative about future fundamentals, as measured by incidence of COVID-19 cases. At the same time, markets do not respond to dimensions of news that prove to be informative about future COVID-19 cases. This raises the possibility that a sizable amount of asset price movements during the hypersensitive phase were, in a sense, unwarranted.

Aspects of this dynamic are suggested by the theoretical analysis in Glasserman, Mamaysky, and Shen (2020), who argue that extreme media scrutiny can push markets into a high-volatility, low-price regime, which in turn invites more media scrutiny. In the model, such regimes are characterized by prices that are extremely sensitive to fundamentals. Part of the market's response to the global coronavirus pandemic appears to stem from this kind of news-induced extreme sensitivity.

Structural break tests, using the Andrews (1993, 2003) distribution for the maximal Chow statistic, suggest that a break in the contemporaneous news-markets relationship occurred around the middle of March 2020. This was a time of extreme market stress, as well as of extreme responses by central banks and fiscal authorities. I speculate that such policy actions contributed to the structural break in the news-markets relationship, suggesting an additional role for monetary and fiscal policy in averting adverse market outcomes: namely, non-virtuous news-market cycles can potentially be broken by appropriate policy responses. The hypersensitivity dynamic is much more prevalent in the pre-break subsample. After the structural break occurs, there is very little evidence of hypersensitivity. News have the same effect on contemporaneous asset price changes during

Covid Economics 38, 16 July 2020: 68-128

Covid Economics 38, 16 July 2020: 68-128

high- and low-VIX periods in the post-break regime.

I document the presence of lead-lag relationships between news and markets. Today's market moves Granger cause tomorrow's topical sentiment. Generally today's news flow is worse when yesterday's market returns were poor, even after controlling for yesterday's news flow. This suggests that news coverage, even of topics that are not directly market-related, is tainted by past poor market performance. Interestingly, today's topical sentiment also Granger causes tomorrow's market moves. This is true even after controlling for today's and yesterday's market moves. The direction of news forecastability for tomorrow's market moves overwhelmingly suggests overreaction, i.e. a price move tomorrow which reverses a portion of today's price response to a piece of news. The Granger causality network is considerably denser in the pre-break than in the post-break regime. Furthermore most of the price overreaction occurs in the pre-break regime, and almost all of the overreaction occurs for hypersensitive news-market pairs. Hypersensitive price reactions apparently are quickly, at least partially, reversed. In the post-break subsample, there is much less overreaction, and none of it happens in response to hypersensitive news-market pairs.

The substantially different contemporaneous and lead-lag relationships that prevail in the pre- and post-break regimes suggest that the markets response to the coronavirus crisis consisted of two distinct phases: an early hypersensitive phase, and a late relatively calm phase. In the hypersensitive phase, markets hung on every mention of coronavirus and COVID-19 as they tried to assess the economic impact of the pandemic. In the post-break phase, the sensitivity to coronavirus news flow diminishes. The strong evidence of overreaction in the early, hypersensitive phase suggests a portion of contemporaneous market response to coronavirus news flow was unwarranted.

The rest of the paper proceeds as follows. Section 2 describes the text analytics. Section 3 describes the non-text data sets. Section 4 studies the contemporaneous relationship between news flow and asset returns. Section 5 analyzes the lead-lag structure between news and market returns. Section 6 offers some interpretations and discusses a mechanism that may explain some of the findings. Section 7 concludes. Section A.2 of the appendix discusses some robustness tests.

1.1 Connection to the literature

The paper most closely related to the present work is Baker et al. (2020), who find that the majority of stock market moves greater than 2.5% from February 24 to March 24, 2020 were attributed by newspapers to news about COVID-19. I look more closely at the dynamics of the news-markets relationship, and do not focus only on large price move days. Furthermore, I analyze other markets (HY, Treasuries, VIX) besides the SP500 stock index. Loughran and McDonald (2020) document that in 2018 only 21% of firms mention pandemic-related risks in their 10-K filings, and wonder why this number wasn't higher. Gormsen and Koijen (2020) use equity dividend futures and the level of the stock market to extract changes in investor expectations about economic growth in response to the coronavirus pandemic. Altman (2020) investigates the impact of the COVID-19 crisis on the credit cycle. My work is related to the two latter papers in that I document the evolution of the narrative about the coronavirus pandemic, and show how the narrative and markets interact with each other; this interaction can in turn impact how investors think about future economic growth and corporate creditworthiness.

Roll's (1988) famous observation that monthly stock returns are difficult to explain even with contemporaneous influences precipitated much interest in the finance profession to understand which events can be associated with stock price movements. Fair (2002) identifies sixty-nine events that led to large SP500 moves between 1982 and 1999, and finds that 53 of these were associated with monetary policy. Tetlock, Saar-Tsechansky, and Macskassy (2008) show that contemporaneous market returns are more sensitive to news sentiment when the word "earn" is mentioned in the news story. I find that market-wide reactions at the asset class level to contemporaneous news about coronavirus are larger in high-VIX periods. Furthermore, the R^2s in my regressions of market returns on contemporaneous news measures are quite high, suggesting that in the early part of the COVID-19 crisis markets movements were, to a meaningful extent, driven by news about coronavirus. Boudoukh, Feldman, Kogan and Richardson (2018) regress daily squared stock returns on an indicator variable for the presence of contemporaneous intraday and overnight news, and conclude that the presence of news increases daily return variance. My results suggest that asset-level return variance during high-VIX days is at least partially caused by increased sensitivity of asset returns to news flow, and not because of the arrival of more extreme news.

The evidence on overreaction in this paper is in contrast to the finding that news sentiment can predict aggregate stock returns (Tetlock 2007), and especially during recessions (Garcia 2013). Both Tetlock (2007) and Garcia (2013) find evidence of index-level underreaction to news. My analysis differs from theirs because I consider multiple asset classes, and not only the US stock index, and because I focus on one particular high-stress market episode whereas they analyze a much longer time-series that is not solely focused on market crises.

Covid Economics 38, 16 July 2020: 68-128

CEPR PRESS

Covid Economics 38, 16 July 2020: 68-128

2 Text analysis

The news articles used in this study are obtained from the Thomson-Reuters news archive, and include all English-language articles in 2019 and through the end of April 2020 that mention the terms "coronavirus" or "COVID-19". There is only one article that mentions coronavirus all of 2019.[1] The first article mentioning coronavirus in 2020 is on January 8, and the coverage begins in earnest starting on January 17, the date that I start my analysis. Figure 1 shows the daily count of articles mentioning coronavirus rises steadily from January 17, 2020 onwards. When there are multiple articles in a given Primary News Access Code chain,[2] the last one is selected. Day t articles must have timestamps prior to or equal to 4pm NY time; day t articles with later timestamps are classified as day $t+1$ articles. Post-4pm Friday, as well as Saturday and Sunday articles, are counted as occurring on the subsequent Monday. Articles occurring after 4pm NY time on April 30, 2020 are excluded from the analysis. The final corpus contains 72,263 articles.

Article j's sentiment is given by

$$Sent_j = \frac{Pos_j - Neg_j}{Total_j},$$

where Pos_j and Neg_j are the number of positive and negative Loughran-McDonald (2011) words appearing in the article, and $Total_j$ is the total number of words after excluding stop words. I employ the Das and Chen (2007) algorithm to mark negated words,[3] and negated sentiment words are ignored in the above counts. Daily aggregate sentiment $Sent_t$ is the average $Sent_j$ of all articles classified as day t articles. This is the series labeled *sent* in Figure 1. The average value of daily sentiment in the corpus is -0.019 (Table 1). Glasserman, Li, and Mamaysky (2020) show that, in a corpus of approximately 1.4 million articles about SP500 firms from 1996 to 2018, the average article sentiment is -0.011. The sentiment in the present corpus is lower, though not dramatically so. I also calculate the daily standard deviation of article-level sentiment; this series is labeled *sent_sd*.

The topic model is obtained by running latent Dirichlet allocation (LDA) on the document-term matrix obtained by first stemming all the words in a news article, then dropping stop words and several other commonly occurring words,[4] and then keeping

[1] It is on February 27, 2019 and describes the Coalition for Epidemic Preparedness Innovations presciently partnering with a German firm to create a faster process for anti-pandemic vaccine production.

[2] This is how Reuters identifies a collection that includes the initial article, and subsequent revisions.

[3] Implemented via the `mark_negation` function of Python's `NLTK` package.

[4] These are: *said, thomsonreut, https, tmsnrt, www, reuter, coronavirus, com, nl.*

Covid Economics 38, 16 July 2020: 68-128

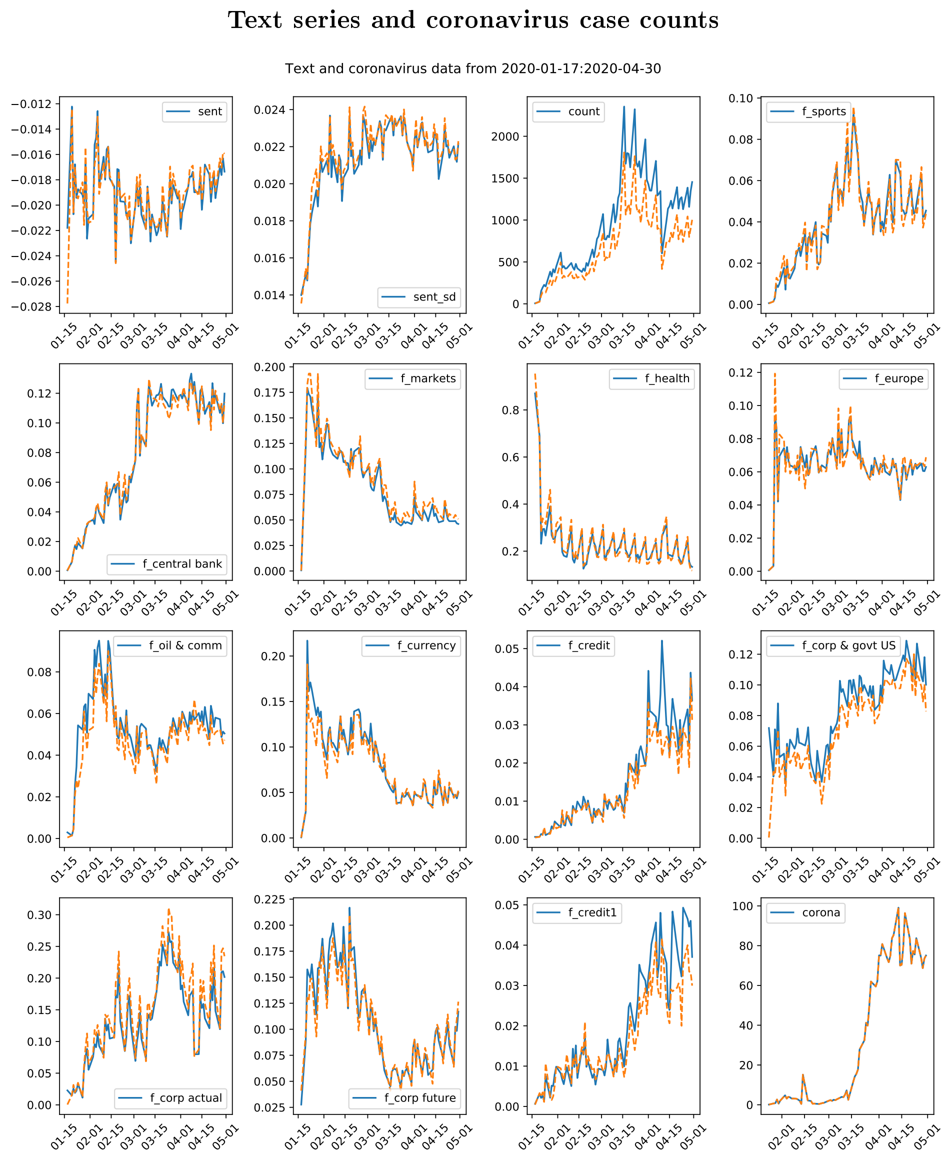


Fig. 1. Summary of sentiment (*sent*), article count (*count*), topic frequencies (marked *f_*), and coronavirus confirmed cases. Data are daily. The text measures are from all articles that mention "coronavirus" or "COVID-19." The solid lines correspond to text measures obtained from all articles in the corpus, and the dashed lines correspond to the same text measures derived after all non-weekend intraday (timestamp from 9:30am–4pm NY time) articles have been removed.

Covid Economics 38, 16 July 2020: 68-128

Table 1
Summary statistics for the variables used in the analysis. All data are daily. The labels *sent*, *sent_sd*, and *count* refer respectively to the daily mean and standard deviation of article sentiment, as well as the daily number of articles. The *f_[topic]* and *s_[topic]* series refer to topical frequency and sentiment respectively. The coronavirus case counts, labeled *corona*, are obtained from the Johns Hopkins Coronavirus Resource Center, and are reported in thousands. N shows the numbers of days for which there are observations. The AR(1) (first-order autocorrelation) coefficient is shown by default for the level of each series, but if marked with r (d) it is shown for one-day returns (differences). The markets series are labeled as follows: *sp500* is the SP500 stock index; *vix* is the VIX volatility index; *gt2* and *gt10* are the 2- and 10-year Treasury yields respectively; and *hy* refers to the FTSE US high-yield index. The series *vix_avg* refers to the average level of *vix* over the prior ten trading days.

Summary statistics for daily variables

Variable	N	Mean	StdDev	5%	25%	Median	75%	95%	AR(1)
sent	75	-0.019	0.002	-0.023	-0.021	-0.019	-0.018	-0.015	0.268
sent_sd	75	0.021	0.002	0.017	0.021	0.022	0.023	0.024	0.806
count	75	960.840	553.533	203.100	439.500	993.000	1348.500	1793.000	0.888
f_sports	75	0.040	0.019	0.008	0.027	0.042	0.050	0.070	0.841
f_central bank	75	0.082	0.041	0.015	0.045	0.102	0.119	0.127	0.953
f_markets	75	0.084	0.039	0.046	0.050	0.073	0.115	0.163	0.854
f_health	75	0.223	0.108	0.143	0.170	0.187	0.252	0.312	0.644
f_europe	75	0.065	0.014	0.051	0.061	0.065	0.070	0.087	0.361
f_oil & comm	75	0.054	0.017	0.030	0.048	0.054	0.061	0.091	0.796
f_currency	75	0.083	0.042	0.037	0.048	0.073	0.113	0.149	0.783
f_credit	75	0.016	0.013	0.001	0.006	0.010	0.025	0.037	0.882
f_corp & govt US	75	0.084	0.025	0.042	0.062	0.090	0.103	0.118	0.847
f_corp actual	75	0.138	0.065	0.023	0.092	0.136	0.181	0.242	0.835
f_corp future	75	0.109	0.047	0.053	0.069	0.098	0.155	0.188	0.825
f_credit1	75	0.020	0.014	0.002	0.008	0.014	0.033	0.046	0.921
s_sports	75	-0.001	0.000	-0.001	-0.001	-0.001	-0.000	-0.000	0.839
s_central bank	75	-0.002	0.001	-0.003	-0.002	-0.002	-0.001	-0.000	0.938
s_markets	75	-0.002	0.001	-0.003	-0.002	-0.001	-0.001	-0.001	0.767
s_health	75	-0.004	0.002	-0.006	-0.005	-0.004	-0.003	-0.002	0.546
s_europe	75	-0.001	0.000	-0.002	-0.001	-0.001	-0.001	-0.001	0.436
s_oil & comm	75	-0.001	0.000	-0.001	-0.001	-0.001	-0.001	-0.001	0.643
s_currency	75	-0.002	0.001	-0.003	-0.002	-0.001	-0.001	-0.001	0.689
s_credit	75	-0.000	0.000	-0.001	-0.000	-0.000	-0.000	-0.000	0.863
s_corp & govt US	75	-0.002	0.001	-0.002	-0.002	-0.002	-0.001	-0.001	0.783
s_corp actual	75	-0.003	0.001	-0.005	-0.004	-0.003	-0.002	-0.000	0.840
s_corp future	75	-0.002	0.001	-0.004	-0.003	-0.002	-0.001	-0.001	0.657
s_credit1	75	-0.000	0.000	-0.001	-0.001	-0.000	-0.000	-0.000	0.914
corona	70	32.092	34.827	0.528	2.216	9.303	71.505	83.783	0.980
sp500	73	88.969	10.066	72.769	82.022	88.287	99.374	101.874	r -0.425
vix	73	36.666	19.561	12.952	16.050	37.760	47.300	73.388	d -0.397
hy	76	93.154	6.834	80.993	88.445	93.675	99.699	100.381	r 0.336
gt2	76	0.784	0.548	0.204	0.241	0.530	1.414	1.530	d 0.014
gt10	76	1.108	0.437	0.599	0.705	0.982	1.584	1.771	d -0.124
vix_avg	76	34.846	19.238	12.734	15.654	37.593	49.560	67.167	d 0.859

all words occurring more than 10 times in any month's set of coronavirus articles.[5] An LDA topic is a probability distribution over, in this case stemmed, words. The interested reader can consult Steyvers and Griffiths (2007) for a good introduction to topic models. Figure 2 shows word clouds that summarize the identified topics, where each word's size is drawn in proportion to its probability weight in a given topic. The topic labels reflect my qualitative judgment about the meaning of each topic. A twelve topic model is optimally chosen, for reasons that I will discuss shortly. For example: the *central bank* topic contains words such *bank*, *govern*, and *billion*; the *credit* topic contains words like *credit*, *fitch* (the rating agency), and *rate*; there is an additional credit-related topic, which I label *credit1*.[6]

LDA assigns to each document a probability distribution over each of the topics in the model. For example, a news article that discusses the credit impact of central bank policy may have a 75% allocation to the *central bank* topic, a 20% allocation to the *credit* topic, with the remaining 5% allocated to the remaining topics. To get a better sense of the types of articles that fall into each topic, Table 5 in the appendix shows eight headlines of articles that are representative of each topic: four each for the highest and lowest sentiment articles among representative articles for each topic. An article is said to be *representative* about topic i if its document-topic allocation to i is above 70%. In the above example, the article would be representative of the *central bank* topic. As can be seen from the table, most of the representative articles about the *central bank* topic are appropriately classified, as are all of the *credit* and *credit1* articles. Most other representative article headlines also appear appropriate. Finally, sentiment allocation of the headlines appear intuitive: higher sentiment headlines generally convey more positive news than negative sentiment ones.

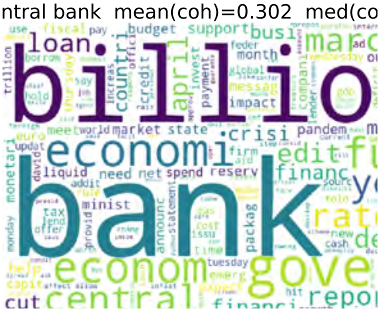
[5] I use the LDA implementation in Python's `scikit-learn` package.

[6] As discussed below, I tried many independent topic estimation runs, and many of these identified two distinct credit-related topics. Some runs identified more than two.

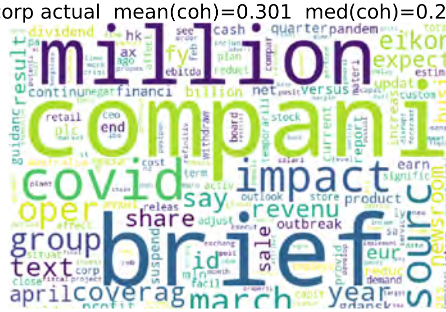
Covid Economics 38, 16 July 2020: 68-128

Covid Economics 38, 16 July 2020: 68-128

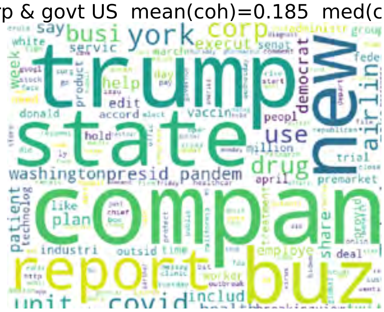
central bank mean(coh)=0.302 med(co



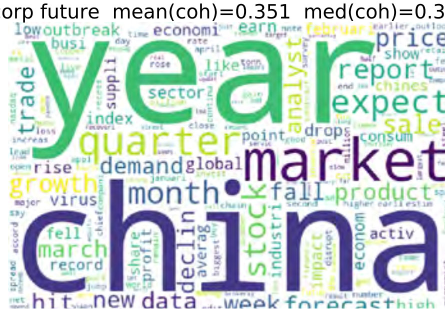
corp actual mean(coh)=0.301 med(coh)=0.278



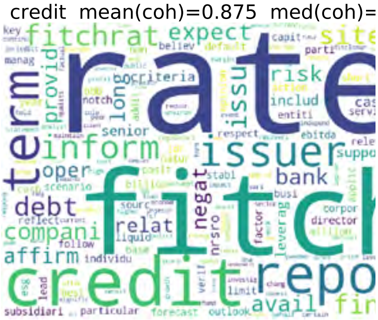
corp & govt US mean(coh)=0.185 med(c



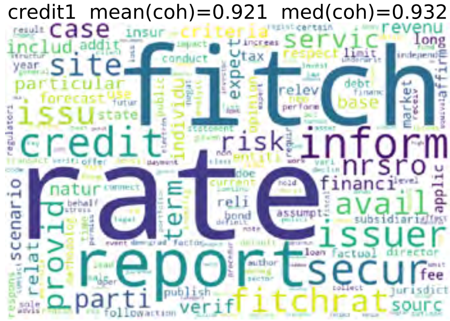
corp future mean(coh)=0.351 med(coh)=0.333



credit mean(coh)=0.875 med(coh)=



credit1 mean(coh)=0.921 med(coh)=0.932



Covid Economics 38, 16 July 2020: 68-128

Fig. 2. Word clouds describing topics associated with articles that mention "coronavirus" or "COVID-19." Each word cloud is labeled with the topic name, as well as the mean and median coherence for the topic (as described in Section 2.1).

CEPR PRESS

Table 2
Averages of document-topic distributions (the first twelve rows), sentiment (labeled *sent*), and article counts (labeled *Number*), grouped by time period: *intra* refers to articles that come out from 9:30am–4pm on non-weekend days, and *over* refers to overnight and weekend articles. By construction, the first twelve entries in each column sum to one.

Average document-topic distribution by time-period

bucket	intra	over
sports	0.049	0.047
central bank	0.107	0.100
markets	0.053	0.075
health	0.201	0.203
europe	0.065	0.066
oil & comm	0.063	0.049
currency	0.068	0.066
credit	0.027	0.017
corp & govt US	0.117	0.085
corp actual	0.122	0.185
corp future	0.090	0.089
credit1	0.036	0.020
sent	-0.020	-0.019
Number	21,640	50,623

For day t, I aggregate the document-topic distributions of all articles classified as belonging to that day into a daily measure of *topic frequency*. Letting $f_{j,k}$ be the probability allocation of article j to topic k, the daily topic frequency measure is

$$f_{t,k} = \frac{1}{N_t} \sum_{j \in \{\text{day } t \text{ articles}\}} f_{j,k}, \tag{1}$$

where N_t is the number of articles in day t. Table 2 shows average document-topic distributions, sentiment, and article counts for *intraday* (9:30am–4pm, non-weekend) and *overnight* periods (all others). The most prevalent topics are: *health* (top words include *case*, *report*, *virus*); two corporate topics *corp & govt US* and *corp actual* (top words include *brief*, *company*, *report*, *buzz*, *corp*); and *central bank*. The largest discrepancy between within day topic incidence is in *corp actual* which is more likely to occur overnight than intraday. These series are labeled *f_[topic]* in subsequent analysis. I discuss the

Covid Economics 38, 16 July 2020: 68-128

Covid Economics 38, 16 July 2020: 68-128

evolution of topic prevalence in Section 2.2.

Calomiris and Mamaysky (2019) find that topical sentiment is an important forecaster of country-level returns. *Topical sentiment* is the product of topic frequency with sentiment at the daily level. This measure captures the extent to which negative or positive news flow is specific to a given topic. The daily topical sentiment series for topic k is

$$Sent_{t,k} = Sent_t \times f_{t,k}. \tag{2}$$

For example, if day t has very negative daily sentiment $Sent_t$, and has articles predominantly about *central bank* and *credit* topics, the topical sentiment of those two topics on day t would be very negative, while the topic sentiment of the non-prevalent topics would be close to zero. These series are labeled $s\_[topic]$ in subsequent analysis.

2.1 Topic model selection and coherence

It is well known that there can be multiple LDA topic models that result in either identical or very similar log likelihoods for a given text corpus.[7] Because of this indeterminacy care must be taken to identify an appropriate topic model for a particular application. To select an optimal model, I estimate 10 independent runs for eight different choices of number of topics: 3, 6, 9, 12, 15, 18, 21 and 24. This results in 80 topic models estimated using independent runs of `scikit-learn`'s LDA algorithm, each of which starts at a different random seed and ends at different model estimates. Intuitively, the optimization here involves choosing the maximum number of topics, which will be helpful in identifying particular market-topic relationships, while ensuring that each of the topics is "sensible" and not overly esoteric.

The natural language processing literature has identified a useful measure of topic quality, called *coherence*. Newman et al. (2010) show that various algorithmic measures of topic coherence fit well with human evaluation of topic quality. Following their approach, I use a coherence measure that utilizes cosine similarities among the top 10 most frequent words in a given topic. Let $\mathcal{D}$ be the subset of the document-topic matrix for the present corpus that is restricted to the top 1,000 most frequently occurring words.[8] $\mathcal{D}$ is then a

[7] See Michal et al. (2004) for a convenient form of the LDA likelihood function for a text collection. Ke, Montiel Olea, and Nesbit (2020) give results on non-identification of LDA topic models.

[8] This restriction only applies to the coherence calculation, and not to the LDA estimation which uses the entire document topic matrix.

CEPR PRESS

Covid Economics 38, 16 July 2020: 68-128

72,263 × 1,000 matrix. The cosine similarity of two words i and j is defined as

$$c_{i,j} = \frac{\mathcal{D}_i^\top \mathcal{D}_j}{\|\mathcal{D}_i\|\|\mathcal{D}_j\|},$$

where $\mathcal{D}_i$ is the i^{th} column of $\mathcal{D}$. $c_{i,j}$ is a number from 0 to 1 and measures the tendency of two words to co-occur. When $c_{i,j} = 1$ then in every document in which i appears, so does j, and in the same relative proportion; and if $c_{i,j}$ is zero, then i and j never appear in the same document. If either i or j is not in the top 1,000 words (which happens very infrequently in the analysis), then $c_{i,j}$ is set to zero.

The coherence of a given topic is then the mean or median of the $\binom{10}{2} = 45$ possible cosine similarities $c_{i,j}$'s among the top 10 words in a topic. High coherence topics consist of top words that co-occur frequently in the corpus. Figure 2 shows the mean and median coherence of each of the twelve topics in the model. *credit* and *credit1* are the two highest coherence topics – the top words in each, *fitch rate*, *credit* – tend to co-occur very frequently. On the other hand, *corp & govt US* is a low coherence topic. Its top words, *compani*, *trump*, *buzz*, *new*, *report*, are less closely related and tend to co-occur much less frequently than the *credit* and *credit1* topics' top words. *sports* is a medium coherence topic, whose top words *sport*, *event*, *game*, *postpon*, co-occur relatively frequently.

While the mean and median topic coherence were, in all cases I examined, very similar, I chose the median as my base measure because it is less sensitive to outlier pairwise cosine similarities. To rank a given topic model with K topics, I calculate the average median coherence across the K topics. I refer to this as the *coherence* of a topic model, and write it as $\mathcal{C}_K$. In the present corpus, there is evidence of a power law that prevails between topic model coherence and number of topics. In a regression of $\log \mathcal{C}_K$ on $\log K$ (recall there are 10 runs for each K), I find that

$$\log \mathcal{C}_K = -0.3967 - 0.2613 \log K + \text{noise},$$

where the White standard error of -0.2613 is 0.007 and the adjusted R^2 of the regression is 0.901. Given the high R^2 of the fit, the quantity $\mathcal{C}_K K^{0.25}$, or the *scaled coherence* of the model, should equal a constant times a small error term. This quantity captures the trade-off between having a higher number of topics (desirable for finding useful market-topic relationships) and the observation that in the present corpus a higher number of topics decreases average topic coherence (undesirable).

Figure 3 summarizes the results of this analysis. It plots the scaled topic coherence of

Covid Economics 38, 16 July 2020: 68-128

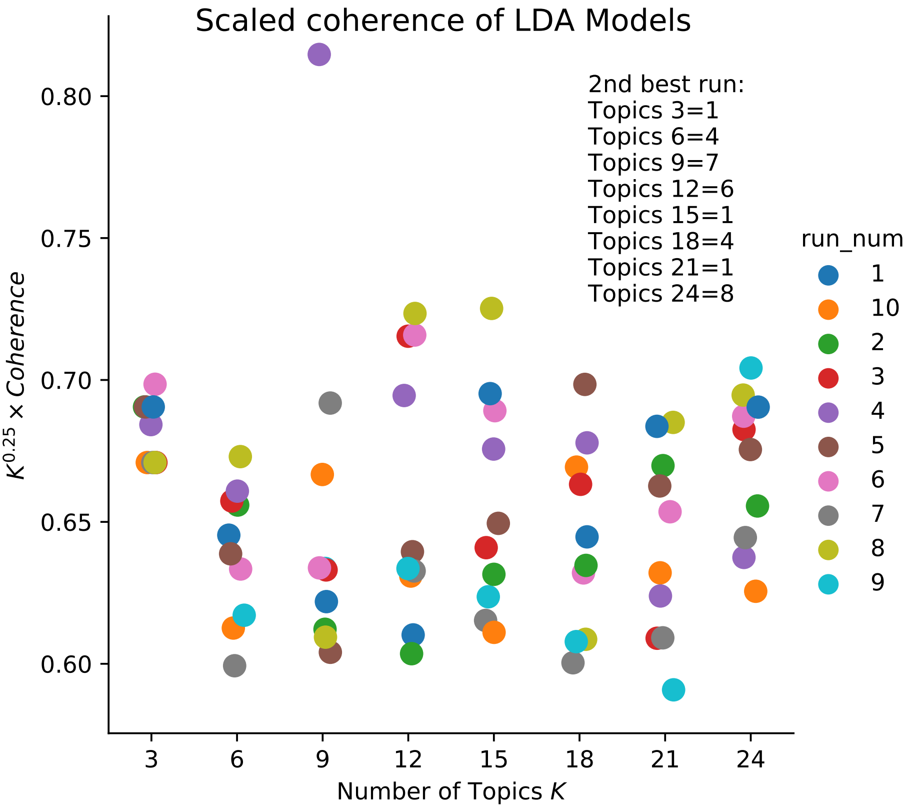


Fig. 3. This figure shows the scaled average of median coherence, as explained in Section 2.1, for all topics across different topic configurations and different model runs. The x-axis shows the number of topics for a given set of model runs. The y-axis shows the scaled average coherence for a given topic model, as a function of the run number and number of topics in the model. The table shows the run with the second best scaled average coherence for each topic configuration.

all 10 runs for each K. As can be seen, except for model run 4 in the 9-topic set of models, the highest scaled average coherence is obtained in the 12- and 15-topic runs. The 9-topic run number 4 has a very high coherence because the model chooses three credit topics instead of the usual two, and credit topics have high coherence as already mentioned. To mitigate the effect of such outlier topic choices, I focus instead on the second best model run across the Ks. The winning model run based on the second best criterion is run 6 in the 12-topic model. This is the topic model that I use in the remaining analysis.

It should be noted that the main conclusion of the paper are not sensitive to the choice of number of topics or to which run is chosen for a given topic number. All model run and topic configuration variants I have tried lead to the same qualitative conclusions. That being said, more coherent topics increase the intepretability of the results. Ultimately, the economic intuition behind the markets-news relationship depends on having topics that are easy to understand and interpret. For this reason my model selection approach should prove useful for future research. Users of topic models in economics should explore the model space and find a suitable model for their particular application. The usual approach of running a single LDA with a guess as to the appropriate topic number is unlikely to yield the best-possible model.

2.2 Evolution of the narrative

Figure 1 shows the evolution of the daily topical frequencies $f_{t,k}$ in equation (1) from January 17 through the end of April of 2020, the daily sentiment $Sent_t$, the daily standard deviation of article-level sentiment ($sent\_sd$), as well daily article counts ($count$). The solid lines in the figure correspond to text measures derived from all articles in the corpus. The orange, dashed lines correspond to text measures which exclude all intraday articles; these will serve as a robustness test, as I discuss in Section A.2.

The aggregate sentiment series exhibits a lot of volatility in the early part of the sample and achieves its minimum level around the middle of March of 2020. The overall sentiment of all coronavirus articles improves from this low point in March through the end of April. The daily standard deviation of article-level sentiment $sent\_sd$ starts very low in January, peaks in the middle of March, and begins to slightly decline through April; though it remains at very elevated levels relative to the start of the sample. The article count starts at almost zero in early January, peaks at over 2,000 articles per day in the middle of March, and then settles into the 1,000-1,500 range by the end of April.

Table 2 shows the average topic prevalence in the intraday and overnight periods, and Figure 1 shows that topic incidence has varied meaningfully over the span of the crisis.

Covid Economics 38, 16 July 2020: 68-128

Covid Economics 38, 16 July 2020: 68-128

The news stories in the first few days of the crisis fall into the *health* topic, whose values on January 17th and 18th were 87% and 69% respectively, before falling and staying in the 20% range for the rest of the sample. On January 17th, there was virtually no news coverage of markets, though this quickly changed by the next week, when 17% of all news coverage regarding coronavirus falls into the *markets* topic. There was a spike in news coverage about oil and commodities in early February, and this series then remains in the mid-single digits for the rest of the period . Much of the early news coverage of the pandemic dealt with its effect on currency markets, though the *currency* topic becomes less prominent towards the end of the sample. The share of coverage of the European impact from coronavirus remains steady at around the 6% level throughout the crisis.

The two corporate topics exhibit somewhat different trends. The *corp future* topic, which is largely about anticipation of future corporate impacts from the crisis, peaks in the early part of the sample and then settles into the 5% to 10% range. The *corp actual* topic, which deals with the realized effect of the pandemic on corporations, starts at a very low level, and then climbs to be the most prevalent topic towards the middle and end of the sample, as anticipated outcomes become realized. The least coherent topic, *corp & govt US*, which seems to be about government and corporate interactions around the coronavirus crisis, start at 5% and grows to roughly 10% of the news share. Given the topic's low coherence, it is perhaps a catch-all topic for stories that do not neatly fit into the other categories.

The *sports* topic starts out very low, as most early news coverage of the pandemic does not deal with sports, then peaks at close to 10% of the news flow in early March as all major professional sports suspend their seasons, and as talk about postponing the Tokyo Olympics intensifies. The *central bank* topic is almost never mentioned in coronavirus articles at the beginning of the sample, and then steadily grows until it represents roughly 12% of news flow by early March, a level it maintains until the end of the sample. The two credit topics are also absent in the early crisis news flow and then grow steadily as the profound impact of quarantine on corporate and consumer credit becomes apparent. In the latter part of the sample the two credit topics, *credit* and *credit1*, represent just under 10% of the news flow.

In summary, the twelve topic model does a good job of tracking the evolution of news flow about the crisis, from the early articles dealing with the health impact of coronavirus to the later articles dealing with central bank interventions and credit impacts. The hope, therefore, is that these news measures will allow us to understand how markets and news flow about the coronavirus evolve throughout the crisis.

CEPR PRESS

Table 1 gives summary statistics for these text series, as well as for topical sentiment, $Sent_{t,k}$, which is the product of daily topic frequency and sentiment. The text series are all fairly persistent as measured by their daily AR(1) coefficients. The aggregate sentiment series $Sent_t$ is less so.

3 Other data

I collect daily price data on the SP500 index, the VIX index, the FTSE US High-Yield Market index (HY), which tracks the performance of high-yield corporate bonds, and US 2- and 10-year Treasury yields, labeled respectively *gt2* and *gt10*. I include high-yield bonds because they are a fairly liquid and particularly economically sensitive asset class. The data are from Bloomberg, and go from January through the end of April 2020.

I obtain data on global confirmed COVID-19 cases from the Johns Hopkins Coronavirus Resource Center. These data are updated daily, and start in 1/22/2020. According to the `GitHub` page which stores the data,[9] the update frequency is: "Once a day around 23:59 (UTC)."[10] For example, the data for 4/1/2020 would be released at 23:59 UTC, which is 7:59pm NY time. This is after the index calculation time on 4/1/2020, which is 3pm for Treasuries and 4pm for SP500 and VIX, and perhaps an hour or two later for high-yield.[11] For these markets, the reaction to the 4/1/2020 COVID-19 case counts would not happen until the next trading day's close, or 4/2/2020. For this reason I use the day t increase in the global COVID-19 case count as the day $t+1$ value of my *corona* series. This aligns the case counts with the days on which markets would have been able to react. For Mondays, the one-day increase comes from Sunday's case counts. Assigning to Mondays the cumulative case count increase from Friday, Saturday and Sunday would introduce a day-of-week effect, which using only the Sunday increase avoids.

Figure 1 shows the *corona* series, reported in thousands of cases. The daily COVID-19 case counts start relatively low, and experience exponential growth starting in mid-March. The spike in the series in February corresponds to a one-time restatement of case counts in China. Table 1 gives summary statistics for the markets and *corona* series. For the table, the HY and SP500 indexes are normalized to a value of 100 on January 16, 2020. The *corona* series, though it represents the daily increase in case counts, is highly persistent with an AR(1) coefficient of 0.980. The markets series, whose AR(1) coefficients measure

[9] I use the file `time_series_covid19_confirmed_global.csv` from `github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series`.

[10] *UTC* stands for Coordinated Universal Time, the successor standard to Greenwich Mean Time.

[11] According to the *FTSE Fixed Income Index Guide* most indexes use closing prices from 3pm–6pm.

Covid Economics 38, 16 July 2020: 68-128

the autocorrelation of either returns (labeled with r) or differences (labeled with d), are also unusually persistent. The AR(1) coefficient for daily SP500 returns, for example, is -0.425, suggesting a very large degree of mean-reversion over the sample. For context, the AR(1) coefficient for daily SP500 returns in 2019 was -0.09. The VIX index is also strongly mean-reverting over this sample, while HY index returns are strongly positively auto-correlated. Interestingly Treasury yield changes have very low autocorrelations.

3.1 A first look at the data

News flow (to an extent) and financial markets are both forward looking. From Figure 1, it is clear that the media began to discuss the impacts of coronavirus long before the COVID-19 case count began its exponential growth. This leads to the question of how well and with that lead time did news flow and markets anticipate future case incidence?

Figure 4 investigates this question. The text series used in the subsequent analysis are the daily topical sentiment series $Sent_{t,k}$ from (2). I use topical sentiment instead of topic frequency following Calomiris and Mamaysky (2019), who find topical sentiment useful for forecasting country-level stock returns. Topical sentiment may be particularly relevant because it captures both the prevalence of news flows about a particular topic, as well as the sentiment of that news flow.

For each text and market series, I find the lag, between 0 and 21 trading days, at which that series is most correlated (in absolute terms) with daily COVID-19 case counts. For example, the level of 2-year yields (*gt2* in the figure) when lagged 15 trading days has a correlation of -97.4% with the *corona* series. All other lags of *gt2* have a correlation that is smaller, in absolute terms. The two series are plotted in the upper left-hand corner of Figure 4. The *corona* series is shown as a dashed line, and the 15-day lag of *gt2*, shown as the solid line, is scaled to have the same range as *corona* and then multiplied by the sign of its correlation to *corona*. 2-year Treasury yields begin falling 15-days prior to the sharp upturn in COVID-19 case counts and the scaled, lagged *gt2* series closely tracks the daily incidence of COVID-19 case counts.

The figure shows the markets and text series sorted from highest correlation in the upper-left corner to lowest correlation in the bottom-right corner. Of the top seven most correlated series, five are the markets series, and the remaining two are *s_central bank* and *s_credit1*, the topical sentiment for the *central bank* and *credit1* topics respectively. The average optimal lag for the markets series is 15.4 trading days. Thus financial markets appear to have anticipated the increased incidence of COVID-19 case counts by roughly three calendar weeks. Of the remaining text series there are several whose optimal lags

Covid Economics 38, 16 July 2020: 68-128

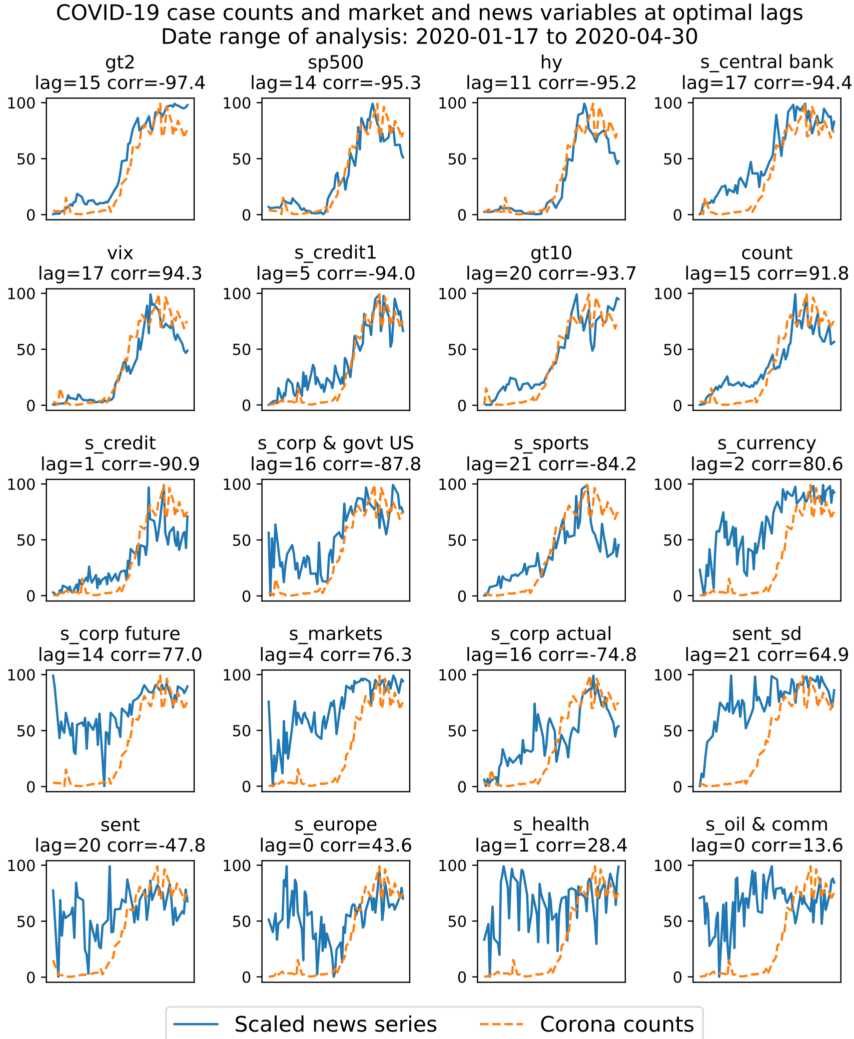


Fig. 4. This figure shows correlations between daily COVID-19 case counts and the markets and news series at optimal lags. For each candidate series, a lag is chosen from 0 to 21 trading days to maximize the correlation between daily COVID-19 case counts and the lagged market or news series. The series are plotted in order of highest (upper left) to lowest (lower right) absolute correlations. The optimal correlation and lag are shown for each series. The lagged market or text series is scaled to have the same range as the case count series *corona* and then multiplied by -1 (+1) if the correlation between it and *corona* is negative (positive). The scaled markets or text series are shown as solid, blue lines in the figure. The *corona* series is shown in each subplot as the dashed, orange line. The y-axis corresponds to units of the *corona* series, in the thousands.

Covid Economics 38, 16 July 2020: 68-128

CEPR PRESS

Covid Economics 38, 16 July 2020: 68-128

have a relatively high correlation with *corona*, though generally the lags are shorter than 15.4 days. Some news series, for example *s_ oil & comm* and, surprisingly, *s_ health*, have very low correlations with future disease incidence. Interestingly, aggregate sentiment $Sent_t$ (*sent* in Figure 4) has a relatively low correlation with future case counts, suggesting that the decomposition of sentiment into topical categories conveys important information.

The general takeaways of this analysis are: (1) markets are highly anticipatory of future COVID-19 incidence, and (2) some of the news series are as well, though less so. Given that markets and news anticipate future coronavirus outcomes, it is natural to examine how the two interact. I now turn to this analysis.

4 Contemporaneous relationships

I assume the contemporaneous relationship between markets and news is captured by

$$h_{t+1} = c + b^\top w_{t+1} + e_{t+1}, \tag{3}$$

where h_{t+1} is the asset return from day t to $t+1$, c is a constant, w_{t+1} is a $k \times 1$ vector derived from news flow and other information available to the econometrician, b is a $k \times 1$ constant vector, and e_{t+1} consists of unobservable information. This contemporaneous markets-news specification has been used in the literature, notably in Tetlock et al. (2008) and in Glasserman, Li, and Mamaysky (2020) for returns and in Boudoukh et al. (2018) for squared returns. Nevertheless a theoretical justification for (3) has not yet been established, to the best of my knowledge. In the appendix, I show that such a justification obtains from the Campbell and Shiller (1988) and Campbell (1991) result that asset returns can be decomposed into changes in investor beliefs about future dividend growth and future discount rates. As I show, the information vector w_{t+1} can then be interpreted as innovations to state variables that are useful for forecasting future cash flows or discount rates. In the rest of the paper, I do not formally separate the contribution of w_{t+1} to one of these two channels, though that is an interesting question for future work.

4.1 Empirical implementation

I specialize (3) as follows:

$$h_{t+1} = c + b_1 h_t + b_2 h_{t-1} + b_3 N_{t+1} + b_4 N_{t+1}(VIX_t^{10} - \overline{VIX}^{10}) + b_5 VIX_t^{10} + e_{t+1}. \tag{4}$$

CEPR PRESS

The observable information includes two lagged returns, h_t and h_{t-1}, which control for the dependence of current news flow on past returns, and for the auto-correlation properties of the dependent variable. None of the results change materially if h_t and h_{t-1} are excluded. In addition w_{t+1} contains N_{t+1}, which can be one of: aggregate daily sentiment $Sent_t$, the daily standard deviation of article-level sentiment *sent_sd*, one of the twelve topical sentiment series $Sent_{t,k}$, or the COVID-19 case count series *corona*, which tests the degree to which markets react to the most recently announced COVID-19 case counts. I refer to these fifteen variables as the *news series* in the remainder of the paper, expanding the term "news" to refer also to COVID-19 case counts. All standard errors use Newey-West with three lags.

The theoretical analysis in Glasserman, Mamaysky, and Shen (2020) suggests information shocks can push the economy into a high-information production, high-volatility regime, where asset prices become both depressed and extremely sensitive to changes in fundamentals. Section 6 discusses this mechanism in more detail. The $N_{t+1}(VIX_t^{10} - \overline{VIX}^{10})$ term tests if the effect of N_{t+1} on returns differs depending on whether markets are currently in a high- or low-volatility period. The volatility regime is proxied by VIX_t^{10}, which is the day t level of the rolling 10 trading-day average of the VIX index, and $\overline{VIX}^{10}$ is the average of VIX_t^{10} in the time period over which the regression is being estimated. I use the 10-day average of the VIX to smooth out high-frequency variation; and I use a one-day lag to avoid endogeneity issues.

For the SP500 and HY indexes, h_{t+1} equals the day t to day $t+1$ total return. For the VIX, h_{t+1} measures the day-over-day difference in the VIX index. At high frequency, this is very similar to the return of investing in a VIX futures contract. For 2- and 10-year Treasuries, h_{t+1} is the day-over-day change in yields. Assuming Treasury duration did not change materially during the sample period, which is reasonable given the relatively low starting and ending interest rate levels from January to April of 2020, Treasury returns are to a first-order linear in yield changes. Therefore, for all asset classes, returns either equal h_{t+1} or are approximately linear in h_{t+1}; I will refer to h_t as *returns*.

For each of the five dependent variable, I therefore run fifteen different version of (4), one for each of the possible N_{t+1}'s. I don't include multiple N_{t+1}'s in the specification at once because given the small sample size, coefficient estimates would become unreliable.

The results of this analysis are described below. I discuss some econometric issues with the present analysis, as well as robustness tests, in Section A.2 of the appendix.

Covid Economics 38, 16 July 2020: 68-128

Covid Economics 38, 16 July 2020: 68-128

4.2 Structural break tests

A concern with running the regression in (4) over the entire January through April sample is that the markets-news relationship may change over time. Markets in February and March of 2020 were extraordinarily stressed, and perhaps the markets-news relationship that prevailed in this period of extreme stress did not last the entire duration of even this short sample. If this is true, then running any analysis over the full period risks simply averaging across two very different regimes. To test for this possibility, I check whether there is evidence of a structural break in any of the 75 markets–news pairs that I analyze (five markets and fifteen news series for each market).

When testing for a structural break at a known break point $0 < t < T$, one can calculate the Chow test statistic given by

$$\phi(t) = \frac{(SSR - SSR_e - SSR_l)/k}{(SSR_e + SSR_l)/(N_e + N_l - 2k)}, \tag{5}$$

where e (early) refers to dates prior to or equal to t and l (late) refers to those dates after t, SSR refers to the sum of squared residuals over the entire sample, SSR_e (SSR_l) refers to the sum of squared residuals over the early (late) part of the sample, N_e (N_l) refers to the number of observations in the early (late) part of the sample, and k refers to the number of regressors. As the number of observations grows, $k\phi(t)$ approaches a χ^2 distribution with k degrees of freedom. Andrews (1993, 2003) tabulates the distribution of the maximal $k\phi(t)$ over all possible ts in some interval of the data. Specifically, for an unknown break point π which is drawn from an interval $[\pi_0, 1-\pi_0] \subset (0,1)$ of the data, Andrews (1993,2003) tabulates the distribution of

$$\sup_{\pi \in [\pi_0, 1-\pi_0]} k\phi(\pi T). \tag{6}$$

under the null hypothesis that there is no structural break. The tabulated distribution depends on k, which here equals 6, and π_0 which here is set to 0.3. Note the choice of time period over which to estimate $\overline{VIX}^{10}$ affects only the b_3 coefficient in (4) but leaves the residuals of the regression unchanged; therefore $\phi(t)$ in (5) is unaffected by this choice.

For each of the five market variables, Figure 5 shows a histogram of the maximal break points $t = \pi T$ that maximize (6) for the fifteen specifications that I analyze. In addition, for every break point, I tabulate the number of market–news pairs that are significant at least at the 10% level according to the distribution tables in Andrews (2003). The bottom-right panel of the figure shows the distribution of the break points for all 75 tests

Covid Economics 38, 16 July 2020: 68-128

Break dates using all articles

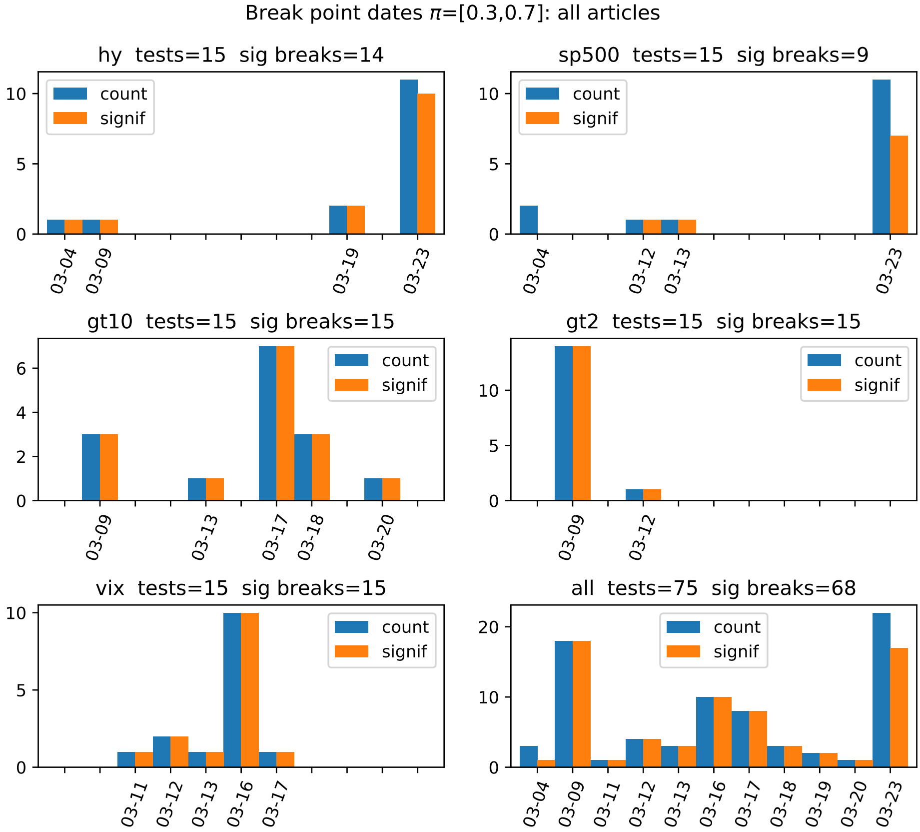


Fig. 5. For each market variable, there are 15 regressions which are tested for a break: overall sentiment, the standard deviation of sentiment, the twelve topical sentiment series, and the COVID-19 case counts series. The data start on January 17, 2020. The starts show number of break points that are significant at the 10% level or better using the Andrews (2003) distribution for the maximal Chow statistic with $\pi_0 = 0.3$.

Covid Economics 38, 16 July 2020: 68-128

conducted, as well as the number of break points that are significant.

All five asset classes exhibit strong evidence of a regime break sometime in March. For the VIX and the two Treasury series, all 15 news variables tested show evidence of a significant break. For HY 14 of 15 do, and for the SP500 9 out of 15 tests show evidence of a significant break. The break dates are concentrated on March 9th, March 16th and 17th, and March 23rd. March 16th and 17th are the Monday and Tuesday that follow an emergency rate cut announced by the Fed on Sunday, March 15th. And the March 23rd break date follows another emergency Fed meeting on Sunday, March 22nd when the Fed removed quantitative guidance from its announced emergency programs and simply vowed to purchase Treasury and mortgage-backed securities "in the amounts needed." The markets reacted favorably to the Fed's, and other central banks', announcements which were perceived to be expansive and bold (see Hartley and Rebucci 2020).

The structural break tests suggest that the markets–news relationships that had prevailed in the early parts of the crisis changed around the middle of March. Going forward, I use Sunday, March 15th as a cutoff date between the early- and late-parts of the sample period, and run all of my analyses separately in these two subsamples. Allowing the break between the early and late parts of the sample to equal, for each market–news pair, the optimal break from (6) produces even stronger results. For this reason, my choice of a fixed break point for all markets–news pairs is conservative.

4.3 The early subperiod

Table 3 summarizes the results of estimating (4) in the early (before March 15) part of the sample. The column groupings correspond to a particular market variable, and the rows correspond to one of the fifteen different news variables. b_3 (column *EV*) is normalized to report the effect of a one standard deviation change in the news variable in units of standard deviations of the market variable; b_4 (column *EV*VIXl1*) shows the change in the normalized b_3 for a one unit increase in the lagged VIX[10]. For example, the value of -0.701 for the effect of aggregate sentiment on the VIX means when aggregate sentiment increases by one standard deviation, the VIX experiences a contemporaneous decline of 0.701 standard deviations of daily changes. I will discuss the *EVl1* column in Section 5.

Table 3

Summary of the contemporaneous and lead-lag model results for daily market changes or returns. The column groupings correspond to different market variables, and the rows correspond to the fifteen different explanatory variables. The first two entries for every market variable show the b_3 (*EV*, for explanatory variable) and b_4 (*EV*VIX*) coefficients in (4) that are significant at the 10% level or better. *EV* column shows the impact of a one standard deviation change in the explanatory variable in units of standard deviation of the market variable. *EV*VIXl1* column shows the impact of a unit increase in VIX^{10} on the value of *EV*. The last entry for each market variable indicates the the c_3 (*EVl1*) coefficient from (7) that are significant at the 10% level or better, where (7) is run using the lagged all-article text measures. The summary statistics underneath the table are as follows: *Jnt sig* (joint significance) is the number of specifications where b_3 and b_4 are both significant; *Hyper* (hypersensitivity) is the number of times that both b_3 and b_4 are significant and have the same sign; *Lead* (lead-lag relationship) is the number of times that c_3 is significant; *Under* (underreaction) is the number of times b_3 and c_3 have the same sign; *Over* (overreaction) is the number of times b_3 and c_3 have opposite signs; and *Over+Hyper* (overreaction and hypersensitivity) is the number of time b_3 and c_3 have opposite signs while b_3 and b_4 have the same sign.

Summary of analysis for the all-article corpus: early subsample

	SP500			VIX			HY			GT2			GT10		
	EV	EV*VIXl1	EVl1	EV	EV*VIXl1	EVl1	EV	EV*VIXl1	EVl1	EV	EV*VIXl1	EVl1	EV	EV*VIXl1	EVl1
sent	0.761***	0.085***		−0.701***	−0.081***		0.603***	0.104***	−0.347**	0.616***	0.093**		0.574***	0.092*	−0.218*
sent_sd	−0.742**	−0.106**	1.137***	0.857***	0.117***	−1.098***	−1.100**	−0.180***	1.181***			0.906***			1.231***
sports	0.940**		−0.613*	−1.041***			0.645**	0.035*	−1.333***	1.129**			0.891*		
central bank										0.615*					
markets	0.857***	0.075***		−0.798***	−0.069***		1.079***	0.110***	−0.442*	0.426*	0.078***	−0.892***	0.960***	0.139***	−1.077**
health	1.053***	0.104**		−1.099***	−0.113***		1.245***	0.171***	−0.892***				0.940*		
europe	0.985***	0.078*		−0.829***			0.676***	0.101***							−0.617**
oil&comm	0.663***	0.215***		−0.460***	−0.187***		1.073***	0.223***		0.817***	0.169***	−0.808***	1.110***	0.232***	−0.863***
currency	0.771**			−0.681**						1.262***	0.148***		1.127***	0.139***	
credit										0.576*	0.061*				
corp&govt US							0.338**		0.379*					−0.077**	
corp actual			0.297**			−0.403**									
corp future										0.531**	0.089***		0.457***	0.095***	
credit1															
corona	−0.187**	−0.083***	0.830***	0.163*	0.079***	−0.782***	−0.222***	−0.076***	0.209*			0.403**			0.519***

Jnt sig = 32, Hyper = 32, Lead = 24, Under = 1, Over = 16, Over+Hyper = 15

Covid Economics 38, 16 July 2020: 68-128

All b_3 coefficients for sentiment series in specifications involving the SP500, HY, and 2- and 10-year Treasury yields that are significant have positive signs. So negative news as conveyed by either aggregate sentiment or any of the topical sentiment series is associated with a contemporaneous decline in the SP500 and HY indexes, and with a contemporaneous decrease in Treasury yields. For the SP500 and HY, the sign of b_3 for *sent_sd* is negative suggesting a contemporaneous market drop associated with a higher standard deviation of article-level sentiment. For VIX, all of these signs are reversed. Bad news, as proxied by aggregate or topical sentiment, or a high *sent_sd*, is associated with a contemporaneous increase in the VIX. Finally, the SP500 and HY indexes fall on days when the prior day's COVID-19 case count (the *corona* series) is high. And the VIX rises on such days. As mentioned earlier, the prior day's case counts come out after that day's market close, and thus markets react to these with a one-day lag.

Another striking feature of Table 3 is that in all 32 cases when both b_3 and b_4 are significant, they have the same sign. On high volatility days, measured by an elevated level of the 10-day average VIX on the prior day, the effects of all news measures are larger than they are during normal-volatility days. Interestingly, the b_4 coefficient for the *corona* series also has the same sign as b_3, suggesting that the SP500, VIX and HY indexes all react *more* to the same increase in case counts during high-volatility than low volatility days. I refer to these phenomena as *hypersensitivity* to contemporaneous news: A portion of the high volatility that asset markets experience in high-volatility states is not due to an increased volatility of news flow, but results instead from an increased sensitivity of markets to similarly volatile news. This is the prediction of Glasserman, Mamaysky, and Shen (2020) for markets in the high-information, high-volatility regime.

Tables 6 – 10 in the appendix show the full regression estimates of the model in (4) for any specification with a significant b_3 or b_4 coefficient. The rows of the table correspond to model estimates for a particular news variable. Note that the large negative b_1 coefficients for the SP500 and VIX indexes in Tables 6 and 7 respectively are consistent with the large negative AR(1) coefficients for these series documented in Table 1. Furthermore, the R^2s in these tables are very high, but should be interpreted with caution because of the small sample size and persistent regressors. The break date column in these tables corresponds to the date t^* which maximizes (6) and the associated $k\phi(t^*)$ value is shown in square brackets. The stars indicate significance relative to the Andrews (2003) tabulated values for the maximal Chow statistic.

4.4 The late subperiod

CEPR PRESS

Covid Economics 38, 16 July 2020: 68-128

Table 4

Summary of the contemporaneous and lead-lag model results for daily market changes or returns. The column groupings correspond to different market variables, and the rows correspond to the fifteen different explanatory variables. The first two entries for every market variable show the b_3 (*EV*, for explanatory variable) and b_4 (*EV*VIX*) coefficients in (4) that are significant at the 10% level or better. *EV* column shows the impact of a one standard deviation change in the explanatory variable in units of standard deviation of the market variable. *EV*VIXl1* column shows the impact of a unit increase in VIX^{10} on the value of *EV*. The last entry for each market variable indicates the the c_3 (*EVl1*) coefficient from (7) that are significant at the 10% level or better, where (7) is run using the lagged all-article text measures. The summary statistics underneath the table are as follows: *Jnt sig* (joint significance) is the number of specifications where b_3 and b_4 are both significant; *Hyper* (hypersensitivity) is the number of times that both b_3 and b_4 are significant and have the same sign; *Lead* (lead-lag relationship) is the number of times that c_3 is significant; *Under* (underreaction) is the number of times b_3 and c_3 have the same sign; *Over* (overreaction) is the number of times b_3 and c_3 have opposite signs; and *Over+Hyper* (overreaction and hypersensitivity) is the number of time b_3 and c_3 have opposite signs while b_3 and b_4 have the same sign.

Summary of analysis for the all-article corpus: late subsample

	SP500			VIX			HY			GT2			GT10		
	EV	EV*VIXl1	EVl1	EV	EV*VIXl1	EVl1	EV	EV*VIXl1	EVl1	EV	EV*VIXl1	EVl1	EV	EV*VIXl1	EVl1
sent	0.750***				0.044**	0.184**	0.804***								
sent_sd	−0.246*	0.023**	0.200**	0.407**		−0.188*		0.037***	0.246**						
sports	0.259**			−0.387***	0.037*		0.437***								
central bank	0.677***	0.034***				0.215***	0.675***	0.026**							
markets	0.507***			−0.268*		0.245***	0.518***								
health	0.481***				0.039**		0.447***								
europe	0.700***	0.050***	0.364**	−0.369**	0.044*	0.301***	0.652***	0.036**	0.482***						
oil&comm										−0.305**		−0.217**			−0.219**
currency				−0.304*			0.340***								
credit	−0.274**	0.046***		0.450**	−0.044**			0.043***				−0.182***			
corp&govt US				0.215***			0.280**		0.142*	−0.339**					
corp actual												0.204**			
corp future		0.018*				0.176**		0.024**							
credit1	−0.275***			0.450***	−0.039**		−0.233*								
corona	0.348***			−0.400**	0.053**		0.390***	−0.038*				0.215**			

Jnt sig = 12, Hyper = 4, Lead = 16, Under = 4, Over = 4, Over+Hyper = 0

Covid Economics 38, 16 July 2020: 68-128

Table 4 summarizes the results of estimating (4) in the late (after March 15) part of the sample. Tables A1 – A5 in the Internet Appendix show the full regression results for all explanatory variables where at least one of the b_3 and b_4 coefficients is significant. The patterns in the late subsample are starkly different from those in the early subsample. Of the 75 (five markets times fifteen news variables) combinations, there are now only 12 times when b_3 and b_4 are jointly significant. In only four of these cases do b_3 and b_4 have the same sign. Compare these numbers to the 32 jointly significant market-news pairs, of which all 32 have the same sign, in the early part of the sample. Furthermore, only two news variable come in significant for either of the Treasury series. When it is significant, topical sentiment in the late subsample generally has the same sign as in the early subsample, with positive news associated with contemporaneous increases in SP500 and HY, and a drop in the VIX. The *credit* and *credit1* series are an exception. However, the sign of the b_3 coefficient in the *corona* specification is opposite in the late subsample compared to the early one; in the late subsample higher case counts are associated with positive SP500 and HY returns, and with drops in the VIX.

Overall, the late subsample looks very different from the early one. Markets are no longer as responsive to news series, and markets are rarely hypersensitive. The mid-March structural break in the markets-coronavirus-news relationship is economically important, and occurs right around the time of major policy actions by the Fed. Though the timing may be coincidental, I speculate that the Fed may have effected the structural break in the markets-news relationship. This suggests a new role for central banks – they might be uniquely positioned to influence the markets-news interaction, and to nudge markets out of their hypersensitive phase.

4.5 Variance decomposition

It is hard to verify how much of market volatility is due to hypersensitivity. Instead, I analyze the following counterfactual. I reestimate (4) but after removing the hypersensitivity term, i.e. $N_{t+1}(VIX_t^{10} - \overline{VIX}^{10})$. I refer to this as the restricted specification. The difference in the adjusted R^2 of the baseline specification versus that of the restricted specification captures the contribution of the hypersensitivity term to the explained variance of SP500, HY, VIX or Treasury returns. In a counterfactual where the hypersensitivity effect is absent, presumably this difference in adjusted R^2s measures how much lower return variance would be. Tables 6–10 (for the early subsample) and Internet Appendix Tables A1–A5 (for the late subsample) show this R^2 difference in the *Chg R2* column.

Figure 6 summarizes the results of this analysis for the early subsample. Each panel

Covid Economics 38, 16 July 2020: 68-128

R^2 summaries for contemporaneous regressions: All-article, early subsample

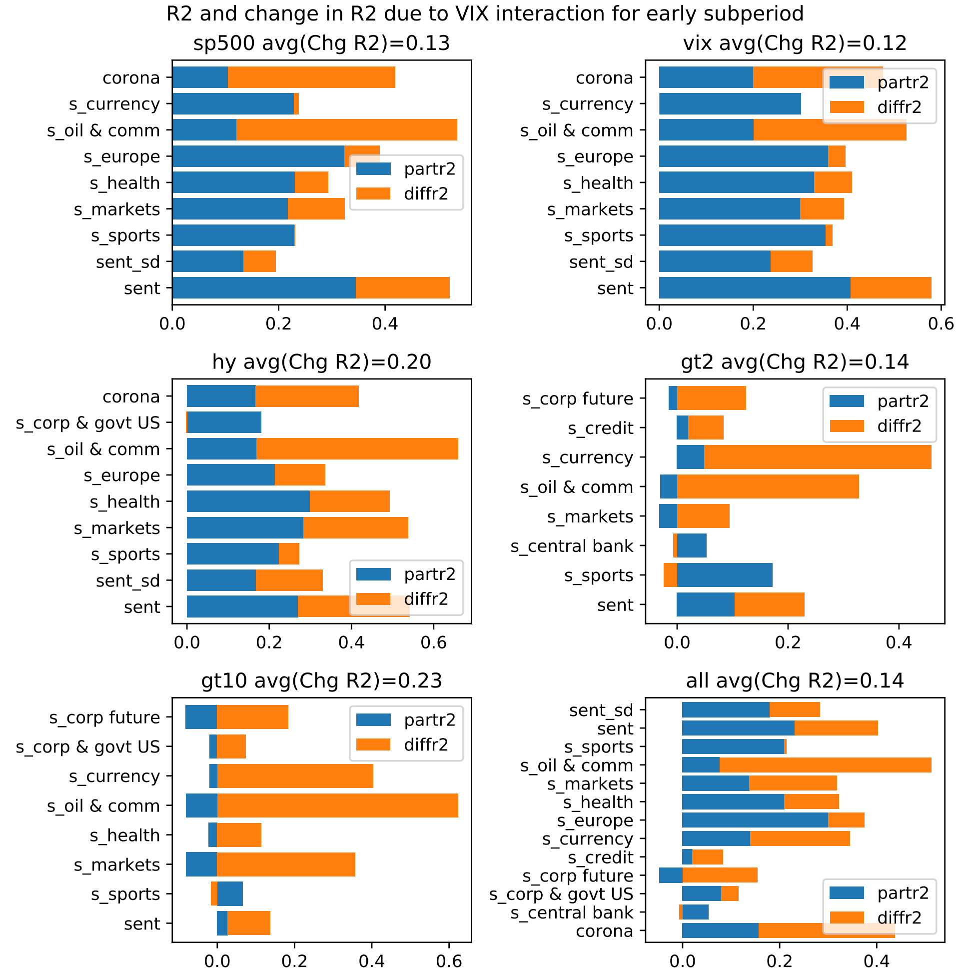


Fig. 6. The figures summarize the R^2s of the contemporaneous regressions in (4), using the all-article text measures for the early subsample. The panels show R^2s for each market variable for which at least one of the b_3 or b_r coefficients in (4) is significant at the 10% level or better. Some market variables may be missing in some specifications if there were no significant b_3s or b_4s for any explanatory variable. Each panel shows the adjusted R^2 (labeled *partr2*) for a version of (4) that excludes the VIX-explanatory variable interaction, i.e. for which $b_4 = 0$, and the difference between this and the adjusted R^2 of the full model (labeled *diffr2*). The panel labeled *all* reports the average of these two measures across all market variables.

shows the change in adjusted R^2s for each of the market-news pairs where either the the news variable or its VIX^{10} interaction (the b_3 or b_4 coefficients in equation 4) is significant at the 10% level or better. The blue, striped bars show the adjusted R^2 of the restricted specification, and the orange bars show the change in R^2 when moving to the baseline (unconstrained) model. The bottom-right panel shows the R^2 decomposition, averaged across all markets-news pairs. Each panel is labeled with the average increase in R^2s when adding back in the hypersensitivity variable to the contemporaneous regression.

For example, for SP500 returns and the *s_oil & comm* explanatory variable, the restricted R^2 is 12%, and this increases by 41.5% when the $N_{t+1}(VIX_t^{10} - \overline{VIX}^{10})$ term is added back in (Table 6 shows the exact numbers). The counterfactual is that the variance of daily SP500 returns would be 41.5% lower if the hypersensitivity effect was absent. The average across all significant news variables for the SP500 is 13%, and the average hypersensitivity term contribution across all significant specifications is 14% (bottom right panel).

Figure 7 shows the results of this analysis for the late subperiod. First, there are no significant instances of news variables or their VIX^{10} interactions for 10-year Treasury yield changes, which is absent from the figure. Second, for the SP500, HY, VIX, and 2-year Treasury series, the average R^2 contribution of the hypersensitivity term is very small. Across all three markets the average increase is only 3%, versus the 14% in the early subsample.

A version of this analysis that looks at all fifteen news variables for each market series reaches the same qualitative conclusions. There is strong evidence that hypersensitivity contributes to the variance of market returns in the early subsample (average R^2 increase of 9%), and almost no such evidence in the late subsample (average R^2 increase of 1%).

Of course, counterfactuals, especially in the absence of a structural model, are highly speculative. But the difference between the R^2 contribution of the hypersensitivity term in the early (14%) and late (3%) subsamples does suggest that some portion of the early subperiod volatility was due to the presence of the hypersensitivity effect.

5 Lead-lag relationships

In addition to analyzing the contemporaneous relationship between news flow and markets, it is natural to analyze the lead-lag relationship between the two. Do markets lead news? Or do news lead markets? To gain insight into these questions, I conduct a series of Granger causality tests. The general specification of these tests mirrors the contempo-

Covid Economics 38, 16 July 2020: 68-128

R^2 summaries for contemporaneous regressions: All-article, late subsample

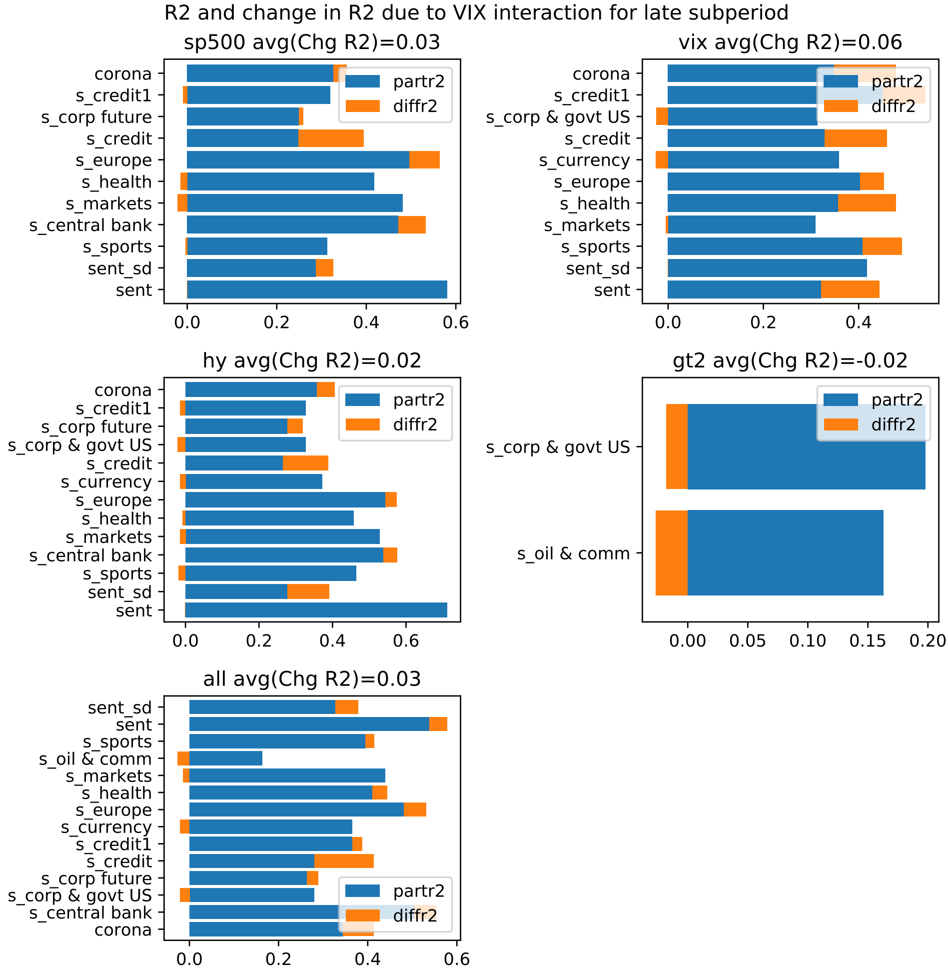


Fig. 7. The figures summarize the R^2s of the contemporaneous regressions in (4), using the all-article text measures for the late subsample. The panels show R^2s for each market variable for which at least one of the b_3 or b_r coefficients in (4) is significant at the 10% level or better. Some market variables may be missing in some specifications if there were no significant b_3s or b_4s for any explanatory variable. Each panel shows the adjusted R^2 (labeled *partr2*) for a version of (4) that excludes the VIX-explanatory variable interaction, i.e. for which $b_4 = 0$, and the difference between this and the adjusted R^2 of the full model (labeled *diffr2*). The panel labeled *all* reports the average of these two measures across all market variables.

CEPR PRESS

raneous regression in (4):

$$\rho_{t+1} = c_0 + c_1\rho_t + c_2\rho_{t-1} + c_3\tau_t + c_4\tau_t(VIX_t^{10} - \overline{VIX}^{10}) + c_5 VIX_t^{10} + e_{t+1}. \qquad (7)$$

Here ρ_{t+1} is the day $t+1$ response variable, and τ_t is the day t test variable. The two lags of the response variable control for contemporaneous correlations with the test variable, and for the possibility that the test variable is itself Granger caused by the lagged response variable. I say that τ Granger causes ρ if the c_3 coefficient above is significant at the 10% level or better. Standard errors are calculated using Newey-West with three lags.

5.1 Early subsample

I first look at the early subsample (dates prior to March 15). Table 11 in the appendix shows the results of estimating (7) with daily changes in the five market series as the test variable, and the next day's change in the fifteen news series as the response variables. For each market variable, the columns of the table show the c_3 and c_4 coefficient estimates from (7) for each future news variable. c_3 and c_4 are reported in standard deviations of the response variable per a standard deviation change in the test variable; c_4 reports the change in c_3 for a unit change in VIX^{10}. Figure 8 shows the Granger causality relationships from the table as a graph. Every significant link from h_t (time t market variable) to N_{t+1} (time $t+1$ news variable or *corona* value) is shown as an arrow. The arrow is a blue, solid (red, dashed) line when the c_3 coefficient in (7) is significant and positive (negative).

For example, there is a blue line from SP500 to *sent* because the one-day lagged SP500 return positively affects the next day's aggregate sentiment *sent*. On the other hand, there is a red, dashed line from the lagged VIX change to aggregate sentiment, because a day t increase in the VIX causes a day $t+1$ decline in aggregate sentiment. In all cases, day t increases in the SP500 and HY indexes, and in the two Treasury series, Granger cause positive day $t+1$ topical sentiment increases. But day t increases in the VIX forecast next day decreases in all sentiment series, except for *sent_sd*; not surprisingly, a high VIX today forecasts a high *sent_sd* or standard deviation of article-level sentiment tomorrow. No other market variable Granger causes *sent_sd*. Also no market variable Granger causes the COVID-19 case count series *corona* at a one-day lag, though as shown in Section 3.1 all market series are highly correlated with case counts two to four weeks in the future.

In the markets to news variables tests, there are 30 instances where c_3 and c_4 from (7) are both significant; in none of these cases do they have the same sign. The market effect

Covid Economics 38, 16 July 2020: 68-128

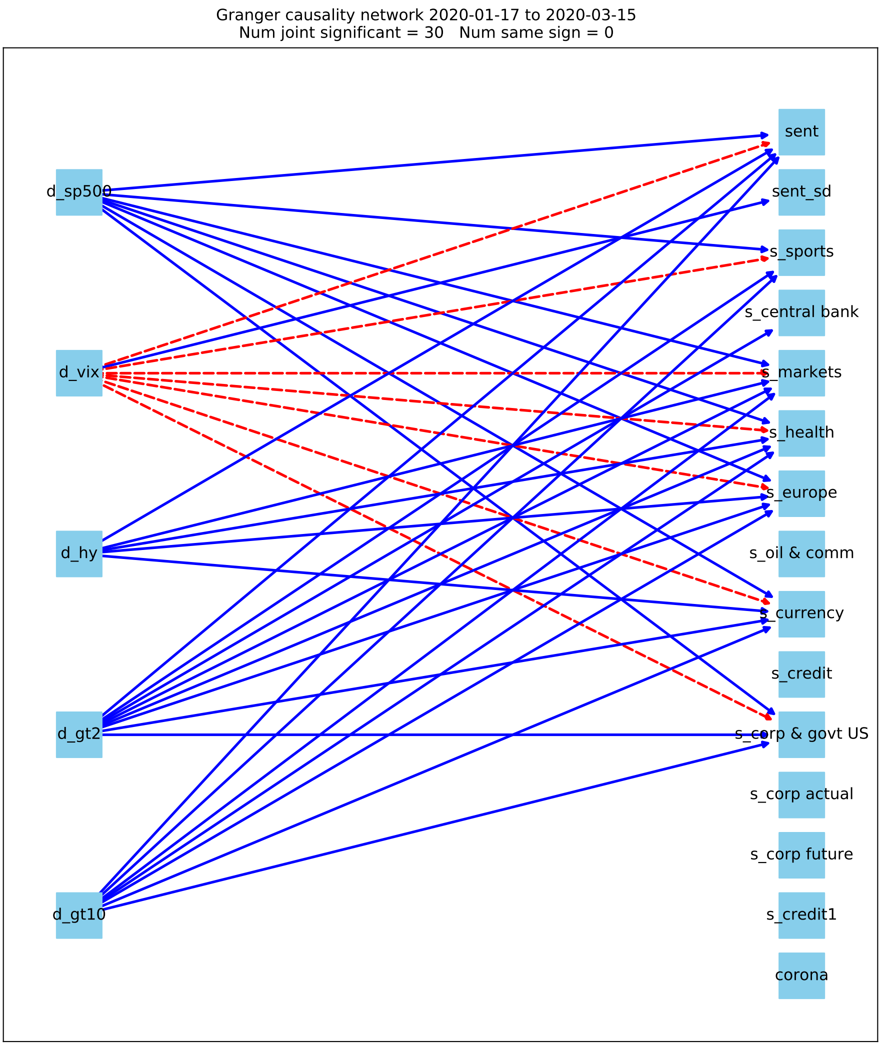


Fig. 8. Granger causality tests, using Newey-West standard errors with three lags. The $d\_[market]$ variables refer to daily returns or changes in the particular market series. A link is shown in the figure if the lagged test variable is significant at the 10% or better level in equation (7). A blue, solid (red, dashed) line indicates the coefficient c_3 from the test to the response variable in (7) is positive (negative). The top of the figure shows the number of links in the graph that are associated with a significant c_4 coefficients from (7), labeled *Num joint significant*. Also shown are the number of times when both c_3 when c_4 are significant and have the same sign, labeled *Num same sign*.

Covid Economics 38, 16 July 2020: 68-128

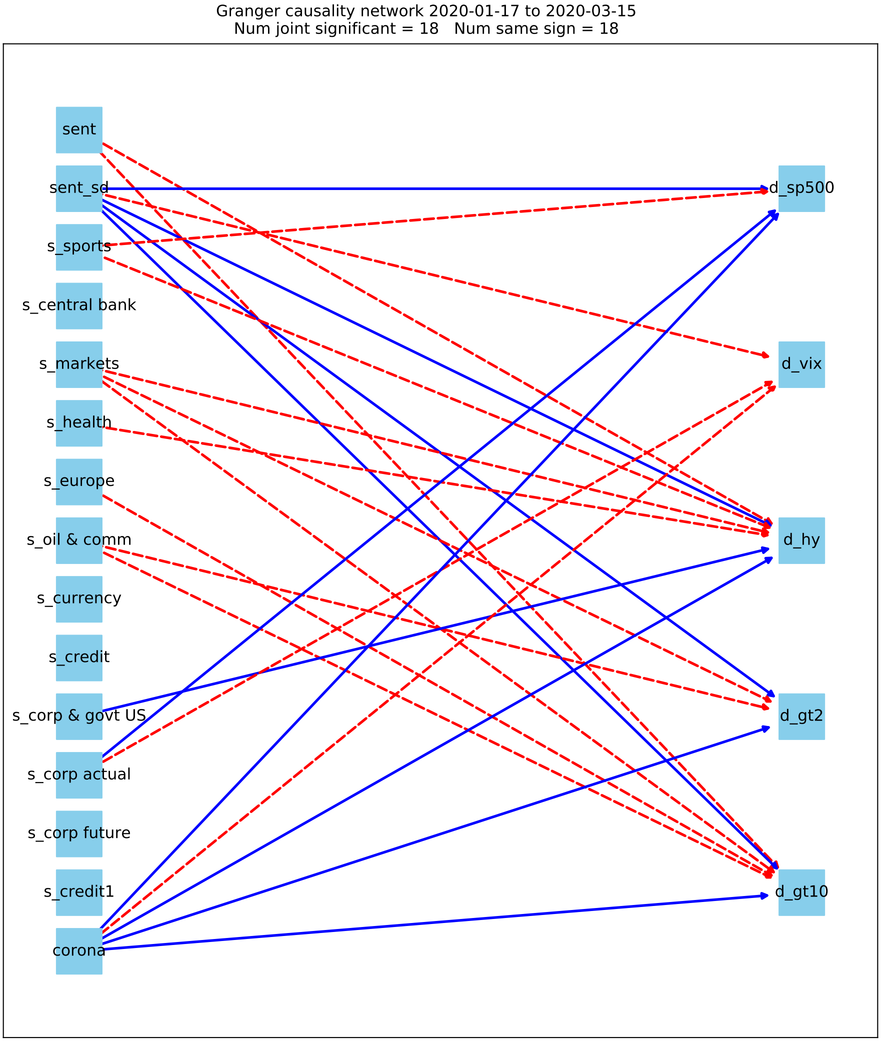


Fig. 9. Granger causality tests, using Newey-West standard errors with three lags. The $d\_[market]$ variables refer to daily returns or changes in the particular market series. A link is shown in the figure if the lagged test variable is significant at the 10% or better level in equation (7). A blue, solid (red, dashed) line indicates the coefficient c_3 from the test to the response variable in (7) is positive (negative). The top of the figure shows the number of links in the graph that are associated with a significant c_4 coefficients from (7), labeled *Num joint significant*. Also shown are the number of times when both c_3 when c_4 are significant and have the same sign, labeled *Num same sign*.

CEPR PRESS

on future news variables is therefore mitigated when the VIX^{10} is high. This suggests an information loss in volatile markets – when markets are extremely volatile, the same unit of market price move becomes less important for future news flow.

Table 12 in the appendix shows Granger causality tests in the other direction for the early subsample. Each row of the table corresponds to one of the fifteen time t news variables, and each column corresponds to one of the five market returns at time $t+1$. Figure 9 shows the graph representation of this table. The arrows in the graph have the same interpretation as before, with blue solid arrows representing significant positive links, and red dashed arrows representing significant negative links. Granger causality from the time t news to time $t+1$ markets is a less dense graph than the markets to news one. Generally the signs of these relationships are opposite to what they were in the contemporaneous regressions. There are many red dashed arrows suggesting that low time t topical sentiment Granger causes positive time $t+1$ market price action. I return to this observation momentarily.

Finally, note that there are 18 cases of jointly significant c_3 and c_4 coefficients in the time t news series to time $t+1$ markets Granger causality tests. In all 18 case, c_3 and c_4 have the same sign. This suggests that hypersensitivity is not only a contemporaneous phenomenon. When time t VIX^{10} is high, the effect of time t news on time $t+1$ markets becomes more pronounced. This observation, and the prior one about the difference in signs between the lagged and contemporaneous news-variable-to-markets relationships, suggest that the lead-lag phenomena of equation (7) are closely related to the contemporaneous relationships of equation (4). Again, I return to this momentarily.

5.2 Late subsample

Table A6 in the Internet Appendix shows the results of estimating (7) for the markets to news Granger causality tests in the late subsample (dates after March 15). Figure 10 shows these results in graph form. The Granger causality network from market prices to news is very sparse in the late part of the sample. Along with the other changes that occur post the mid-March structural break, there is relatively little evidence that markets Granger cause news flow. There are now 6 cases of jointly significant c_3 and c_4 coefficients in (7), down from 30 in the early subsample, and again none of them have the same sign.

Table A7 in the Internet Appendix shows the results of estimating (7) for the news to markets Granger causality tests in the late subsample. Figure 11 shows these results in graph form. The Granger causality network from time t news variables to time $t+1$ market responses is again much more sparse than in the early part of the sample. The

Covid Economics 38, 16 July 2020: 68-128

Covid Economics 38, 16 July 2020: 68-128

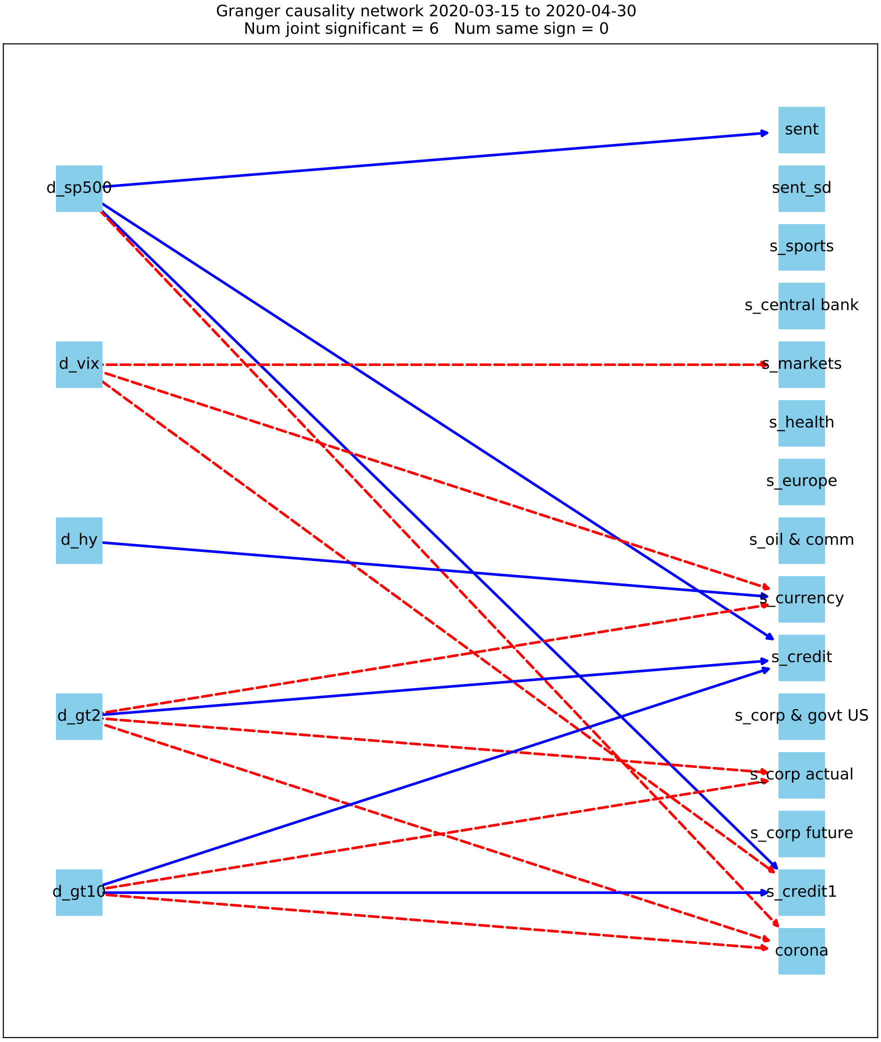


Fig. 10. Granger causality tests, using Newey-West standard errors with three lags. The *d*_[*market*] variables refer to daily returns or changes in the particular market series. A link is shown in the figure if the lagged test variable is significant at the 10% or better level in equation (7). A blue, solid (red, dashed) line indicates the coefficient c_3 from the test to the response variable in (7) is positive (negative). The top of the figure shows the number of links in the graph that are associated with a significant c_4 coefficients from (7), labeled *Num joint significant*. Also shown are the number of times when both c_3 when c_4 are significant and have the same sign, labeled *Num same sign*.

Covid Economics 38, 16 July 2020: 68-128

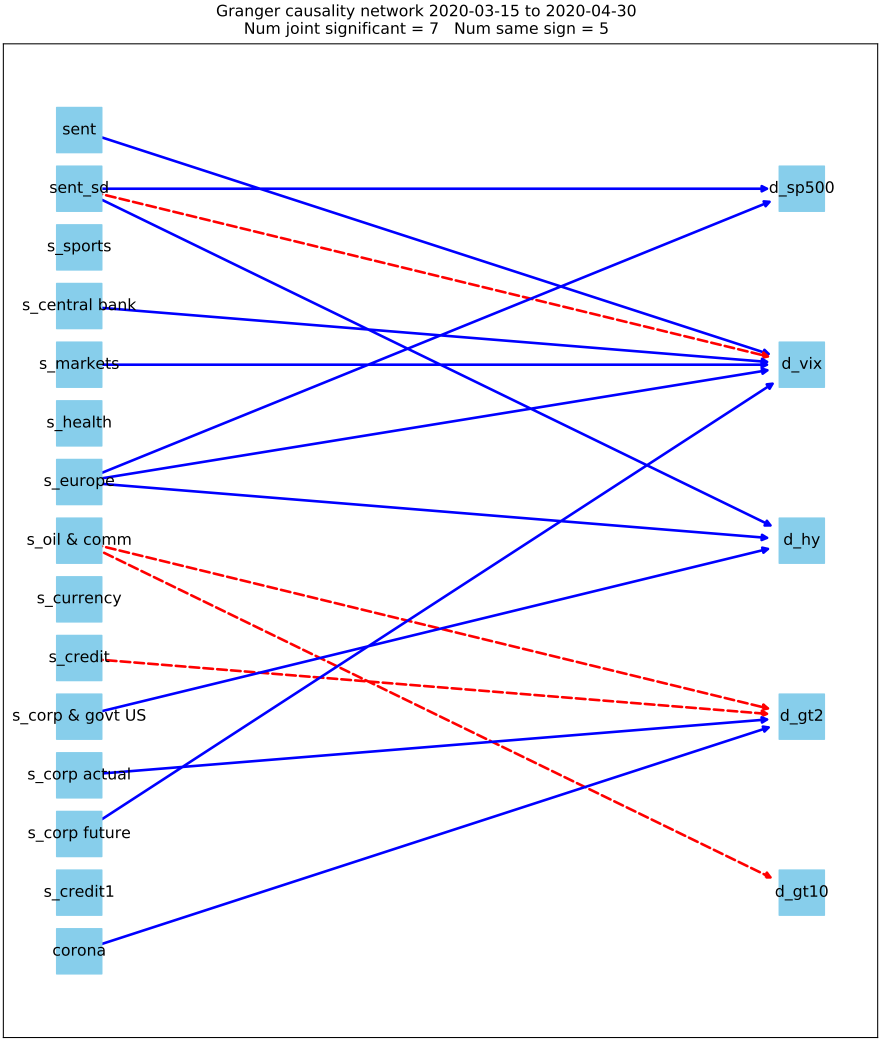


Fig. 11. Granger causality tests, using Newey-West standard errors with three lags. The $d\_[market]$ variables refer to daily returns or changes in the particular market series. A link is shown in the figure if the lagged test variable is significant at the 10% or better level in equation (7). A blue, solid (red, dashed) line indicates the coefficient c_3 from the test to the response variable in (7) is positive (negative). The top of the figure shows the number of links in the graph that are associated with a significant c_4 coefficients from (7), labeled *Num joint significant*. Also shown are the number of times when both c_3 when c_4 are significant and have the same sign, labeled *Num same sign*.

signs of the relationships are mixed, with some instances of positive time t sentiment Granger causing positive time $t+1$ market reactions, as well as some instances where the sign changes. There are now seven jointly significant c_3 and c_4 coefficients, of which five have the same sign, and this compares to the 18 jointly significant relationships in the early part of the sample, of which all 18 had the same sign.

5.3 Feedback loops

Figure 8 makes clear that day t market action has a profound effect on day $t+1$ news coverage of the impact of coronavirus across a broad range of topical categories. These results are reminiscent of Garcia (2018), who shows that future New York Times and Wall Street Journal news are especially responsive to lagged negative stock returns. Given that the price action during the early subsample is very bearish, the early part of the coronavirus crisis is an example of a regime when news articles are extraordinarily sensitive to lagged market performance. The evidence in the late subsample suggests that lead-lag relationships between markets and news are then much less pronounced than in the early subsample; in particular, in Figure 10 news flow becomes much less sensitive to lagged price action, as Garcia's (2018) results suggest.

A consequence of the dense, early subsample networks of Figures 8 and 9 is the possibility of feedback loops, from markets to news back to markets, that may exacerbate volatility. One form of feedback is time $t-1$ market action affecting time t news, which in turn affects time t markets via (4). A second form is from time $t-1$ markets to time t news to time $t+1$ markets. The specification in (7) controls for the possibility that, for example, time $t-1$ SP500 returns Granger cause time t *markets* sentiment, which in turn Granger causes time $t+1$ SP500 returns; such direct loops are precluded by including two lags of the response variable in (7). It is possible, however, that time t *markets* sentiment, having been Granger caused by times $t-1$ SP500 returns, goes on to Granger cause time $t+1$ HY returns. In this way, tight lead-lag networks might create excessive volatility unrelated to COVID-19 case incidence. Recall that case incidence, *corona*, is not Granger caused by any market series in the early subsample, as seen in Figure 8.

The late subsample networks of Figures 10 and 11 are much less dense, and thus less susceptible to such feedback loops.

Covid Economics 38, 16 July 2020: 68-128

CEPR PRESS

Covid Economics 38, 16 July 2020: 68-128

6 Interpretation

Thus far, I have presented evidence that markets react more strongly to news and COVID-19 case counts in high volatility periods. Removing this effect from contemporaneous regressions of market returns on news and case counts meaningfully lowers the R^2, suggesting a counterfactual where the hypersensitivity effect is missing and return variances are lower. This hypersensitivity effect is present in the early subsample, but is largely missing from the late subsample. Furthermore, in tests of lead-lag relationships between market and news variables, I find the Granger networks to be much more dense in the early than the late subsample, and speculate that this is more likely to lead to feedback loops in the early part of the sample.

I now check whether the news series that are associated with contemporaneous market responses are also the news series that are informative, in the sense that they forecast future incidence of COVID-19 cases. I then check whether hypersensitivity and the lead-lag relationships of Section 5 are related. Finally, I propose a theoretical mechanism that can explain part of my findings.

6.1 Hypersensitivity and the information content of news

As I showed in Section 3.1, the market price series are very informative about two to four week ahead incidence of COVID-19 cases, as are several news series. During the early phase of the coronavirus crisis, markets were keenly concerned with understanding the repercussions of COVID-19 incidence for future economic and societal outcomes. Therefore, it would seem natural that markets should have been particularly responsive to news series that would prove to be informative about future incidence of COVID-19. Figure 4 shows that *central bank*, *credit1*, *credit*, *corp & govt US* and *sports* topical sentiment were particularly informative for future case counts, with average lead times of twelve trading days. Were these, therefore, the news topics that were most strongly contemporaneously related to market price moves in equation (4)?

The early subperiod summary results in Table 3 show *credit1* was not significant in any of the five market specifications. The *central bank* and *credit* series were each significant once, for 2-year Treasuries. *corp & govt US* was significant only for high-yield and the *sports* series was significant for all five markets, and interestingly associated with hypersensitivity only once. The other news series that figure prominently in the contemporaneous early subsample regressions are: aggregate sentiment *sent*, article-level sentiment standard deviation *sent_sd*, *markets*, *health*, *europe*, *oil & comm*, and *currency*. However, again looking at Figure 4, with the possible exception of *currency*, all of these news series are particularly poor forecasters of future case outcomes. Furthermore, the vast majority of these news series are associated with hypersensitivity in the early subsample.

The late subsample summary results for equation (4) in Table 4 show that the two credit topics now enter with significant b_3 coefficients in five markets-news pairs, that *sports* enters as significant for three market variables, and *corp & govt US* and *central bank* together have significant b_3 coefficients five times. Many of the significant news series from the early sample remain significant here, though without any clear patterns. Series that are more informative about future COVID-19 incidence play a larger role in the late subsample, where there is little evidence of hypersensitivity, than in the early one. Among the informative news series, *sports* stands out as the one most contemporaneously related to market action in both subsamples.

In the early subsample, hypersensitivity occurred with regard to news series that were not informative about future COVID-19 incidence, though admittedly this would not have been known to market participants at the time. News series that did forecast future outcomes were less important for contemporaneous market movements during the early phase of the crisis. Furthermore the more informative (about future cases) news series in the early subsample tended not to be hypersensitive. It is highly unlikely, therefore, that hypersensitivity effects were due to the high information content of news series in high volatility periods. In fact, the news series that were associated with hypersensitivity were not generally informative about future COVID-19 outcomes. To the extent that such price reaction would prove to be unwarranted, I now check to see whether hypersensitivity is associated with overreaction.

6.2 Hypersensitivity and overreaction

For each market-news pair, the $EVl1$ column of Tables 3 and 4 reports all significant (at the 10% level or better) c_3 coefficients from the Granger causality regressions of future markets on current news in equation (7), along with the significant b_3 or b_4 coefficients from the contemporaneous regression in (4). There are 24 (16) cases of news variables Granger causing next-day market variables, as indicated by a significant entry in the $EVl1$ column, in the early (late) subsample. These numbers are shown in the captions underneath the two tables.

I define *underreaction* as a day $t+1$ market move in response to day t news that goes in the same direction as the day t market response, the implication being that the market did

Covid Economics 38, 16 July 2020: 68-128

not react "enough" in day t to day t news.[12] In the present context, underreaction would occur when b_3 (EV column) and c_3 ($EVl1$ column) have the same sign. This happens once in the early subsample, and four times in the late subsample. Three of these instances involve the HY index, and one each involves the SP500 and 2-year Treasuries. There is relatively little evidence of underreaction.

Overreaction occurs when the contemporaneous news coefficient, b_3 from (4), and the lagged news coefficient, c_3 from (7), have opposite signs. This indicates that the time t market effect of a news variable is partially or fully reversed at time $t+1$. Note the c_3 coefficient reported in column $EVl1$ is scaled in the same way as coefficients b_3 and b_4 in the EV and $EV * VIXl1$ columns; it is reported in standard deviations of the market variable per unit of standard deviation of the news variable. Thus the magnitudes of coefficients in the EV (contemporaneous regression) and $EVl1$ (Granger causality regression) columns of Tables 3 and 4 are directly comparable.

For example, in the early subsample, *sent_sd*'s effect on SP500 contemporaneous return is a decrease of 0.742 standard deviations of SP500 returns for a one standard deviation in increase in *sent_sd*. The coefficient of lagged *sent_sd* for future SP500 returns suggests a one standard deviation increase in *sent_sd* forecasts a next-day positive SP500 return of 1.137 standard deviations. This completely reverses the contemporaneous SP500 reaction to the standard deviation article-level sentiment.

Counting instances of b_3 and c_3 coefficients with opposite signs indicates that this occurs 16 times in the early subsample, and four times in the late subsample (counts shown in captions underneath the tables). So while there is little evidence of underreaction of markets to news, especially in the early subsample, there is strong evidence of overreaction of markets to news in the early subsample, with somewhat weaker evidence of the same effect in the late subsample. Interestingly, of the 16 instances of overreaction in the early subsample, 15 occur in the presence of hypersensitivity in the contemporaneous markets-news relationship. Recall that hypersensitivity in Table 3 occurs when the coefficients b_3 and b_4 from (4), shown in columns *EV* and *EV*VIXl1*, have the same sign. The fact that of the 16 instances of early subsample overreaction, 15 of those instances happen for markets-news pairs that are hypersensitive suggests that hypersensitivity is associated with overreaction. Importantly, in Table 3 there is evidence of overreaction and hypersensitivity not only for markets-news pairs, but also for the markets-*corona* combination. During high-volatility periods, markets experience excessive sensitivity to COVID-19 case counts, which then gets reversed in the next trading day. In the late subsample, none of

[12] Glasserman, Li, and Mamaysky (2020) explores stock-level over- and under-reaction to lagged news.

Covid Economics 38, 16 July 2020: 68-128

the four instances of overreaction is associated with hypersensitivity.

There is little evidence of underreaction during the coronavirus crisis. There is, however, compelling evidence for overreaction. Furthermore, overreaction appears strongly associated with hypersensitivity.

6.3 Theoretical mechanism

The theoretical mechanism proposed in Glasserman, Mamaysky and Shen (GMS 2020) yields important insights into price action around the coronavirus crisis. The paper develops a dynamic, asymmetric information model in the spirit of Grossman and Stiglitz (1980) where the information state is persistent from one time period to the next. An application of the model to the coronavirus crisis would require that agents can acquire, at a cost, information about the pandemic that is not widely available. Given that many professional investors had extensive conversations with medical experts in the early stages of the crisis, where access to those experts was not widely available, supports this assumption. Though in reality, as in theory, there is spillover from this news acquisition to prices. Indeed, as I showed in Section 3.1, market prices have been very informative about future fundamentals, here measured as COVID-19 case counts.

The model yields an economy where markets can be in either a low- or a high-information regime. Markets endogenously transition from one regime to the other in response to information shocks, and each regime is highly persistent.[13] In the high-information regime, a news production sector generates a large volume of news because it anticipates high investor demand for news. In the high-information state, investors understand that tomorrow's news production is likely to be very high, and therefore that tomorrow's prices will be very volatile. To protect against this volatility, investors demand a very high risk premium today, and their high risk sensitivity makes today's prices highly responsive to today's news flow. Therefore when markets enter the high volatility regime, stock prices fall (and risk premia rise), and realized price volatility is very high.

In GMS (2020) such high-volatility, low-price states can arise without any change in fundamentals; in particular, high market price volatility is not associated with high fundamental volatility. This is because, in the high information state, prices become hypersensitive, in the sense that the coefficients in the linear price function become larger

[13] Of course, the world is far more complex than what the model can capture, and aspects of the model are unrealistic. One thing the model gets wrong relative to the present crisis is that the high-information state in the base calibration of the model is much longer-lived than the two-months or so of the hypersensitive regime in the early subsample.

CEPR PRESS

in magnitude and identical supply or dividend shocks have a larger price impact. Clearly, the coronavirus pandemic is characterized by a large change in fundamentals, but the extreme sensitivity of market prices to news in high-VIX[10] periods suggests one dimension of the crisis is the market being in a high-information state. News production today begets news production tomorrow. Prices respond to news as I've shown empirically, and the fear of high news production tomorrow causes prices today to be low and extremely sensitive to today's news flow. The model and empirical results suggest the high volatility experienced by financial markets during the coronavirus pandemic is, at least partially, caused by extreme market responses to news, rather than by extreme news flow itself.

7 Conclusion

I have documented that a twelve topic model captures the news narrative about the coronavirus crisis of 2020, and optimally balances the trade-off between having a higher number of topics and topic coherence. Using structural break tests, I show that the contemporaneous relationship between market returns and news undergoes a regime break around the middle of March 2020. In the early subsample, I show that the contemporaneous effect of news on markets is characterized by a hypersensitive state, where the same quantum of news causes larger market price reactions on high volatility days. Such hypersensitivity is associated with excessive volatility, as well as with market overreaction to contemporaneous news, which gets systematically reversed the next trading day. Finally, while there are several news series that are informative about future fundamentals, as measured by COVID-19 case incidence, these are not the news series that are most associated with contemporaneous price moves. Therefore much of market hypersensitivity to news is largely unwarranted by future fundamentals. However, market prices themselves *are* able to effectively forecast two to four week ahead COVID-19 case counts.

Overall, the hypersensitive market state is not beneficial. It makes prices overly volatile, and is associated with price overreaction to news. Furthermore, the analysis in Glasserman, Mamaysky, and Shen (2020) suggests that hypersensitive markets are associated with low prices and high risk premia, which is also not beneficial to society as a persistently high cost of capital would inhibit investment. The evidence of a structural break in mid-March, right around the time of extensive intervention in markets by the Fed and other central banks, suggests another role central banks can play in the economy. They can nudge markets out of hypersensitive states into more normal ones.

Covid Economics 38, 16 July 2020: 68-128

References

Altman, E., 2020, "COVID-19 and the credit cycle," *Journal of Credit Risk*, forthcoming.

Andrews, D., 1993, "Tests for parameter instability and structural change with unknown change point," *Econometrica*, 61 (4), 821–856.

Andrews, D., 2003, "Tests for parameter instability and structural change with unknown change point: A corrigendum," *Econometrica*, 71 (1), 395–397.

Baker, S., N. Bloom, S. Davis, K. Kost, M. Sammon and T. Viratyosin, 2020, "The unprecedented stock market reaction to COVID-19," *Covid Economics*, 1, 33–42.

Barclay, M. and T. Hendershott, 2004, "Liquidity externalities and adverse selection: Evidence from trading after hours," *Journal of Finance*, 59 (2), 681–710.

Boudoukh, J., R. Feldman, S. Kogan and M. Richardson, 2018, "Information, trading, and volatility: Evidence from firm-specific news," *Review of Financial Studies*, 32 (3), 992-1033.

Calomiris, C. and H. Mamaysky, "How news and its context drive risk and returns around the world," 2019, *Journal of Financial Economics*, 133 (2), 299–336.

Campbell, J., 1991, "A variance decomposition for stock returns," *The Economic Journal*, 101 (405), 157–179.

Campbell, J. and R. Shiller, 1988, "The dividend-price ratio and expectations of future dividends and discount factors," *Review of Financial Studies*, 1 (3), 195–228.

Das, S. and M, Chen, 2007, "Yahoo! for Amazon: Sentiment extraction from small talk on the web," *Management Science*, 53 (9), 1375–1388.

Fair, R., 2002, "Events that shook the market," *Journal of Business*, 75 (4), 713–731.

Garcia, D., 2013, "Sentiment during recessions," *Journal of Finance*, 68 (3), 1267–1300.

Garcia, D., 2018, "The kinks of financial journalism," working paper.

Glasserman, P., H. Mamaysky, and Y. Shen, 2020, "Dynamic information regimes in financial markets," working paper.

Glasserman, P., F. Li, and H. Mamaysky, 2020, "Time variation in the news-returns relationship," working paper.

Gormsen, N. and R. Koijen, 2020, "Coronavirus: Impact on stock prices and growth expectations," working paper.

Covid Economics 38, 16 July 2020: 68-128

Grossman, S. and J. Stiglitz, 1980, "On the impossibility of informationally efficient markets," *The American Economic Review*, 70 (3), 393–408.

Hartley, J. and A. Rebucci, 2020, "An event study of COVID-19 central bank quantitative easing in advanced and emerging economies," *NBER working paper 27339*.

Ke, S., J.L. Montiel Olea, and J. Nesbit, 2020, "A robust machine learning algorithm for text analysis," working paper.

Loughran, T. and B. McDonald, 2011, "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks," *Journal of Finance*, 66, 35–65.

Loughran, T. and B. McDonald, 2020, "Management disclosure of risk factors and COVID-19," working paper.

Michal, R.-Z., T. Griffiths, M. Steyvers, and P. Smyth, 2004, "The author-topic model for authors and documents," *Proceedings of the 20th conference on Uncertainty in artificial intelligence*, 487–494.

Newman, D., J. Han Lau, K. Grieser, and T. Baldwin, 2010, "Automatic evaluation of topic coherence," *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the ACL*, 100–108.

Roll, R., 1988, "R2", *Journal of Finance*, 43 (3), 541–566.

Steyvers, M. and T. Griffiths, 2007, "Probabilistic topic models," in T. K. Landauer, D. S. McNamara, S. Dennis, and W. Kintsch (Eds.), *Handbook of Latent Semantic Analysis*, pp. 427-448, Lawrence Erlbaum Associates Publishers.

Tetlock, P., 2007, "Giving content to investor sentiment: The role of media in the stock market," *Journal of Finance*, 62, 1139–1168.

Tetlock, P., M. Saar-Tsechansky, and S. Macskassy, 2008, "More than words: Quantifying language to measure firms' fundamentals," *Journal of Finance*, 63 (3), 1437–1467.

A Appendix

A.1 Derivation of equation (3)

Campbell (1991) shows that h_{t+1} can be approximated as

$$h_{t+1} \approx c + (E_{t+1} - E_t) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} - (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j h_{t+1+j}, \tag{8}$$

Covid Economics 38, 16 July 2020: 68-128

where $\rho < 1$ is a constant discounting factor, c equals $E_t h_{t+1}$ which is assumed to be approximately constant, d_t is the time t log dividend, E_t is an expectation taken over the investor's information set, and the change-in-beliefs operator $(E_{t+1} - E_t)X$ is shorthand for $E_{t+1}X - E_tX$.

As in Campbell (1991), assume that an M-dimensional state vector z_t follows a VAR process

$$z_{t+1} = Az_t + \epsilon_{t+1}.$$

The one-lag specification is without loss of generality since z_t can be augmented with lagged state variables. Let the first two elements of z_t be the dividend growth Δd_t and the return h_t, with the remaining elements representing other information useful for forecasting future returns and dividends. The change in beliefs from t to $t+1$ about z_{t+T} is given by

$$(E_{t+1} - E_t)z_{t+T} = (A^{T-1}z_{t+1} - A^T z_t) = A^{T-1}\epsilon_{t+1}.$$

Therefore h_{t+1} from (8) can be written as

$$h_{t+1} \approx c + e1^\top \epsilon_{t+1} + e2^\top(\rho A + \rho^2 A^2 + \cdots)\epsilon_{t+1} = c + \left(e1^\top + e2^\top \rho A(I - \rho A)^{-1}\right)\epsilon_{t+1} \quad (9)$$

where I is the $M \times M$ identity matrix, $e1$ is an $M \times 1$ vector with a 1 in the first element, and zeros in all the others, and $e2$ is an $M \times 1$ vector with a 1 in the first element, -1 in the second, and zeros everywhere else. The linear combination of ϵ_{t+1} in (9) can be decomposed into a part that loads on w_{t+1}, the information set of the econometrician, and an orthogonal part e_{t+1}. With this, equation (3) follows.

A.2 Robustness

Sample size and persistence

The present analysis, unavoidably, involves a small number of daily observations. The main research question is how markets and news interact during the coronavirus crisis, and the crisis is (hopefully) a short-lived episode. The small sample size is thus a fundamental constraint of the setting. An issue is that some of the regressors, for example *corona*, are very persistent. All of the paper's regressions use daily data, and none of the observations are overlapping. The persistence therefore is not mechanical, but rather a time-series property of the underlying data. To control for this, I use two lags of the dependent variables in the contemporaneous specification in (4), and also use the same two lags in the lead-lag specification in (7). I also tried running the regression in (4) in first differences, thus reducing persistence in all variables, and the results were not materially different. Of course, all standard errors also use the Newey-West HAC estimator.

A mitigating consideration is that I run all the analysis separately for the early- and late-parts of the sample. The small-sample and persistence concerns are present in both parts of the sample, and yet the early- and late-subsample results are fundamentally different. This difference supports the assertion that neither set of results obtains from small-sample biases.

Furthermore, I run all the paper's results at the level of 75 markets-news pairs. There

Covid Economics 38, 16 July 2020: 68-128

are important differences between results for different market-news pairs, and yet all specifications are subject to the same econometric issues. The differences in markets-news relationships across the 75 such combinations (5 market variables and 15 news variables) in the contemporaneous and lead-lag regressions suggests that results are capturing fundamental aspects of the data, and are not influenced by a specific bias. Finally, the fact that all 32 jointly significant b_3 and b_4 coefficients in the early subsample (Table 3) have the same sign, and that only four of the 12 jointly significant b_3 and b_4 coefficients in the late subsample do (Table 4), is highly unlikely to stem from small sample issues.

Endogeneity

The contemporaneous relationship in (4) is not necessarily indicative of a causal relationship from news to market prices, though this is the likeliest explanation. It is possible that news articles are written in response to market prices, in which case news flow clearly does not cause price changes. This concern is most pronounced for topics explicitly discussing market prices, such as the *markets* topic. However, there are several mitigating considerations. First, other hypersensitive markets-news pairs are explicitly *not* about financial markets. For example, the *sports* and *health* topics are concerned with issues other than market price action, as evidenced by the headlines in Table 5. Second, *sent_sd* is not in any obvious way related to contemporaneous market price action, and indeed reflects the dispersion of sentiment across hundreds or thousands of daily articles, most of which are not about market activity. And *sent_sd* comes in as an important market-news pair for SP500, HY, and VIX. Third, the *corona* series, which is obviously fully exogenous from contemporaneous market prices, also comes in as a significant – and hypersensitive – markets-news pair for the SP500, HY, and VIX indexes.

Furthermore my analysis leans heavily on the interaction of the b_3 and b_4 coefficients from (4) and there is no reason why reverse causality would induce the two coefficients to always have the same sign in the early subsample, and to never do so in the late subsample. And the two subsamples are subject to the same endogeneity considerations.

Furthermore, none of the analysis involving the lead-lag specification in (7) is subject to this concern, because here the explanatory variables are lagged by one-day relative to the dependent variables. In the news to markets versions of these regressions, in the early subsample, out of the 18 cases of significant c_3 and c_4 coefficients, the two always have the same sign (see Figure 9 and Table 12). But in the late subsample, there are only 5 news-markets pairs that evidence hypersensitivity (see Figure 11 and Table A7 in the Internet Appendix.) Therefore, hypersensitivity is present in the lead-lag specification in (7) which is free of any endogeneity concerns.

Also the finding in Section 6.2 that hypersensitivity is closely associated with overreaction does not follow in any mechanical way from potential endogeneity in (4).

Dropping intraday articles

To further address endogeneity concerns with regard to the contemporaneous specification in (4), I rerun all the analysis in the paper but exclude any non-weekend intraday news articles. These articles are defined as Monday through Friday news that come out between 9:30am and 4pm NY time; I refer to the complement of this set as the *overnight* articles. I recalculate $Sent_t$, *sent_sd*, and each of the twelve topical sentiment measures using only the overnight articles. The topical sentiment series use the same topic model described in Section 2, and which is estimated using all articles in the corpus. Of course, the *corona* series is unaffected by this.

Barclay and Hendershott (2004) show that after hours trading (from 4-6:30pm and then from 8-9:30am) contains "less than 1/20 as many trades per unit time" as does trading during regular market hours. This vastly diminishes the probability that overnight news stories are written in response to contemporaneous market action. Of course, news stories that reflect past market action are not problematic. Figure 1 shows each of the overnight series as the orange, dashed line. As can be seen, the overnight version of the series are very close to the all-article versions. The number of articles, the *count* series from the figure, is of course lower for the overnight articles.

The Internet Appendix shows the results of this analysis. Figure A1 shows the structural break tests for the overnight news series. Figures A2 and A3 show the R^2 model summaries discussed in Section 4.5. Tables A8 and A9 summarize the results of (4) for the markets-overnight news pairs. Tables A10 – A14 show the results for the early subsample using the overnight news measures. Tables A15 – A19 show the results from the same analysis in the late subsample. In all cases, these results are qualitatively very similar to those that use the all-article version of the news series. The small differences that arise likely stem from excluding intraday articles that are not endogenous, but informative. I conclude endogeneity is unlikely to have an effect on the main takeaways of the paper.

A.3 Additional results

Covid Economics 38, 16 July 2020: 68-128

Covid Economics 38, 16 July 2020: 68-128

Table 5
Headlines of extreme sentiment (top two most negative and positive) news stories whose document-topic weight for a given topic is above 0.70.

Headlines for representative articles by topic

label	headline	sent	Date
sports	Rugby-Champions Cup, Challenge Cup quarter-finals postponed due to coronavirus	-0.119048	2020-03-16
sports	Olympics-Judo qualifiers scrapped through end-April over coronavirus fears	-0.108108	2020-03-09
sports	Rugby-Ireland's Healy to miss rest of Six Nations with hip injury	-0.099237	2020-02-26
sports	Soccer-Barcelona directors quit, throwing club into crisis	-0.096774	2020-04-10
sports	Tom Hanks' son confident dad will make full recovery	0.047297	2020-03-12
sports	UPDATE 1-Soccer-Juventus's Dybala tests positive for coronavirus	0.051095	2020-03-21
sports	Soccer-Juve's France midfielder Matuidi tests positive for coronavirus - club	0.062500	2020-03-17
sports	Soccer-Premier League convenes emergency club meeting after Arteta positive test	0.090909	2020-03-12
central bank	Insurers warns on forced payouts for uncovered coronavirus losses	-0.169014	2020-04-06
central bank	Bank of England cancels annual stress test of banks	-0.112245	2020-03-20
central bank	Much more serious than the 2008 crisis, former BoE Governor King says	-0.104895	2020-03-23
central bank	Hungary govt expands tax relief measures, suspends evictions - PM Orban	-0.103448	2020-03-23
central bank	Kuwait central bank ready to take all necessary measures to ensure stability	0.041667	2020-03-30
central bank	BRIEF-Gulf Central Banks Governors Say Regional Financial Sector Strong Enough to Face Coronavirus - Wam	0.042254	2020-03-16
central bank	Swedish central banker tests positive for new coronavirus	0.050000	2020-03-09
central bank	BRIEF-Canada's move to reduce drop in revenue requirement for businesses to claim emergency wage subsidy will only be valid in March	0.058824	2020-04-08
markets	U.S. stock futures tumble at open on coronavirus contagion fears	-0.089552	2020-03-08
markets	BUZZ-Australian tech stocks drop most in over 4 years on virus fears	-0.082569	2020-03-08
markets	Australia shares likely to dip at open, NZ up	-0.080000	2020-04-07
markets	China shares swoon, Hong Kong plummets amid global virus panic	-0.078431	2020-03-12
markets	BUZZ-Australian tech stocks hit record high as U.S. peers surge	0.041237	2020-02-04
markets	CANADA STOCKS-TSX opens higher on reports of effective coronavirus drugs	0.045455	2020-02-05
markets	BUZZ-Aussie energy index sees best day in over a month as oil prices jump	0.061224	2020-02-05
markets	BUZZ-Australia's gold stocks hit near 2-wk peak as virus fears boost safe-haven demand	0.062500	2020-02-10
health	Vietnam arrests seven over procurement 'wrongdoings' during coronavirus crisis	-0.171429	2020-04-22
health	Coronavirus crisis delays opening of Netanyahu trial	-0.140625	2020-03-15
health	Kremlin denies EU allegations Russia is feeding coronavirus disinformation campaign	-0.134615	2020-03-18
health	Brazil court OKs investigating allegations Bolsonaro tried to interfere with police	-0.134259	2020-04-27
health	Britain making good progress with antibody tests - junior minister	0.068966	2020-04-27
health	PM Johnson fully able to run coronavirus response despite positive test - spokesman	0.069767	2020-03-30
health	BRIEF-Positive Coronavirus Case Detected In India's Capital New Delhi-Health Ministry Statement	0.071429	2020-03-02
health	UK PM making very good progress in COVID-19 recovery, office says	0.076923	2020-04-11
europe	BRIEF-Coronavirus Is Not A Major Concern For British Insurer Rsa CEO	-0.094340	2020-02-27
europe	Reuters Insider - Global auto industry threatened by virus crisis	-0.092105	2020-02-07
europe	Reuters Insider - Medic unions say Spain ignored doctors' warnings	-0.090909	2020-04-02
europe	Reuters Insider - Lebanese protesters and soldiers clash in Tripoli	-0.090909	2020-04-30

Covid Economics 38, 16 July 2020: 68-128

Table 5

Headlines of extreme sentiment (top two most negative and positive) news stories whose document-topic weight for a given topic is above 0.70.

label	headline	sent	Date
europe	Reuters Insider - Doctor who showed Putin round hospital tests positive	0.035714	2020-03-31
europe	Reuters Insider - Reuters Today: European shares open lower, Tesco says coronavirus costs could top £925 million	0.036145	2020-04-08
europe	LIVE MARKETS-Morning call: Positive thinking	0.038835	2020-04-09
europe	LIVE MARKETS-Morning call: Clinging to the rebound	0.041379	2020-01-29
oil & comm	France's Total rejects force majeure notice from Chinese LNG buyer	-0.081395	2020-02-06
oil & comm	UPDATE 1-Speculators cut U.S. crude oil net longs-CFTC	-0.080292	2020-01-31
oil & comm	FACTBOX-U.S. ethanol plants slash output as coronavirus hits fuel demand	-0.078740	2020-03-26
oil & comm	Marathon Petroleum to idle 26,000-bpd Gallup refinery - source	-0.070423	2020-04-08
oil & comm	CNPC to boost Xinjiang oilfield daily crude oil output to 36,000 T by end-June	0.019608	2020-03-25
oil & comm	U.S. wheat futures rise with demand for bread strong	0.020513	2020-03-18
oil & comm	CBOT wheat closes firm on strong demand	0.028571	2020-03-20
oil & comm	GRAINS-Soybeans, wheat climb on currency; corn follows crude oil	0.028736	2020-04-09
currency	BUZZ-US jobless claims back in focus as warning sign, Philly Fed tumbles	-0.121212	2020-03-19
currency	BUZZ-COMMENT-Euro suppressing factors are now weaker	-0.111111	2020-03-05
currency	FOREX-Dollar erases early losses on global economy worries	-0.099237	2020-04-15
currency	BUZZ-COMMENT-AUD/JPY vulnerable as coronavirus fears escalate	-0.098266	2020-03-15
currency	BUZZ-AUD/USD-Moves higher after better than expected Aus GDP	0.041667	2020-03-03
currency	BUZZ-COMMENT-It's good to think positively, better to play safe	0.042017	2020-04-02
currency	REFILE-BUZZ-EUR/USD-Bears maintain their grip despite risk-on	0.046512	2020-04-06
currency	Turkish lira firms 0.7% to strongest level in a week	0.068493	2020-03-25
credit	Fitch Downgrades South African Banks to 'BB', Negative Outlook on Coronavirus Impact	-0.047221	2020-03-31
credit	Fitch Takes Actions on Colombian FIs & Related Entities After Sovereign Downgrade	-0.046079	2020-04-08
credit	Fitch Takes Action on 14 Italian Banking Groups On Coronavirus Disruption	-0.042446	2020-03-24
credit	Fitch Takes Actions on 13 Mexican Banks Due to Operating Environment Deterioration	-0.036833	2020-04-21
credit	Fitch Rates Texas Instruments' $750 Million of Five-Year Senior Notes 'A+'	0.010546	2020-03-03
credit	Fitch Rates KLA Corp.'s Senior Notes Offering 'BBB+'	0.011581	2020-02-19
credit	Fitch Affirms Sri Lanka's Dialog Axiata at 'AAA(lka)'; Outlook Stable	0.012177	2020-03-31
credit	Fitch Affirms L'Oreal at 'F1+'; Withdraws Ratings	0.012441	2020-02-27
corp & govt US	BRIEF-HHS Launches Covid-19 Uninsured Program Portal	-0.113208	2020-04-27
corp & govt US	U.S. House Speaker Pelosi calls Trump WHO decision senseless, dangerous	-0.111111	2020-04-15
corp & govt US	Pelosi calls on Trump to speed production of critical medical equipment	-0.111111	2020-03-19
corp & govt US	BRIEF-FDA & FTC Warn 7 Companies Selling Fraudulent Products That Claim To Treat Or Prevent COVID-19	-0.106383	2020-03-09
corp & govt US	BRIEF-CEPI Collaborates With The Institut Pasteur In A Consortium To Develop Covid-19 Vaccine	0.047619	2020-03-19
corp & govt US	BUZZ-Borqs set for best day ever after receiving $150 mln line of credit	0.053763	2020-04-13
corp & govt US	Washington, D.C., has first 'presumptive positive' test for coronavirus	0.055556	2020-03-07
corp & govt US	BRIEF-GSK Announces Collaboration With CEPI To Develop Coronavirus Vaccine	0.061224	2020-02-02

CEPR PRESS

Covid Economics 38, 16 July 2020: 68-128

Table 5

Headlines of extreme sentiment (top two most negative and positive) news stories whose document-topic weight for a given topic is above 0.70.

label	headline	sent	Date
corp actual	BRIEF-Galaxy Gaming Suspends Billing During Casino Closures	-0.151515	2020-03-16
corp actual	BRIEF-SilkAir Says To Suspend Operations To Hiroshima	-0.147059	2020-03-03
corp actual	BRIEF-Portland General Electric Says Suspending Disconnections, Late Fees For Customers During Coronavirus Crisis	-0.142857	2020-04-06
corp actual	Australia Stock Exchange release from FLETBUILD <FBU.AX>: Covid-19 update cancels interim dividend suspends buyback	-0.133333	2020-03-24
corp actual	BRIEF-Luna Innovations Provides Business Update On Covid-19	0.108108	2020-04-14
corp actual	BRIEF-Inspiration Health Says 15% Rev Growth To £17.8 Mln In FY	0.123967	2020-04-21
corp actual	Australia Stock Exchange release from SOMNOMED <SOM.AX>: Satisfactory Q3 despite COVID-19 affected March	0.142857	2020-04-26
corp actual	Australia Stock Exchange release from FRONTIER DIGITAL <FDV.AX>: Improved profitability and balance sheet despite COVID-19	0.200000	2020-04-29
corp future	UPDATE 1-German investor morale worsens on coronavirus fears - ZEW	-0.095000	2020-02-18
corp future	BUZZ-Hershey: Falls on Q1 profit miss, massive sales decline in China	-0.091954	2020-04-23
corp future	U.S. weekly jobless claims surge to a record 3.28 mln as coronavirus spurs mass layoffs	-0.087912	2020-03-26
corp future	Italy think-tank sees recession in Q1 over coronavirus	-0.084507	2020-02-27
corp future	BUZZ-Nvidia down slightly after upbeat Susquehanna report	0.053191	2020-03-31
corp future	BUZZ-Analog Devices: Up as brokerages remain positive on future growth	0.055046	2020-02-20
corp future	BUZZ-Micron lifts as Raymond James upgrades to "strong buy" on demand growth	0.055172	2020-02-06
corp future	BUZZ-Brokerage sees Advanced Micro Devices gaining share from Intel, upgrades AMD	0.057143	2020-03-03
credit1	Fitch Downgrades ClearBridge Funds' Notes to 'A' & MRPS to 'BBB'; Rating Watch Negative	-0.047769	2020-03-17
credit1	Fitch Downgrades Kayne Anderson Closed End Funds' Notes to 'A' & MRPS to 'BBB'; Rating Watch Neg.	-0.043635	2020-03-17
credit1	Fitch Ratings Defines Coronavirus Scenarios for U.S. Toll Roads	-0.040597	2020-03-24
credit1	Fitch Ratings: USPF Housing Defines Coronavirus Scenarios for Loan Program Models	-0.040170	2020-04-30
credit1	Fitch Affirms Jackson County Schneck Memorial Hospital's (IN) Revs at 'AA-'; Outlook Stable	0.011084	2020-03-20
credit1	Fitch Affirms Franciscan Communities, Inc. (IL) at 'BBB-'; Outlook Remains Positive	0.013284	2020-03-19
credit1	Fitch Affirms FirstHealth of the Carolinas, NC's Bonds at 'AA'; Outlook Stable	0.014033	2020-03-24
credit1	Fitch Affirms Chubb's 'AA' IFS Ratings; Outlook Stable	0.014760	2020-04-03

Covid Economics 38, 16 July 2020: 68-128

Table 6
Contemporaneous regression of daily changes in SP500 index on text variables and *corona*. Rows correspond to each explanatory variable for which either the b_3 or b_4 coefficient in (4) is significant at the 10% level or better. *EV* column shows the impact of a one standard deviation change in the explanatory variable in units of standard deviation of the SP500 index. *EV*VIXl1* column shows the impact of a unit increase in VIX^{10} on the value of *EV*. The $VIXl1$ column shows the effect of a one standard deviation increase in VIX^{10} in units of standard deviation of the market variable. Numbers in parentheses are t-statitics, and the numbers in square brackets underneath the adjusted R^2s are joint F-tests. Both sets of tests use Newey-West standard errors with 3 lags. The *R2* column shows adjusted R^2s, and *Chg R2* shows the difference in adjusted R^2s between the specification in (4) and a version of (4) which sets $b_4 = 0$. The *Break Date* column shows the structural break date of (4) using the procedure from Section 4.2, and the maximal Chow statistic is shown in square backets, with siginficance levels obtained from the Andrews (1993, 2003) distribution indicated via the *s. Significance at the 10%, 5%, or 1% levels is indicated via *,**,*** respectively.

Contemp. changes in SP500 index: early subperiod, VIX interact

	const	h_t	h_{t-1}	EV	EV*VIXl1	VIXl1	R2	Chg R2	Break Date
sent	-23.464***	-0.745***	-0.106	0.761***	0.085***	5.595***	0.521	0.176	2020-03-12
	(-2.71)	(-4.56)	(-0.64)	(5.72)	(4.43)	(4.16)	[0.000]		[22.15**]
sent_sd	-41.755	-0.743**	-0.215	-0.742**	-0.106**	8.795**	0.194	0.061	2020-03-23
	(-1.60)	(-2.27)	(-0.87)	(-2.21)	(-2.24)	(2.24)	[0.032]		[23.94***]
s_sports	-3.391	-0.733***	-0.327	0.940**	0.023	0.953***	0.231	0.002	2020-03-23
	(-1.34)	(-3.16)	(-1.42)	(2.54)	(1.17)	(2.81)	[0.001]		[14.93]
s_markets	-0.774	-0.698***	-0.239	0.857***	0.075***	1.087**	0.324	0.107	2020-03-23
	(-0.15)	(-3.08)	(-0.96)	(3.72)	(4.85)	(2.18)	[0.000]		[20.64**]
s_health	-4.113	-0.693***	-0.321	1.053***	0.104**	1.151**	0.293	0.062	2020-03-04
	(-1.35)	(-2.82)	(-1.33)	(2.58)	(2.41)	(2.10)	[0.048]		[12.79]
s_europe	-10.194	-0.498***	-0.330	0.985***	0.078*	2.817*	0.390	0.066	2020-03-23
	(-1.03)	(-3.57)	(-1.61)	(5.03)	(1.86)	(1.78)	[0.000]		[18.43*]
s_oil & comm	-29.775***	-0.321*	0.011	0.663***	0.215***	4.728***	0.536	0.415	2020-03-23
	(-4.11)	(-1.94)	(0.06)	(4.05)	(4.83)	(4.53)	[0.000]		[33.80***]
s_currency	2.197	-0.541***	-0.088	0.771**	0.049	0.499	0.238	0.009	2020-03-23
	(0.46)	(-3.31)	(-0.32)	(2.48)	(1.05)	(0.60)	[0.001]		[13.51]
corona	-6.182***	-0.428***	-0.048	-0.187**	-0.083***	0.883***	0.419	0.314	2020-03-13
	(-2.98)	(-2.92)	(-0.39)	(-2.03)	(-7.09)	(2.98)	[0.000]		[21.50**]

Covid Economics 38, 16 July 2020: 68-128

Table 7

Contemporaneous regression of daily changes in VIX index on text variables and *corona*. Rows correspond to each explanatory variable for which either the b_3 or b_4 coefficient in (4) is significant at the 10% level or better. *EV* column shows the impact of a one standard deviation change in the explanatory variable in units of standard deviation of the VIX index. *EV*VIXl1* column shows the impact of a unit increase in VIX^{10} on the value of EV. The $VIXl1$ column shows the effect of a one standard deviation increase in VIX^{10} in units of standard deviation of the market variable. Numbers in parentheses are t-statitics, and the numbers in square brackets underneath the adjusted R^2s are joint F-tests. Both sets of tests use Newey-West standard errors with 3 lags. The *R2* column shows adjusted R^2s, and *Chg R2* shows the difference in adjusted R^2s between the specification in (4) and a version of (4) which sets $b_4 = 0$. The *Break Date* column shows the structural break date of (4) using the procedure from Section 4.2, and the maximal Chow statistic is shown in square backets, with siginficance levels obtained from the Andrews (1993, 2003) distribution indicated via the *s. Significance at the 10%, 5%, or 1% levels is indicated via *,**,*** respectively.

Contemp. changes in VIX index: early subperiod, VIX interact

	const	h_t	h_{t-1}	EV	EV*VIXl1	VIXl1	R2	Chg R2	Break Date
sent	42.398***	-0.768***	-0.206	-0.701***	-0.081***	-5.227***	0.579	0.171	2020-03-16
	(2.84)	(-5.43)	(-1.52)	(-5.22)	(-4.82)	(-4.42)	[0.000]		[54.29***]
sent_sd	84.147**	-0.825***	-0.245	0.857***	0.117***	-9.662***	0.326	0.089	2020-03-16
	(2.23)	(-2.59)	(-1.23)	(2.64)	(3.57)	(-3.62)	[0.001]		[28.04***]
s_sports	7.050*	-0.810***	-0.438**	-1.041***	-0.025	-1.003***	0.368	0.014	2020-03-16
	(1.69)	(-3.33)	(-2.18)	(-3.23)	(-1.59)	(-3.30)	[0.001]		[20.25**]
s_markets	0.772	-0.737***	-0.357*	-0.798***	-0.069***	-0.892**	0.393	0.093	2020-03-16
	(0.09)	(-3.44)	(-1.66)	(-3.56)	(-4.93)	(-2.08)	[0.000]		[24.65***]
s_health	8.645**	-0.752***	-0.451**	-1.099***	-0.113***	-1.181***	0.410	0.080	2020-03-16
	(2.18)	(-3.20)	(-2.26)	(-3.38)	(-3.91)	(-3.45)	[0.016]		[43.62***]
s_europe	14.929	-0.563***	-0.347	-0.829***	-0.064	-2.205	0.397	0.037	2020-03-11
	(0.82)	(-3.64)	(-1.62)	(-4.54)	(-1.56)	(-1.45)	[0.000]		[17.66*]
s_oil & comm	50.372***	-0.435***	-0.167	-0.460***	-0.187***	-4.011***	0.526	0.325	2020-03-13
	(3.53)	(-2.60)	(-0.94)	(-2.63)	(-4.05)	(-3.77)	[0.000]		[36.76***]
s_currency	-5.863	-0.621***	-0.216	-0.681**	-0.038	-0.204	0.301	-0.000	2020-03-17
	(-0.83)	(-3.57)	(-0.93)	(-2.15)	(-0.96)	(-0.32)	[0.006]		[16.74*]
corona	10.840***	-0.432***	-0.153	0.163*	0.079***	-0.769***	0.476	0.276	2020-03-16
	(3.11)	(-3.32)	(-1.60)	(1.73)	(6.79)	(-2.89)	[0.000]		[40.40***]

CEPR PRESS

Table 8
Contemporaneous regression of daily changes in US high-yield index on text variables and *corona*. Rows correspond to each explanatory variable for which either the b_3 or b_4 coefficient in (4) is significant at the 10% level or better. *EV* column shows the impact of a one standard deviation change in the explanatory variable in units of standard deviation of the US high-yield index. *EV*VIXl1* column shows the impact of a unit increase in VIX^{10} on the value of *EV*. The $VIXl1$ column shows the effect of a one standard deviation increase in VIX^{10} in units of standard deviation of the market variable. Numbers in parentheses are t-statitics, and the numbers in square brackets underneath the adjusted R^2s are joint F-tests. Both sets of tests use Newey-West standard errors with 3 lags. The *R2* column shows adjusted R^2s, and *Chg R2* shows the difference in adjusted R^2s between the specification in (4) and a version of (4) which sets $b_4 = 0$. The *Break Date* column shows the structural break date of (4) using the procedure from Section 4.2, and the maximal Chow statistic is shown in square backets, with siginficance levels obtained from the Andrews (1993, 2003) distribution indicated via the *s. Significance at the 10%, 5%, or 1% levels is indicated via *,**,*** respectively.

Contemp. changes in US high-yield index: early subperiod, VIX interact

	const	h_t	h_{t-1}	EV	EV*VIXl1	VIXl1	R2	Chg R2	Break Date
sent	-9.038***	-0.184	-0.002	0.603***	0.104***	6.551***	0.542	0.272	2020-03-04
	(-5.49)	(-1.13)	(-0.03)	(10.65)	(7.38)	(6.94)	[0.000]		[41.69***]
sent_sd	-19.528**	-0.669	0.063	-1.100**	-0.180***	14.641**	0.329	0.162	2020-03-19
	(-2.57)	(-1.47)	(0.51)	(-2.49)	(-2.59)	(2.57)	[0.011]		[35.46***]
s_sports	-0.646	-0.299	-0.154**	0.645**	0.035*	0.663	0.272	0.049	2020-03-23
	(-0.79)	(-1.38)	(-2.08)	(2.46)	(1.66)	(1.24)	[0.001]		[14.86]
s_markets	-0.555	-0.374*	-0.048	1.079***	0.110***	1.515***	0.538	0.254	2020-03-23
	(-0.66)	(-1.72)	(-0.42)	(5.49)	(5.45)	(3.26)	[0.000]		[31.05***]
s_health	-2.147*	-0.313	-0.117	1.245***	0.171***	1.812**	0.493	0.194	2020-03-19
	(-1.84)	(-1.54)	(-1.19)	(3.80)	(3.18)	(2.40)	[0.000]		[25.82***]
s_europe	-3.948*	-0.146	-0.436***	0.676***	0.101***	2.995**	0.336	0.122	2020-03-23
	(-1.87)	(-0.64)	(-5.68)	(4.40)	(2.99)	(2.20)	[0.000]		[27.21***]
s_oil & comm	-6.551***	0.060	0.126	1.073***	0.223***	4.565***	0.659	0.490	2020-03-23
	(-5.93)	(0.55)	(1.43)	(6.36)	(6.91)	(7.01)	[0.000]		[39.53***]
s_corp & govt US	-0.741	-0.171	-0.194*	0.338**	0.031	0.743	0.178	-0.003	2020-03-23
	(-0.46)	(-0.60)	(-1.67)	(2.01)	(1.26)	(0.72)	[0.040]		[22.63**]
corona	-0.697	-0.089	-0.219***	-0.222***	-0.076***	0.370	0.418	0.251	2020-03-23
	(-1.45)	(-0.49)	(-2.86)	(-2.99)	(-7.47)	(1.17)	[0.000]		[28.31***]

Covid Economics 38, 16 July 2020: 68-128

Covid Economics 38, 16 July 2020: 68-128

Table 9

Contemporaneous regression of daily changes in 2-year Treasury yield on text variables and *corona*. Rows correspond to each explanatory variable for which either the b_3 or b_4 coefficient in (4) is significant at the 10% level or better. *EV* column shows the impact of a one standard deviation change in the explanatory variable in units of standard deviation of the 2-year Treasury yield. *EV*VIXl1* column shows the impact of a unit increase in VIX^{10} on the value of EV. The $VIXl1$ column shows the effect of a one standard deviation increase in VIX^{10} in units of standard deviation of the market variable. Numbers in parentheses are t-statitics, and the numbers in square brackets underneath the adjusted R^2s are joint F-tests. Both sets of tests use Newey-West standard errors with 3 lags. The *R2* column shows adjusted R^2s, and *Chg R2* shows the difference in adjusted R^2s between the specification in (4) and a version of (4) which sets $b_4 = 0$. The *Break Date* column shows the structural break date of (4) using the procedure from Section 4.2, and the maximal Chow statistic is shown in square backets, with siginficance levels obtained from the Andrews (1993, 2003) distribution indicated via the *s. Significance at the 10%, 5%, or 1% levels is indicated via *,**,*** respectively.

Contemp. changes in 2-year Treasury yield: early subperiod, VIX interact

	const	h_t	h_{t-1}	EV	EV*VIXl1	VIXl1	R2	Chg R2	Break Date
sent	-0.597*	-0.050	0.587***	0.616***	0.093**	6.439**	0.229	0.125	2020-03-09
	(-1.76)	(-0.64)	(2.74)	(3.65)	(2.18)	(2.14)	[0.003]		[37.47***]
s_sports	-0.045	-0.167	0.231*	1.129**	-0.006	0.820	0.148	-0.023	2020-03-09
	(-0.64)	(-1.32)	(1.78)	(2.53)	(-0.28)	(1.02)	[0.000]		[29.60***]
s_central bank	0.046	-0.112	0.174	0.615*	-0.029	-0.211	0.046	-0.007	2020-03-09
	(0.53)	(-0.71)	(0.88)	(1.66)	(-0.93)	(-0.25)	[0.236]		[37.22***]
s_markets	-0.187**	-0.141	0.278	0.426*	0.078***	1.714**	0.062	0.094	2020-03-09
	(-2.48)	(-0.87)	(1.55)	(1.82)	(2.90)	(2.57)	[0.000]		[36.28***]
s_oil & comm	-0.388***	0.087	0.375***	0.817***	0.169***	3.776***	0.298	0.328	2020-03-09
	(-4.53)	(0.81)	(2.69)	(4.45)	(5.37)	(5.05)	[0.000]		[44.46***]
s_currency	-0.187***	-0.355***	0.296***	1.262***	0.148***	2.627***	0.458	0.409	2020-03-12
	(-2.79)	(-2.58)	(2.67)	(5.64)	(4.69)	(3.87)	[0.000]		[44.29***]
s_credit	-0.169*	0.113	0.457**	0.576*	0.061*	1.784*	0.083	0.063	2020-03-09
	(-1.94)	(1.28)	(2.04)	(1.94)	(1.78)	(1.87)	[0.326]		[45.22***]
s_corp future	-0.128	-0.171	0.166	0.531**	0.089***	1.292**	0.109	0.124	2020-03-09
	(-1.46)	(-1.01)	(0.90)	(2.46)	(2.72)	(2.17)	[0.045]		[40.15***]

Covid Economics 38, 16 July 2020: 68-128

Table 10

Contemporaneous regression of daily changes in 10-year Treasury yield on text variables and *corona*. Rows correspond to each explanatory variable for which either the b_3 or b_4 coefficient in (4) is significant at the 10% level or better. *EV* column shows the impact of a one standard deviation change in the explanatory variable in units of standard deviation of the 10-year Treasury yield. *EV*VIXl1* column shows the impact of a unit increase in VIX^{10} on the value of EV. The $VIXl1$ column shows the effect of a one standard deviation increase in VIX^{10} in units of standard deviation of the market variable. Numbers in parentheses are t-statitics, and the numbers in square brackets underneath the adjusted R^2s are joint F-tests. Both sets of tests use Newey-West standard errors with 3 lags. The *R2* column shows adjusted R^2s, and *Chg R2* shows the difference in adjusted R^2s between the specification in (4) and a version of (4) which sets $b_4 = 0$. The *Break Date* column shows the structural break date of (4) using the procedure from Section 4.2, and the maximal Chow statistic is shown in square backets, with siginficance levels obtained from the Andrews (1993, 2003) distribution indicated via the *s. Significance at the 10%, 5%, or 1% levels is indicated via *,**,*** respectively.

Contemp. changes in 10-year Treasury yield: early subperiod, VIX interact

	const	h_t	h_{t-1}	EV	EV*VIXl1	VIXl1	R2	Chg R2	Break Date
sent	-0.850	-0.060	0.310*	0.574***	0.092*	6.493*	0.137	0.111	2020-03-17
	(-1.44)	(-0.39)	(1.95)	(3.04)	(1.66)	(1.67)	[0.006]		[22.38**]
s_sports	-0.043	-0.176	-0.054	0.891*	-0.015	0.581	0.048	-0.017	2020-03-18
	(-0.38)	(-1.35)	(-0.38)	(1.70)	(-0.58)	(0.61)	[0.185]		[26.52***]
s_markets	-0.393***	-0.408***	-0.016	0.960***	0.139***	3.143***	0.277	0.357	2020-03-17
	(-3.47)	(-3.07)	(-0.12)	(3.90)	(4.52)	(4.03)	[0.000]		[23.82***]
s_health	-0.312	-0.178*	0.070	0.940*	0.141	2.234	0.092	0.114	2020-03-17
	(-1.48)	(-1.87)	(0.70)	(1.83)	(1.48)	(1.48)	[0.075]		[31.37***]
s_oil & comm	-0.742***	0.000	0.181*	1.110***	0.232***	5.292***	0.543	0.623	2020-03-13
	(-6.25)	(0.00)	(1.85)	(6.64)	(8.30)	(7.43)	[0.000]		[22.33**]
s_currency	-0.286***	-0.279**	0.110	1.127***	0.139***	2.693***	0.383	0.404	2020-03-17
	(-3.64)	(-2.33)	(0.79)	(5.50)	(4.43)	(4.47)	[0.000]		[36.30***]
s_corp & govt US	0.422**	-0.171	-0.201	0.153	-0.077**	-2.522*	0.054	0.074	2020-03-20
	(2.07)	(-1.15)	(-1.09)	(0.74)	(-2.21)	(-1.85)	[0.005]		[28.99***]
s_corp future	-0.258**	-0.211	-0.092	0.457***	0.095***	1.661***	0.103	0.184	2020-03-17
	(-2.45)	(-1.42)	(-0.45)	(3.02)	(3.67)	(3.45)	[0.000]		[18.95**]

Covid Economics 38, 16 July 2020: 68-128

Table 11
Test of whether the variable in a given row Granger causes the variables in the columns of the table. The table shows c_3 (EV) and c_4 ($EV*VIX$) from the specifications in (7). The c_3 coefficient is normalized to show standard deviation changes in the response variable due to a single standard deviation change in the test variable. c_4 shows changes in c_3 due to a unit change in the VIX. T-statistics are shown in parentheses. The $d\_[market]$ variables refer to daily returns or changes in the particular market series. Standard errors use Newey-West with three lags. Significance at the 10%, 5%, or 1% levels is indicated via *,**,*** respectively. Data are daily.

Granger causality tests: Market changes to text measures in early subsample

		sent	sent_sd	s_sports	s_central bank	s_markets	s_health	s_europe
d_sp500	EV	0.882***	-0.115	0.292**	0.150	0.515***	0.344***	0.531***
	(EV t)	(5.22)	(-1.04)	(2.53)	(1.49)	(2.84)	(2.86)	(2.74)
	EV*VIX	-0.044***	-0.001	-0.010	-0.015***	-0.018*	-0.017**	-0.035***
	(EV*VIX t)	(-4.76)	(-0.22)	(-1.21)	(-4.00)	(-1.80)	(-2.54)	(-3.60)
d_vix	EV	-0.788***	0.196*	-0.225**	-0.091	-0.537***	-0.256**	-0.470***
	(EV t)	(-5.68)	(1.73)	(-1.98)	(-1.20)	(-3.16)	(-2.44)	(-3.25)
	EV*VIX	0.038***	-0.003	0.007	0.013***	0.019**	0.012**	0.030***
	(EV*VIX t)	(5.36)	(-0.47)	(1.09)	(5.53)	(2.22)	(2.30)	(4.27)
d_hy	EV	0.803***	-0.073	0.169	0.026	0.661***	0.354**	0.532**
	(EV t)	(3.14)	(-0.52)	(0.82)	(0.29)	(2.87)	(2.27)	(2.48)
	EV*VIX	-0.041***	-0.003	-0.009	-0.010**	-0.024**	-0.017**	-0.030***
	(EV*VIX t)	(-3.37)	(-0.38)	(-0.94)	(-2.28)	(-2.17)	(-2.44)	(-2.83)
d_gt2	EV	0.627***	-0.013	0.145**	0.140**	0.344**	0.281***	0.357***
	(EV t)	(4.98)	(-0.14)	(2.00)	(1.99)	(2.53)	(4.36)	(3.28)
	EV*VIX	-0.039***	-0.008	-0.003	-0.010**	-0.022**	-0.018***	-0.032***
	(EV*VIX t)	(-4.76)	(-1.32)	(-0.53)	(-2.00)	(-2.53)	(-4.40)	(-4.30)
d_gt10	EV	0.854***	0.016	0.153*	0.134	0.572***	0.429***	0.436***
	(EV t)	(5.32)	(0.15)	(1.81)	(1.45)	(3.90)	(6.22)	(2.71)
	EV*VIX	-0.049***	-0.009	-0.006	-0.010**	-0.031***	-0.025***	-0.032***
	(EV*VIX t)	(-5.43)	(-1.41)	(-0.99)	(-2.19)	(-4.06)	(-6.78)	(-3.61)

CEPR PRESS

Table 11
Test of whether the variable in a given row Granger causes the variables in the columns of the table. The table shows c_3 (EV) and c_4 ($EV*VIX$) from the specifications in (7). The c_3 coefficient is normalized to show standard deviation changes in the response variable due to a single standard deviation change in the test variable. c_4 shows changes in c_3 due to a unit change in the VIX. T-statistics are shown in parentheses. The $d\_[market]$ variables refer to daily returns or changes in the particular market series. Standard errors use Newey-West with three lags. Significance at the 10%, 5%, or 1% levels is indicated via *,**,*** respectively. Data are daily.

Granger causality tests: Market changes to text measures in early subsample

		s_oil & comm	s_currency	s_credit	s_corp & govt US	s_corp actual	s_corp future	s_credit1	corona
d_sp500	EV	0.013	0.282***	0.126	0.279**	0.090	0.086	-0.085	-0.017
	(EV t)	(0.19)	(2.60)	(1.01)	(2.30)	(0.54)	(0.55)	(-0.37)	(-0.11)
	EV*VIX	-0.001	-0.013**	-0.009	-0.018**	-0.007	-0.002	0.016	0.007
	(EV*VIX t)	(-0.14)	(-2.32)	(-1.52)	(-2.56)	(-0.87)	(-0.27)	(1.37)	(0.70)
d_vix	EV	-0.098	-0.342***	-0.165	-0.256**	-0.049	-0.110	-0.029	-0.000
	(EV t)	(-1.61)	(-3.67)	(-1.58)	(-2.03)	(-0.42)	(-0.64)	(-0.15)	(-0.00)
	EV*VIX	0.004	0.015***	0.010**	0.016**	0.005	0.003	-0.011	-0.010
	(EV*VIX t)	(1.02)	(3.69)	(2.10)	(2.37)	(0.94)	(0.37)	(-1.14)	(-1.17)
d_hy	EV	0.148	0.520**	-0.083	0.132	-0.199	-0.007	-0.174	0.259
	(EV t)	(0.94)	(2.49)	(-0.54)	(0.67)	(-0.97)	(-0.04)	(-0.66)	(0.98)
	EV*VIX	-0.011	-0.024***	-0.003	-0.013	0.007	0.001	0.011	-0.008
	(EV*VIX t)	(-1.20)	(-2.63)	(-0.33)	(-1.34)	(0.65)	(0.07)	(0.82)	(-0.51)
d_gt2	EV	-0.008	0.330***	0.046	0.269***	-0.052	0.022	0.054	0.252
	(EV t)	(-0.08)	(2.72)	(0.54)	(3.95)	(-0.45)	(0.24)	(0.54)	(1.55)
	EV*VIX	-0.003	-0.017**	-0.009*	-0.013**	-0.005	0.001	-0.001	-0.008
	(EV*VIX t)	(-0.54)	(-2.43)	(-1.70)	(-2.44)	(-0.71)	(0.09)	(-0.11)	(-0.87)
d_gt10	EV	0.078	0.507***	0.116	0.230*	-0.142	0.039	0.048	0.187
	(EV t)	(0.75)	(4.09)	(1.18)	(1.87)	(-0.92)	(0.31)	(0.34)	(1.64)
	EV*VIX	-0.007	-0.027***	-0.013**	-0.013*	0.001	-0.001	-0.005	-0.003
	(EV*VIX t)	(-1.19)	(-4.25)	(-2.56)	(-1.72)	(0.16)	(-0.12)	(-0.59)	(-0.38)

Covid Economics 38, 16 July 2020: 68-128

Table 12
Test of whether the variable in a given row Granger causes the variables in the columns of the table. The table shows c_3 (EV) and c_4 (EV^*VIX) from the specifications in (7). The c_3 coefficient is normalized to show standard deviation changes in the response variable due to a single standard deviation change in the test variable. c_4 shows changes in c_3 due to a unit change in the VIX. T-statistics are shown in parentheses. The $d\_[market]$ variables refer to daily returns or changes in the particular market series. Standard errors use Newey-West with three lags. Significance at the 10%, 5%, or 1% levels is indicated via *,**,*** respectively. Data are daily.

Granger causality tests: Text measures to market changes in early subsample

		d_sp500	d_vix	d_hy	d_gt2	d_gt10
sent	EV	-0.129	-0.085	-0.347**	-0.174	-0.218*
	(EV t)	(-0.80)	(-0.53)	(-2.48)	(-1.38)	(-1.66)
	EV*VIX	-0.083***	0.058**	-0.078***	-0.041**	-0.057***
	(EV*VIX t)	(-2.90)	(2.24)	(-3.08)	(-2.38)	(-2.76)
sent_sd	EV	1.137***	-1.098***	1.181***	0.906***	1.231***
	(EV t)	(6.31)	(-5.57)	(3.07)	(3.25)	(3.01)
	EV*VIX	0.175***	-0.169***	0.155***	0.135***	0.174***
	(EV*VIX t)	(5.50)	(-5.49)	(2.82)	(3.45)	(3.27)
s_sports	EV	-0.613*	0.355	-1.333***	0.068	-0.364
	(EV t)	(-1.67)	(0.87)	(-3.20)	(0.13)	(-0.69)
	EV*VIX	-0.067**	0.055**	-0.090***	-0.056**	-0.075***
	(EV*VIX t)	(-2.47)	(2.55)	(-4.48)	(-2.19)	(-2.81)
s_central bank	EV	0.096	-0.169	-0.481	0.384	0.284
	(EV t)	(0.19)	(-0.40)	(-1.49)	(0.89)	(0.84)
	EV*VIX	-0.003	-0.003	-0.002	-0.020	-0.039*
	(EV*VIX t)	(-0.15)	(-0.15)	(-0.09)	(-0.94)	(-1.86)
s_markets	EV	-0.253	0.174	-0.442*	-0.892***	-1.077**
	(EV t)	(-1.06)	(0.97)	(-1.78)	(-3.54)	(-2.56)
	EV*VIX	-0.061***	0.054***	-0.078***	-0.108***	-0.133**
	(EV*VIX t)	(-3.08)	(3.48)	(-2.76)	(-3.31)	(-2.44)
s_health	EV	-0.163	0.111	-0.892***	-0.714	-0.834
	(EV t)	(-0.65)	(0.65)	(-3.93)	(-1.43)	(-1.29)
	EV*VIX	-0.097**	0.088***	-0.176***	-0.100	-0.140
	(EV*VIX t)	(-2.06)	(2.71)	(-3.80)	(-1.20)	(-1.31)
s_europe	EV	0.136	-0.306	0.012	-0.295	-0.617**
	(EV t)	(0.29)	(-0.74)	(0.04)	(-1.35)	(-2.14)
	EV*VIX	-0.013	-0.000	0.000	-0.051	-0.114**
	(EV*VIX t)	(-0.33)	(-0.00)	(0.01)	(-1.60)	(-2.06)
s_oil & comm	EV	-0.509	0.345	0.019	-0.808***	-0.863***
	(EV t)	(-1.18)	(1.28)	(0.04)	(-3.20)	(-2.84)

Covid Economics 38, 16 July 2020: 68-128

Covid Economics 38, 16 July 2020: 68-128

Table 12

Test of whether the variable in a given row Granger causes the variables in the columns of the table. The table shows c_3 (EV) and c_4 (EV*VIX) from the specifications in (7). The c_3 coefficient is normalized to show standard deviation changes in the response variable due to a single standard deviation change in the test variable. c_4 shows changes in c_3 due to a unit change in the VIX. T-statistics are shown in parentheses. The $d\_[market]$ variables refer to daily returns or changes in the particular market series. Standard errors use Newey-West with three lags. Significance at the 10%, 5%, or 1% levels is indicated via *,**,*** respectively. Data are daily.

Granger causality tests: Text measures to market changes in early subsample

		d_sp500	d_vix	d_hy	d_gt2	d_gt10
s_oil & comm	EV*VIX	-0.050	0.019	0.022	-0.119***	-0.128**
	(EV*VIX t)	(-0.52)	(0.31)	(0.22)	(-2.85)	(-2.42)
s_currency	EV	0.120	-0.190	-0.162	0.420	0.546
	(EV t)	(0.36)	(-0.62)	(-0.49)	(1.02)	(1.52)
	EV*VIX	-0.062	0.050	-0.046	0.046	0.068
	(EV*VIX t)	(-1.46)	(1.28)	(-0.98)	(0.96)	(1.49)
s_credit	EV	0.035	-0.182	-0.006	0.085	0.025
	(EV t)	(0.10)	(-0.61)	(-0.02)	(0.26)	(0.08)
	EV*VIX	-0.014	-0.002	0.010	-0.011	-0.019
	(EV*VIX t)	(-0.33)	(-0.06)	(0.28)	(-0.26)	(-0.43)
s_corp & govt US	EV	0.312	-0.320	0.379*	-0.187	0.118
	(EV t)	(1.15)	(-1.09)	(1.78)	(-0.68)	(0.34)
	EV*VIX	0.035	-0.039	0.051	-0.025	-0.018
	(EV*VIX t)	(0.88)	(-1.00)	(1.61)	(-0.86)	(-0.60)
s_corp actual	EV	0.297**	-0.403**	0.162	0.451	0.269
	(EV t)	(2.51)	(-2.52)	(0.94)	(1.64)	(0.98)
	EV*VIX	0.013	-0.026	0.027	0.042	0.020
	(EV*VIX t)	(0.64)	(-1.30)	(0.99)	(1.03)	(0.43)
s_corp future	EV	0.017	-0.166	0.040	0.308	0.507
	(EV t)	(0.07)	(-0.98)	(0.19)	(1.08)	(1.64)
	EV*VIX	-0.004	-0.001	0.005	0.066	0.104*
	(EV*VIX t)	(-0.10)	(-0.02)	(0.12)	(1.38)	(1.91)
s_credit1	EV	-0.066	0.093	-0.074	0.092	0.122
	(EV t)	(-0.30)	(0.44)	(-0.38)	(0.47)	(0.52)
	EV*VIX	-0.033	0.034	-0.013	-0.002	-0.009
	(EV*VIX t)	(-1.43)	(1.55)	(-0.49)	(-0.09)	(-0.35)
corona	EV	0.830***	-0.782***	0.209*	0.403**	0.519***
	(EV t)	(3.09)	(-2.88)	(1.73)	(2.17)	(3.25)
	EV*VIX	-0.007	0.011	0.026	0.032*	0.069***
	(EV*VIX t)	(-0.33)	(0.64)	(1.30)	(1.86)	(3.22)

CEPR PRESS

Covid Economics **Issue 38, 16 July 2020**

COVID-19, race, and redlining[1]

Graziella Bertocchi[2] and Arcangelo Dimico[3]

Date submitted: 7 July 2020; Date accepted: 9 July 2020

Using individual, race-disaggregated, and georeferenced death data collected by the Cook County Medical Examiner, we look at the impact of COVID-19 on African Americans and at its determinants. First, we provide evidence that - as of June 16, 2020 - blacks in Cook County are dying from COVID-19 at a rate 1.3 times higher than their population share. Second, by combining the spatial distribution of mortality with the redlining maps for the Chicago area, we establish that - after the epidemic outbreak - historically lower-graded neighborhoods display a sharper increase in mortality, driven by blacks. Thus, we uncover a persistence influence of the racial segregation induced by the lending practices of the 1930s, by way of a diminished resilience of African Americans to the COVID-19 shock. Such influence is channeled through socioeconomic status and household composition, and magnified in combination with a higher black share.

Covid Economics 38, 16 July 2020: 129-195

1 We would like to thank Chris Colvin, Guido Friebel, Marianne Wanamaker, and seminar participants at Queen's University Belfast and Goethe University for helpful comments.
2 Professor of Economics, University of Modena and Reggio Emilia, EIEF, CEPR, CHILD, Dondena, GLO, and IZA.
3 Senior Lecturer in Economics, Queen's University Belfast, GLO, IZA, CHaRMS, and QUCEH.

Covid Economics 38, 16 July 2020: 129-195

1 Introduction

Ever since when The Atlantic magazine (Kendi, 2020), on March 4, 2020, first launched a cry for attention to the disproportionate impact of COVID-19 on African Americans, the issue has taken center stage in the debate on the socioeconomic implications of the pandemic. Nevertheless, a race-disaggregated analysis of individual data has so far been lacking due to data unavailability.

The urgency of the racial issue and the need for data collection to account for ethnic and racial factors has been widely recognized also within the medical and epidemiological literature. Preliminary evaluations suggest that the high risk of COVID-19 death for minority ethnic groups can be explained by pre-existing health conditions, such as diabetes, obesity, hypertension, and asthma, that are more common among these groups, possibly because of genetic and biological factors. However, the emerging consensus is that the race differential in the prevalence of COVID-19 is also associated with socioeconomic correlates. As argued by Yancy (2020), a large share of the black population in the US lives in poor areas characterized by high unemployment, low housing quality, and unhealthy living conditions, making low socioeconomic status a critical risk factor. Furthermore, the higher prevalence of comorbidities in blacks is also itself associated with highly persistent socioeconomic factors. Other relevant health-related and behavioral risk factors, such as smoking, drinking, and drug abuse, are deeply ingrained in cultural norms that are also driven by social inequalities. African Americans suffer for further disadvantage in their ability to adhere to social distancing norms, as working from home, avoiding public transportation, and finding refuge in second homes away from crowded cities, are indeed privileges that are denied to the majority of them. Anecdotal evidence also suggests the possibility of a different response by health practitioners to black individuals showing COVID-19 symptoms. All the above considerations, as summarized by Yancy (2020) resonate with the long-standing debate on the racial disparities that are deeply entrenched in US history and point to a need to account for race in COVID-19 research and to investigate the role not only of biological factors but also of socioeconomic ones, while acknowledging that the latter may be rooted on inequalities that have been long neglected.

In this paper, we take advantage of an unexplored and extraordinarily detailed source of information on daily deaths from COVID-19 that includes race among a wide array of individual characteristics. The data are collected by the Medical Examiner's Officer and made available by the Government of Cook County, Illinois, the county that hosts—among others—the City of Chicago.[1]

[1]Since April 15, after some states started to report data, the COVID Tracking Project at The Atlantic

Covid Economics 38, 16 July 2020: 129-195

On the basis of the Cook County data we can provide, to start with, evidence of how race affects COVID-19 individual outcomes. Furthermore, we can illustrate the correlates and dig into the historical determinants of such outcomes. The higher COVID-19 death toll paid by black Americans has been linked to the "redlining" policies introduced by the Home Owners Loan Corporation (HOLC) in the 1930s.[2] These policies are believed to have favored the development of segregated neighborhoods plagued by unemployment, low housing quality, and unhealthy living conditions. In a public speech following the death of George Floyd, in the midst of the pandemic, former President Obama has also attributed the explosion of protests to the history of racial discrimination, including redlining.[3] By combining the information on COVID-19 fatalities in Cook County with the redlining maps for the Chicago area, we can assess the explanatory power of the induced vulnerabilities that are rooted in history.

In more detail, the first contribution of the paper is to document how race affects COVID-19 mortality in Cook County, on the basis of individual data that also account for age, gender, and pre-existing health conditions. The analysis confirms that, as reported by media and public authorities, blacks are overrepresented in terms of COVID-19 related deaths since—as of June 16, 2020—they constitute 35 percent of the deaths against a black population share of only 27 percent in our sample. In other words, blacks are dying at a rate 1.3 times higher than their population share. This figure is lower than those previously reported at a early stage, because blacks were the first to be hit by the virus, while other groups followed up only later on.

Second, by exploiting the fact that the Cook County Medical Examiner also provides the geographical coordinates of the home address of each individual that died from COVID-19, we combine the spatial distribution of mortality with the redlining maps for the Chicago area, to assess whether the higher vulnerability of blacks to the disease can indeed be found in socioeconomic inequalities rooted in history, rather than in biological conditions. Using cross-sectional information about individual deaths from COVID-19, we show that the probability that an individual who died from COVID-19 is black is much larger in lower graded areas, that is, those that were historically redlined and yellowlined, even after controlling for demographic and socioeconomic factors as well as pre-existing

has been providing updated state-level race and ethnicity data on cases and deaths, with information limited to race and ethnicity. See https://covidtracking.com/race. On July 5, 2020 The New York Times (Oppel et al., 2020) reported data on 640,000 individual cases, by race, ethnicity and home county, collected through May 28. The data were acquired after filing a Freedom of Information Act suit against the Centers for Disease Control and Prevention.

[2]See Eligon et al. (2020) in The New York Times.

[3] *"They are the outcomes not just of the immediate moments in time, but they are the result of a long history of slavery and Jim Crow and redlining and institutionalized racism that too often have been the plague, the original sin, of our society."* See https://www.youtube.com/watch?v=q_qB6SsErpA.

Covid Economics 38, 16 July 2020: 129-195

conditions. However, where we control for the black population share, it fully absorbs the influence of the HOLC policies, suggesting that the latter did induce segregation along racial lines. In other words, the black share fully captures the effect of residence in a low-graded area. A warning is in order since the cross-sectional analysis, being based solely on information about those that died from COVID-19, is severely biased because of sample selection. This limitation motivates our next set of results.

Third, to establish a causal link between the vulnerabilities induced by historical redlining and the racial gaps in COVID-19 outcomes, and to disentangle socioeconomic factors from other determinants which normally correlate with African Americans, we rely on an event study analysis based on a weekly balanced panel of deaths from any cause that we assemble at census block group level over the period from January 1, 2020 to June 16, 2020. In this way, we can identify the impact of the asymmetric shock introduced by COVID-19 on lower-grade block groups—the treated group—after treatment initiation, that is, after the epidemic outbreak. Furthermore, we can test for pre-treatment differences while controlling for block group fixed effects which should filter out the effect of invariant socioeconomic factors. Consistent with the cross-sectional evidence, to determine the treatment we refer to block groups that belong to HOLC areas graded either C or D, that is, either yellow or redlined. We find that, while no pre-treatment differences are detected, mortality in treated neighborhoods increases sharply after the epidemic shock, and the increase is driven by the death toll of African Americans. The implied magnitude of the estimated effect is an increase in the number of black deaths by almost 60 percent.

We also carry out a heterogeneity analysis focused on the degree of vulnerability to shocks, that we measure using data provided by the Centers for Disease Control and Prevention (CDC). The analysis reveals that the channels through which historical redlining influences COVID-19 outcomes are socioeconomic characteristics, in particular personal income and the population share below poverty, and household composition, in particular the population shares of the elderly and of single parent households. Strikingly, however, the influence of these factors manifests itself only in combination with a higher black share. We complete our investigation with a battery of robustness checks, extensions, and falsification tests.

Overall, the evidence we collect points to a persistence influence of the racial segregation introduced by the discriminatory lending practices of the 1930s, that operates by way of an asymmetric effect of the epidemic shock, which is in turn channeled through a diminished resilience of African Americans. Far from being determined by genetic and biological factors, their vulnerability can be linked to socioeconomic status and household composition, as channels through which the legacy of the past manifests itself.

Covid Economics 38, 16 July 2020: 129-195

The paper is organized as follows. Section 2 presents related literature. Section 3 summarizes background information on historical redlining policies and Cook County. Section 4 describes the data. Section 5 presents cross-sectional evidence. Section 6 introduces an event study approach and the corresponding baseline results, while Section 7 is devoted to robustness checks and Section 8 to the heterogeneity analysis. In Section 9 we extend the investigation to the Hispanic population. Section 10 presents two falsification tests respectively based on deaths in the years 2017 to 2019 and Spanish flu deaths in 1918. Section 11 concludes. The Appendix contains additional figures and tables.

2 Related literature

We contribute to several streams of the literature. The first is the literature on the racial impact of COVID-19. The so far largest epidemiological study on the racial impact of COVID-19 has been performed in the UK based on the medical records of more than 17 million individuals (OpenSAFELY Collaborative, 2020, working on behalf of NHS England).[4] The study confirms that pre-existing medical conditions such as diabetes or deprivation are linked to a higher likelihood of in-hospital death, but also that clinical risk factors alone cannot explain the observed disparities. Again with a focus on the UK, Bhala et al. (2020) also support structural, socioeconomic and environmental explanations, rather than biological ones, of the racial differences in COVID-19 susceptibility. Attention is called to the fact that ethnic/racial minorities hold highly-exposed jobs in health and social care, retail, and public transport, and to cultural habits including the approach to worship and the multigenerational family structure. On the basis of medical data from COVID-19 patients at a hospital in Louisiana, Price-Haywood et al. (2020) find that blacks were overrepresented among patients and fatalities, but did not show higher in-hospital mortality than whites after controlling for differences in sociodemographic (i.e., insurance type and zip code of residence) and clinical characteristics on admission.

Despite the above mentioned growing body of contributions within medicine and epidemiology, within the economics field the literature on the racial impact of COVID-19 is still limited. Borjas (2020), Schmitt-Grohe et al. (2020), and Almagro and Orane-Hutchinson (2020) account for the racial dimension while looking at the demographic and socioeconomic correlates of the COVID-19 epidemics, with a focus on testing incidence and infections, but not deaths, across New York City zip codes. Across US counties, Desmet and Wacziarg (2020) find a positive correlation between the shares of African Americans and Hispanics and both the number of cases and the number of deaths, a

[4]In the UK, the racial issue came to public attention when the first eleven doctors who died from COVID-19 were all reported to belong to black, Asian, and minority ethnic communities (Kirby, 2020).

finding which is confirmed by McLaren (2020) for African Americans' deaths even after controlling for education, occupation, and commuting patterns. Using the 2017 wave of the Panel Study of Income Dynamics to examine the prevalence of specific health conditions, Wiemers et al. (2020) show evidence of large disparities across race-ethnicity and socioeconomic status in the prevalence of conditions which are associated with the risk of severe complications from COVID-19. Using CPS microdata on unemployment, Couch et al. (2020) show that African Americans were only slightly disproportionately impacted by COVID-19, while Latin Americans were hardly hit. With reference to the UK, Platt and Warwick (2020), Sa (2020), and White and Nafilyan (2020) report descriptive evidence on vulnerability factors, infections, and deaths, for ethnic minorities including Black Caribbean and Black African.[5]

Our second contribution is to research on the long-term influence of redlining policies, which have been the object of investigation not only within the field of economics but also within medicine, history, and law. Zenou and Boccard (2000), Appel and Nickerson (2016), Krimmel (2018), Aaronson et al. (2017), Mitchell and Franco (2018), and Anders (2019) respectively look at the effects on unemployment, home prices, homeownership, racial segregation, inequality, and crime. Krieger et al. (2020a, 2020b) and Nardone et al. (2020) respectively associate historical redlining with higher risk of cancer, preterm birth, and asthma, suggesting that this discriminatory practice might be contributing to racial and ethnic disparities.[6] Jackson (1980, 1985), Hillier (2003, 2005), and Greer (2012, 2014) provide historical accounts of the activities of the HOLC and its influence on American cities. Schill and Wachter (1995) and Nier (1999) offer a legal interpretation of redlining and the induced spatial bias of federal housing law and policy.[7]

A third research stream that is relevant to our approach has looked at other pandemics. The long-term determinants of the HIV/AIDS pandemic, with special attention to women, who represent within Africa the most vulnerable group, have been studied by Anderson (2018), who links the higher female HIV rates to the tradition of common law, while Bertocchi and Dimico (2019), Loper (2019), and Cage and Rueda (2020) respectively refer to the slave trade and the associated diffusion of polygyny, the practice of matrilinearity in ancestral societies, and the influence of the Christian missions

[5] For the much broader literature on racial discrimination in the US, see Lang and Kahn-Lang Spitzer (2020) for a survey and Nunn (2008), Bertocchi and Dimico (2012, 2014, 2020a), Bertocchi (2016), and references therein, on the legacy of slavery as its long-term determinant.

[6] On residential segregation and cancer see also Landrine et al. (2017). An extensive literature confronts racial and ethnic disparities in health care and health outcomes. See Institute of Medicine (2003) and Orsi et al. (2010).

[7] On the history of the HOLC, see Fishback et al. (2013). On the broader determinants, other than redlining, of race segregation in the metropolitan areas of the US, see Boustan (2011) and Rothstein (2017). On the influence of zoning policies, with a focus on 1923 Chicago, see Shertzer et al. (2016).

Covid Economics 38, 16 July 2020: 129-195

CEPR PRESS

established during the colonial period. The long-term economic, social and cultural consequences of the "Spanish" influenza have been studied by Almond (2006), Karlsson et al. (2014), Lin and Liu (2014), Helgertz and Bengtsson (2019), Aassve et al. (2020), and Guimbeau et al. (2020), while Richardson and McBride (2009), Voitglander and Voth (2012), Jedwab et al. (2016, 2019), and Alfani and Murphy (2017) have studied the Black Death.

3 Historical background

3.1 Redlining

The Home Owners Loan Corporation (HOLC) was created in June 1933 by the US Congress, in the aftermath of the Great Depression and within the first 100 days of the Roosevelt administration, as part of a key package of New Deal policies aimed at rescuing the housing and banking sectors through actions on the mortgage lending market. In the general effort to revive the economy, housing policies were viewed as critical and were therefore assigned a major role. The task of the HOLC was to refinance mortgages in default to prevent foreclosures, as a response to the banking sector turmoil and the drastic fall in home loans and ownership (Harriss, 1951). In 1934 the National Housing Act established the Federal Housing Administration (FHA) to reinforce previous measures and boost the market for single-family homes. With the goal of improving the accuracy of real-estate appraisal and in turn standardizing the process of mortgage lending, credit worthiness assessment, and mortgage support assignment, in 1935 the HOLC was asked to create "Residential Security Maps" of 239 cities that ranked areas on the basis of default risk. The ranking encompassed four levels. The safest areas, mostly consisting of newly-build suburban neighborhoods, were labelled as "Best", assigned to Type A, and outlined in color green. "Still Desirable" areas were assigned to Type B and outlined in blue. The next two levels included "Definitely Declining" areas, assigned to Type C and outlined in yellow, and "Hazardous" areas assigned to Type D and outlined in red. Because of the color used to highlight to the worst-assessed neighborhoods, those that ended up being de facto denied any mortgage financing, the process came to be known as "redlining".

The HOLC rankings were based on meticulous assessments and recording of neighborhood characteristics including population growth, class and occupation of the inhabitants, and block-by-block quality of the buildings (type, size, construction material, age, need for repair, occupancy rate, owner-occupancy rate, past and predicted property prices, rents, and sales and rental demand trends). Notably, the share of foreign and black families and the degree to which the neighborhood was deemed to be "infiltrated" were also

Covid Economics 38, 16 July 2020: 129-195

accurately recorded. Figure A1 in the Appendix reports the area description drafted by the HOLC in 1939 for a D-rated neighborhood in South Chicago, summarily described as *"A 100 per cent negro development. (...) A blighted section"*.

To assemble the mapping, the HOLC trained thousands of home appraisers and, in the process, set standards for the development of a new approach to mortgage lending, which were adopted and further refined in the Underwriting Manuals compiled by the FHA (1938). The task for which the HOLC was created was undoubtedly fulfilled, as the agency had a major impact on the subsequent expansion of real estate investments. As documented by Harriss (1951), between 1933 and 1935 the HOLC received almost 1.9 million applications for home mortgage refinancing.[8] Out of the 54 percent of them that were accepted, the majority involved one- or two-family homes of modest size and value and borrowers of relatively limited income. In the New York region, 44 percent of the properties whose purchase was supported with a loan were located in neighborhoods described as "native white"and 42 percent in "native white and foreign."The fact that only 1 percent of the applications covered properties in neighborhoods described as "native Negro"is attributed by Harriss to the low percentage of applicants from such areas, which *"doubtless reflects the fact that most Negroes (...) lived in rented quarters and did not, therefore, fall within the limits of the HOLC programs."*

The direct and indeed intended consequences of redlining were to channel credit and investment away from poorer areas and toward more affluent ones. As a result, the slums deteriorated even farther. Over time, the practice is widely believed to have contributed to the exacerbation and persistence of initial inequalities (Douglas Commission, 1968). After the Second World War, racial segregation further intensified with the "white flight"from the inner cities to the suburbs (Boustan, 2011). It was only with the Fair Housing Act of 1968, a provision of the Civil Rights Act, that housing segregation was outlawed, while specific legislation to establish fair lending practices was only enacted in the 1970s with the Equal Credit Opportunity Act (1974) and the Community Reinvestment Act (1977).

Throughout the process, the HOLC maps were deliberately hidden from public view, even though they may have been shared selectively with realtors and lenders (Greer, 2012, 2014). The existence of the maps emerged later and became the subject of investigation of the National Commission on Urban Problems (Douglas Commission, 1968), created by President Johnson in 1965 *"to study building codes, housing codes, zoning, local and Federal tax policies and development standards"* in order to *"provide knowledge that would be useful in dealing with slums, urban growth, sprawl and blight, and to insure decent and durable housing."* But it was only much later that Jackson (1980, 1985), an urban

[8]Subsequently, the activities of the HOLC were devoted to loan management and repayments. Operations were officially ceased in 1951 and termination was ordered in 1954.

Covid Economics 38, 16 July 2020: 129-195

historian, discovered the HOLC Residential Security Maps in the National Archives, documenting what he describes as a system designed to apply *"notions of ethnic and racial worth to real-estate appraising on an unprecedented scale."* His discovery spurred a renewed research effort aimed at identifying redlining as a key factor in perpetuating racial disparities that are still observed up to the present day.

3.2 Cook County, Illinois

With over five million residents in 2019, Cook County is the most populous county in Illinois and the second-most-populous county in the US after Los Angeles County, California. Over 40 percent of all residents of Illinois live in Cook County. The largest of the county's 135 municipalities is the City of Chicago—the third-most-populous US city—followed by the City of Evanston. Overall, Cook County is highly urbanized and densely populated. According to the United States Census Bureau,[9] in 2019 non-Hispanic whites, Hispanics, African Americans, and Asians were the most represented racial and ethnic groups, respectively with 42.0, 25.6, 23.8, and 7.9 percent of the population,[10] A 21.6 percent share of the population was under age 18 and 15.1 percent was above 65. A 13.8 fraction of the population, higher than the national average, was below the poverty line. Health status disparities between black and white populations widened in Chicago between 1990 and 2005 (Orsi et al., 2010). Politically, the county is heavily Democratic-leaning, with a 73.9 share of the votes being cast for the Democratic Party in the 2016 presidential elections.

According to the Johns Hopkins University & Medicine Coronavirus Resource Center,[11] as of June 16, 2020 Cook County ranked third among US counties, after Kings and Queens in New York, in terms of COVID-19 related deaths, with over 4,000 deaths, and first in terms of confirmed cases, with over 80,000 cases. A strict stay-at-home order was issued by Illinois Governor Pritzker on March 20 (effective March 21), four days after the first COVID-19 related death, when the death toll was still limited to five.

Turning to the urban history of Cook County in the aftermath of the Great Depression, in his 1933 dissertation on the evolution of land values in Chicago Hoyt (1933)—before joining the FHA in 1934 as Principal Housing Economist—provided a ranking of races and nationalities with respect to their beneficial effect upon land values. While acknowledging that such an effect may have been reflecting racial prejudice, he ranked Anglo-Saxons and Northern Europeans at the top and Negroes, followed by Mexicans, at the bottom.

[9] See https://www.census.gov/quickfacts/fact/table/cookcountyillinois/PST120219.

[10] Racial and ethnic data are based on self-identification. Reported figures are "alone", that is, refer to those individuals that self-classify themselves as belonging to a single racial or ethnic group.

[11] See https://coronavirus.jhu.edu/.

Covid Economics 38, 16 July 2020: 129-195

Covid Economics 38, 16 July 2020: 129-195

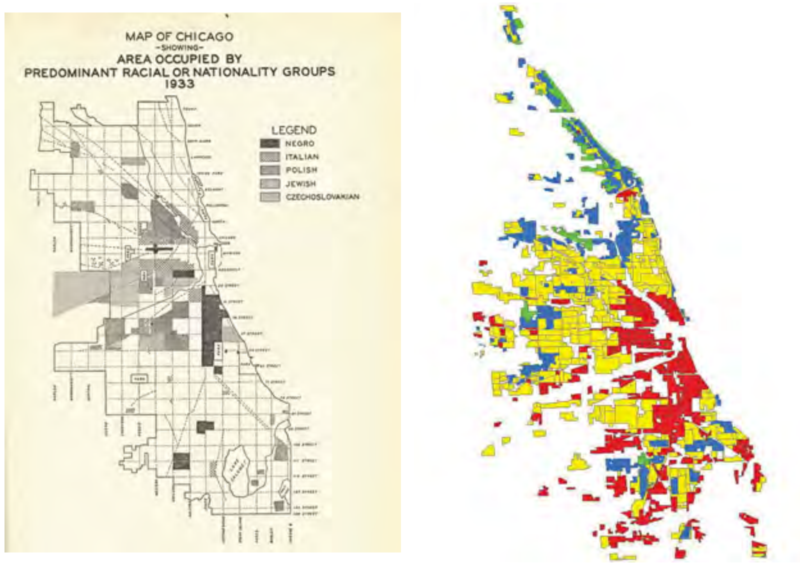


Figure 1: Historical Maps of the Chicago Area

Note: The figure shows, on the left panel, the "Map of Chicago Showing Area Occupied by Predominant Racial or Nationality Groups, 1933"(Hoyt, 1933) and, on the right panel, the HOLC maps for the Chicago area (Nelson et al., 2020), with green, blue, yellow, and red denoting respectively grade A, B, C, and D neighborhoods.

He also produced a map of Chicago (Figure 1, left panel) reporting the areas occupied by predominant groups among the most recent immigrant waves, namely Negro, Italian, Polish, Jewish, and Czechoslovakian newcomers. As the figure shows, blacks were concentrated in the so-called "Black Belt"on Chicago's South Side, where they were forced to settle from the beginning of the Great Migration, facing squalid housing conditions and extremely high population densities (Greer, 2014).

By 1940, a large portion of Cook County was mapped by the HOLC. On the right panel, Figure 1 shows the HOLC areas of the Chicago area as rendered in the Mapping Inequality: Redlining in New Deal America 1935-1940 dataset by the University of Richmond.[12] As in the other American cities, the geography of redlining had a clear racial connotation. The figure reveals that the same areas inhabited by blacks in the Hoyt map were assigned the lowest grade and highlighted in red. Indeed those blocks were characterized by houses lacking basic amenities such as access to water and heating (Greer, 2014). Krimmel (2018) estimates that, by 1940, 98 percent of the relatively small share of blacks living in the city, about 8 percent, were redlined.

The programs that were implemented by the agency were welcome with wide support from the Chicago white press, and were originally only weakly opposed by the black

[12] See Nelson et al. (2020) and https://dsl.richmond.edu/panorama/redlining/.

CEPR PRESS

press, including the Chicago Defender. Nevertheless, the disinvestment induced by the HOLC policies had consequences that extended well past the Second World War, with white families easily obtaining mortgage insurance to move to the suburbs, and black ones being relegated in blighted neighborhoods where the only financing opportunity consisted in exploitative installment land contracts (Greer, 2012, 2014). As a result, as of 1967, in the words of the Douglas Commission (1968) Chicago was hosting the most notorious slums in the country.

4 Data

The Cook County Medical Examiner's Officer[13] has been reporting individual COVID-19 related deaths daily since March 16, 2020, the date of the first fatality recorded in the county.[14] The declared goal of the initiative is to provide direct access to critical facts that can allow to identify communities that are most severely impacted by the virus. We employ data recorded up to June 16, that we downloaded on June 18.

The Medical Examiner's Office reports those deaths that are under its jurisdiction, including among others those due to diseases constituting a threat to public health.[15] The information on COVID-19 fatalities coincides with that provided by the Johns Hopkins University & Medicine Coronavirus Resource Center.[16]

The race, gender, age, comorbidities, and residence (home address, city, zip code, and geographical coordinates) of each dead individual are also provided by the Medical Examiner.[17] Overall, 4,325 individuals—1,491 of whom black (that is, 34.47 percent)—

[13]See Cook County Medical Examiner COVID-19 Related Deaths at https://datacatalog.cookcountyil.gov/Public-Safety/Medical-Examiner-Case-Archive-COVID-19-Related-Dea/3trz-enys.

[14]Data refer to deaths for which COVID-19 is reported among either primary or secondary causes. Operationally, the Medical Examiner's Officer looks for references to COVID-19 in any of these fields: Primary Cause, Primary Cause Line A, Primary Cause Line B, Primary Cause Line C, or Secondary Cause.

[15]Namely, the Medical Examiner's Office investigates any human death that falls within any or all of the following categories: criminal violence; suicide; accident; suddenly when in apparent good condition; unattended by a practicing licensed physician; suspicious or unusual circumstances; unlawful fetal death as provided in Public Act 101-0013 of the 101st General Assembly of Illinois; poisoning or attributable to an adverse reaction to drugs and/or alcohol; disease constituting a threat to public health; injury or toxic agent resulting from employment; during some medical diagnostic or therapeutic procedures; in any prison or penal institution; when involuntarily confined in jail prison hospitals or other institutions or in police custody; when any human body is to be cremated, dissected or buried at sea; when a dead body is brought into a new medico-legal jurisdiction without proper medical certification. Overall, each year, more than 16,000 deaths are reported to the Cook County Medical Examiner. See https://www.cookcountyil.gov/agency/medical-examiner.

[16]Data may temporarily differ from those provided by the departments of public health because of time lags in notification.

[17]Data are obtained from vital records, hospitals, and families.

Covid Economics 38, 16 July 2020: 129-195

Covid Economics 38, 16 July 2020: 129-195

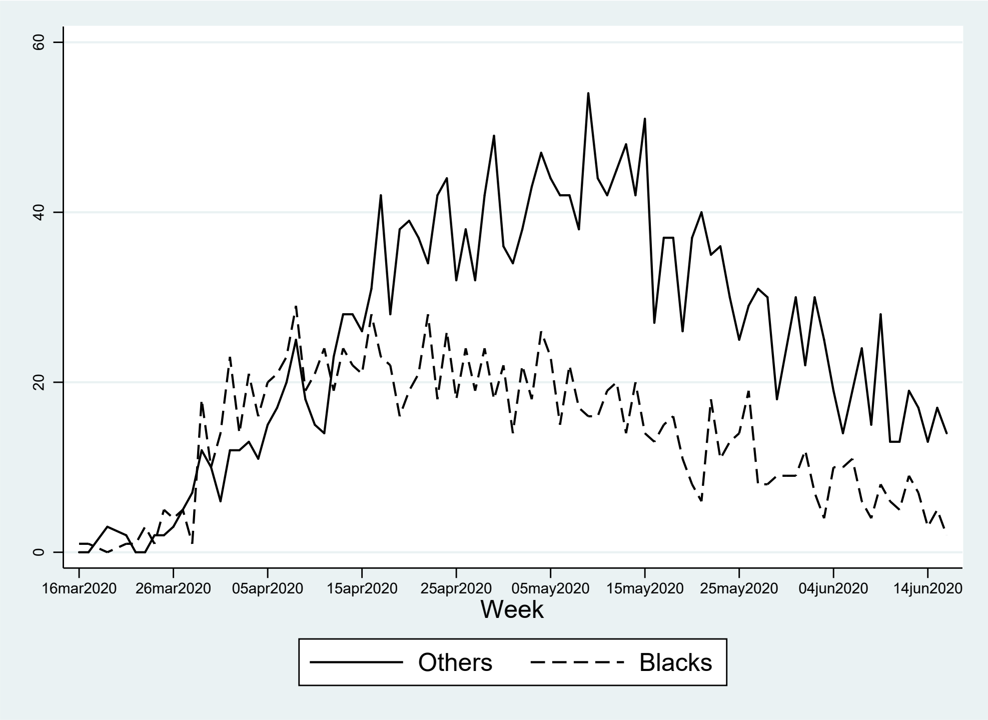


Figure 2: COVID-19 Deaths, by Race - Cook County, March 16-June 16, 2020

Note: The figure reports the number of COVID-19 related deaths by day, separately for blacks and all other races combined.

have died of COVID-19 in Cook County between March 16, 2020 and June 16, 2020. However, geographical coordinates are missing for 704 individuals. This leaves us with a sample of 3,621 deaths—1280 (35.35 percent) of blacks—recorded over the over 14-week span.[18]

Figure 2 plots the number of COVID-19 deaths in our sample for each day, separately for blacks and all other races combined. From March 16 to April 9 the daily number of blacks dying from COVID-19 is above the number for all other races, although the share of the black population is only about one fourth of the total population. By April 9, the cumulative share of blacks who have died from COVID-19 represents almost 58 percent of the total COVID-19 deaths. The daily number of deaths among blacks keeps increasing until mid-April and then starts decreasing at a slow pace. The number of deaths among other races, on the other hand, keeps increasing until mid-May. By May 16, the cumulative share of black COVID-19 deaths is down to about 39 percent, to decrease further to 35.3 percent by June 16. In other words, cumulatively to June 16,

[18]The racial distribution remains very similar, whether or not unreferenced individuals are included.

CEPR PRESS
Covid Economics 38, 16 July 2020: 129-195

blacks in our sample are dying at a rate 1.3 times higher than their 27 percent sample share in the population.

The above data clearly show that blacks are overrepresented in terms of COVID-19 deaths. Furthermore, the data document that blacks likely became infected, and eventually died, before the rest of the population, with a consequent decline in the share of cumulative black deaths as the epidemic followed its course—a trend that has been overlooked by media and public bodies reports.[19]

Figure A2 shows the distribution of deaths for blacks and all other races combined, by age group.[20] Blacks display a much larger number of deaths in their 60s and 70s, and a lower number past their 80s. Nevertheless, the age distribution below age 60 is quite similar. Overall, the graph confirms a large number of fatalities within the elderly population, a phenomenon which has been largely documented. Figure A3 shows the breakdown by gender among the two groups. Compared to other groups, blacks have a much higher probability of death among women (almost 48 percent against 39 percent).[21]

We extract information on comorbidities by generating a set of 14 dummy variables that take value one (and zero otherwise) when an individual who died from COVID-19 was affected by diabetes and/or asthma, liver disease, cancer, hypertension, kidney disease, obesity, respiratory diseases, neuro-cardiac diseases, neuro-respiratory diseases, asplenia, immunodeficiency, transplant, and heart diseases.[22] Figure A4 shows the distribution of deaths by comorbidity and race. Diabetes and hypertension are by far the two most common comorbidities. For both, blacks are more likely to suffer from them than the other races combined.

We spatially merge the death data from the Medical Examiner with census block group boundary files[23] and with the redlining maps produced by HOLC and georeferenced by the University of Richmond.[24] Figure 3 shows the result of the spatial merge. Each individual COVID-19 death is mapped into a specific block group and HOLC area using the georeferenced home address of the deceased. The map highlights block group boundaries,

[19]On April 7, on the basis of the Medical Examiner's data, the Chicago Tribune (Reys et al. 2020)—echoed by the Journal of the American Medical Association (Yancy, 2020) and the Lancet (Bhala, 2020)—reported that 68 percent of the dead in the City of Chicago involved African Americans, who represent about 30 percent of the city's population. As of June 16, African Americans account for less than 42 percent of the deaths in Chicago. Thus, the same trend can be detected both for the City of Chicago and Cook County as a whole.

[20]Only four deaths are reported below age 20, for three blacks aged 19, 18 and below 1, respectively, and a white aged 12. Figures are not normalized by the size of the population in each age group.

[21]Again, figures are not normalized by the degree of feminization of the population.

[22]Disease groupings followed those employed by OpenSAFELY Collaborative (2020). Groups are not mutually exclusive.

[23]A block groups represents a combination of census blocks and a subdivision of a census tract. A block group is defined to contain between 600 and 3,000 individuals.

[24]See Nelson et al. (2020) and https://dsl.richmond.edu/panorama/redlining/.

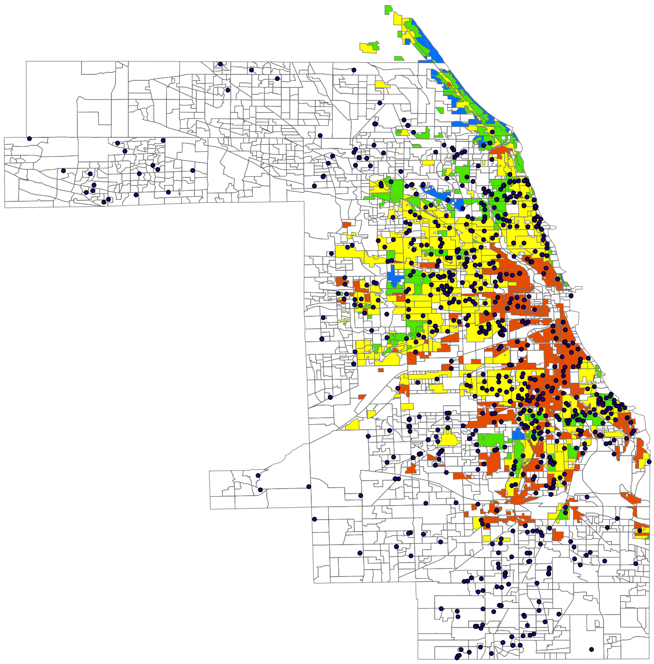


Figure 3: Cook County Map and Total COVID-19 Deaths, March 16-June 16, 2020

Note: The map reports census block group boundaries and HOLC areas, with green, blue, yellow, and red denoting respectively grade A, B, C, and D neighborhoods.

while HOLC areas are identified by the colour, with green, blue, yellow, and red denoting respectively grade A, B, C, and D.

From the Cook County Government[25] we also obtain census tract level data on socioeconomic characteristics (i.e., age, education, personal income, unemployment, population, and racial groups), averaged over the period 2014-2018. We match these data using for each block group the information available for the corresponding tract.

Figure A5 illustrates the share of black deaths by HOLC grade, only for census tracts where a COVID-19 death has been reported. In D-ranked neighborhoods blacks represent the highest share of the dead, while no black death is reported in those ranked A. Figure A6 illustrates the share of blacks in the population by HOLC grade. With the only exception of A-ranked neighborhoods, the share of the black population is lower than the corresponding share of black deaths shown in Figure A5. This preliminary evidence supports a role for redlining in determining segregation.

We complement the dataset with distance (measured in degrees) from a hospital and

[25]See Cook County Government Open Data at https://datacatalog.cookcountyil.gov/GIS-Maps/2010-U-S-Census-Mail-Return-Rates-and-Demographics/mpyu-4jqk. Data come from the American Community Survey estimates and are averages over the period 2014-2018 available by census tract

Covid Economics 38, 16 July 2020: 129-195

a nursing home.[26] We interpret the latter as a proxy for the likelihood that a COVID-19 death occurred in a nursing home.[27]

Lastly, for the heterogeneity analysis in Sub-section 8.2, we employ the CDC's Social Vulnerability Index (SVI) dataset,[28] created by the Geospatial Research, Analysis & Services Program (GRASP) run by the Agency for Toxic Substances and Disease Registry (ATSDR) with the scope of helping emergency response planners and public health officials in the face of a hazardous event, such as a natural disaster (e.g., a tornado or epidemic) or a human-made event (e.g., an oil spill). Social vulnerability refers to those factors that may affect the resilience of a geographical area to such an event. The source of the data used to obtain these indices is the American Community Survey, that provides averages for geographic areas over the period 2014-2018 at a census tract level. The SVI dataset include 15 characteristics, grouped into four indices: Socioeconomic Status (comprising income, poverty, employment, and education variables), Household Composition/Disability (comprising age, single parenting, and disability variables), Minority Status/Language (comprising race, ethnicity, and English language proficiency variables), and Housing/Transportation (comprising housing structure, crowding, and vehicle access variables). A general measure of social vulnerability, which summarizes the above four indices, is also provided.[29]

Tables A1 and A2 in the Appendix reports variable definitions and sources, for the cross section and the panel respectively.

5 Cross-sectional evidence

We start our analysis by exploiting cross-sectional information about individual deaths from COVID-19 occurring in Cook County between March 16, 2020 and June 16, 2020. Summary statistics for the cross-sectional sample are reported in Table A3.

A preliminary warning is in order, since the sample is clearly self-selected, as it includes only individuals who have died of COVID-19 and therefore exhibiting specific characteristics, which are precisely those that tend to be more prevalent among blacks.

[26]We take georeferenced hospital location from the Cook County Government (see Cook County Health and Hospitals Facilities at https://datacatalog.cookcountyil.gov/Economic-Development/Cook-County-Health-and-Hospitals-Facilities/jdix-z6uf) and georeferenced nursing home location from the Medicare Nursing Home Compare dataset (see https://data.medicare.gov/data/nursing-home-compare).

[27]The Illinois Department of Public Health estimates that about 40 percent of the COVID-19 deaths in Cook County occurred in a long-term facility. See http://www.dph.illinois.gov/covid19/long-term-care-facility-outbreaks-covid-19.

[28]See https://svi.cdc.gov/data-and-tools-download.html and Flanagan et al. (2011).

[29]Some of the variables, e.g., personal income and unemployment, coincide with those obtained through the Cook County Government from the same source, i.e., the American Community Survey.

Covid Economics 38, 16 July 2020: 129-195

Keeping in mind the sample selection problem afflicting the cross-sectional sample, in order to assess whether residence in a historically redlined area is a predictor of the probability that an individual that dies from COVID-19 is black, we employ as the outcome variable a dummy taking value one if the individual that died from COVID-19 in Cook County between March 16 and June 16 is reported to be black, and zero otherwise. Table A3 shows that the probability of a black death is 35.3 percent.

Table A3 also indicates, for each HOLC area, the probability that a COVID-19 death has occurred in that area. For instance, the probability that a COVID-19 death has occurred in A-graded neighborhoods is 0.2 percent.[30] Mean age is between 60 and 70 and 42 percent of the dead are females. Among pre-existing conditions, the disease with highest prevalence is hypertension, that affects 52 percent of the sample, followed by diabetes with nearly 41 percent.

The empirical model aims at exploiting the cross-sectional variation in the mortality of blacks across HOLC areas. Formally, we estimate the following model:

$$D_{i,d} = \lambda_d + \sum_{h=1}^{5} \beta_i \mathbb{1}\left(H_i = h\right) + X_i'\pi + Z_c'\rho + \mu_{i,d} \tag{1}$$

where $D_{i,d}$ is a dummy taking value one if individual i that died from COVID-19 in day d is black (and zero otherwise); λ_d represent day fixed effects that are meant to capture the daily variation in the number of deaths; $\sum_{h=1}^{5} \mathbb{1}\left(H_i = h\right)$ are a full set of dummies denoting the four HOLC-graded areas, from A to D, plus the ungraded area, where individual i used to reside; the vector X_i' includes a set of individual characteristics (age group, gender, comorbidities, and distance from the closest hospital and nursing home); the vector Z_c' includes a set of socioeconomic characteristics at tract level (the logarithmic of population and the shares of the population of age 18-64, over 65, without a high school diploma, and black); $\mu_{i,d}$ is the error term which we cluster at day level.

Table 1 reports OLS estimates for eight variants of Equation 1. In Model 1 we only control for HOLC (the omitted area is the ungraded area, No HOLC) and day fixed effects. In Model 2 we add demographic information on age and gender, in Model 3 comorbidities, in Model 4 population, in Model 5 the shares of the population aged 18-64 and over 65, in Model 6 the share of the population without a high school diploma, and in Model 7 distance from hospital and nursing home. In all models, the fact that an individual that died from COVID-19 used to live in HOLC areas D and C is positively and significantly associated with the probability that the individual is black, while the

[30]The reason why probabilities across the four HOLC areas do not sum up to one is that some neighborhoods of Cook County were not mapped by HOLC. 40 percent of the COVID-19 deaths occurred in ungraded neighborhoods of the county.

Covid Economics 38, 16 July 2020: 129-195

Covid Economics 38, 16 July 2020: 129-195

Table 1: Black COVID-19 Death, Cross-Sectional Results - Cook County, March 16-June 16, 2020

	Black COVID-19 Death							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HOLC A	-0.3808***	-0.3652***	-0.3244***	-0.3261***	-0.3247***	-0.3636***	-0.3514***	-0.1250***
	(0.0640)	(0.0720)	(0.0757)	(0.0800)	(0.0811)	(0.0798)	(0.0751)	(0.0424)
HOLC B	0.0389	0.0360	0.0399	0.0373	0.0355	0.0303	0.0072	0.0169
	(0.0329)	(0.0334)	(0.0327)	(0.0323)	(0.0327)	(0.0322)	(0.0316)	(0.0250)
HOLC C	0.0800***	0.0817***	0.0830***	0.0677***	0.0716***	0.1176***	0.0848***	-0.0146
	(0.0177)	(0.0185)	(0.0181)	(0.0175)	(0.0202)	(0.0216)	(0.0234)	(0.0182)
HOLC D	0.1752***	0.1769***	0.1745***	0.0833***	0.0852***	0.1232***	0.0922***	0.0106
	(0.0233)	(0.0238)	(0.0231)	(0.0255)	(0.0266)	(0.0281)	(0.0297)	(0.0222)
Age Groups	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Female	No	Yes	Yes	Yes	Yes	No	Yes	Yes
Comorbidities	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Population (log)	No	No	No	Yes	Yes	Yes	Yes	Yes
Share Aged 18-64	No	No	No	No	Yes	Yes	Yes	Yes
Share Aged 65+	No	No	No	No	Yes	Yes	Yes	Yes
Share With No High School	No	No	No	No	No	Yes	Yes	Yes
Distance From Hospital	No	No	No	No	No	No	Yes	Yes
Distance From Nursing Home	No	No	No	No	No	No	Yes	Yes
Black Share	No	No	No	No	No	No	No	Yes
Day FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj.R-squared	0.056	0.069	0.090	0.111	0.111	0.132	0.144	0.475
Observations	3618	3616	3616	3616	3616	3616	3616	3616

Note: The dependent variable is a dummy variable that takes value one if an individual who died from COVID-19 is black, an zero otherwise. The omitted area is the one ungraded by the HOLC. Robust standard errors clustered at a day level in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

opposite is true for area A. In other words, black mortality is much larger in low-graded areas. To be noticed is that the effect is statistically equally significant for areas C and D, even though the size of the coefficient is larger for D.[31] In terms of magnitudes, relative to the sample mean (35.3), the probability that an individual that died from COVID-19 is black is over 24 percent larger in area C and 26 percent larger in area D. However, when in Model 8 we also add the share of blacks in the census tract, having been a resident in such areas is no longer a predictor of the dead individual's racial group, suggesting that indeed historical redlining induced segregation along racial lines, that fully absorbs its influence in the present day. In other words, the black share fully captures the effect of residence in a low-graded area, because of the correlation between black share and the latter.

It is instructive to report how the dependent variable is affected by other covariates (all coefficients are reported in Table A4). For instance, even in the full specification where we control for the black share, its likelihood is higher for females, which means that, within the black population, women have a relatively higher chance to die from COVID-19, relative to the chance they have within the total population. Furthermore, pre-existing conditions, in particular hypertension and kidney and respiratory diseases among those with high prevalence, exert a significantly positive effect on the dependent variable. Distance from hospital decreases the probability of a black death although it becomes

[31]The similarity between the effects of red and yellowlining is consistent with the evidence presented by Aaronson et al. (2017) over a sample of US cities.

CEPR PRESS

Covid Economics 38, 16 July 2020: 129-195

insignificant once we control for the share of blacks. Distance from a nursing home, that is, a lower probability that a death has occurred in a nursing home, increases the probability of a black death but, once race is accounted for, the sign of the association is reversed, which can be attributed to a negative correlation between the presence of nursing homes and the black share of a neighborhood. Nevertheless, neither comorbidities nor the other demographic and socioeconomic factors fully absorb the influence of redlining, at least until the black share is accounted for.

Overall, the cross-sectional results confirm that yellow and redlined areas are associated with a higher incidence of COVID-19 deaths among blacks, and that the difference between redlined and yellowlined areas is blurred. However, since blacks are concentrated in the areas where mortality is higher, the cross-sectional analysis solely based on information about those that died from COVID-19 is severely biased because of sample selection. This limitation motivates the event study approach we introduce in the next section.

6 An event study approach

6.1 Empirical strategy

The cross-sectional approach only provides simple correlations with no causal implications because of issues related to sample selection and omitted variable bias, which may confound pre-existing differences between areas with and without COVID-19 mortality. To reduce such biases, we assemble a weekly balanced panel at block group level over the period from January 1, 2020 to June 16, 2020.[32] For these 24 weeks, we collect information on all types of deaths, that is, from COVID-19 and other causes under the jurisdiction of the Cook County Medical Examiner's Officer. Between January 1 and June 16, 6,753 deaths—of which 5,492 after March 16[33]—are reported, each associated with home residence, geographical coordinates, and individual characteristics.[34] Again

[32]The merged dataset includes over 3,992 block groups and 1,318 census tracts.

[33]This implies that between March 16 and June 16 there are more than 1,871 deaths in addition to COVID-19 related deaths, which is consistent with the 1,256 deaths reported before March 16.

[34]As we examine how mortality evolves before and after the epidemic outbreak, we should keep in mind that, generally speaking, officially reported COVID-19 deaths do not match excess deaths, as measured by the gap between observed deaths after the epidemic outbreak and deaths observed during the same period in normal years. However, since our source is the Medical Examiner, the discrepancy between recorded COVID-19 and excess deaths is greatly alleviated, because of the specific nature of the deaths under his jurisdiction. Consequently, the deaths data we use in the cross-sectional analysis and those we use in the event study are directly comparable, provided that it remains possible that deaths from other causes under the Medical Examiner's jurisdiction may also have risen or declined for various reasons (e.g., more individuals may have died because they did not receive care for other diseases, but fewer may have died for car accidents due to the lockdown). Indeed the Medical Examiner reports, to June 16,

Covid Economics 38, 16 July 2020: 129-195

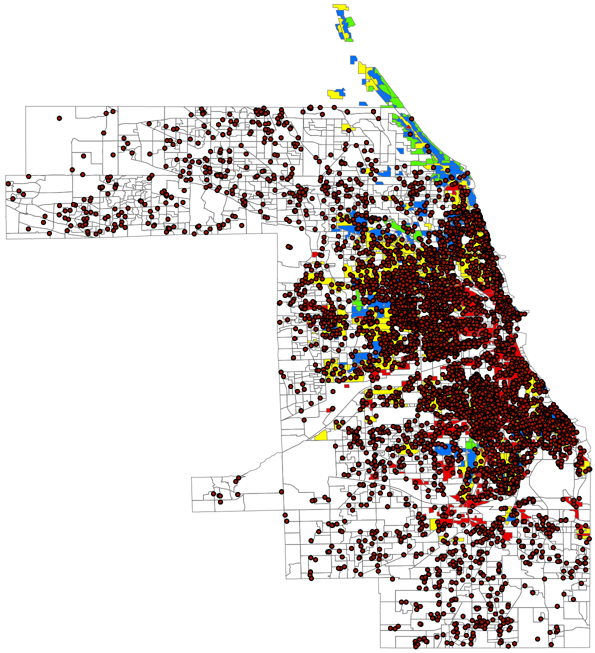


Figure 4: Cook County Map and Total Deaths, January 1-June 16, 2020

Note: The map reports census block group boundaries and HOLC areas, with green, blue, yellow, and red denoting respectively grade A, B, C, and D neighborhoods.

we map individual deaths into census block groups and HOLC areas and then we aggregate them at a census block group and by week. Therefore, for each block group-week, we gather information on reported number of deaths (if any) for a given block group in any of the 24 weeks from January 1 to June 16. Figure 4 illustrates the spatial merge. The data is then merged with the previously described socioeconomic characteristics provided at census tract level by the Cook County Government and with proxies for social vulnerability to shocks from the CDC.

Aggregating at a block group level has clear advantages in terms of identification, since it allows to test for pre-treatment differences while controlling for block group fixed effects which should filter out the effect of all socioeconomic factors which do not change over this 24-week period. However, the aggregation implies that a given block group may overlap multiple HOLC neighborhoods, possibly assigned to different grades. Figure A7 zooms into the map and displays as an example four block groups, each including neighborhoods that belong to three different HOLC types. In each block group, two neighborhoods are graded C and outlined in yellow, and one is ungraded (the white

2020, over 3,000 excess deaths compared with the same period in 2019, a figure that closely matches the COVID-19 deaths that were recorded.

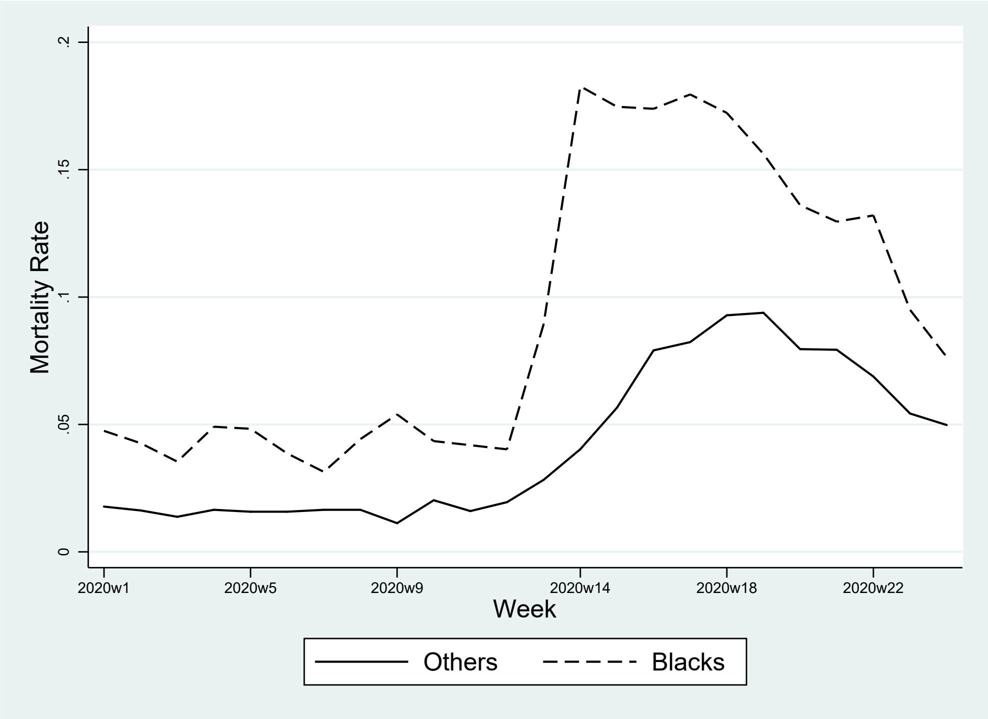


Figure 5: Mortality Rate, by Race - Cook County, January 1-June 16, 2020

Note: The figure reports mortality rates from any cause of death by week, separately for blacks and all other races combined.

one). The distribution of HOLC neighborhoods by block group is shown in Figure A8. Almost 80 percent of the block groups in our sample include no more than two HOLC neighborhoods (potentially with the same HOLC grade) and almost 95 percent of the block groups include no more than three, with the rest of the block groups including up to ten different HOLC neighborhoods. Thus, in order to determine treated block groups, we will focus on those for which the majority of the surface area falls into HOLC areas graded either C or D. This is consistent with the cross-sectional evidence, according to which redlined and yellowlined areas exert a similar effect on the probability that an individual that died from COVID-19 is black, and also with the evidence reported by Aaronson et al. (2017), who stress the relevance of yellowlining for racial segregation. The treatment therefore will capture whether resilience to external shocks, namely, COVID-19, is affected by the historical policies which have favored racial segregation and economic discrimination.

Figure 5 plots the mortality rate from any cause, separately for blacks and all other

Covid Economics 38, 16 July 2020: 129-195

CEPR PRESS

Covid Economics 38, 16 July 2020: 129-195

races combined, for each week in the sample.[35] From the beginning of the sample period, the mortality rate is higher for blacks. For the first 10 weeks, before the epidemic outbreak, average mortality for blacks fluctuates between 0.04 and 0.05 deaths per thousand versus around 0.02 per thousand for other races. Starting from the eleventh week (i.e., the week of March 13, that is the week when the first COVID-19 death is recorded on March 16), mortality soars among both groups, but much more steeply so for blacks. Both curves seem to be reaching a plateau by week 18 and start converging toward to pre-COVID mortality rates by week 24.

Summary statistics for the panel dataset, including socioeconomic covariates, are reported in Table A5. On average, for each block group-week in the sample, 0.070 deaths—0.029 (41.4 percent) of blacks—are recorded during the period under examination. The probability of a death in a given block group-week (measured as a dummy variable) is 5.7 percent, while the probability of a black death is 2.5 percent. Treated block groups (i.e., those predominantly graded C or D) represent almost 62 percent of the sample. Social vulnerability indices represent percentile ranking of geographical areas depending on 15 categories and are bounded between 0 and 1, with 1 representing the maximal level of vulnerability (the last percentile). A socioeconomic vulnerability of 0.54 therefore denotes that the average census tract within Cook County falls approximately within the 50th percentile of the distribution of vulnerability.

Figure A9 summarizes a number of demographic and economic characteristics by HOLC grade. Personal income in red and yellowlined neighborhoods is well below $40,000 and even lower in the latter. A similar pattern emerges for the share of population with no high school diploma, which is highest in C-graded neighborhoods. The share of the black population is highest in redlined neighborhoods (close to 40 percent), while the rate of black mortality peaks in yellowlined ones, with an average of two deaths per thousand of blacks. However, in reading mortality statistics one must consider the fact that the denominator (the black population) is not constant across the four areas. Figure A10 plots the weekly mortality rate for blacks and the other races by HOLC grade. Mortality increases with the outbreak of the COVID-19 epidemic in all areas. The spikes we observe for black mortality in A-graded neighborhoods are due to the very small share of the black population. In neighborhoods belonging to the other three areas, black mortality is always higher than the one for other races and particularly high in yellowlined areas, although one must consider that the black population in yellowlined areas is much smaller than in redlined ones, as shown in the previous figure.

Our goal is to capture the impact of the asymmetric shock introduced by COVID-19 on

[35]To compute mortality rates, i.e., number of deaths over population, we first sum weekly deaths at a census tract level (because population data is at a tract level) and then we collapse by week.

historically segregated areas, that is, whether treated block groups react differently after treatment initiation. In other words, we aim at establishing whether, after the outbreak of the epidemic, in majority C and D block groups deaths deviate more significantly from those recorded in the pre-treatment period. Thus, we estimate variants of the model below:

$$Y_{i,t} = \delta_i + \gamma_t + \sum_{k=-10}^{13} \beta_k \mathbb{1}\left(K_{i,t} = k\right) + \varepsilon_{i,t} \tag{2}$$

where $Y_{i,t}$ represents the number of deaths in block group i and week t;[36] δ_i and γ_t are block group and week fixed effects meant to control for block group characteristics which are fixed over the 24-week period and to capture natural fluctuations in mortality as well as policies that affect the county uniformly (namely, the lockdown); $\mathbb{1}(K_{i,t} = k)$ denote treated block groups for $k_{i,t} = -10, -9, ..., 13$ periods (i.e., weeks) before and after the treatment kicks in, with the β_ks for $k < 0$ corresponding to pre-treatment effects and the β_ks for $k \geq 0$ corresponding to the dynamic effects k periods relative to the event. We omit the period before the treatment kicks in, i.e., week 10. The error term, $\varepsilon_{i,t}$, is clustered at block group level.

The event study approach outlined in Equation 2 allows to alleviate some of the shortcomings affecting the cross-sectional analysis. The inclusion of the pre-treatment periods allows to test for the parallel trend assumption and therefore for potential secular differences between treated and control groups, as well as the occurrence of self-selection, that may lead to different rates of disease transmission between treated and untreated group. Issues related to sample selection will also be ruled out, since the sample includes the universe of the block groups in the county. Learning and adaptation to the treatment before it kicks in, as it occurs with staggered treatment, is also unlikely given that the treatment period is constant. The only potential source of bias is therefore likely to be related to measurement error, since not all the deaths that occur in Cook County are reported to the Medical Examiner. Measurement error may be more severe in areas with higher mortality but, since this is likely to cause an attenuation bias, the resulting estimator would produce more conservative estimates.

[36] The alternative would be to use the mortality rate as a dependent variable. However, we have racially-disaggregated data on population only at a census tract level rather than at block group level. In addition, block group fixed effects should absorb differences in population, given that the latter should remain constant over the span of only 24 weeks. In other words, using mortality data would amount to divide deaths by a constant, which should be picked up by the fixed effect.

Covid Economics 38, 16 July 2020: 129-195

Covid Economics 38, 16 July 2020: 129-195

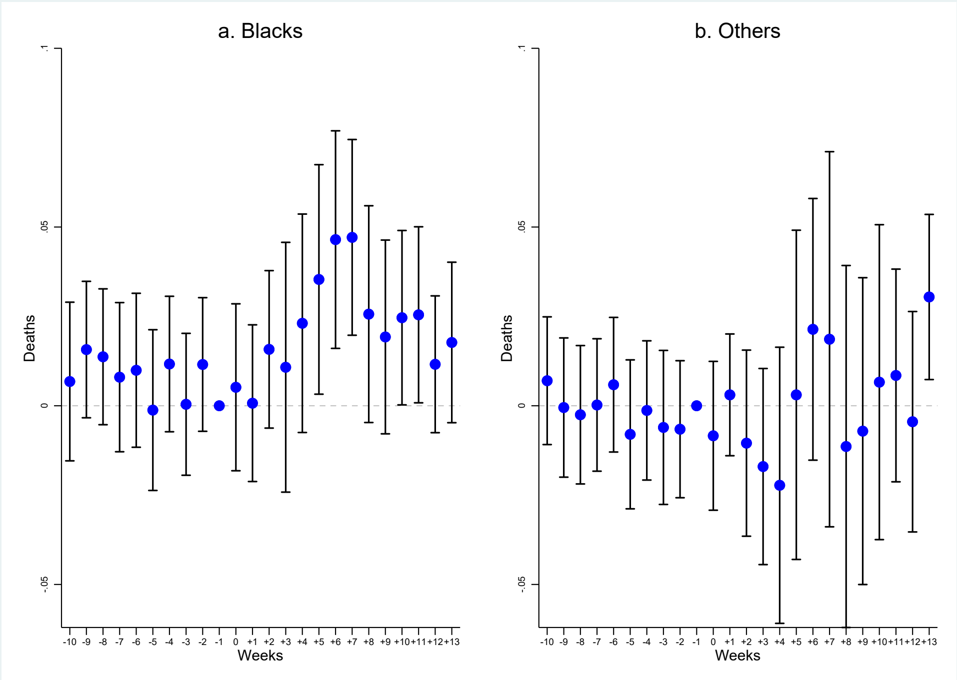


Figure 6: Dynamic Effect of the Treatment on Deaths, by Race - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of deaths, of blacks (Panel a) and of other groups (Panel b). The coefficients are least-squares estimates of the β_ks. Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

6.2 Baseline results

As shown in Figure 1, Cook County was only partially mapped by the HOLC in the 1930s, since the Corporation focused on cities with a population in 1930 above 40,000. It is therefore possible that those neighborhoods which were mapped had completely different characteristics from those which were not, and that these differences in pre-existing conditions may be somewhat correlated with the treatment. To minimize the possibility of biases arising from comparing areas that were already different, we focus on block groups within municipalities which were mapped by the HOLC.[37]

Figure 6 illustrates the dynamic treatment effect on the number of deaths for blacks (Panel a) and all other races (Panel b). The coefficients are least-squares estimates of the β_ks and vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. Before the COVID-19 outbreak, the average number of

[37]A robustness check that includes ungraded neighborhoods actually leads to even sharper results, reported in Figure A11 that are omitted for brevity and available upon request.

deaths for blacks and all other races in treated block groups is not significantly different from the average number of deaths in non-treated ones. However, after the epidemic shock, mortality in treated block groups increases quite sharply and this is particularly evident among African Americans, for whom the number of deaths in treated block groups increases up to 0.05 per week, an over 17 percent increase with respect to the average number of black deaths by week and by block group. The effect starts picking up in period 4 (toward mid April, when overall mortality sharply increases). Four weeks after the outbreak, the number of deaths among African Americans increases by almost 0.03 per week and continues rising for the next few weeks before it starts fading away toward the end of the 24-week period. By contrast, deaths for other races are not significantly affected by the treatment.[38]

To quantify the overall effect of the treatment, in Table A6 we estimate the Average Treatment Effect on the Treated (ATET) for blacks and the other races using a simple-difference-in differences method. Overall, we find that the number of black deaths in treated block groups after the epidemic shock (relative to non-treated blocks) increases by 0.015, while the effect for the other races is not statically significant (and negative). If compared to the average number of deaths for the untreated sample in the post-treatment period (i.e., 0.025), the implied magnitude of the estimated effect is an increase in the number of black deaths by almost 60 percent.

7 Robustness

7.1 Allowing for differential trends

The evidence presented in the previous section reveals an asymmetric effect of the epidemic shock and suggests that the relative resilience to it depends on the level of historical racial segregation induced by loan market discrimination as a result of the HOLC Residential Security Maps. An alternative explanation for the positive effect we detect for the treatment points to potential differential trends which may have placed treated and untreated neighborhoods on different trajectories that then surfaced only when the shock struck. To control for such differential trends, we use tract level socioeconomic and de-

[38]Similar results obtain, as shown in Figure A12, if we replace the dependent variable with a dummy variable capturing the occurrence of a death. This is also addressing the issue raised in footnote 36, since the definition of the dependent variable as a dummy does not reflect the distinction between number of deaths and mortality rate. Other robustness checks which we do not report for brevity include slightly changing the definition of the treatment group (Majority C & D may include some A and B neighborhoods, so we also redefine the treatment group by excluding As, thus making the criterion more stringent) and using distributed and error lag models to respectively account for spillovers and serial correlation in the error.

Covid Economics 38, 16 July 2020: 129-195

Covid Economics 38, 16 July 2020: 129-195

mographic controls, that we interact with week dummies in order to provide the time variation we aim at exploiting. The result of this exercise is reported in Figure A13 for black deaths. In Panel a we interact with week dummies the share of the population aged 65 and above. In Panel b we further add interactions with the shares of Chinese speakers and Hispanics. In Panel c we keep adding differential trends depending on unemployment and income. In the last three plots we insert interactions with the share with no high school diploma (Panel d), the share of blacks (Panel e), and distance from hospital and nursing home (Panel f). Differential trends in income and unemployment partially offset the effect of the treatment, but we can still observe for it a sizeable (although diminished) effect on the number of deaths.

7.2 Robustness to aggregation

As mentioned above, aggregating at a block group level allows us to tighten the identification of the effects. However, the definition of the treatment may be blurred by the fact that a block group can overlap multiple HOLC areas (as in the example shown in Figure A7). To explore the extent of issues that may be raising due to our aggregation approach, we perform a sensitivity test and split the sample between block groups including a single HOLC neighborhood and block groups including up to two, up to three, and up to four HOLC neighborhoods. Block groups that only include a single HOLC neighborhood are not affected by treatment definition issues due to aggregation, since each of them falls entirely within a given HOLC category. As a result, comparing results obtained for these block groups with those for block groups that include up to two, three and four neighborhoods will allow us to understand whether aggregation is truly a problem. Reassuringly, the effect of the treatment does not change sensibly whenever we include blocks which include multiple HOLC neighborhoods (Figure A14).

8 Heterogeneity

8.1 Heterogeneity by HOLC grade

In order to evaluate factors that are potentially correlated with the partitioning of cities by the HOLC and may explain the effect we found, this section goes on to investigate how the effect of the treatment varies with specific characteristics. We start with a heterogeneity analysis by HOLC grade. Specifically, we consider sub-samples of neighborhoods, defined on the basis of their HOLC ranking.

Covid Economics 38, 16 July 2020: 129-195

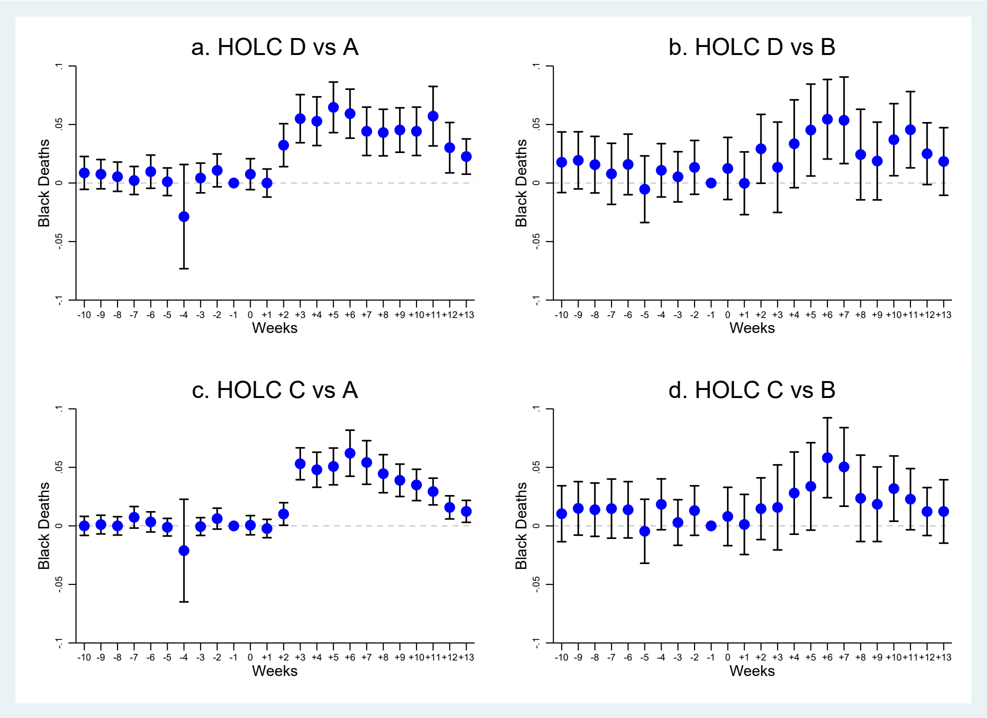


Figure 7: Heterogeneity by HOLC Grade - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of black deaths. The coefficients are least-squares estimates of the β_ks over samples of block groups for which the majority of the surface area falls in either a grade A or a grade D area (Panel a), in either a grade B or a grade D area (Panel b), in either a grade A or a grade C area (Panel c), and in either a grade B or a grade C area (Panel d). Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

Results are shown in Figure 7. In Panel a we focus only on block groups for which the majority of the surface area falls in either a grade A or a grade D area.[39] In other words, we compare the best and the worst neighborhoods (according to the HOLC grading scheme). In Panel b we compare the second best neighborhoods (HOLC grade B) with the worst (grade D). In Panel c we compare the best neighborhoods (grade A) with the second worst (grade C) and in Panel d the second best (grade B) with the second worst (grade C). The effect of the treatment is much larger when we compare the worst and the second worst neighborhoods with the best (respectively Panels a and c), and becomes significant from period 2 in Panel a and from period 3 in Panel c. However, even when we compare the worst and the second worst neighborhoods with the second best (respectively Panel b and Panel d), we still observe a quite significant effect of the treatment for several weeks.

[39]This means that we drop block groups for which the majority of the surface area falls in grade B and C areas.

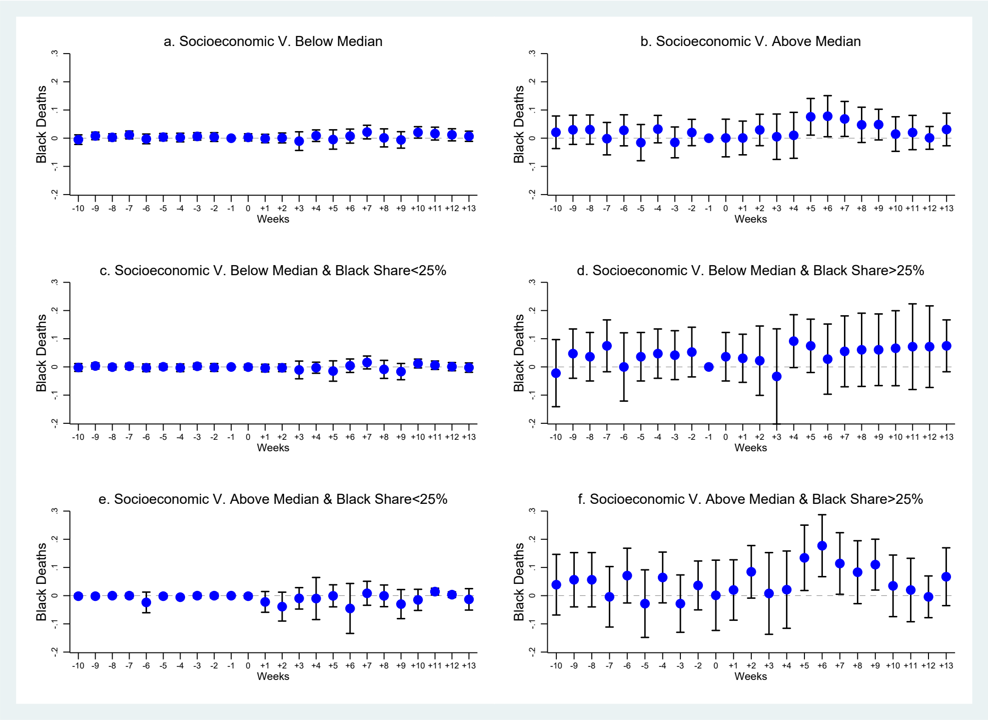


Figure 8: Heterogeneity by Socioeconomic Status Vulnerability - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of black deaths. The coefficients are least-squares estimates of the β_ks over samples of block groups with socioeconomic status vulnerability below median (Panel a), above median (Panel b), below median and with black share below 25 percent (Panel c), below median and with black share above 25 percent (Panel d), above median and with black share below 25 percent (Panel e), and above median and with black share above 25 percent (Panel f). Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

8.2 Heterogeneity by social vulnerability

The social vulnerability indices collected at a census tract level by the CDC allow us to carry out a heterogeneity analysis depending on the level of vulnerability to shocks of specific areas. Thus, in this section we group neighborhoods on the basis of the four distinct dimensions of social vulnerability, as well as some of their components.

Figure 8 shows heterogeneity results by socioeconomic status vulnerability, an index that comprises four components, reflecting percentile scores for personal income, poverty, unemployment, and education. In Panels a and b we split the sample between block groups with a value of the socioeconomic vulnerability index below and above the median. In the next two panels (Panels c and d) we split block groups below median socioeconomic vulnerability between those with a population black share respectively below and above

Covid Economics 38, 16 July 2020: 129-195

Covid Economics 38, 16 July 2020: 129-195

25 percent.[40] In the two bottom panels (Panels e and f) we repeat the same exercise for block groups with above median socioeconomic vulnerability. Unsurprisingly, the effect of the treatment is much larger (and significant) in block groups with above median socioeconomic vulnerability (Panel b). The estimated dynamic treatment effect in the sample of block groups below median is nearly zero and not significant (at a 5 percent level), despite the fact that standard errors are exceptionally small. When we split the two sub-samples between those with a share of blacks below (Panels c and e) and above (Panels d and f) 25 percent, some noteworthy differences emerge. The treatment effect is much larger in neighborhoods with a black share above 25 percent, which is not surprising given that the black share is this sub-sample is larger. However, among these neighborhoods, the treatment effect changes quite significantly depending on the level of socioeconomic vulnerability (Panel d vs Panel f). For areas below median vulnerability (Panel d), the post-treatment effect of being red or yellowlined is not that different from the pre-treatment, and in any case not significantly so at a 5 percent level. When, on the other hand, we look at the sample above median vulnerability, that is, at neighborhoods with worse performances, we observe quite a strong treatment effect, which resembles the one we found over the whole sample. These dissimilarities point to a much stronger impact of socioeconomic determinants, rather than genetic and biological one, of black mortality. In other words, neighborhoods with the same black share exhibit different effects of the treatment, and therefore different level of resilience to external shocks, depending on the level of socioeconomic vulnerability.

To better understand which component of the socioeconomic vulnerability index drives our findings, in Figures A15-A18 we replicate the same analysis focusing, one by one, on its four components, that is personal income, poverty, unemployment, and education. Overall, the effect of the treatment is much stronger in neighborhoods that underperform in all four dimensions and at the same time exhibit a black share above 25 percent. However, when we focus on neighborhoods with a black share above 25 percent, the difference between the worst and the best performing neighborhoods is especially striking for the income and poverty dimensions. The treatment effect for neighborhoods below median income (Figure A15, Panel d) is large and significant, while above median income (Panel f) the treatment effect is not significant in most post-treatment periods. The same occurs for poverty (Figure A16), with a relatively small (and significant in one period only) average treatment effect in the post-treatment period for neighborhoods with a population share below the poverty line smaller than the median (Panel d), and a sizeable effect for those with a larger one (Panel f).

[40] We choose this threshold because it is close to the average black share in the county. Furthermore, a higher threshold would greatly reduce the size of the sub-sample displaying a relatively higher black share.

Covid Economics 38, 16 July 2020: 129-195

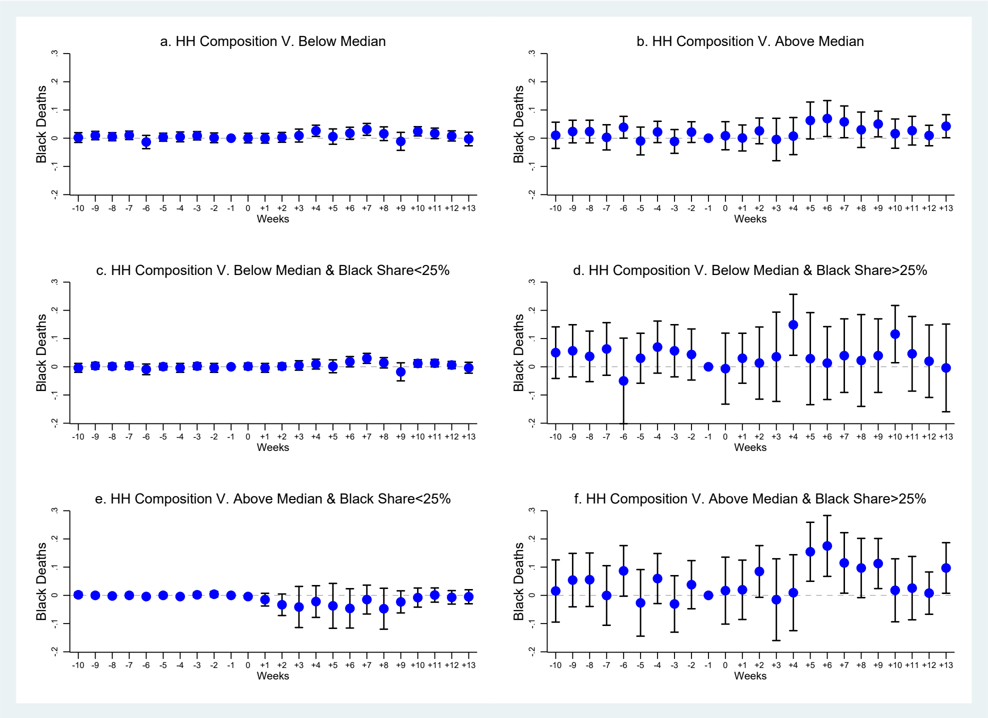


Figure 9: Heterogeneity by Household Composition Vulnerability - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of black deaths. The coefficients are least-squares estimates of the β_ks over samples of block groups with household composition vulnerability below median (Panel a), above median (Panel b), below median and with black share below 25 percent (Panel c), below median and with black share above 25 percent (Panel d), above median and with black share below 25 percent (Panel e), and above median and with black share above 25 percent (Panel f). Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

In Figure 9 we repeat the same analysis with a focus on the household composition vulnerability index, which includes four components, that is the shares of the population over 65, below 17, older than 5 with disabilities, and of single parents. Again the effect is much more marked when we focus on block groups with above median household vulnerability and a black share above 25 percent (Panels d and f). Once again, to understand which component of the index drives the effect, in Figure A19 we split the sample between block groups with a share of population aged 65+ below and above the median, while in Figure A19 we do so according to the share of single parents. As expected, aging is quite an important factor. Consistent with the pattern reported for blacks in Figure A2, fatalities among the elderly are more numerous when we look at the sample with a black share above 25 percent (Panels d and f). For the share of single parents in Figure A20, the difference in the estimated effect between Panel d and Panel

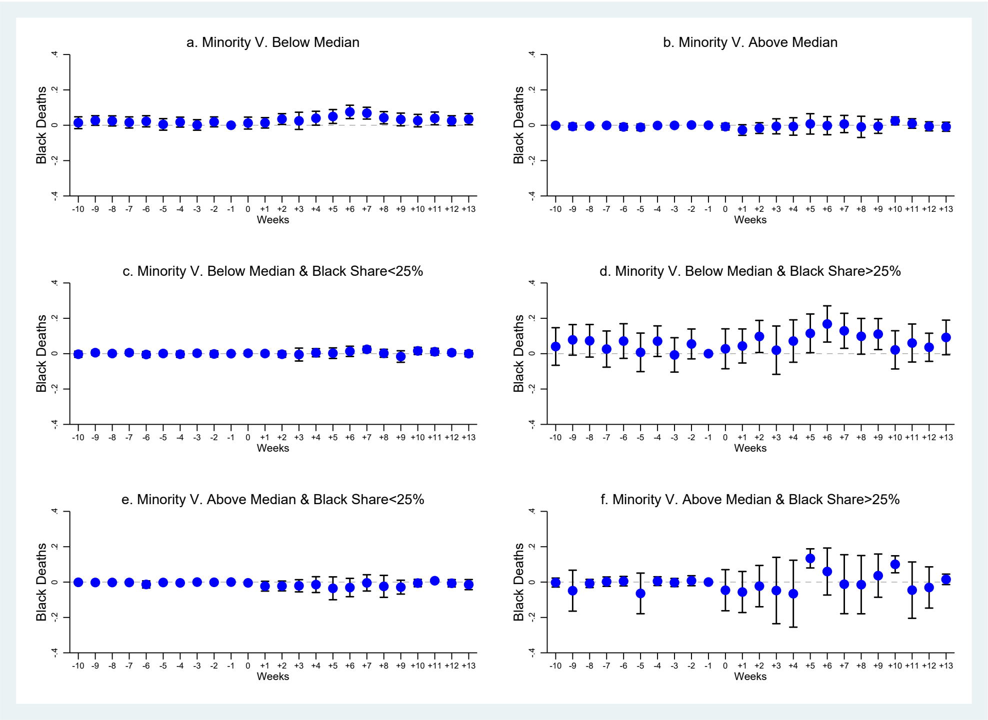


Figure 10: Heterogeneity by Minority Status Vulnerability - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of black deaths. The coefficients are least-squares estimates of the β_ks over samples of block groups with minority status vulnerability below median (Panel a), above median (Panel b), below median and with black share below 25 percent (Panel c), below median and with black share above 25 percent (Panel d), above median and with black share below 25 percent (Panel e), and above median and with black share above 25 percent (Panel f). Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

f is sizeable. In other words, the effect of yellow and redlining is most detrimental when a high share of single parents is combined with a high black share.

In Figure 10 we repeat the heterogeneity analysis by vulnerability depending on minority status (the index also comprises a measure of English language proficiency) and in Figure 11 by housing/transportation vulnerability (the index comprises percentages of multi-unit structures, mobile homes, a measure of crowding based on the presence of more people than rooms, households with no vehicle, and households in group quarters). Differences in the treatment depending on vulnerability along these two dimensions, and combined with the black share, are still detectable but less pronounced than those we found for socioeconomic status and household composition vulnerability.

Covid Economics 38, 16 July 2020: 129-195

Covid Economics 38, 16 July 2020: 129-195

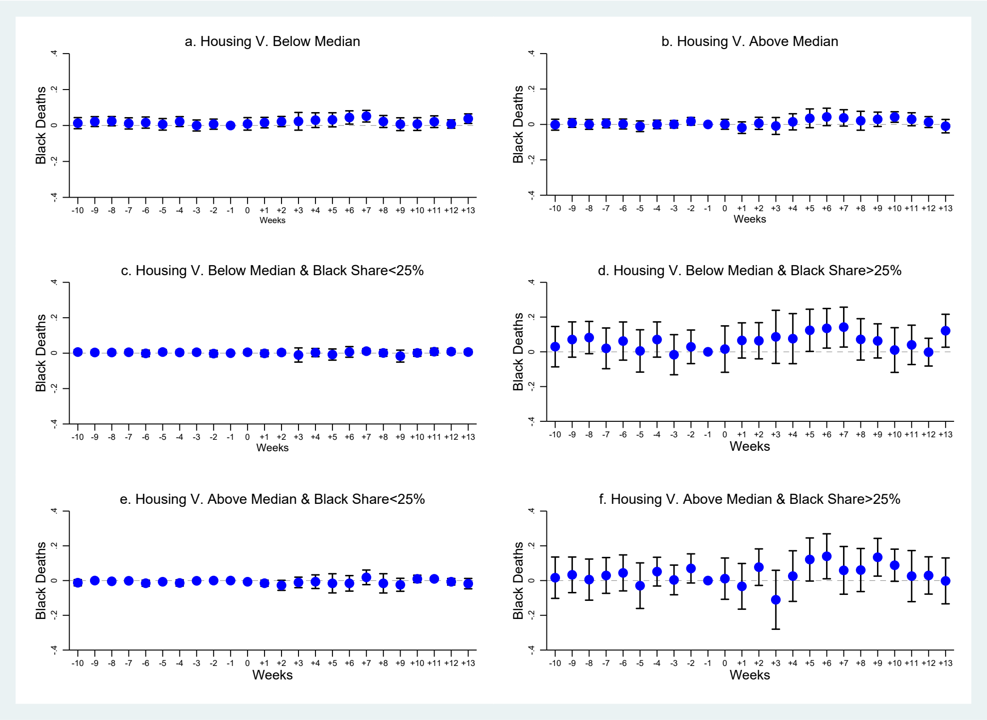


Figure 11: Heterogeneity by Housing/Transportation Vulnerability - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of black deaths. The coefficients are least-squares estimates of the β_ks over samples of block groups with housing/transportation vulnerability below median (Panel a), above median (Panel b), below median and with black share below 25 percent (Panel c), below median and with black share above 25 percent (Panel d), above median and with black share below 25 percent (Panel e), and above median and with black share above 25 percent (Panel f). Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

8.3 Heterogeneity by nursing home location

Lastly, the high levels of mortality recorded in nursing homes have been stressed both by the media and the medical literature. To explore the relevance of this potential channel, using the information on nursing home location available from Medicare[41] we generate the minimal centroid distance of a block group from a nursing home and we replicate the heterogeneity analysis focusing on block groups located within 0.009 degrees (i.e., 1 km) from a nursing home. Our aim is to capture the probability that a recorded death has occurred in a nursing home. Figure 12 shows that the treatment effect is much stronger in neighborhoods further away from nursing homes and that this is especially true when the black share is above 25 percent. Thus, we can confidently exclude potential concerns related to the possibility that we could have captured the effect of deaths in nursing

[41] See https://data.medicare.gov/data/nursing-home-compare.

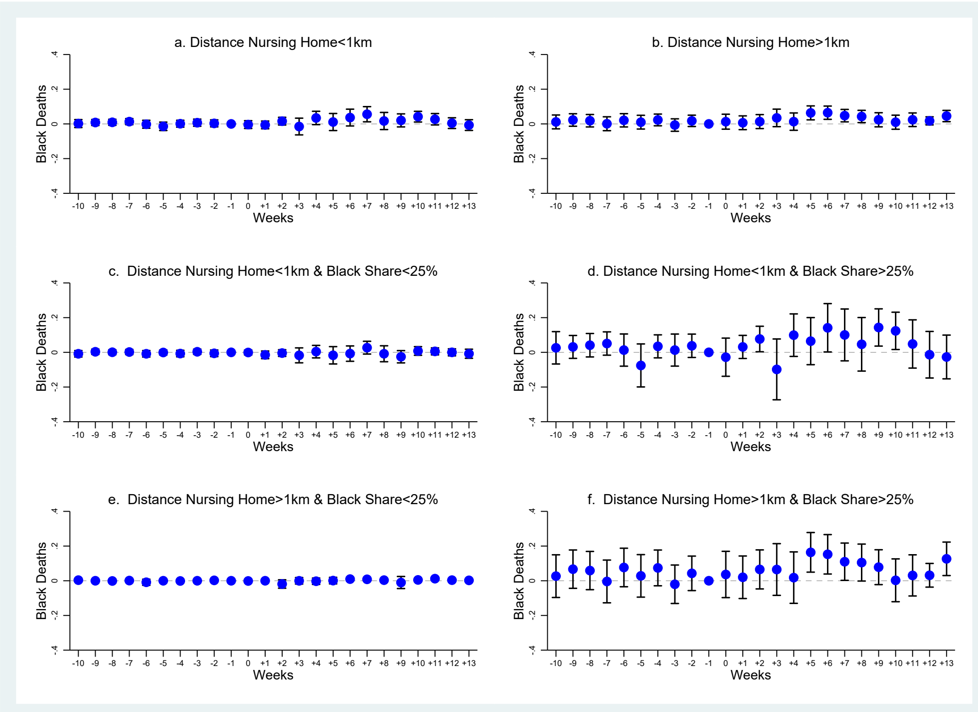


Figure 12: Heterogeneity by Distance from Nursing Home - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of black deaths. The coefficients are least-squares estimates of the β_ks over samples of block groups with distance from nursing home below 1 km (Panel a), above 1 km (Panel b), below 1 km and with black share below 25 percent (Panel c), below 1 km and with black share above 25 percent (Panel d), above median and with black share below 25 percent (Panel e), and above median and with black share above 25 percent (Panel f). Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

homes.

9 The Hispanic population

While our main focus so far has been on how the African American population has been hit by COVID-19, the Hispanic population has also been the subject of concern, both in the media and the medical literature.[42] Therefore, in this section, we extend the previous analysis of COVID-19 outcomes to the white Hispanic population of Cook County.[43]

[42]See, for instance, Singh and Koran (2020) and Yancy (2020).

[43]The Medical Examiner's racial classifications are based on US Census Bureau categories, according to which Latino, or equivalently Hispanic, is defined as ethnicity, and can therefore belong to any racial group. This section focuses on white Hispanics, who in Cook County represent the vast majority of Hispanics. To June 16, only 13 deaths were reported for black Hispanics, i.e., 0.09 percent of the black

Covid Economics 38, 16 July 2020: 129-195

CEPR PRESS

The history of Latino immigration to Cook County starts at least as early as in the period 1916-1928, when a steady and large flow of Mexicans moved to Chicago to find work in the railroad and steel industries. Another wave took place in the period 1942-1964. Their settlement pattern was similar to that of blacks and they were also historically affected by segregation and redlining (Betancur, 1996). In fact Hoyt (1933) placed Mexicans last, after blacks, in his ranking of the influence of ethnic groups on property values.

Figure A21 is a replica of Figure 2 that also reports Hispanic COVID-19 deaths, as well as blacks and the remaining groups combined, from March 16 to June 16. Even though the number of Hispanic deaths does increase in the initial weeks, it stays below the number of black deaths. Figure A22 is a replica of Figure 5 that plots the Hispanic overall mortality rate from January 1 to June 16. Strikingly, before the epidemic, the mortality rate for Hispanic is smaller than that of blacks and even of other groups combined. It does increase with the COVID-19 outbreak, but remains relatively contained if compared to other groups.

The fact that Hispanics exhibit lower levels of mortality than the rest of the population is actually a well-known fact, that has been referred to as the "Latino paradox", as Hispanics tend to display relatively favorable health outcomes despite their low socioeconomic status (Markides and Coreil, 1986; Abraido-Lanza et al., 1999).

Nevertheless, when in Figure A23 we replicate Figure 6 with the baseline event study, including results for Hispanics, we find that they are also affected by the treatment, even though the effect kicks in later, and lasts longer, if compared to blacks. This confirms that HOLC practices carry long-term implications, in terms of the induced resilience to the epidemic shock, also for Hispanics.

10 Falsification tests

10.1 From 2017 to 2019

There are still additional potential threats to identification. To test the extent of such threats, in this section we will carry out two sorts of falsification tests.

First, it might be the case that the treatment captures endemic annual trends in mortality that would have anyway occurred and that have nothing to do with the COVID-19 epidemic. To test the extent of such threat to identification, we carry out falsification tests using the number of deaths in the corresponding time frame for the three years prior to the epidemic, that is, from January 1 and June 16 in 2017, 2018, and 2019. The

deaths, and 1.6 percent of the Hispanic deaths.

Covid Economics 38, 16 July 2020: 129-195

Covid Economics 38, 16 July 2020: 129-195

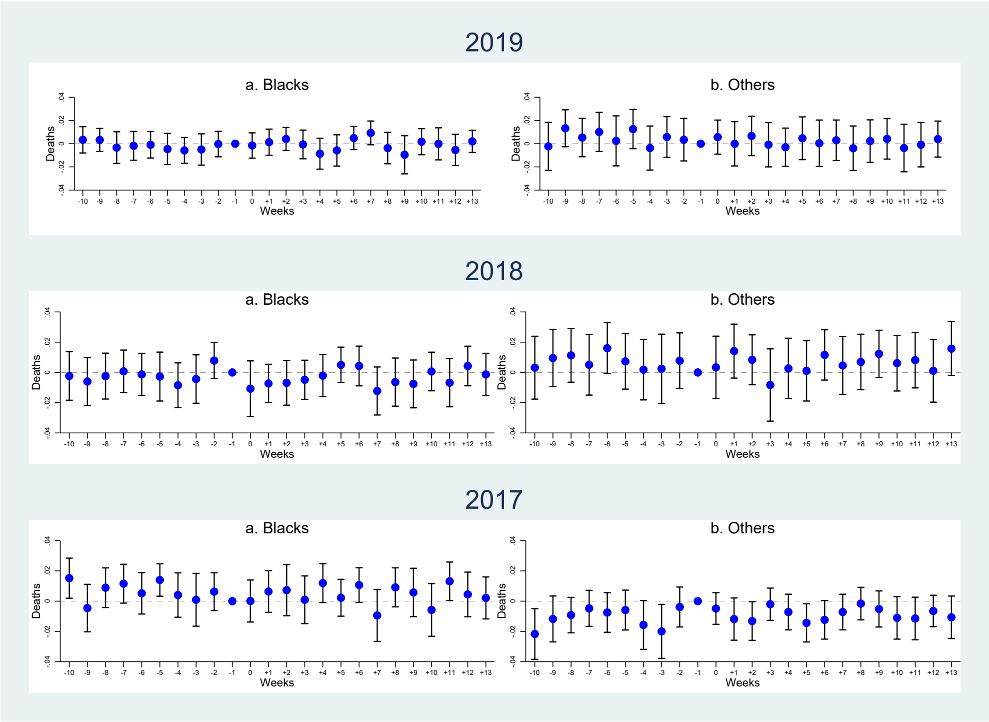


Figure 13: Dynamic Effect of the Treatment on Deaths, by Race - Cook County, January 1-June 16, 2019, 2018, and 2017.

Note: The dependent variable is number of deaths, of blacks (Panels a) and of other groups (Panels b), in the period from January 1 to June 16, in 2019 (top panels), 2018 (middle panels), and 2017 (bottom panels). The coefficients are least-squares estimates of the β_ks. Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

goal is to gauge the possibility that the treatment is capturing annual trends in mortality related to the diffusion of other diseases (e.g., the flu).[44] By replicating the analysis over the previous years, we shall be able to ascertain that we are not merely capturing a yearly trend having nothing to do with the COVID-19 epidemic.

Figure 13 shows that there is no effect of the treatment on black or other deaths in either of the three previous years over the same time frame, which confirms that the effect we found is attributable to the 2020 COVID-19 epidemic.

10.2 The Spanish flu

A further threat to identification is due to the fact the transmission rate of diseases (not mortality) may have always been larger in redlined neighborhoods, so that the 1930s

[44]Data for the years 2017-2019 are also provided by the Medical Examiner's Office.

Table 2: Spanish Flu - Chicago, 1918

	Deaths From Spanish Flu			
	(1)	(2)	(3)	(4)
Majority C & D	0.3020*	0.1039	0.1257	0.1070
	(0.1588)	(0.1525)	(0.1498)	(0.1479)
Population Density (log)	Yes	Yes	Yes	Yes
Illiteracy Rate	No	Yes	Yes	Yes
Home Ownership Rate	No	No	Yes	Yes
Unemployment Rate	No	No	No	Yes
Week FE	Yes	Yes	Yes	Yes
Adj.R-squared	0.252	0.281	0.284	0.287
Observations	3416	3416	3416	3416

Note: The dependent variable is the number of deaths from Spanish flu in Chicago in 1918. Robust standard errors clustered at a census tract level in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Covid Economics 38, 16 July 2020: 129-195

rankings may have merely formalized existing conditions. In other words, it may be the case that low-graded neighborhoods may have already been subject to higher transmission of viral diseases, even prior to redlining. To make sure that our results are not affected by this kind of threat, we focus on Spanish flu mortality.

By 1918, Chicago had already experienced the first wave of the Great Migration, with the black population rising from about 34,000 to 92,000 between 1910 and 1920, and a parallel increase of racial segregation (Tuttle, 1970). Chicago was badly hit by the Spanish flu, but blacks were actually hit less harshly than whites, as in the rest of the country, a fact that is still largely unexplained (Gamble, 2010; Okland and Mamelund, 2019). From our perspective, the fact that blacks were already segregated, by and large, in the same areas of the city, while at the same time they were somewhat protected from the disease, suggests that those areas were not per se more unhealthy.

In order to perform a falsification test, we use census tract level data on Spanish flu mortality in Chicago in 1918 to test whether the transmission of epidemics in low-graded areas was already higher before they were assessed by the HOLC. Data are provided by the Infectious Disease Dynamics Group.[45] The dataset contains census tract location and week of epidemic of 8,031 influenza and pneumonia deaths as well as sociodemographic data (including population density, the illiteracy rate, the home ownership rate, and the unemployment rate) for 496 census tracts within the City of Chicago. However, race-disaggregated information is not reported.[46]

Table 2 reports four variants of a model where the number of Spanish flu deaths in 1918 is the dependent variable. In Model 1, where we only control for majority C or D neighborhoods and week fixed effect, we find a marginally significant effect of the

[45]See http://www.ufiddynamics.org/data. The sources of the data are the 1920 Census and the 1920 annual report of City of Chicago Department of Health.

[46]See Table A7 for summary statistics.

control variable on deaths. However, when in the following models we add (the log of) population density, illiteracy, home ownership, and unemployment, no residual influence remains. Thus, reassuringly, we find no evidence that the HOLC areas that we found to be associated with higher COVID-19 mortality were subject to higher infections rates prior to the implementation of the redlining policies of the 1930s.

11 Conclusion

Not only the United States is registering the worldwide highest number of fatalities from the COVID-19 pandemic but, within the country, the death toll on African Americans has been disproportionately large. Up to now, however, lack of race-disaggregated data has prevented a rigorous assessment of this phenomenon and of its determinants.

Using so far unexploited individual and georeferenced death data collected by the Cook County Medical Examiner, we provide first evidence that race does affect COVID-19 outcomes. The data confirm that in Cook County blacks are overrepresented in terms of COVID-19 related deaths since—cumulatively in the period that goes from the outbreak of the epidemic on March 16, 2020 until June 16, 2020—they constitute 35 percent of the dead, which implies that they have been dying at a rate 1.3 times higher than their population share.

Furthermore, by combining the spatial distribution of mortality with the redlining maps for the Chicago area, we obtain a block group level panel dataset of weekly deaths from all causes, over the period January 1, 2020-June 16, 2020, to which we apply an event study design, where the treated neighborhoods are those that were historically either yellow or redlined and treatment initiation coincides with the outbreak of the COVID-19 epidemic. We show that, while no pre-treatment differences are detected, after the outbreak of the epidemic on March 16, 2020 historically low-graded neighborhoods display a sharper increase in mortality, which is driven by blacks. Thus, we uncover a persistence influence of the racial segregation induced by the discriminatory lending practices introduced in the 1930s.

We also establish that this influence runs by way of an asymmetric effect of the epidemic shock, which is in turn channeled through a diminished resilience of the black population to the shock represented by the COVID-19 outbreak. Far from being determined by genetic and biological factors, such vulnerability can be linked to socioeconomic status and household composition, as the likely channels through which the legacy of the past manifests itself.

To conclude, one of the stylized facts emerging from this paper, and that deserves

Covid Economics 38, 16 July 2020: 129-195

further attention, is that not only blacks are disproportionately hit by COVID-19, but also that they started to succumb to it earlier than other groups. Several explanations may be behind this phenomenon. On the one hand, it is possible that blacks become infected as much as the rest of the population, but they experience a faster progression through the stages of the disease, because of pre-existing medical conditions and/or access to health care. It may also be the case that blacks were more exposed from the beginning of the outbreak, because of their occupations and living conditions, so that once stay-at-home orders where issued they benefitted from them more thoroughly. The age and gender composition of each racial group also may also play a role.[47] The fact remains that the evolution of the epidemiological curve reveals for the US an extraordinary degree of racial and ethnic segregation, with different groups displaying profoundly distinct patterns even in the timing of their exposure to the epidemic.

REFERENCES

Aaronson D., Hartley D., Mazumder B. (2017), The effects of the 1930s HOLC "redlining" maps. Federal Reserve Bank of Chicago Working Paper No. 2017-12.

Abraido-Lanza A.F., Dohrenwend B.P; Ng-Mak D.S., Turner J.B. (1999). The Latino mortality paradox: A test of the "salmon bias" and healthy migrant hypotheses. American Journal of Public Health, 89, (10): 15431548.

Alfani G., Murphy T. (2017), Plague and lethal epidemics in the pre-industrial World. Journal of Economic History, 77, 314–343.

Almond, D. (2006). Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 U.S. population. Journal of Political Economy, 114, 672–712.

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

Almagro M., Orane-Hutchinson A. (2020), The determinants of the differential exposure to COVID-19 in New York City and their evolution over time. Covid Economics, 13, 31–50.

Anders J. (2019), The long run effects of de jure discrimination in the credit market: How redlining increased crime. Mimeo.

Anderson S. (2018), Legal origins and female HIV. American Economic Review, 108, 1407–1439.

Appel I., Nickerson J. (2016), Pockets of poverty: the long-term effects of redlining. Mimeo.

[47] In a companion paper (Bertocchi and Dimico, 2020b) we examine COVID-19 outcomes along the gender and work dimensions.

Bertocchi G. (2016), The legacies of slavery in and out of Africa. IZA Journal of Migration, 5, 1–19.

Bertocchi G., Dimico A. (2020a), Bitter sugar: Slavery and the black family. CEPR Discussion Paper No. 14837.

Bertocchi G., Dimico A. (2020b), COVID-19, gender, and work. Mimeo.

Bertocchi G., Dimico A. (2019), The long-term determinants of female HIV infection in Africa: The slave trade, polygyny, and sexual behavior. Journal of Development Economics, 140, 90–105.

Bertocchi G., Dimico A. (2014), Slavery, education, and inequality. European Economic Review, 70, 197–209.

Bertocchi G., Dimico A. (2012), The racial gap in education and the legacy of slavery. Journal of Comparative Economics, 40, 581–595.

Betancur J.J. (1996), The settlement experience of Latinos in Chicago: Segregation, speculation, and the ecology model. Source: Social Forces, 74, 1299–1324.

Bhala N., Curry G., Martineau A.R., Agyemang C., Bhopal R. (2020). Sharpening the global focus on ethnicity and race in the time of COVID-19. Lancet. Comment. May 8.

Borjas G.J. (2020), Demographics determinants of testing incidence and Covid-19 infections in New York City neighborhoods. Covid Economics, 3, 12-39.

Boustan L.P. (2011), Racial residential segregation in American cities. In Brooks N., Donaghy K., Knaap G.-J. (eds.), Oxford Handbook of Urban Economics and Planning. New York: Oxford University Press, 318–339.

Cage J., Rueda V. (2020), Sex and the mission: The conflicting effects of early Christian investments on sub-Saharan Africa's HIV epidemic. Journal of Demographic Economics, forthcoming.

Desmet K., Wacziarg R. (2020), Understanding spatial variation in COVID-19 across the United States. CEPR Discussion Paper No. 14842.

Douglas Commission [National Commission on Urban Problems] (1968), Building the American City. Washington: GPO.

Eligon J., Burch A.D.S., Searcey D., Oppel R.A.Jr. (2020), Black Americans face alarming rates of coronavirus infection in some states. The New York Times, April 7.

Couch K., Fairlie R.W., Xu H. (2020), The impacts of COVID-19 on minority unemployment: First evidence from April 2020 CPS microdata. IZA Discussion Paper No. 13264.

Federal Housing Administration (1938), Underwriting Manual: Underwriting and Valuation Procedure Under Title II of the National Housing Act. Washington: Federal Housing Administration.

Covid Economics 38, 16 July 2020: 129-195

Fishback P., Rose J., Snowden K. (eds.) (2013), Well Worth Saving: How the New Deal Safeguarded Home Ownership. Chicago: University of Chicago Press.

Flanagan B., Gregory E., Hallisey E., Heitgerd J., Lewis B. (2011), A social vulnerability index for disaster management. Journal of Homeland Security and Emergency Management, 8, 1–22.

Gamble V.N. (2010), "There Wasnt a Lot of Comforts in Those Days:"African Americans, Public Health, and the 1918 Influenza Epidemic. Public Health Reports, Supplement 3, 125 , 114–121.

Greer J.L. (2014), Historic home mortgage redlining in Chicago. Journal of the Illinois State Historical Society, 107, 204–233.

Greer J.L. (2012), Race and mortgage redlining in the United States. Mimeo.

Guimbeau A., Menon N., Musacchio A. (2020), The Brazilian bombshell? The long-term impact of the 1918 influenza pandemic the South American Way. NBER Working Paper No. 26929.

Harriss C.L. (1951), History and Policies of the Home Owners' Loan Corporation. New York: NBER.

Helgertz J., Bengtsson T. (2019), The long-lasting influenza: The impact of fetal stress during the 1918 influenza pandemic on socioeconomic attainment and health in Sweden, 1968-2012. Demography, 56, 1389–1425.

Hillier A. (2005), Residential Security Maps and neighborhood appraisals: The Home Owners Loan Corporation and the case of Philadelphia. Social Science History, 29, 207–233.

Hillier A. (2003), Redlining and the Home Owners Loan Corporation. Journal of Urban History, 29, 394–420.

Hoyt H. (1933), One Hundred Years of Land Values in Chicago: The Relationship of the Growth of Chicago to the Rise of Its Land Values, 1830-1933. Chicago: University of Chicago Press.

Hunt D.B. (2009), Blueprint for Disaster: The Unraveling of Chicago Public Housing. Chicago: University of Chicago Press.

Institute of Medicine (2003), Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. Washington: National Academies Press.

Kendi, I.X. (2020), Stop looking away from the race of COVID-19 victims. The Atlantic, April 1.

Jackson K. (1985), Crabgrass Frontier: The Suburbanization of the United States. New York: Oxford University Press.

Jackson K. (1980), Race, ethnicity, and real estate appraisal: The Home Owners Loan Corporation and the Federal Housing Administration. Journal of Urban History, 6, 419–

Covid Economics 38, 16 July 2020: 129-195

452.

Jedwab R., Johnson N.D., Koyama M. (2019), Negative shocks and mass persecutions: Evidence from the black death. Journal of Economic Growth, 24, 345–395.

Jedwab R., Johnson N.D., Koyama M. (2016), Bones, bacteria and break points: The heterogeneous spatial effects of the Black Death and long-run growth. GMU Working Paper in Economics No. 16-30.

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

Kirby T. (2020), Evidence mounts on the disproportionate effect of COVID-19 on ethnic minorities. Lancet Respiratory Medicine. News. May 8.

Krieger N., Wright E., Chen J.T., Waterman P.D., Huntley E.R., Arcaya M. (2020), Cancer stage at diagnosis, historical redlining, and current neighborhood characteristics: Breast, cervical, lung, and colorectal cancer, Massachusetts, 2001-2015. American Journal of Epidemiology, in press.

Krieger N., Van Wye G., Huynh M., Waterman P.D., Maduro G., Li W., Gwynn R.C., Barbot O., Bassett M.T. (2020), Structural racism, historical redlining, and risk of preterm birth in New York City, 20132017. American Journal of Public Health, in press.

Krimmel J. (2018), Persistence of prejudice: Estimating the long term effects of redlining. Mimeo.

Landrine H., Corral I., Lee J.G.L., Efird J.T., Hall M.B., Bess J.J. (2017), Residential segregation and racial cancer disparities: A systematic review. Journal of Racial and Ethnic Health Disparities, 4, 1195–1205.

Lang K., Kahn-Lang Spitzer A. (2020), Race discrimination: An economic perspective. Journal of Economic Perspectives, 34, 68–89.

Lin, M.-J., Liu e.M. (2014), Does in utero exposure to illness matter? The 1918 influenza epidemic in Taiwan as a natural experiment. Journal of Health Economics, 37, 152–163.

Loper J. (2019), Womens position in ancestral societies and female HIV: The long-term effect of matrilineality in sub-Saharan Africa. Mimeo.

McLaren J. (2020), Racial disparity in COVID-19 deaths: Seeking economic roots with census data. NBER Working Paper No. 27407.

Markides K.S:, Coreil J. (1986), The health of Hispanics in the southwestern United States: An epidemiologic paradox. Public Health Reports, 101, 253–265-

Nardone A., Casey J.A., Morello-Frosch R., Mujahid M., Balmes J.R., Thakur N. (2020), Associations between historical residential redlining and current age-adjusted rates of emergency department visits due to asthma across eight cities in California: An

Covid Economics 38, 16 July 2020: 129-195

ecological study. Lancet Planet Health, 4, e24–31.

Nelson R.K., Winling L., Marciano R., Connolly N., et al. (2020), Mapping inequality: Redlining in New Deal America. In Nelson R.K., Ayers E.L. (eds.), American Panorama: An Atlas of United States History. University of Richmond: Digital Scholarship Lab.

Nier, C.L.III (1999), Perpetuation of segregation: Toward a new historical and legal interpretation of redlining under the Fair Housing Act. John Marshall Law Review, 32, 617–665.

Nunn N. (2008), Slavery, inequality, and economic development in the Americas: An examination of the Engerman-Sokoloff hypothesis. In Helpman E. (ed.), Institutions and Economic Performance. Cambridge: Harvard University Press, 48–180.

Okland H., Mamelund S.-E. (2019), Race and 1918 influenza Pandemic in the United States: A review of the literature. International Journal of Environmental Research and Public Health, 16, 1–18.

OpenSAFELY Collaborative (2020), OpenSAFELY: Factors associated with COVID-19-related hospital death in the linked electronic health records of 17 million adult NHS patients. medRxiv, May 7.

Oppel R.J.Jr., Gebeloff R., Lai K.K.R., Wright W., Smith M. (2020), The fullest look yet at the racial inequity of coronavirus. The New York Times, July 5.

Orsi J.M., Margellos-Anast H., Whitman S. (2010), Black-white health disparities in the United States and Chicago: A 15-year progress analysis. American Journal of Public Health, 100, 349–356.

Platt L., Warwick R. (2020), Are Some Ethnic Groups more Vulnerable to COVID-19 than Others? London: Institute for Fiscal Studies.

Price-Haywood E.G., Burton J., Fort D., Seoane L. (2020), Hospitalization and mortality among black patients and white patients with Covid-19. The New England Journal of Medicine, 382, 2534–2543.

Reyes C., Husain N., Gutowski C., St. Clair S., Pratt G. (2020), Chicago's coronavirus disparity: Black Chicagoans are dying at nearly six times the rate of white residents, data show. Chicago Tribune, April 6.

Richardson G., McBride M. (2009), Religion, longevity, and cooperation: The case of the craft guild. Journal of Economic Behavior & Organization, 71, 172–186.

Rothstein R. (2017), The Color of Law: A Forgotten History of How Our Government Segregated America. New York: Liveright.

Schill M.H., Wachter S.M. (1995), The spatial bias of federal housing law and policy: Concentrated poverty in urban America. University of Pennsylvania Law Review, 143, 1285–1342.

Schmitt-Grohe S., Teoh K., Uribe M. (2020), COVID-19: Testing inequality in New

Covid Economics 38, 16 July 2020: 129-195

York City. Covid Economics, 8, 27-44

Shertzer A., Twinam T., Walsh R. (2016), Race, ethnicity, and discriminatory zoning. American Economic Journal: Applied Economics, 8, 217246.

Singh M., Koran M. (2020), 'The virus doesn't discriminate but governments do': Latinos disproportionately hit by coronavirus. The Guardian, April 18.

Tuttle W.M.Jr. (1970), Race Riot: Chicago in the Summer of 1919. New York: Atheneum.

Voitglander N., Voth H. (2012), Persecution perpetuated: The medieval origins of anti-Semitic violence in Nazi Germany. Quarterly Journal of Economics, 127, 1339–1392.

White C., Nafilyan V. (2020). Coronavirus (COVID-19) related deaths by ethnic group, England and Wales: 2 March 2020 to 10 April 2020. Office for National Statistics.

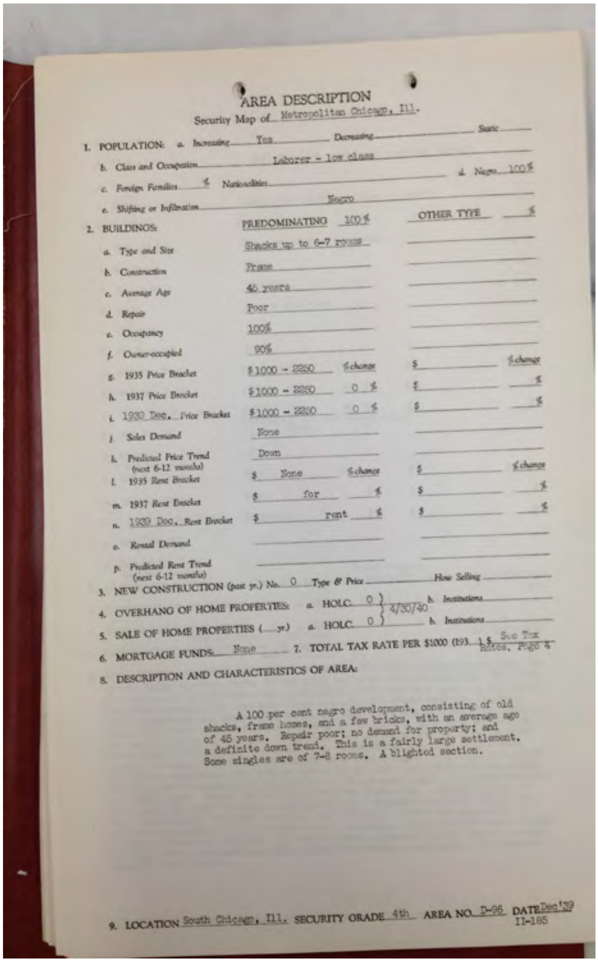
Wiemers E.E., Abrahams S., AlFakhri M., Hotz V.J., Schoeni R.F., Seltzer J.A. (2020). Disparities in vulnerability to severe complications from Covid-19 in the United States. NBER Working Paper No. 27294.

Yancy C.W. (2020), COVID-19 and African Americans. Journal of the American Medical Association. Opinion. April 15.

Zenou Y., Boccard N. (2000), Racial discrimination and redlining in cities. Journal of Urban Economics, 48, 260–285.

Covid Economics 38, 16 July 2020: 129-195

APPENDIX



AREA DESCRIPTION

Security Map of Metropolitan Chicago, Ill.

1. POPULATION: a. Increasing Yes Decreasing Static
   - b. Class and Occupation Laborer - low class
   - c. Foreign Families % Nationalities d. Negro 100%
   - e. Shifting or Infiltration Negro

2. BUILDINGS:

	PREDOMINATING 100%		OTHER TYPE %	
a. Type and Size	Shacks up to 6-7 rooms			
b. Construction	Frame			
c. Average Age	45 years			
d. Repair	Poor			
e. Occupancy	100%			
f. Owner-occupied	90%			
g. 1935 Price Bracket	$1000 - 2250	% change	$	% change
h. 1937 Price Bracket	$1000 - 2250	0 %	$	%
i. 1939 Dec. Price Bracket	$1000 - 2250	0 %	$	%
j. Sales Demand	None			
k. Predicted Price Trend (next 6-12 months)	Down			
l. 1935 Rent Bracket	$ None	% change	$	% change
m. 1937 Rent Bracket	$ for	%	$	%
n. 1939 Dec. Rent Bracket	$ rent	%	$	%
o. Rental Demand				
p. Predicted Rent Trend (next 6-12 months)				

3. NEW CONSTRUCTION (past yr.) No. 0 Type & Price How Selling
4. OVERHANG OF HOME PROPERTIES: a. HOLC 0 } 4/30/40 b. Institutions
5. SALE OF HOME PROPERTIES (yr.) a. HOLC 0 } b. Institutions
6. MORTGAGE FUNDS: None 7. TOTAL TAX RATE PER $1000 (193) $ See Tax Rates, Page 4
8. DESCRIPTION AND CHARACTERISTICS OF AREA:

A 100 per cent negro development, consisting of old shacks, frame homes, and a few bricks, with an average age of 45 years. Repair poor; no demand for property; and a definite down trend. This is a fairly large settlement. Some singles are of 7-8 rooms. A blighted section.

9. LOCATION South Chicago, Ill. SECURITY GRADE 4th AREA NO. D-96 DATE Dec '39

II-185

Figure A1: HOLC Area Description File for Area D96, Chicago Metropolitan Area, 1939

Covid Economics 38, 16 July 2020: 129-195

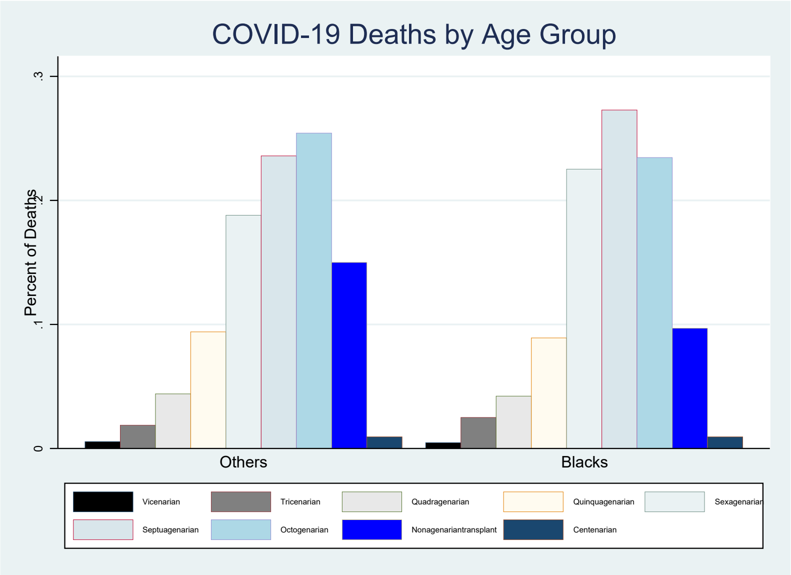


Figure A2: COVID-19 Deaths, by Age Group - Cook County, March 16-June 16, 2020

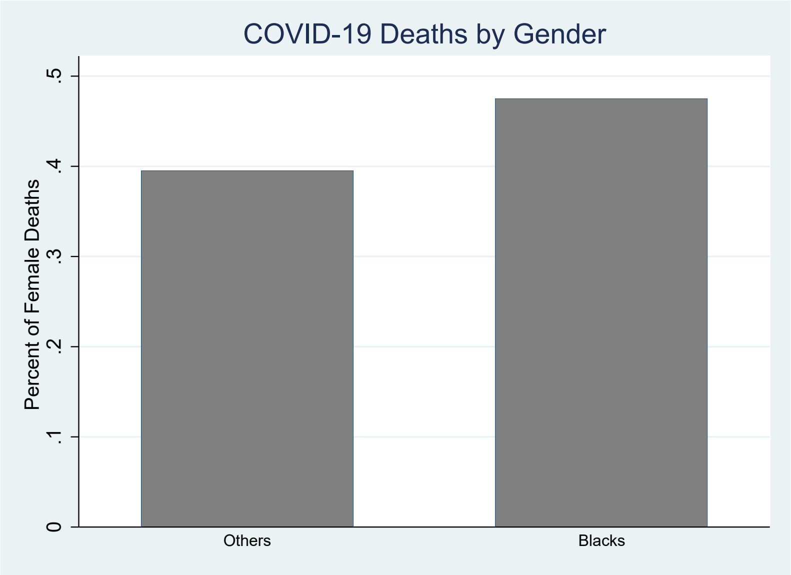


Figure A3: COVID-19 Deaths, by Gender - Cook County, March 16-June 16, 2020

Covid Economics 38, 16 July 2020: 129-195

CEPR PRESS

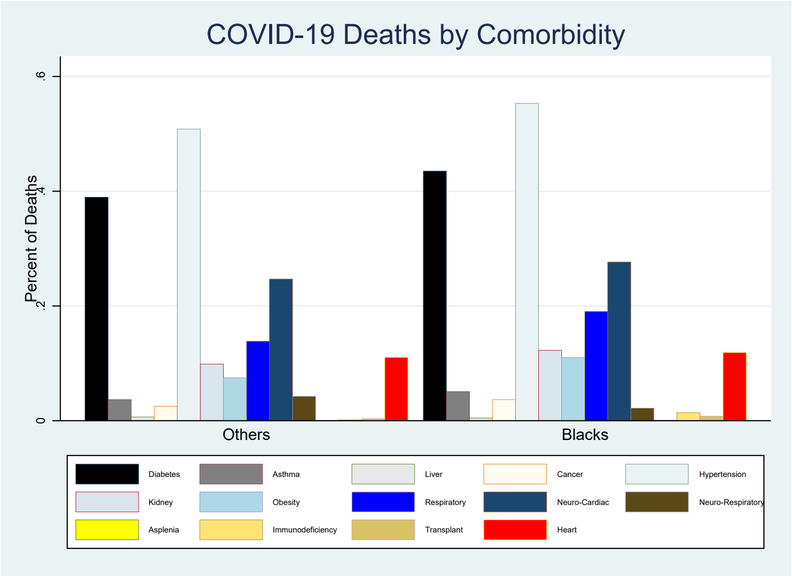


Figure A4: COVID-19 Deaths, by Comorbidity - Cook County, March 16-June 16, 2020

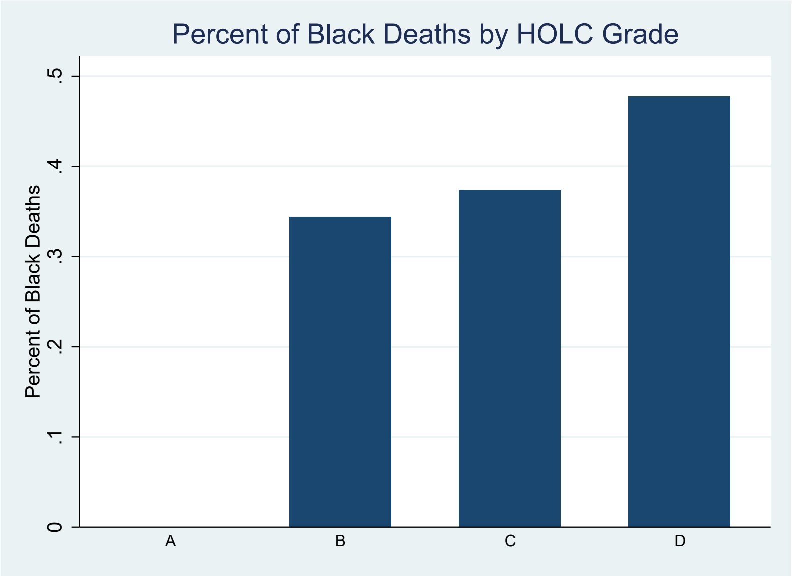


Figure A5: Share of Black COVID-19 Deaths, by HOLC Grade - Cook County, March 16-June 16, 2020

Covid Economics 38, 16 July 2020: 129-195

Covid Economics 38, 16 July 2020: 129-195

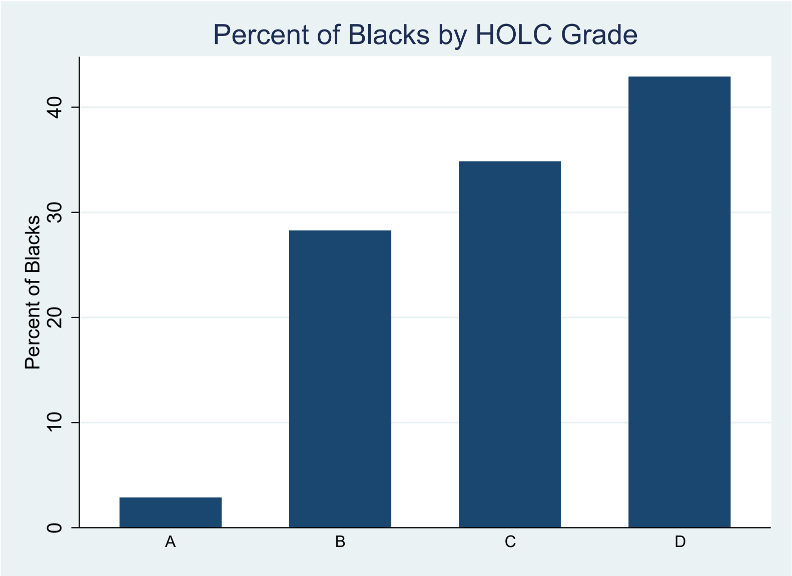


Figure A6: Black Population Share, by HOLC Grade - Cook County, 2010 Census

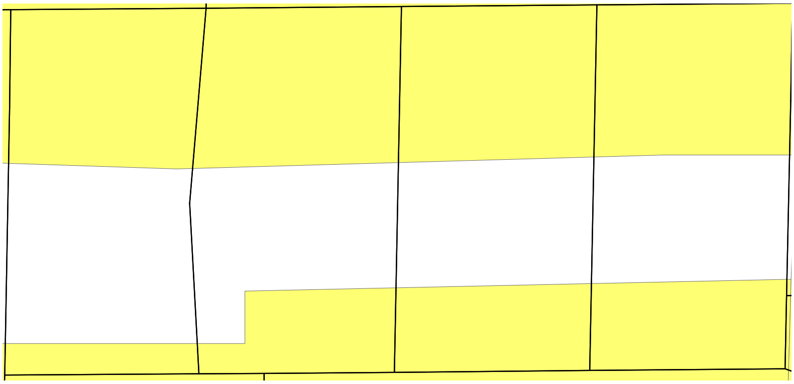


Figure A7: HOLC Neighborhoods by Block Group

Note: The figure reports an example of four block groups, each partitioned into three neighborhoods, two of which belonging to two different HOLC areas, both graded C (in yellow) and one ungraded (in white).

CEPR PRESS

Covid Economics 38, 16 July 2020: 129-195

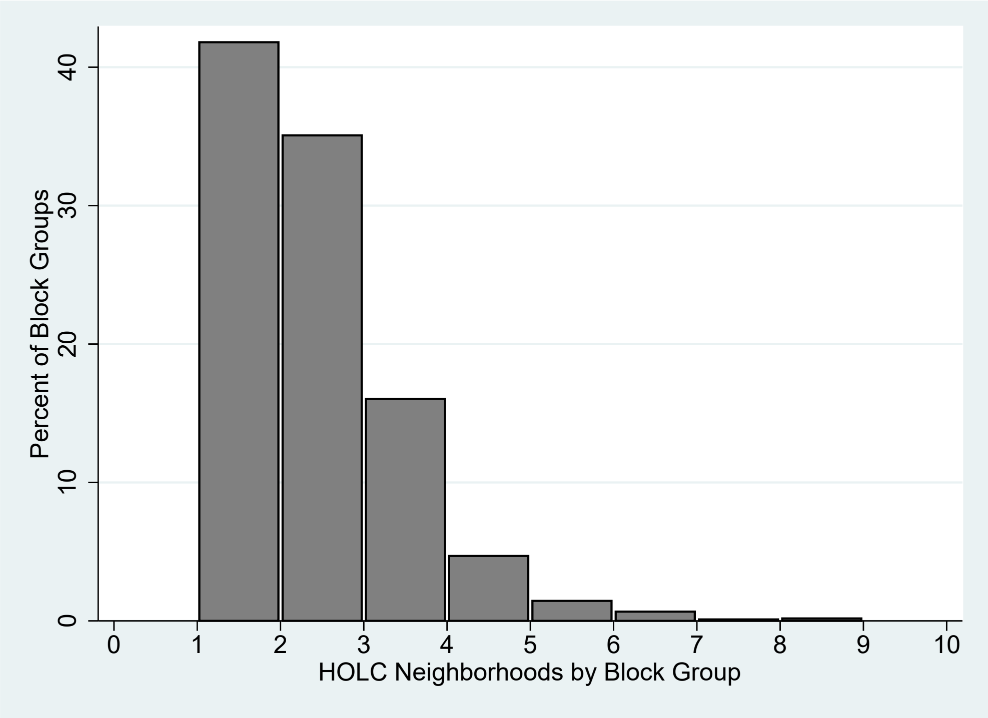


Figure A8: Distribution of HOLC Neighborhoods by Block Group

Note: The figure shows that over 40 percent of the block groups in the sample include only one HOLC neighborhood (potentially with the same HOLC grade), while around 35 percent include two, and so on, with a small percentage including up to ten.

Covid Economics 38, 16 July 2020: 129-195

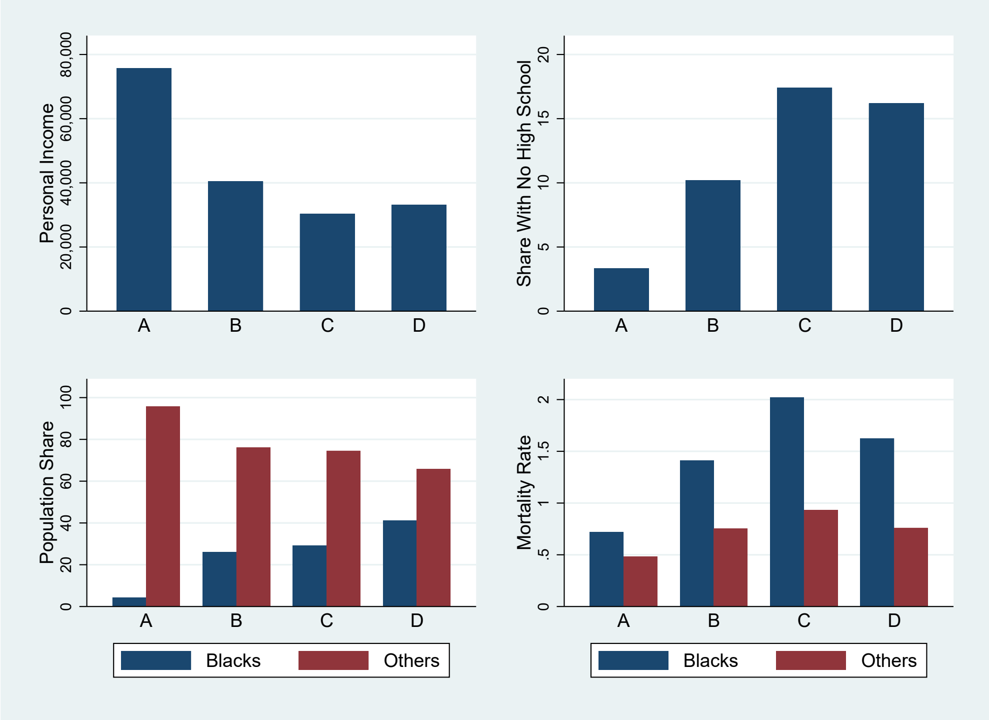


Figure A9: Socioeconomic and Demographic Characteristics by HOLC Grade

Note: Personal income, share with no high school, population share by race, and mortality rate by race, by HOLC area.

Covid Economics 38, 16 July 2020: 129-195

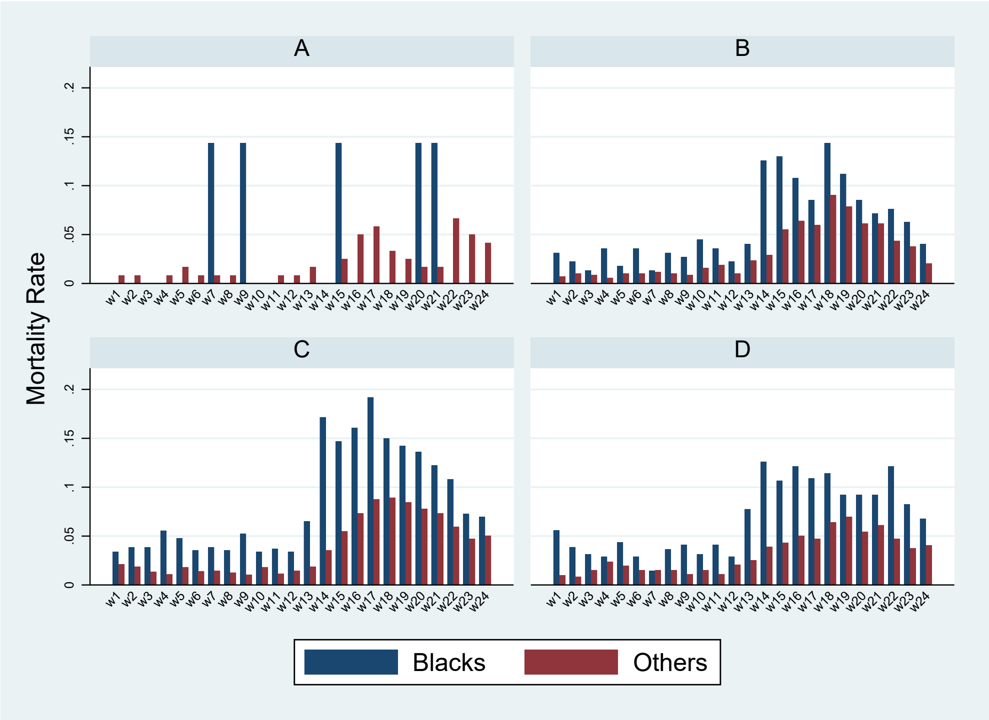


Figure A10: Mortality Rate, by Race, Week, and HOLC Grade - Cook County, March 16-June 16, 2020

Covid Economics 38, 16 July 2020: 129-195

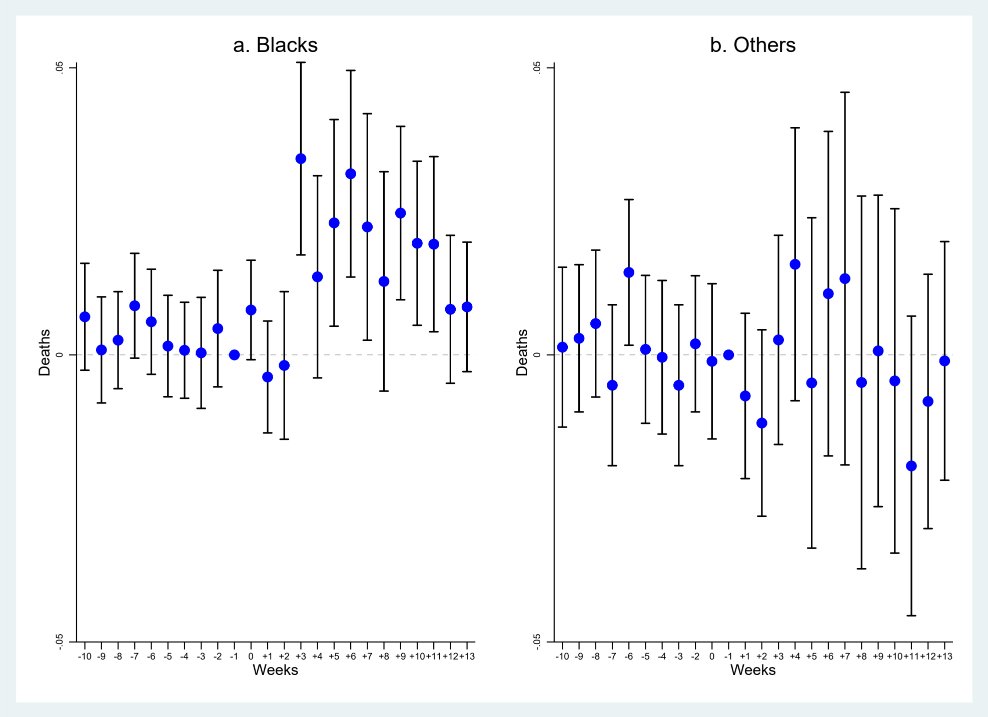


Figure A11: Including Block Groups Ungraded by the HOLCD - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of deaths, of blacks (Panel a) and of other groups (Panel b). The sample includes block groups that were not graded by the HOLC. The coefficients are least-squares estimates of the β_ks in regressions sequentially including additional interactions between week dummies and the variables indicated for each panel. Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

Covid Economics 38, 16 July 2020: 129-195

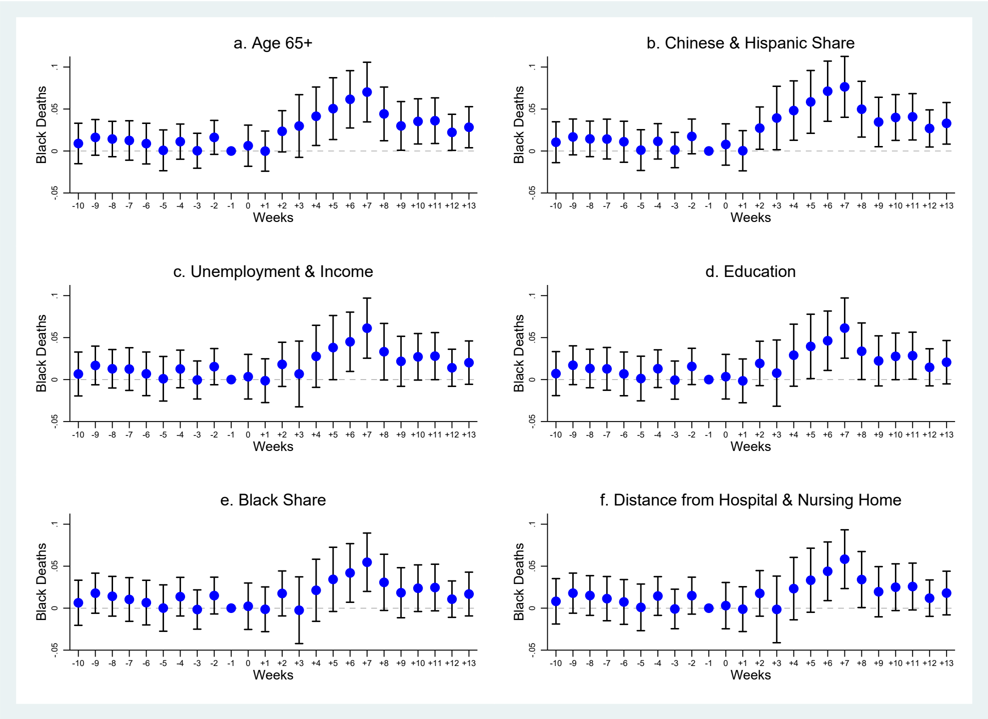


Figure A12: Dummy Death as Alternative Dependent Variable - Cook County, January 1-June 16, 2020

Note: The dependent variable is a dummy variable taking value one if a death, of blacks (Panel a) and of other groups (Panel b), occurred in a block group-week, and zero otherwise. The coefficients are least-squares estimates of the β_ks in regressions sequentially including additional interactions between week dummies and the variables indicated for each panel. Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

Covid Economics 38, 16 July 2020: 129-195

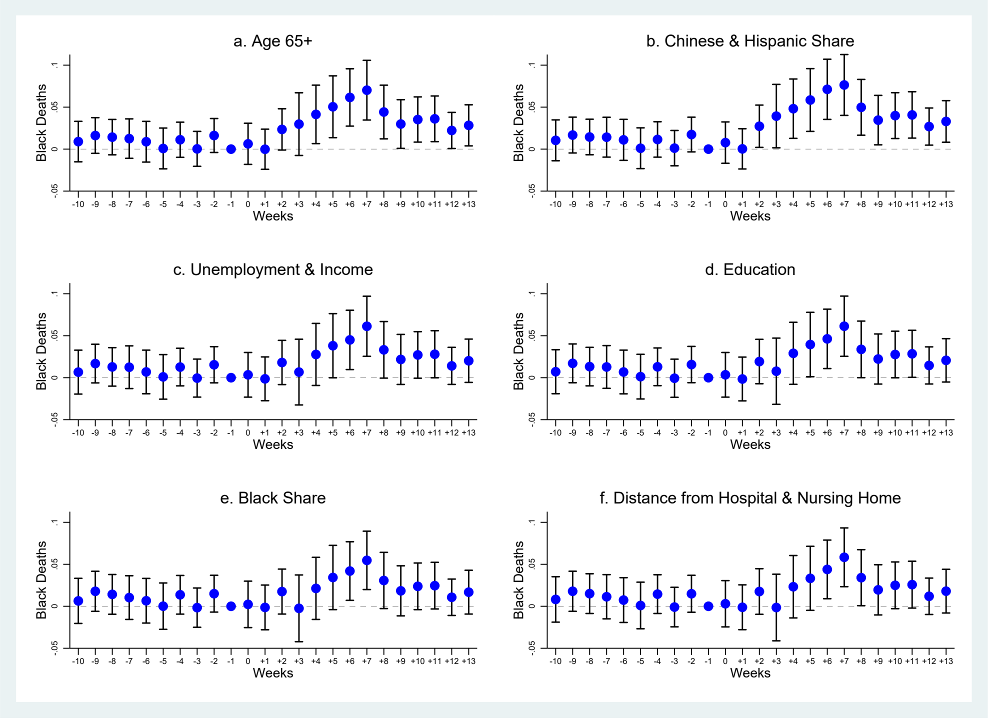


Figure A13: Allowing for Differential Trends - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of black deaths. The coefficients are least-squares estimates of the β_ks in regressions sequentially including additional interactions between week dummies and the variables indicated for each panel. Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

Covid Economics 38, 16 July 2020: 129-195

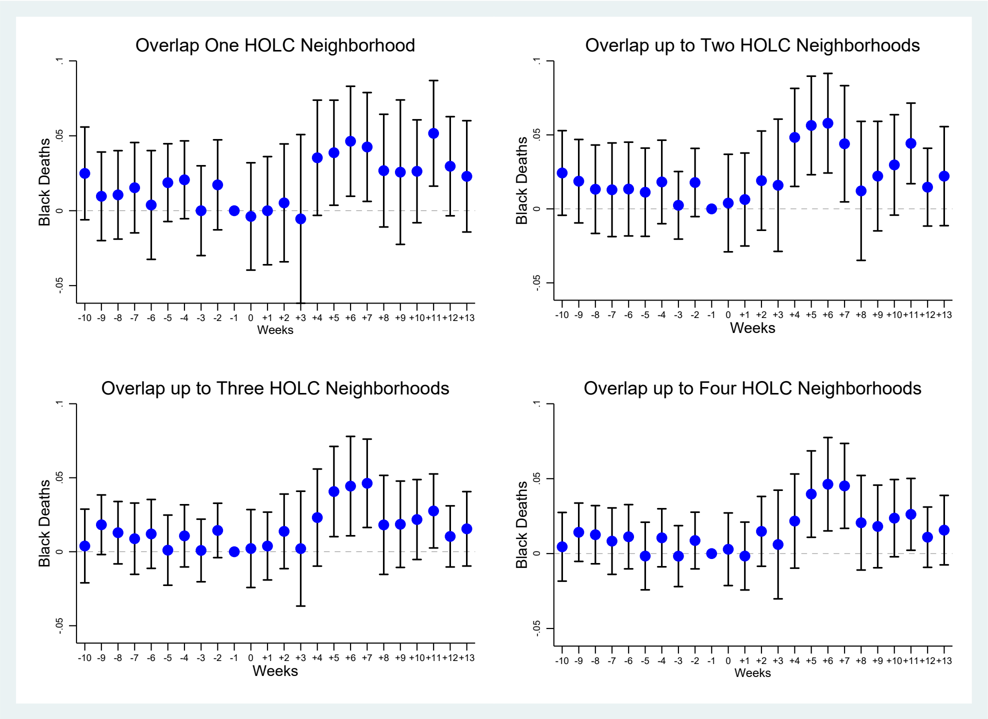


Figure A14: Robustness to Aggregation - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of black deaths. The coefficients are least-squares estimates of the β_ks over samples of block groups including one, up to two, up to three, and up to four HOLC neighborhoods. Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

Covid Economics 38, 16 July 2020: 129-195

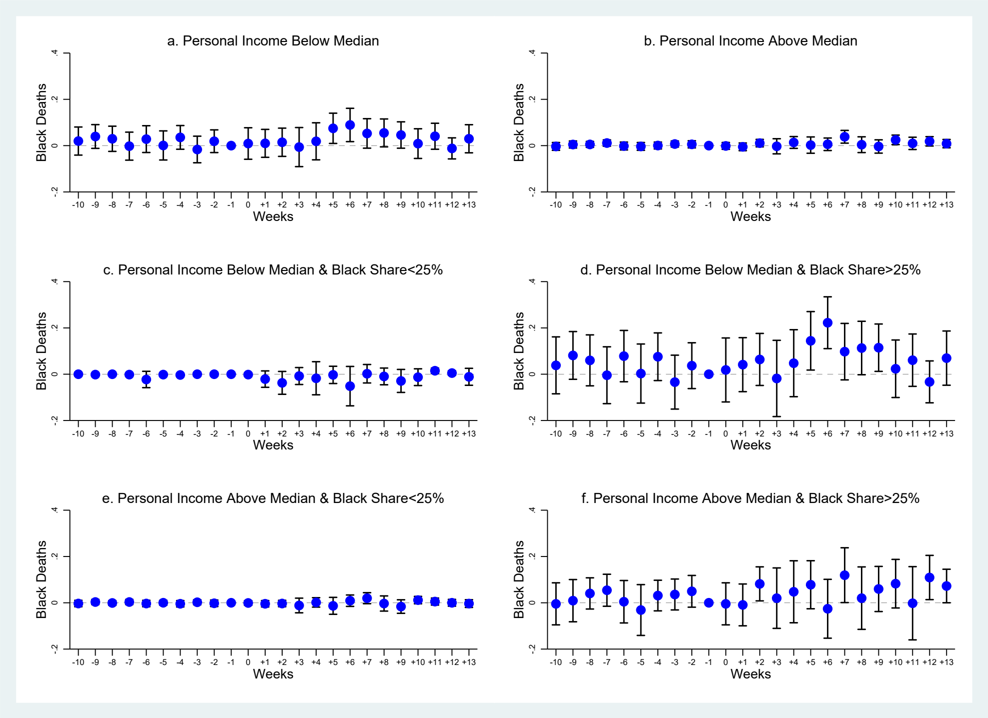


Figure A15: Heterogeneity by Income - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of black deaths. The coefficients are least-squares estimates of the β_ks over samples of block groups with personal income below median (Panel a), above median (Panel b), below median and with black share below 25 percent (Panel c), below median and with black share above 25 percent (Panel d), above median and with black share below 25 percent (Panel e), and above median and with black share above 25 percent (Panel f). Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

Covid Economics 38, 16 July 2020: 129-195

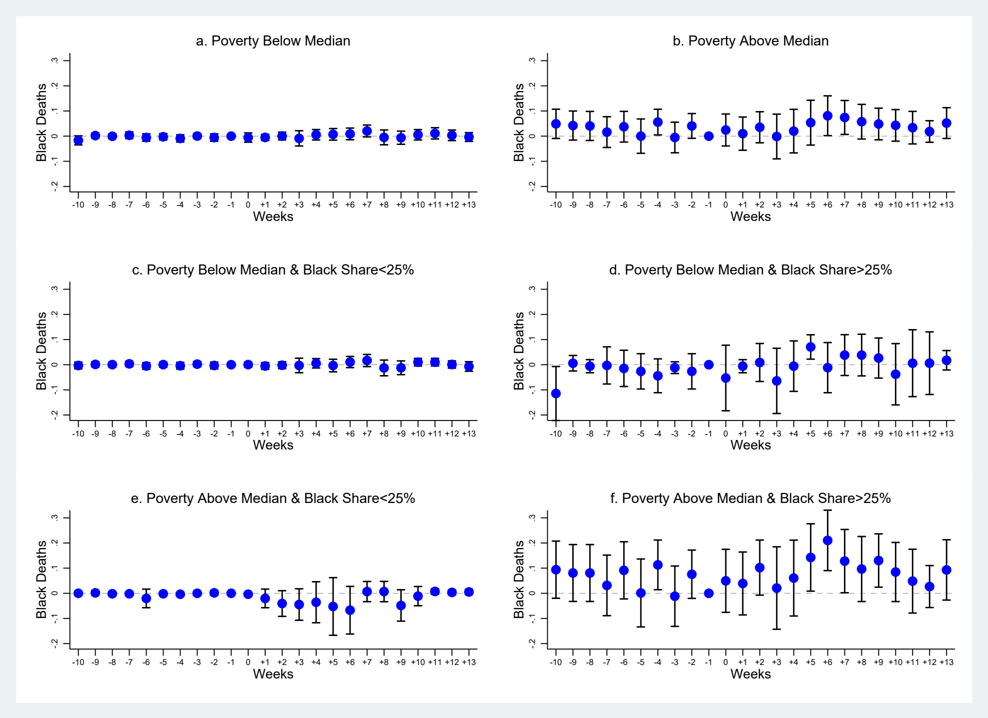


Figure A16: Heterogeneity by Poverty - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of black deaths. The coefficients are least-squares estimates of the β_ks over samples of block groups with population share below the poverty line below median (Panel a), above median (Panel b), below median and with black share below 25 percent (Panel c), below median and with black share above 25 percent (Panel d), above median and with black share below 25 percent (Panel e), and above median and with black share above 25 percent (Panel f). Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

Covid Economics 38, 16 July 2020: 129-195

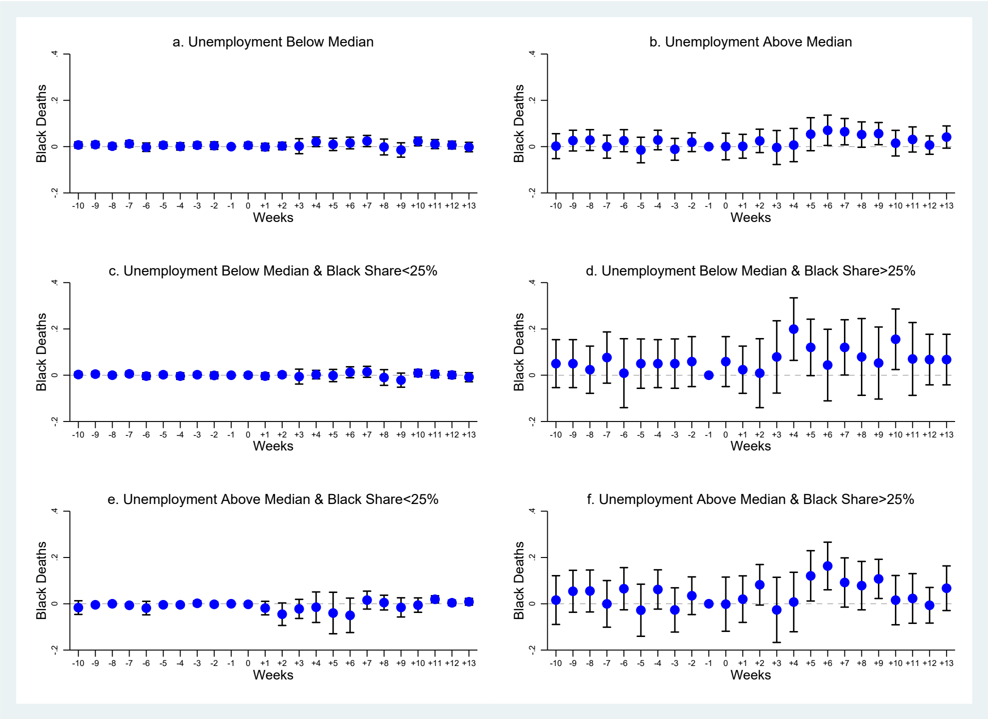


Figure A17: Heterogeneity by Unemployment - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of black deaths. The coefficients are least-squares estimates of the β_ks over samples of block groups with unemployment rate below median (Panel a), above median (Panel b), below median and with black share below 25 percent (Panel c), below median and with black share above 25 percent (Panel d), above median and with black share below 25 percent (Panel e), and above median and with black share above 25 percent (Panel f). Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

Covid Economics 38, 16 July 2020: 129-195

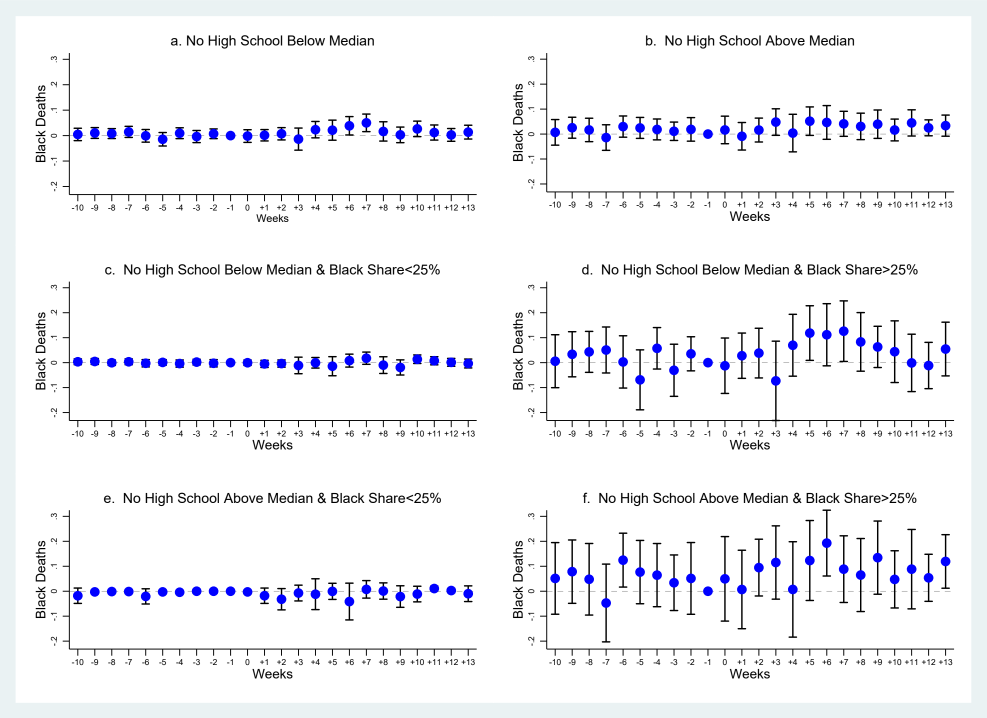


Figure A18: Heterogeneity by Education - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of black deaths. The coefficients are least-squares estimates of the β_ks over samples of block groups with population share with no high school below median (Panel a), above median (Panel b), below median and with black share below 25 percent (Panel c), below median and with black share above 25 percent (Panel d), above median and with black share below 25 percent (Panel e), and above median and with black share above 25 percent (Panel f). Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

CEPR PRESS

Covid Economics 38, 16 July 2020: 129-195

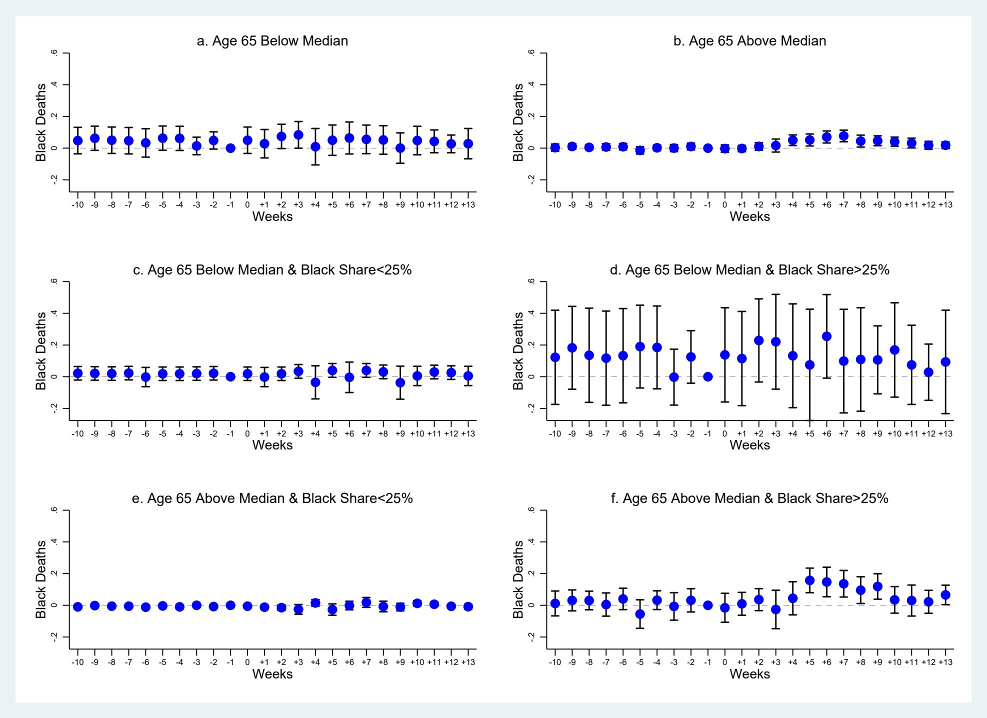


Figure A19: Heterogeneity by Age - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of black deaths. The coefficients are least-squares estimates of the β_ks over samples of block groups with population share aged 65+ below median (Panel a), above median (Panel b), below median and with black share below 25 percent (Panel c), below median and with black share above 25 percent (Panel d), above median and with black share below 25 percent (Panel e), and above median and with black share above 25 percent (Panel f). Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

Covid Economics 38, 16 July 2020: 129-195

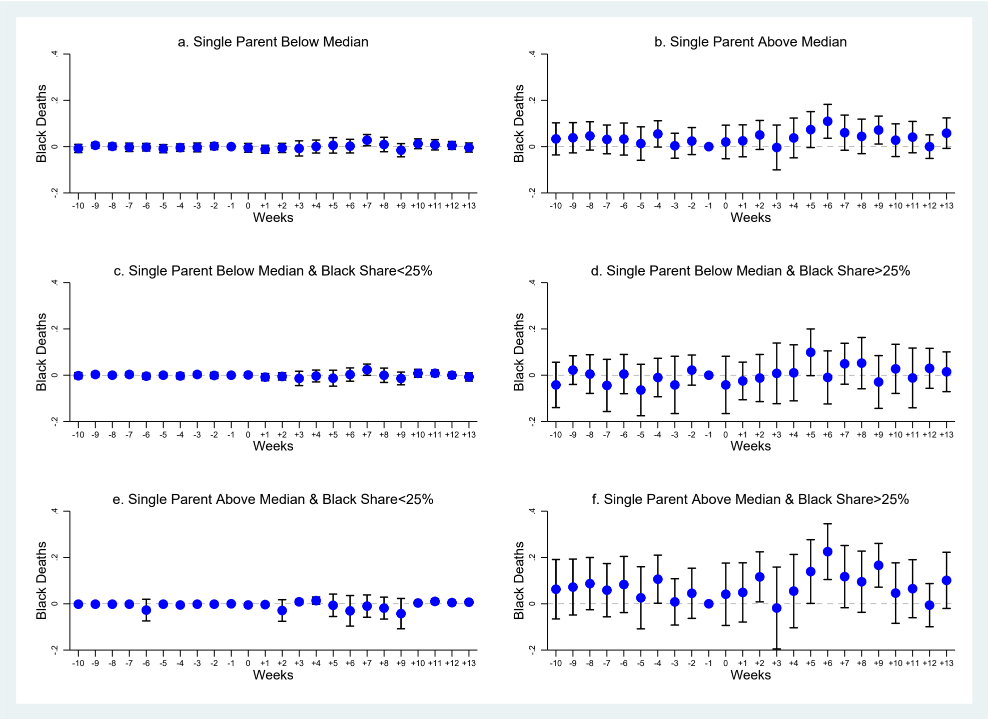


Figure A20: Heterogeneity by Single Parents - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of black deaths. The coefficients are least-squares estimates of the β_ks over samples of block groups with population share of single parents below median (Panel a), above median (Panel b), below median and with black share below 25 percent (Panel c), below median and with black share above 25 percent (Panel d), above median and with black share below 25 percent (Panel e), and above median and with black share above 25 percent (Panel f). Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

Covid Economics 38, 16 July 2020: 129-195

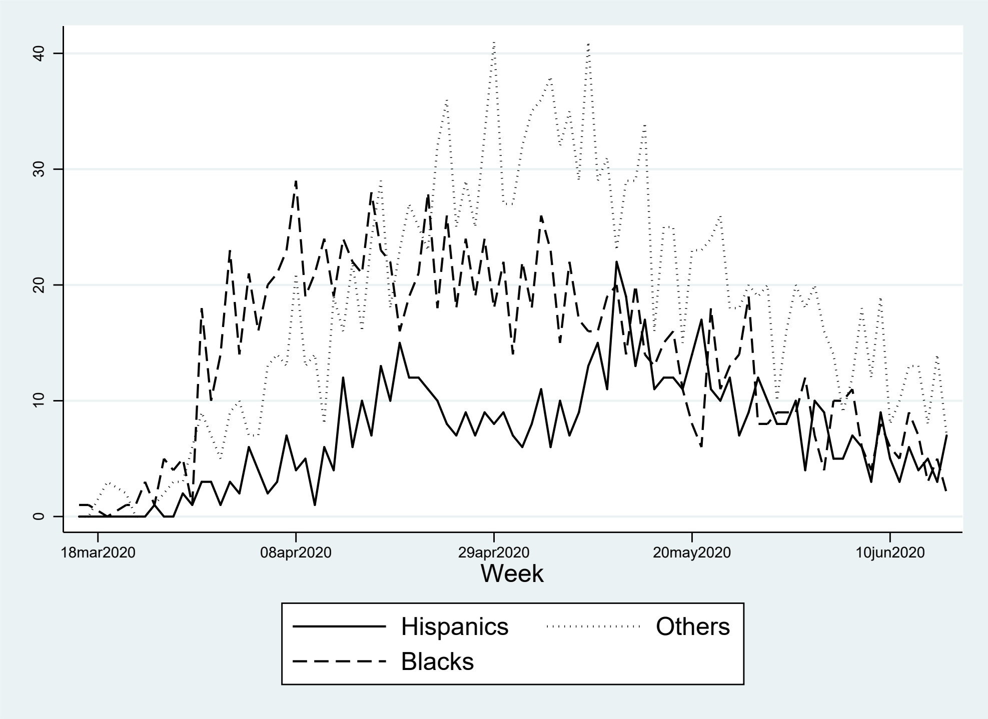


Figure A21: Hispanic, Black, and Other COVID-19 Deaths - Cook County, March 16-June 16, 2020

Note: The figure reports the number of COVID-19 related deaths by day, separately for Hispanics, blacks, and other groups.

Covid Economics 38, 16 July 2020: 129-195

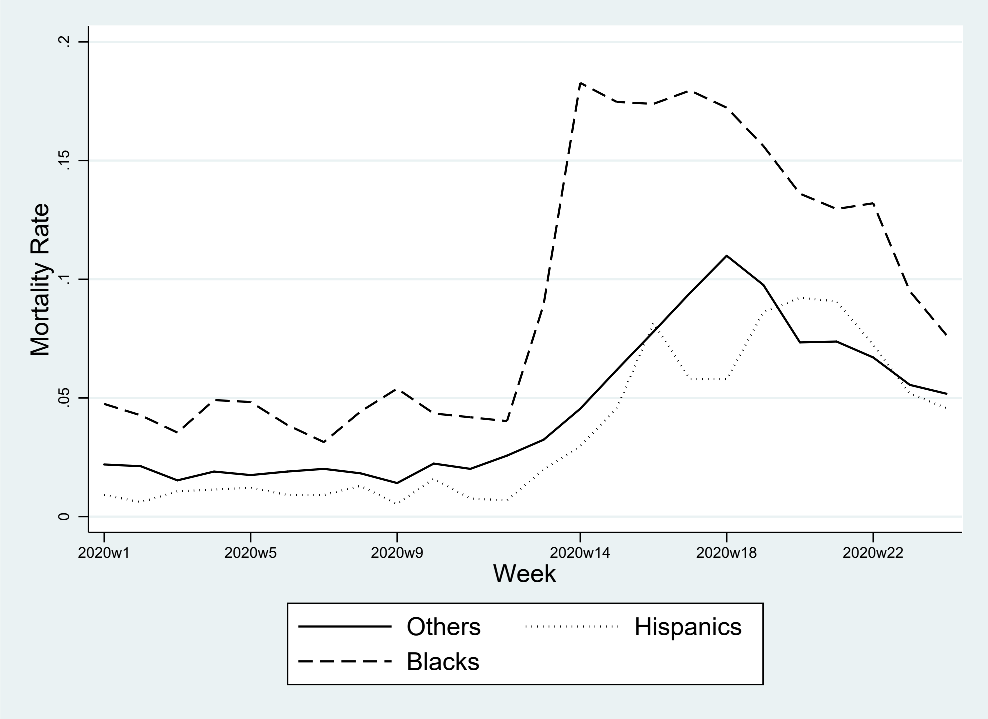


Figure A22: Mortality Rate, Hispanics, Blacks, and Others - Cook County, January 1-June 16, 2020

Note: Mortality rates from any cause of death by week, separately for Hispanics, blacks, and other groups.

Covid Economics 38, 16 July 2020: 129-195

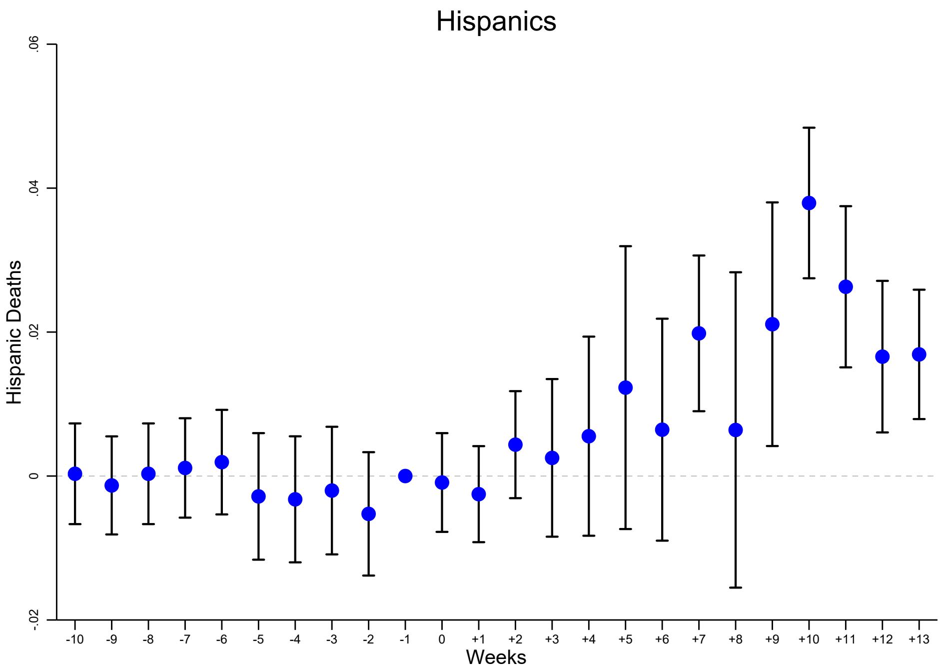


Figure A23: Dynamic Effect of the Treatment on Number of Hispanic Deaths - Cook County, January 1-June 16, 2020

Note: The dependent variable is number of Hispanic deaths. The coefficients are least-squares estimates of the β_ks with $10k9$. Block group and week fixed effects are included. Vertical lines represent 95 percent confidence intervals based on standard errors clustered at at block group level. The omitted period $k = -1$, i.e., week 10.

Covid Economics 38, 16 July 2020: 129-195

Table A1: Variable Definitions and Sources - Cross Section

Variable	Definition	Source
Black COVID-19 Death	Dummy variable taking value one if an individual who died from COVID-19 is black, and zero otherwise	Cook County Medical Examiner's Officer
Hispanic COVID-19 Death	Dummy variable taking value one if an individual who died from COVID-19 is Hispanic, and zero otherwise	Cook County Medical Examiner's Officer
Other COVID-19 Death	Dummy variable taking value one if an individual who died from COVID-19 is other than black, and zero otherwise	Cook County Medical Examiner's Officer
HOLC A	Dummy variable taking value one if an individual who died from COVID-19 lived in HOLC Area A, and zero otherwise	University of Richmond Mapping Inequality: Redlining in New Deal America 1935-1940
HOLC B	Dummy variable taking value one if an individual who died from COVID-19 lived in HOLC Area B, and zero otherwise	University of Richmond Mapping Inequality: Redlining in New Deal America 1935-1940
HOLC C	Dummy variable taking value one if an individual who died from COVID-19 lived in HOLC Area C, and zero otherwise	University of Richmond Mapping Inequality: Redlining in New Deal America 1935-1940
HOLC D	Dummy variable taking value one if an individual who died from COVID-19 lived in HOLC Area D, and zero otherwise	University of Richmond Mapping Inequality: Redlining in New Deal America 1935-1940
Age Groups	Set of eight dummy variables taking value one if an individual who died from COVID-19 was respectively aged from 20-29 (tricenarian) up to 100+ (centenarian), and zero otherwise	Cook County Medical Examiner's Officer
Female	Dummy variable taking value one if an individual who died from COVID-19 was female	Cook County Medical Examiner's Officer
Comorbidities	Set of 14 dummy variables that take value one (and zero otherwise) when an individual who died from COVID-19 was respectively affected by diabetes and/or asthma, liver disease, cancer, hypertension, kidney disease, obesity, respiratory diseases (including cystic fibrosis, pulmonary and lung diseases), neuro-cardiac diseases (including cardiovascular disease, stroke, and dementia), neuro-respiratory diseases (including sclerosis, Parkinson, myastenia, palsy, hemiplegia, quadriplegia, brain and cerebellum diseases), asplenia (including spenectomy, spleen and sickle cell disease), immunodeficiency (including HIV, immunosuppression, and anaemia), transplant, and heart diseases (including valve disease).	Cook County Medical Examiner's Officer
Distance From Hospital	Individual distance (in degrees) from hospital	Cook County Health and Hospitals Facilities
Distance From Nursing Home	Individual distance (in degrees) from nursing home	Medicare Nursing Home Compare
Population	2014-2018 tract level population	American Community Survey through Cook County Government Open Data
Share Aged 18-64	2014-2018 tract level share of the population aged 18-64	American Community Survey through Cook County Government Open Data
Share Aged 65+	2014-2018 tract level share of the population aged 65+	American Community Survey through Cook County Government Open Data
Share With No High School	2014-2018 tract level share of the population without a high school diploma	American Community Survey through Cook County Government Open Data
Black Share	2014-2018 tract level share of the population who is black	American Community Survey through Cook County Government Open Data
Hispanic Share	2014-2018 tract level share of the population who is Hispanic	American Community Survey through Cook County Government Open Data

Covid Economics 38, 16 July 2020: 129-195

Table A2: Variable Definitions and Sources - Panel

Variable	Definition	Source
Deaths	Number of deaths from any cause	Cook County Medical Examiner's Officer
Black Deaths	Number of deaths from any cause of black individuals	Cook County Medical Examiner's Officer
Other Deaths	Number of deaths from any cause of other than black individuals	Cook County Medical Examiner's Officer
Hispanic Deaths	Number of deaths from any cause of Hispanic individuals	Cook County Medical Examiner's Officer
Death Dummy	Dummy variable taking value one if a death occurred in a block group-week, and zero otherwise	Cook County Medical Examiner's Officer
Black Death Dummy	Dummy variable taking value one if a black death occurred in a block group-week, and zero otherwise	Cook County Medical Examiner's Officer
Other Dummy	Dummy variable taking value one if an other than black death occurred in a block group-week, and zero otherwise	Cook County Medical Examiner's Officer
Black Mortality Rate	Number of deaths from any cause of black individuals over tract level black population	Cook County Medical Examiner's Officer and American Community Survey through Cook County Government Open Data
Other Mortality Rate	Number of deaths from any cause of other than black individuals over tract level other than black population	Cook County Medical Examiner's Officer and American Community Survey through Cook County Government Open Data
Hispanic Mortality Rate	Number of deaths from any cause of Hispanic individuals over tract level Hispanic population	Cook County Medical Examiner's Officer and American Community Survey through Cook County Government Open Data
Treated Groups	Majority C & D block groups	University of Richmond Mapping Inequality: Redlining in New Deal America 1935-1940
Share Aged 65+	2014-2018 tract level share of the population aged 65+	American Community Survey through CDC Social Vulnerability Index
Hispanic Share	2014-2018 tract level share of the population who is Hispanic	American Community Survey through CDC Social Vulnerability Index
Share Chinese Primary Language	2014-2018 tract level share of the population speaking Chinese as primary language	American Community Survey through CDC Social Vulnerability Index
Personal Income	2014-2018 tract level personal income	American Community Survey through CDC Social Vulnerability Index
Unemployment Rate	2014-2018 tract level unemployment rate	American Community Survey through CDC Social Vulnerability Index
Share With No High School	2014-2018 tract level share of the population without a high school diploma	American Community Survey through CDC Social Vulnerability Index
Black Share	2014-2018 tract level share of the population who is black	American Community Survey through CDC Social Vulnerability Index
Distance From Hospital	Minimal centroid distance of a block group from hospital	Cook County Health and Hospitals Facilities
Distance From Nursing Home	Minimal centroid distance of a block group from nursing home	Medicare Nursing Home Compare
Socioeconomic Status Vulnerability	Index comprising personal income, share below poverty line, unemployment rate, and share without high school	CDC Social Vulnerability Index
Household Composition Vulnerability	Index comprising share aged 65+, share aged 17-, share aged 5+with disability, share of single-parent households	CDC Social Vulnerability Index
Minority Status Vulnerability	Index comprising population share in minority status and "Speak English Less than Well"	CDC Social Vulnerability Index
Housing Vulnerability	Index comprising housing share of multi-housing structures, mobile homes, and more people per room and population share with no vehicle available and in group quarters	CDC Social Vulnerability Index
Overall Vulnerability	Index based on the above four indices	CDC Social Vulnerability Index
Share Below Poverty	2014-2018 tract level share of the population below poverty line	American Community Survey through CDC Social Vulnerability Index
Share Single Parents	2014-2018 tract level share of the population in single-parent households	American Community Survey through CDC Social Vulnerability Index

CEPR PRESS

Covid Economics 38, 16 July 2020: 129-195

Table A3: Summary Statistics, COVID-19 Deaths, Cross Section - Cook County, March 16-June 16, 2020

	count	mean	sd	min	max
Black COVID-19 Death	3621	0.353	0.478	0.000	1.000
Hispanic COVID-19 Death	3621	0.186	0.389	0.000	1.000
Other COVID-19 Death	3621	0.647	0.478	0.000	1.000
HOLC A	3621	0.002	0.047	0.000	1.000
HOLC B	3621	0.076	0.265	0.000	1.000
HOLC C	3621	0.381	0.486	0.000	1.000
HOLC D	3621	0.141	0.348	0.000	1.000
No HOLC	3621	0.400	0.490	0.000	1.000
Age Groups	3619	5.912	1.509	1.000	9.000
Female	3621	0.423	0.494	0.000	1.000
Diabetes	3621	0.406	0.491	0.000	1.000
Asthma	3621	0.042	0.200	0.000	1.000
Liver Diseases	3621	0.006	0.076	0.000	1.000
Cancer	3621	0.029	0.169	0.000	1.000
Hypertension	3621	0.524	0.499	0.000	1.000
Kidney Disease	3621	0.107	0.309	0.000	1.000
Obesity	3621	0.087	0.282	0.000	1.000
Respiratory Diseases	3621	0.157	0.364	0.000	1.000
Neuro-cardiac Diseases	3621	0.257	0.437	0.000	1.000
Neuro-respiratory Diseases	3621	0.035	0.184	0.000	1.000
Asplenia	3621	0.001	0.023	0.000	1.000
Immunodeficiency	3621	0.006	0.076	0.000	1.000
Transplant	3621	0.005	0.070	0.000	1.000
Heart Diseases	3621	0.113	0.317	0.000	1.000
Population	3621	4666.875	1776.020	592.000	19015.000
Share Aged 18-64	3621	62.219	7.527	38.300	94.000
Share Aged 65+	3621	15.557	8.022	0.000	51.000
Distance From Hospital	3621	0.023	0.016	0.000	0.116
Distance From Nursing Home	3621	0.011	0.012	0.000	0.067
Share With No High School	3621	17.061	11.926	0.000	60.600
Black Share	3621	30.940	36.592	0.000	100.000
Hispanic Share	3621	25.903	28.420	0.000	99.600

Table A4: Black COVID-19 Death, Cross-Sectional Results, All Coefficients - Cook County, March 16-June 16, 2020

	Black COVID-19 Death							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HOLC A	-0.3808***	-0.3652***	-0.3244***	-0.3261***	-0.3247***	-0.3636***	-0.3514***	-0.1250***
	(0.0640)	(0.0720)	(0.0757)	(0.0800)	(0.0811)	(0.0798)	(0.0751)	(0.0424)
HOLC B	0.0389	0.0360	0.0399	0.0373	0.0355	0.0303	0.0072	0.0169
	(0.0329)	(0.0334)	(0.0327)	(0.0323)	(0.0327)	(0.0322)	(0.0316)	(0.0250)
HOLC C	0.0800***	0.0817***	0.0830***	0.0677***	0.0716***	0.1176***	0.0848***	-0.0146
	(0.0177)	(0.0185)	(0.0181)	(0.0175)	(0.0202)	(0.0216)	(0.0234)	(0.0182)
HOLC D	0.1752***	0.1769***	0.1745***	0.0833***	0.0852***	0.1232***	0.0922***	0.0106
	(0.0233)	(0.0238)	(0.0231)	(0.0255)	(0.0266)	(0.0281)	(0.0297)	(0.0222)
Tricenarian		0.1373	0.1069	0.0898	0.0934	0.1113	0.1244	-0.0055
		(0.1324)	(0.1333)	(0.1306)	(0.1318)	(0.1287)	(0.1278)	(0.0981)
Quadragenarian		0.0593	0.0282	0.0278	0.0326	0.0600	0.0728	0.0233
		(0.1262)	(0.1278)	(0.1282)	(0.1294)	(0.1255)	(0.1236)	(0.0961)
Quinquagenarian		0.0559	0.0140	0.0198	0.0252	0.0480	0.0648	0.0121
		(0.1225)	(0.1229)	(0.1230)	(0.1240)	(0.1211)	(0.1192)	(0.0895)
Sexagenarian		0.1348	0.0778	0.0795	0.0857	0.0907	0.1144	0.0603
		(0.1200)	(0.1216)	(0.1207)	(0.1220)	(0.1197)	(0.1178)	(0.0911)
Septuagenarian		0.1149	0.0543	0.0619	0.0678	0.0672	0.0890	0.0288
		(0.1171)	(0.1178)	(0.1177)	(0.1189)	(0.1168)	(0.1145)	(0.0892)
Octogenarian		0.0768	0.0148	0.0155	0.0218	0.0140	0.0409	0.0180
		(0.1188)	(0.1192)	(0.1200)	(0.1212)	(0.1196)	(0.1182)	(0.0891)
Nonagenarian		0.0082	-0.0498	-0.0433	-0.0374	-0.0468	-0.0102	-0.0291
		(0.1226)	(0.1225)	(0.1233)	(0.1248)	(0.1230)	(0.1203)	(0.0921)
Centenarian		0.0986	0.0372	0.0399	0.0445	0.0270	0.0744	0.0061
		(0.1350)	(0.1358)	(0.1337)	(0.1356)	(0.1338)	(0.1311)	(0.1046)
Female		0.1036***	0.0977***	0.0961***	0.0955***	0.0886***	0.0902***	0.0504***
		(0.0156)	(0.0157)	(0.0154)	(0.0154)	(0.0155)	(0.0155)	(0.0115)
Diabetes			0.0040	-0.0004	-0.0003	0.0103	0.0092	0.0005
			(0.0156)	(0.0155)	(0.0154)	(0.0152)	(0.0149)	(0.0122)
Asthma			0.0145	0.0155	0.0147	0.0085	-0.0003	-0.0034
			(0.0440)	(0.0438)	(0.0438)	(0.0427)	(0.0427)	(0.0314)
Liver Diseases			-0.1527	-0.1488*	-0.1498*	-0.1516*	-0.1496*	0.0089
			(0.0948)	(0.0870)	(0.0868)	(0.0820)	(0.0801)	(0.0771)
Cancer			0.0722	0.0747	0.0753	0.0670	0.0580	0.0166
			(0.0478)	(0.0467)	(0.0474)	(0.0480)	(0.0477)	(0.0395)
Hypertension			0.0939***	0.0936***	0.0938***	0.0908***	0.0881***	0.0545***
			(0.0165)	(0.0166)	(0.0166)	(0.0167)	(0.0167)	(0.0142)
Kidney Disease			0.0782***	0.0806***	0.0802***	0.0694**	0.0702**	0.0659***
			(0.0272)	(0.0266)	(0.0268)	(0.0266)	(0.0269)	(0.0213)
Obesity			0.0413	0.0305	0.0306	0.0290	0.0264	0.0118
			(0.0297)	(0.0287)	(0.0289)	(0.0281)	(0.0280)	(0.0234)
Respiratory Diseases			0.0764***	0.0740***	0.0747***	0.0700***	0.0752***	0.0405**
			(0.0214)	(0.0209)	(0.0209)	(0.0203)	(0.0201)	(0.0175)
Neuro-cardiac Diseases			0.0158	0.0166	0.0177	0.0196	0.0236	0.0453***
			(0.0213)	(0.0214)	(0.0217)	(0.0217)	(0.0213)	(0.0165)
Neuro-respiratory Diseases			-0.0997***	-0.0951***	-0.0943***	-0.0986***	-0.0843**	-0.0739***
			(0.0338)	(0.0333)	(0.0330)	(0.0327)	(0.0338)	(0.0261)
Asplenia			0.1501	0.1260	0.1394	0.1531	0.1704	-0.1478
			(0.3863)	(0.3213)	(0.3237)	(0.3323)	(0.3184)	(0.1288)
Immunodeficiency			0.4063***	0.3897***	0.3927***	0.3756***	0.3685***	0.2470**
			(0.0899)	(0.0934)	(0.0935)	(0.0883)	(0.0877)	(0.1020)
Transplant			0.1204	0.1200	0.1177	0.0940	0.0805	0.0136
			(0.1186)	(0.1159)	(0.1166)	(0.1071)	(0.1071)	(0.1024)
Heart Diseases			0.0256	0.0178	0.0180	0.0093	0.0157	0.0081
			(0.0267)	(0.0271)	(0.0271)	(0.0266)	(0.0259)	(0.0197)
Population (log)				-0.1734***	-0.1725***	-0.1780***	-0.1568***	-0.0260
				(0.0210)	(0.0208)	(0.0209)	(0.0207)	(0.0178)
Share Aged 18-64					-0.0016	-0.0065***	-0.0067***	0.0014
					(0.0014)	(0.0018)	(0.0017)	(0.0012)
Share Aged 65+					-0.0007	-0.0065***	-0.0058***	-0.0011
					(0.0013)	(0.0015)	(0.0016)	(0.0011)
Share With No High School						-0.0073***	-0.0073***	-0.0007
						(0.0009)	(0.0008)	(0.0006)
Distance From Hospital							-2.9815***	0.4574
							(0.4945)	(0.4217)
Distance From Nursing Home							3.5242***	-1.5140***
							(0.6538)	(0.5396)
Black Share								0.0086***
								(0.0002)
Constant	0.2957***	0.1637	0.1384	1.6084***	1.6999***	2.2492***	2.0984***	0.1585
	(0.0094)	(0.1185)	(0.1233)	(0.2004)	(0.2262)	(0.2516)	(0.2511)	(0.2039)
Adj.R-squared	0.056	0.069	0.090	0.111	0.111	0.132	0.144	0.475
Observations	3618	3616	3616	3616	3616	3616	3616	3616

Note: The dependent variable is a dummy variable that takes value one if an individual who died from COVID-19 is black, an zero otherwise. The omitted HOLC area is A. The omitted age group is vicenarian. Robust standard errors clustered at a day level in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Covid Economics 38, 16 July 2020: 129-195

Covid Economics 38, 16 July 2020: 129-195

Table A5: Summary Statistics, All Deaths, Panel - Cook County, January 1-June 16, 2020

	count	mean	sd	min	max
Deaths	95808	0.070	0.342	0.000	13.000
Black Deaths	95808	0.029	0.201	0.000	10.000
Other Deaths	95808	0.042	0.257	0.000	11.000
Hispanic Deaths	95808	0.012	0.117	0.000	7.000
Death Dummy	95808	0.057	0.233	0.000	1.000
Black Death Dummy	95808	0.025	0.156	0.000	1.000
Other Death Dummy	95808	0.035	0.184	0.000	1.000
Black Mortality Rate	93576	0.064	3.845	0.000	1000.000
Other Mortality Rate	95544	0.030	0.806	0.000	125.000
Hispanic Mortality Rate	93504	0.029	3.324	0.000	1000.000
White Mortality Rate	94961	0.055	1.909	0.000	333.333
Treated Groups (Majority C & D)	95808	0.619	0.486	0.000	1.000
Post-treatment (Week After 11)	95808	0.583	0.493	0.000	1.000
Treatment	95808	0.361	0.480	0.000	1.000
Share Aged 65+	95736	14.156	6.262	0.000	51.300
Hispanic Share	95808	23.627	25.999	0.000	99.600
Share Chinese Primary Language	95808	1.113	4.582	0.000	78.000
Personal Income	95736	34845.017	20754.386	2530.000	154760.000
Unemployment Rate	95808	9.930	7.949	0.000	92.820
Share With No High School	95736	13.847	10.704	0.000	61.600
Black Share	95808	27.330	35.880	0.000	100.000
Distance From Hospital	95808	0.021	0.016	0.000	0.105
Distance From Nursing Home	95808	0.013	0.011	0.000	0.066
Socioeconomic Status Vulnerability	95736	0.539	0.303	0.000	1.000
Household Composition Vulnerability	95736	0.454	0.286	0.002	0.999
Minority Status Vulnerability	95736	0.674	0.221	0.035	1.000
Housing Vulnerability	95736	0.544	0.270	0.000	1.000
Overall Vulnerability	95736	0.565	0.284	0.001	0.998
Share Below Poverty	95736	15.883	12.047	0.300	77.100
Share Single Parents	95808	9.896	7.413	0.000	51.300

Table A6: Average Treatment Effect

	(1) Black Deaths	(2) Other Deaths
Treatment	0.015**	-0.002
	(0.006)	(0.009)
R-squared	0.090	0.082
Observations	53903	53903
Mean Untreated In Post-treatment	0.025	0.025

Note: The dependent variable is the average number of deaths in a block group in a week. Robust standard errors clustered at a block groups level in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Summary Statistics, Spanish Flu Deaths, Chicago, 1918

	count	mean	sd	min	max
Deaths From Spanish Flu	3472	2.296	2.974	0.000	31.000
Majority C & D	3419	0.848	0.359	0.000	1.000
Population Density (log)	3472	3.488	1.098	-3.640	5.205
Illiteracy Rate	3472	0.040	0.048	0.000	0.279
Home Ownership Rate	3472	0.061	0.038	0.000	0.197
Unemployment Rate	3472	0.346	0.050	0.058	0.495

Unequal consequences of Covid-19 across age and income: Representative evidence from six countries[1]

Michèle Belot,[2] Syngjoo Choi,[3] Egon Tripodi,[4] Eline van den Broek-Altenburg,[5] Julian C. Jamison[6] and Nicholas W. Papageorge[7]

Date submitted: 14 July 2020; Date accepted: 14 July 2020

Covid-19 and the measures taken to contain it have led to unprecedented constraints on work and leisure activities, across the world. This paper uses nationally representative surveys to document how people of different ages and incomes have been affected across six countries (China, South Korea, Japan, Italy, UK and US). We first document changes in economic variables (income and consumption) and leisure. Second, we document self-reported negative and positive non-financial effects of the crisis. We then examine attitudes towards recommendations (wearing a mask in particular) and the approach taken by public authorities. We find similarities across countries in how people of different generations have been affected. Young people have experienced more drastic changes to their lives, and overall they are less supportive of these measures. These patterns are less clear across income groups: while some countries have managed to shield lower income individuals from negative consequences, others have not. We also show that how people have been affected by the crisis (positively or negatively) does little to explain whether or not they support measures implemented by the public authorities. Young people are overall less supportive of such measures independently of how they have been affected.

1 Survey and data collection protocol were approved by the ethics board at the University of Exeter (application id eUEBS003014v2.0). Research funding from the Creative-Pioneering Researchers Program at Seoul National University, and from the European University Institute are gratefully acknowledged. This paper uses data that is publicly available at https://osf.io/aubkc/.
2 Professor of Economics, European University Institute.
3 Professor of Economics, Seoul National University.
4 PhD Candidate, European University Institute.
5 Assistant Professor, University of Vermont.
6 Professor of Economics, University of Exeter.
7 Broadus Mitchell Associate Professor of Economics, Johns Hopkins University.

Covid Economics 38, 16 July 2020: 196-217

1 Introduction

The Covid-19 pandemic has affected almost all countries in the world and has led to unprecedented measures being implemented to contain the virus. Countries have differed in their response to the epidemic. Some adopted stringent measures, such as shelter-in-place order, while others implemented early and widespread testing and tracing procedures.

The adjustments required to contain the epidemic have had a dramatic impact on how we live, on our ability to work and on our leisure activities. A key concern is that the groups that have been affected most by the measures taken are not the ones who face the highest risks of severe illness. Such misalignment between personal incentives and burdens on the one hand, and public health concerns on the other hand, are a main challenge in devising and implementing effective public policies.

Without evidence that improves our understanding on the nature of such misalignment, our ability to contain the epidemic and reduce economic and social damages is limited. This paper offers survey-based evidence from six countries on the heterogeneous nature of the economic and social consequences of Covid-19 along with information on behavioral response to the crisis and attitudes towards government measures that have been implemented. Our work complements preliminary evidence put forward in a few recent studies focusing on specific countries and specific aspects (economic) of the crisis such as Adams et al. (2020), Montenovo et al. (2020), Fairlie et al. (2020), von Gaudecker et al. (2020).

We document how the experience of the epidemic and measures that have been implemented have differed according to two key individual characteristics: age and income. The evidence is based on data collected in the third week of April 2020 on samples of around 6,000 individuals from three Western countries—US, UK and Italy—and three Asian countries—China, Japan and South Korea (Belot et al., 2020). The samples are nationally representative on three dimensions: age, gender and income. We focus in this paper on how age and income gradients relate to the Covid pandemic.

At the time of data collection, countries we examined were at different phases of the epidemic and had implemented different measures.[1] These differences, on top of differences in other factors (such as cultural attitudes), can all contribute to explain the cross-country differences in the nature of the Covid-19 effects. Instead of identifying the causes of such differences, this paper marks a first step in understanding how the pandemic has affected different age and income groups across countries.

First, we document objective changes in two key aspects of life: (1) economic variables (income/consumption); and (2) leisure and social life. In particular, we examine whether people experienced a loss in income and a drop in spending during the pandemic. Next, we look at how people reduced the frequency of different behaviors that have a social

[1]See for the six countries, in Figures A1 and A2 of the Online Supplementary Material, how the data collection time window overlaps with the time series of the number of confirmed infected people and deaths per million inhabitants together with the stringency index of government policies to contain the pandemic, based on the Oxford Covid-19 Government Response Tracker data (Hale et al., 2020).

Covid Economics 38, 16 July 2020: 196-217

component—attending large social gatherings, visiting family and friends, and going to large close or open public spaces.

Second, we examine subjective non-financial consequences that individuals report experiencing. Negative effects we consider include boredom, loneliness, trouble sleeping, anxiety and stress, and conflicts with friends and neighbors. These negative, non-financial effects are potentially important because they speak to the burden of complying with measures to contain the pandemic and also shape incentives for individuals to follow social distance measures. Positive effects we examine include spending more time with family, enjoying more free time, and reductions in pollution and noise.

Third, we look at a measure of a specific *behavior* that has aspects of solidarity and precaution, in particular, wearing a mask. In the presence of strong externality, the benefits of wearing a mask are understood to be not only to protect oneself but also to reduce chances of transmitting the corona virus to other people (Chu et al., 2020; Howard et al., 2020).

Finally, we examine measures of support for the approach taken by each country's government, and examine to what extent differences in support can be explained by differences in the impact of the pandemic on individuals.

2 Materials and Methods

2.1 Data

We use publicly available data that was collected by Belot et al. (2020) between April 15 and April 23. This dataset includes 6,082 respondents; roughly 1000 from each of six countries. Three Asian countries (China, Japan and South Korea) and three western countries (Italy, the United Kingdom and the United States). For each country, the sample is nationally representative along age, gender, and (pre-Covid) household income. In the United States the data includes respondents from the 4 most populous states: California, Florida, New York and Texas. American respondents self-identify their race, and the sample is also nationally representative along this dimension.

2.2 Methods

Our analysis is based on ordinary least square models (or linear probability models when the outcome is binary). The right hand side variables include age and income dummies, as well as additional control variables such as gender, a rural-urban indicator, and regional dummies. The age categories we consider are: below 25 (between 18 and 25), 26-45, 46-65 and above 65. For income, we use the categorical variable indicating the household income quintile as reported by the respondent.

We first examine the extent to which groups of different ages and income quintiles have been affected differently in their economic situation and in their social life. To assess

CEPR PRESS

the economic impact, we consider two key variables: (1) Experience of a fall in household income and (2) Experience of a drop in spending.

To assess the impact on social life and leisure, we construct an index measuring the degree of engagement in different leisure activities that have a social component. Respondents were asked to indicate the frequency of engagement in a series of activities at different points in time: in normal times before the outbreak, at the start of the outbreak and at the time of the survey. The index aggregates information from four variables: participation in large social gatherings, going to large close spaces (such as a museum or a shopping center), going to large open spaces (such as a public park), visiting friends/family.

Second, we examine self-reported positive and negative non-financial effects of the pandemic and measures implemented. Negative effects include anxiety, trouble sleeping, increased conflicts, boredom or loneliness. Positive effects include more time with family, more free time, less pollution or less noise. Survey participants could indicate as many as applicable. We construct two simple indicators of the number of positive and negative effects indicated.

Third, we look at one specific behavior that appears to exhibit large cross-country variation: wearing a facial mask. This behavior is interesting because it is not very costly, and it has a clear element of solidarity, since the main benefit appears to reduce transmission to others (rather than protecting oneself). However, countries did not universally recommended the use of masks in their population, at least not early on.

The final variable of interest is the support for the approach taken by the country's government. Here we will highlight age and income differences in beliefs of effectiveness of measures implemented and in the general support for the approach taken by the government. We explore to what extent these can be explained by the variables capturing the economic and social impact of the pandemic. We present an analysis where we add controls for variables capturing the economic and social impact as described above.

Note that we will interpret significance levels of the coefficients at face value, without implementing corrections for multiple hypotheses testing. Given how little is known on the topic this analysis is necessarily exploratory, and confirmatory research will be needed. Yet, the goal is to see if a coherent story emerges.

3 Results

3.1 Economic and social consequences of the pandemic

We find evidence of a negative age gradient in the probability of having experienced a fall in household income across all countries except South Korea, as shown in Figure 1. The oldest group (65+) is 47 percentage points less likely to have experienced a drop in income in China, relative to the youngest group (18-25), the difference is large but less pronounced in other countries (around 25 percentage points in Japan, Italy and the US, and 35 percentage points in the UK). For income, we find a less clear pattern, except for

Covid Economics 38, 16 July 2020: 196-217

Covid Economics 38, 16 July 2020: 196-217

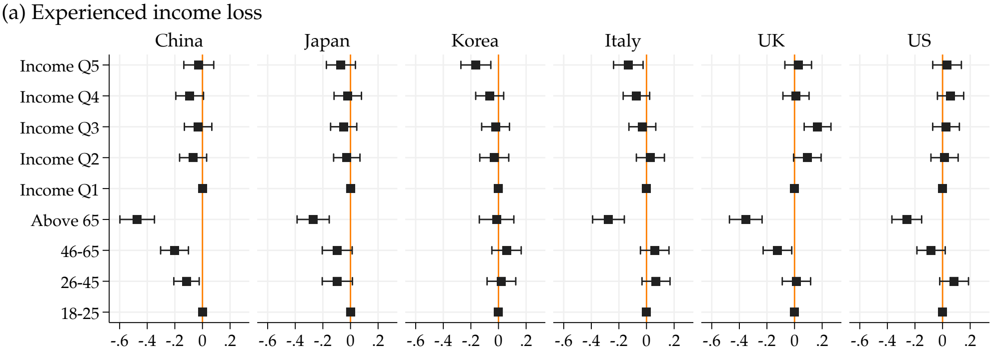
Italy and Korea, where those with incomes in the top 20% are significantly less likely (by 13 and 16 percentage points, respectively) to have experienced a fall in household income.[2]

Note that in all countries, we see a very similar pattern in the probability of teleworking (see Table A3): Younger groups and higher income groups are substantially more likely to be teleworking than those in the bottom 20% income. In China, the 46-65 are 25 percentage points less likely to telework relative to to the 18-25 group. The difference is smaller but remains large in other countries, except for the US and Italy, where there is no significant difference.

Regarding spending, we see a similar pattern as with income loss, according to age. Older groups are less likely to have experienced a fall in consumption. We also see a positive income gradient, but only in the UK and the US: The higher income groups are more likely to have experienced a drop in consumption.[3] The most likely explanation is the closure of shops and leisure-related facilities.

The picture that emerges is heterogeneous: Some countries appear to have succeeded in the early months of the pandemic to shield lower income groups from negative financial effects (like China, Japan, UK and US), while others did not (South Korea, Italy).

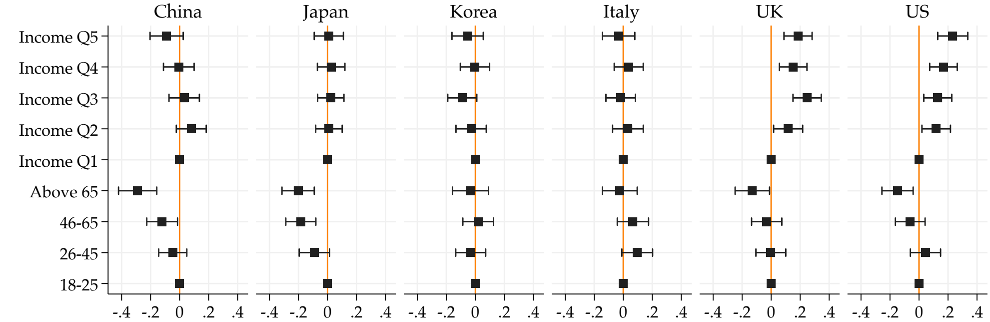
(a) Experienced income loss



[2]We do not present cross-country comparisons of job loss because the financial implications of job loss vary across countries, depending on transfer programs that have been implemented as a result of the crisis. For the US, Papageorge et al. (2020) show an income gradient in the probability of permanent job loss.

[3]This echoes evidence from credit card transaction data reported in Chetty et al. (2020) for the US.

(b) Experienced drop in spending



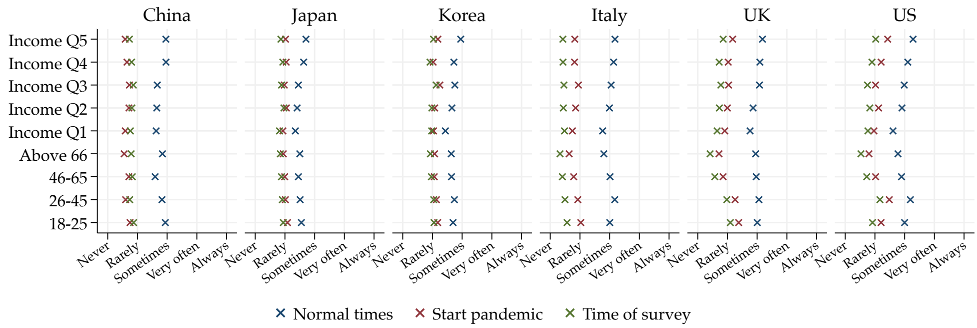
Country average for:	China	Japan	Korea	Italy	UK	US
Experienced income loss (share)	0.603	0.385	0.478	0.579	0.428	0.486
Experienced drop in spending (share)	0.565	0.308	0.489	0.566	0.396	0.431

Note: Point estimates and 95% confidence intervals from a linear probability model of an indicator variable. This indicator denotes loss of household income during the pandemic in panel (a), drop in consumption in panel (b). Covariates include income quintile, age group, gender and geographical controls. Figure based on regression results in Tables A1, A2. 18-25 and Income Q1 are baseline categories for age and income quintile groups, respectively. Country averages of the outcome variables used for the regression are reported below the figure.

Figure 1: Age and income gradients on drop in household income and spending

Turning to social interactions, Figure 2 shows the mean reported levels of our index variable at three points in time—in normal times before the outbreak, at the start of the outbreak and at the time of the survey. In all countries, the younger groups (18-25 or 26-45) are most engaged in social activities. But the older groups appear to have reduced their social life most. There is also a clear income gradient: Higher income groups are more likely to engage in leisure activities with a social component, in all countries. Since those were effectively discouraged or forbidden at the time of the survey, higher income groups by then had experienced a larger negative impact on their social life in most of the countries. This is evidenced by marked income gradients on how bothered they report being for not being able to participate in large social gatherings, go to large (close or open) spaces, and visit friends or family (see Table A4).

Covid Economics 38, 16 July 2020: 196-217



Note: We report group averages of an index that includes frequency of (i) participation in large social gatherings, (ii) visit to large open spaces, (iii) large close spaces, and (iv) visits to friends or family. The index is constructed by averaging (i)-(iv) frequencies on a 1 to 5 scale, where "1" is "Never" and "5" is "Always".

Figure 2: Social interactions over time, by age and income groups

Summarizing and looking across countries, we find that those who experienced the largest negative economic impacts are the young, while older groups and high income groups experienced the largest negative impact in their social life and leisure.

3.2 Psychological costs and the positive side of the pandemic

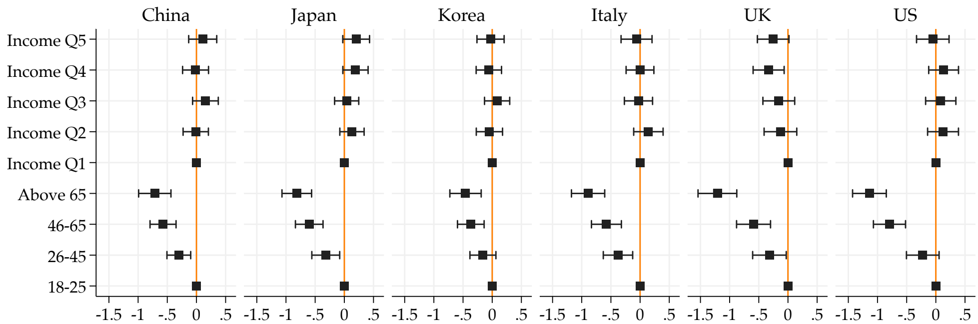
Looking at negative non-financial effects, we find that the younger groups are more likely to report negative effects, in all countries. Understanding the higher psychological costs of the younger groups is important because they may comply less with social distancing measures. Again, the pattern is less clear across income groups: There is no gradient in China, Korea, Italy and the US, but there is a negative income gradient in the UK and a positive one in Japan.

We also find that people report experiencing some benefits from the pandemic—between enjoying more free time, enjoying time with family, cleaner air, and less noise pollution. The older groups are less likely to report positive effects (Table A7). We see a positive income gradient in Japan, Italy, and the US, where people in the lowest income quintile report fewer positive effects from the crisis.

Summarizing, we find that young people are most affected (negatively and positively) in non-financial, psychological terms; all income groups appear to experience negative effects, but positive effects appear concentrated among the higher income groups.

Covid Economics 38, 16 July 2020: 196-217

Covid Economics 38, 16 July 2020: 196-217



	China	Japan	Korea	Italy	UK	US
Country average	1.675	1.228	1.408	1.640	1.659	1.693

Note: Point estimates and 95% confidence intervals from a linear regression model of number of negative non-financial effects due to the pandemic (which include: (i) boredom, (ii) loneliness, (iii) trouble sleeping, (iv) general anxiety and stress, and (v) increased conflicts with friends/family/neighbors) on income quintile, age group, gender and geographical controls. Figure based on regression results in Table A5. 18-25 and Income Q1 are baseline categories for age and income quintile groups, respectively. Country averages of the outcome variable used for the regression are reported below the figure.

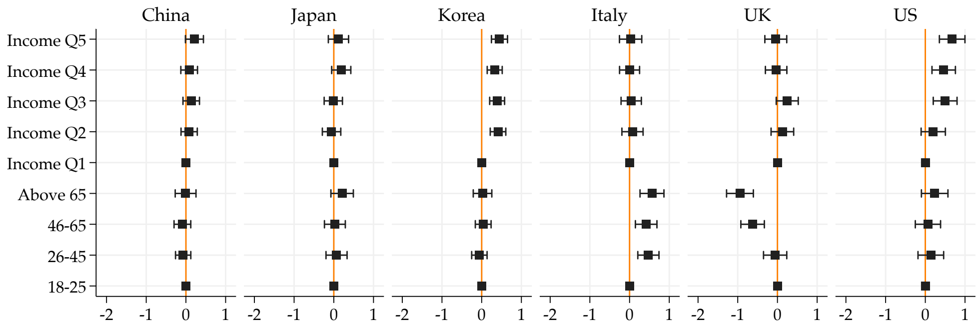
Figure 3: Age and income gradients on negative non-financial effects

3.3 Wearing a facial mask

We now look at the age and income gradients in the probability of wearing a mask Figure 4. This behavior is interesting because it involves a relatively low cost and it has a clear solidarity component, since the benefit appears to accrue mostly to others rather than oneself (Chu et al., 2020; Howard et al., 2020).

Here we find striking differences across countries. First, on average, the willingness to wear a mask is much higher in Asian countries and Italy than in the UK and the US. Second, in Asian countries, there is hardly any gradient along age or income. In Korea, only the bottom 20% income group appears more reluctant to wear a mask. In the US, there is a clear positive income gradient.[4] In the UK differences are most pronounced according to age, with the older groups much more reluctant to wear a mask. In Italy, the age gradient goes in the opposite direction.

[4]In a related paper focusing on the US and using the same data (Papageorge et al. (2020)), we study changes in a wider range of self-protective behaviors in the United States. We find that higher income groups are more likely to adopt self-protective measures in response to the outbreak.



	China	Japan	Korea	Italy	UK	US
Country average	4.411	4.172	4.531	4.252	1.902	3.514

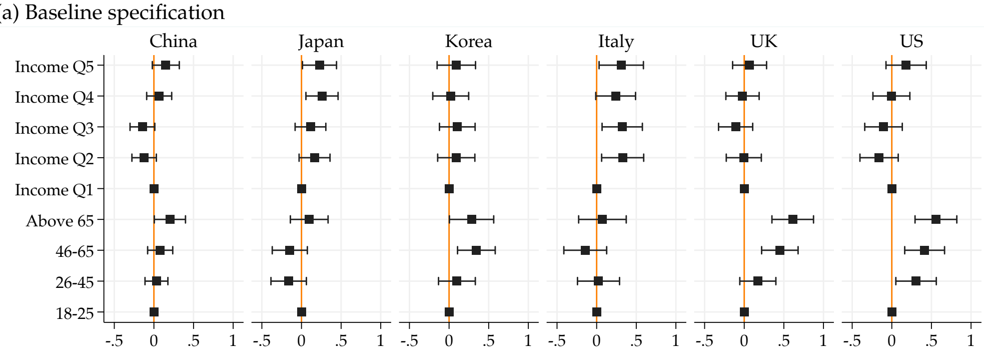
Note: Point estimates and 95% confidence intervals from a linear regression model for frequency of use of face masks on income quintile, age group, gender and geographical controls. Figure based on regression results in Table A8. 18-25 and Income Q1 are baseline categories for age and income quintile groups, respectively. Country averages of the outcome variable used for the regression are reported below the figure.

Figure 4: Age and income gradients on frequency of use face mask

3.4 Support for the government and recommendations

The last question we turn to is the support of the population for the approach taken by their governments (Figure 5). Older individuals tend to be more supportive. Though we do not observe such gradients in China—possibly because of ceiling effects, and Japan—where support for the government is the lowest. The pattern across income is again less clear.

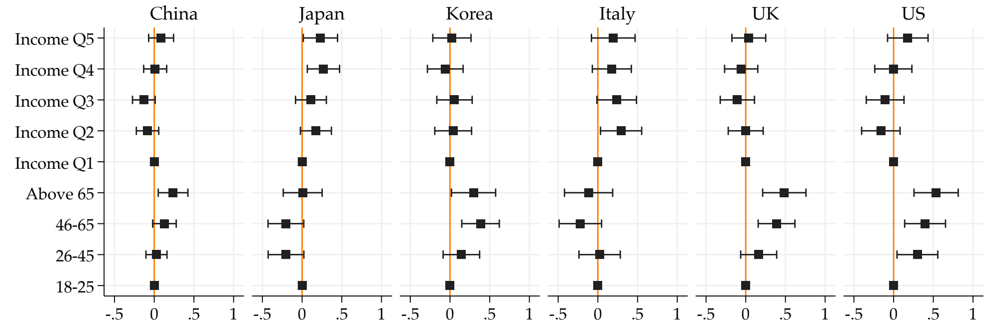
(a) Baseline specification



Covid Economics 38, 16 July 2020: 196-217

Covid Economics 38, 16 July 2020: 196-217

(b) Specification with additional controls



	China	Japan	Korea	Italy	UK	US
Country average	4.517	2.902	3.686	3.727	3.887	3.757

Note: Point estimates and 95% confidence intervals from a linear regression model of government support on income quintile, age group, gender, geographical controls, and additional controls. Additional controls include indicators for having lost the job at least temporarily and having lost household income, as well as count variables of the negative non-financial and positive non-financial effects). 18-25 and Income Q1 are baseline categories for age and income quintile groups, respectively. Country averages of the outcome variable used for the regression are reported below the figure.

Figure 5: Age and income gradients on support of the government's handling of the pandemic

We explore to what extent age gradients can be explained by disproportionate effects of the pandemic on different groups (using the variables presented above). When we control for these variables (see bottom panel of Figure 5), we do not see substantial changes in these age and income gradients, suggesting that support is not directly driven by the economic, social or psychological impact of the outbreak on individuals.

4 Discussion

The epidemic and measures taken in response to it across the world appear to have affected different groups of the population in different ways. As a result, some subgroups of the population are economically and psychologically more vulnerable than other subgroups. Understanding the heterogeneous nature of the impact of Covid-19 impacts is a necessary step toward improving the current set of policy tools, i.e., to encourage compliance with measures that align with societal goals of containing the pandemic while minimizing economic and social damage.

In the six countries we surveyed, we find consistent evidence that younger people have been more negatively affected—both economically and psychologically—and that they appear to be less supportive of their governments' approaches. On the other hand, we find a less clear pattern across income groups. Despite their lower ability to work from home, lower income people have not necessarily experienced the strongest negative income consequences, at least not in all countries. Some countries took early measures to shield the low income groups from the economic consequences of the crisis. However, our findings on income groups are not definitive because the extent of economic impacts was not fully revealed by the time of our survey.

This evidence that younger people are more affected by the pandemic and support less their government response strengthens the case for more differentiated policies that shield the young from the negative consequences of the epidemic and necessary measures. A number of recent papers propose policies that target lockdown policies to the older part of the population (see, for the effects of age-specific policies, Acemoglu et al., 2020; Brotherhood et al., 2020; Favero et al., 2020). The advantage would be that such targeted policies would allow for economic recovery, while shielding those with the highest health risks. However, the consequences of shutting down interactions between the old and the young are not yet well understood. People from different age groups rely on each other for many reasons, and breaking such inter-generational bonds and arrangements may have negative consequences, which are difficult to assess and will require more empirical work. It is also imperative to find ways to match young people's incentives and burdens of complying to public policies.

While our focus is on age and income differences, our data present important systematic patterns indicating that women are disproportionately affected by this economic crisis.[5] For most of the surveyed countries women are less likely to have started teleworking, more likely to be socially isolated because of the pandemic, and more likely to report suffering psychological consequences of the pandemic. These findings echo evidence presented by Alon et al. (2020) who show that women are concentrated in sectors disproportionately affected by the crisis. The data from Belot et al. (2020) do not include some of the questions that are key for understanding the sources of such gender gaps (e.g task allocation within the household), but a cross-country perspective can prove helpful in directing ongoing investigations on the root causes of such gender gaps.

References

Acemoglu, Daron, Victor Chernozhukov, Iván Werning, and Michael D Whinston, "A multi-risk sir model with optimally targeted lockdown," Technical Report, National Bureau of Economic Research 2020.

Adams, Abigail, Teodora Boneva, Christopher Rauh, and Marta Golin, "Inequality in the Impact of the Coronavirus Shock: New Survey Evidence for the UK," 2020, (2023).

Alon, Titan M, Matthias Doepke, Jane Olmstead-Rumsey, and Michèle Tertilt, "The Impact of COVID-19 on Gender Equality," Working Paper 26947, National Bureau of Economic Research April 2020.

Belot, Michèle, Syngjoo Choi, Julian C Jamison, Nicholas W Papageorge, Egon Tripodi, and Eline van den Broek-Altenburg, "Six-Country Survey on Covid-19," 2020.

[5]See also Etheridge et al. (2020) for evidence from the UK.

Covid Economics 38, 16 July 2020: 196-217

Brotherhood, Luiz, Philipp Kircher, Cezar Santos, and Michèle Tertilt, "An economic model of the Covid-19 epidemic: The importance of testing and age-specific policies," 2020.

Chetty, Raj, John N Friedman, Nathaniel Hendren, and Michael Stepner, "How Did COVID-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data," 2020.

Chu, Derek K, Elie A Akl, Stephanie Duda, Karla Solo, Sally Yaacoub, Holger J Schünemann, Amena El-harakeh, Antonio Bognanni, Tamara Lotfi, Mark Loeb et al., "Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis," *The Lancet*, 2020.

Etheridge, Ben, Lisa Spantig et al., "The gender gap in mental well-being during the Covid-19 outbreak: evidence from the UK," Technical Report, Institute for Social and Economic Research 2020.

Fairlie, Robert W, Kenneth Couch, and Huanan Xu, "The Impacts of COVID-19 on Minority Unemployment: First Evidence from April 2020 CPS Microdata," Working Paper 27246, National Bureau of Economic Research May 2020.

Favero, Carlo A, Andrea Ichino, and Aldo Rustichini, "Restarting the economy while saving lives under Covid-19," 2020.

Hale, Thomas, Anna Petherick, Toby Phillips, and Samuel Webster, "Variation in government responses to COVID-19," *Blavatnik School of Government Working Paper*, 2020.

Howard, Jeremy, Austin Huang, Zhiyuan Li, Zeynep Tufekci, Vladimir Zdimal, Helene-Mari van der Westhuizen, Arne von Delft, Amy Price, Lex Fridman, Lei-Han Tang et al., "Face masks against COVID-19: an evidence review," 2020.

Montenovo, Laura, Xuan Jiang, Felipe Lozano Rojas, Ian M Schmutte, Kosali I Simon, Bruce A Weinberg, and Coady Wing, "Determinants of Disparities in Covid-19 Job Losses," Working Paper 27132, National Bureau of Economic Research May 2020.

Papageorge, Nick, Matthew V. Zahn, Michéle Belot, Eline van den Broek-Altenberg, Syngjoo Choi, Julian C. Jamison, and Egon Tripodi, "Socio-Demographic Factors Associated with Self-Protecting Behavior during the Covid-19 Pandemic," 2020.

von Gaudecker, Hans-Martin, Radost Holler, Lena Janys, Bettina Siflinger, and Christian Zimpelmann, "Labour supply in the early stages of the CoViD-19 Pandemic: Empirical Evidence on hours, home office, and expectations," 2020.

Covid Economics 38, 16 July 2020: 196-217

CEPR PRESS

Appendix for Online Publication

A Additional Tables

Table A1: Linear probability model for having experienced household income loss

	China	Japan	Korea	Italy	UK	US	6 countries
Female	-0.039	0.009	0.033	0.084***	-0.018	0.035	0.012
	(0.031)	(0.031)	(0.033)	(0.030)	(0.030)	(0.032)	(0.013)
Age group (*baseline: 18 to 25*)							
26 to 45	-0.116**	-0.095*	0.021	0.069	0.014	0.084	-0.011
	(0.046)	(0.055)	(0.052)	(0.051)	(0.052)	(0.053)	(0.021)
46 to 65	-0.202***	-0.096*	0.058	0.059	-0.124**	-0.083	-0.073***
	(0.051)	(0.055)	(0.054)	(0.052)	(0.052)	(0.052)	(0.021)
Above 66	-0.473***	-0.271***	-0.014	-0.276***	-0.354***	-0.259***	-0.283***
	(0.063)	(0.059)	(0.063)	(0.058)	(0.060)	(0.054)	(0.024)
Income quintile (*baseline: First quintile*)							
Second quintile	-0.068	-0.027	-0.031	0.028	0.093*	0.014	0.003
	(0.049)	(0.048)	(0.053)	(0.051)	(0.050)	(0.050)	(0.021)
Third quintile	-0.031	-0.050	-0.021	-0.030	0.167***	0.025	0.011
	(0.050)	(0.048)	(0.051)	(0.049)	(0.049)	(0.049)	(0.020)
Fourth quintile	-0.093*	-0.020	-0.064	-0.073	0.010	0.057	-0.024
	(0.051)	(0.050)	(0.051)	(0.048)	(0.048)	(0.048)	(0.020)
Fifth quintile	-0.027	-0.069	-0.164***	-0.132**	0.026	0.032	-0.038*
	(0.055)	(0.053)	(0.055)	(0.054)	(0.049)	(0.052)	(0.021)
Current living area (*baseline: Urban*)							
Semi-urban	0.110***	-0.108***	-0.048	-0.012	-0.052	-0.027	-0.032**
	(0.039)	(0.038)	(0.048)	(0.035)	(0.035)	(0.034)	(0.015)
Country-side	0.205***	-0.094**	-0.141*	-0.123***	0.076	-0.132**	-0.040*
	(0.061)	(0.043)	(0.083)	(0.043)	(0.050)	(0.053)	(0.021)
Constant	0.609***	0.654***	0.494***	0.552***	0.420***	0.505***	0.482***
	(0.083)	(0.076)	(0.078)	(0.119)	(0.078)	(0.066)	(0.069)
Regional fixed effects	Y	Y	Y	Y	Y	Y	Y
Country fixed effects	N	N	N	N	N	N	Y
Observations	999	1013	964	1042	1016	1055	6089
adj. R^2	0.101	0.027	0.007	0.091	0.097	0.076	0.071

$^*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: Standard errors in parentheses. All specifications include regional fixed effects for the place of residence of the respondent (relevant administrative level is the *province* in China and South Korea, the *region* in Japan, Italy and the United Kingdom, and the *state* in the United States).

Covid Economics 38, 16 July 2020: 196-217

Table A2: Linear probability model for having experienced drop in spending

	China	Japan	Korea	Italy	UK	US	6 countries
Female	0.034	0.041	0.028	0.006	-0.097***	-0.061*	-0.016
	(0.032)	(0.030)	(0.033)	(0.031)	(0.030)	(0.032)	(0.013)
Age group (*baseline: 18 to 25*)							
26 to 45	-0.047	-0.091*	-0.032	0.096*	-0.002	0.044	-0.011
	(0.049)	(0.053)	(0.052)	(0.053)	(0.052)	(0.053)	(0.021)
46 to 65	-0.120**	-0.184***	0.019	0.067	-0.030	-0.061	-0.055***
	(0.054)	(0.052)	(0.053)	(0.054)	(0.053)	(0.052)	(0.021)
Above 66	-0.289***	-0.203***	-0.034	-0.024	-0.129**	-0.149***	-0.138***
	(0.067)	(0.056)	(0.063)	(0.061)	(0.060)	(0.054)	(0.024)
Income quintile (*baseline: First quintile*)							
Second quintile	0.080	0.010	-0.029	0.032	0.118**	0.119**	0.054***
	(0.052)	(0.046)	(0.053)	(0.053)	(0.051)	(0.050)	(0.021)
Third quintile	0.032	0.022	-0.090*	-0.018	0.248***	0.129***	0.058***
	(0.053)	(0.046)	(0.051)	(0.051)	(0.049)	(0.049)	(0.020)
Fourth quintile	-0.006	0.025	-0.003	0.038	0.152***	0.168***	0.074***
	(0.053)	(0.048)	(0.051)	(0.051)	(0.048)	(0.048)	(0.020)
Fifth quintile	-0.090	0.009	-0.053	-0.032	0.185***	0.232***	0.058***
	(0.058)	(0.051)	(0.055)	(0.056)	(0.049)	(0.052)	(0.021)
Current living area (*baseline: Urban*)							
Semi-urban	0.016	-0.001	-0.069	-0.025	-0.046	0.031	-0.020
	(0.041)	(0.036)	(0.048)	(0.037)	(0.035)	(0.034)	(0.015)
Country-side	-0.014	-0.037	-0.104	-0.086*	0.015	-0.039	-0.044**
	(0.064)	(0.041)	(0.083)	(0.045)	(0.050)	(0.053)	(0.021)
Constant	0.519***	0.380***	0.487***	0.427***	0.325***	0.407***	0.455***
	(0.087)	(0.073)	(0.078)	(0.124)	(0.078)	(0.065)	(0.069)
Regional fixed effects	Y	Y	Y	Y	Y	Y	Y
Country fixed effects	N	N	N	N	N	N	Y
Observations	999	1013	964	1042	1016	1055	6089
adj. R^2	0.027	0.012	0.018	0.017	0.068	0.062	0.055

$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$
Notes: Standard errors in parentheses. All specifications include regional fixed effects for the place of residence of the respondent (relevant administrative level is the *province* in China and South Korea, the *region* in Japan, Italy and the United Kingdom, and the *state* in the United States).

Table A3: Ordinary least squares for having started teleworking

	China	Japan	Korea	Italy	UK	US	6 countries
Female	-0.007	-0.045*	-0.046**	-0.027	-0.102***	-0.080***	-0.057***
	(0.029)	(0.023)	(0.023)	(0.026)	(0.023)	(0.024)	(0.010)
Age group (*baseline: 18 to 25*)							
26 to 45	-0.111**	-0.050	-0.055	0.025	0.015	0.157***	-0.000
	(0.044)	(0.041)	(0.037)	(0.045)	(0.040)	(0.040)	(0.017)
46 to 65	-0.252***	-0.108***	-0.120***	-0.022	-0.089**	0.001	-0.092***
	(0.048)	(0.041)	(0.037)	(0.046)	(0.041)	(0.039)	(0.017)
Above 66	-0.574***	-0.242***	-0.136***	-0.235***	-0.193***	-0.141***	-0.242***
	(0.060)	(0.044)	(0.044)	(0.051)	(0.046)	(0.041)	(0.019)
Income quintile (*baseline: First quintile*)							
Second quintile	0.211***	0.084**	0.034	0.066	0.025	0.080**	0.094***
	(0.047)	(0.036)	(0.037)	(0.045)	(0.039)	(0.038)	(0.016)
Third quintile	0.075	0.156***	0.145***	0.161***	0.095**	0.153***	0.145***
	(0.047)	(0.036)	(0.035)	(0.043)	(0.038)	(0.037)	(0.016)
Fourth quintile	0.124***	0.214***	0.130***	0.229***	0.152***	0.259***	0.201***
	(0.048)	(0.038)	(0.036)	(0.042)	(0.037)	(0.036)	(0.016)
Fifth quintile	0.171***	0.327***	0.162***	0.351***	0.219***	0.304***	0.278***
	(0.052)	(0.040)	(0.038)	(0.047)	(0.038)	(0.040)	(0.017)
Current living area (*baseline: Urban*)							
Semi-urban	-0.118***	-0.010	-0.029	-0.102***	-0.041	-0.041	-0.073***
	(0.037)	(0.028)	(0.034)	(0.031)	(0.027)	(0.026)	(0.012)
Country-side	-0.249***	-0.029	-0.038	-0.127***	-0.059	-0.128***	-0.110***
	(0.058)	(0.032)	(0.058)	(0.038)	(0.039)	(0.040)	(0.017)
Constant	0.674***	0.143**	0.154***	0.222**	0.163***	0.107**	0.508***
	(0.078)	(0.057)	(0.054)	(0.104)	(0.060)	(0.050)	(0.055)
Regional fixed effects	Y	Y	Y	Y	Y	Y	Y
Country fixed effects	N	N	N	N	N	N	Y
Observations	999	1013	964	1042	1016	1055	6089
adj. R^2	0.229	0.149	0.036	0.126	0.110	0.259	0.213

$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: Standard errors in parentheses. All specifications include regional fixed effects for the place of residence of the respondent (relevant administrative level is the *province* in China and South Korea, the *region* in Japan, Italy and the United Kingdom, and the *state* in the United States).

Covid Economics 38, 16 July 2020: 196-217

CEPR PRESS

Table A4: Ordinary least squares for index of dissatisfaction with social distance

	China	Japan	Korea	Italy	UK	US	6 countries
Female	-0.266***	0.366***	0.129	0.024	0.083	0.063	0.067*
	(0.077)	(0.107)	(0.109)	(0.084)	(0.098)	(0.100)	(0.039)
Age group (*baseline: 18 to 25*)							
26 to 45	-0.170	-0.124	-0.202	0.181	-0.268	0.088	-0.102
	(0.116)	(0.190)	(0.173)	(0.144)	(0.169)	(0.165)	(0.065)
46 to 65	-0.256**	-0.324*	-0.226	-0.002	-0.226	-0.098	-0.184***
	(0.128)	(0.188)	(0.177)	(0.147)	(0.171)	(0.163)	(0.065)
Above 66	-0.342**	-0.947***	-0.731***	-0.238	-0.523***	-0.272	-0.519***
	(0.158)	(0.202)	(0.208)	(0.163)	(0.195)	(0.171)	(0.073)
Income quintile (*baseline: First quintile*)							
Second quintile	0.325***	0.253	0.462***	-0.007	0.425***	0.331**	0.286***
	(0.124)	(0.165)	(0.175)	(0.144)	(0.165)	(0.156)	(0.063)
Third quintile	0.359***	0.258	0.668***	0.171	0.430***	0.313**	0.337***
	(0.126)	(0.164)	(0.168)	(0.138)	(0.160)	(0.153)	(0.062)
Fourth quintile	0.337***	0.412**	0.895***	0.367***	0.576***	0.431***	0.500***
	(0.127)	(0.172)	(0.169)	(0.136)	(0.156)	(0.151)	(0.061)
Fifth quintile	0.388***	0.558***	0.803***	0.444***	0.550***	0.386**	0.511***
	(0.137)	(0.182)	(0.181)	(0.152)	(0.159)	(0.165)	(0.065)
Current living area (*baseline: Urban*)							
Semi-urban	0.203**	0.031	-0.334**	0.124	-0.035	-0.006	0.015
	(0.098)	(0.130)	(0.159)	(0.099)	(0.114)	(0.107)	(0.047)
Country-side	-0.053	0.032	-0.672**	-0.166	0.010	0.144	-0.033
	(0.153)	(0.147)	(0.274)	(0.122)	(0.164)	(0.168)	(0.064)
Constant	-0.478**	-0.160	-0.381	0.033	-0.265	-0.132	-0.646***
	(0.207)	(0.261)	(0.257)	(0.334)	(0.255)	(0.206)	(0.212)
Regional fixed effects	Y	Y	Y	Y	Y	Y	Y
Country fixed effects	N	N	N	N	N	N	Y
Observations	999	1013	964	1042	1016	1055	6089
adj. R^2	0.075	0.049	0.052	0.040	0.042	0.016	0.037

$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: Standard errors in parentheses. All specifications include regional fixed effects for the place of residence of the respondent (relevant administrative level is the *province* in China and South Korea, the *region* in Japan, Italy and the United Kingdom, and the *state* in the United States). The outcome is index that is constructed as the first principal component fitting the variables that capture how bothered respondents are for not being able to meet other people in their free time, do leisure activities outside of home, shop non-essentials.

Covid Economics 38, 16 July 2020: 196-217

CEPR PRESS

Covid Economics 38, 16 July 2020: 196-217

Table A5: Ordinary least square for negative non-financial effects

	China	Japan	Korea	Italy	UK	US	6 countries
Female	0.024	0.366***	0.241***	0.281***	0.387***	0.404***	0.277***
	(0.068)	(0.068)	(0.071)	(0.074)	(0.085)	(0.085)	(0.030)
Age group (*baseline: 18 to 25*)							
26 to 45	-0.302***	-0.317***	-0.159	-0.379***	-0.318**	-0.219	-0.284***
	(0.103)	(0.121)	(0.113)	(0.127)	(0.145)	(0.141)	(0.051)
46 to 65	-0.573***	-0.600***	-0.367***	-0.577***	-0.589***	-0.792***	-0.568***
	(0.113)	(0.119)	(0.115)	(0.130)	(0.147)	(0.139)	(0.051)
Above 66	-0.714***	-0.813***	-0.457***	-0.891***	-1.207***	-1.134***	-0.885***
	(0.140)	(0.128)	(0.136)	(0.144)	(0.168)	(0.146)	(0.058)
Income quintile (*baseline: First quintile*)							
Second quintile	-0.012	0.131	-0.048	0.142	-0.131	0.122	0.023
	(0.110)	(0.105)	(0.114)	(0.127)	(0.141)	(0.133)	(0.050)
Third quintile	0.154	0.040	0.084	-0.028	-0.157	0.085	0.002
	(0.111)	(0.104)	(0.109)	(0.122)	(0.138)	(0.131)	(0.048)
Fourth quintile	-0.015	0.192*	-0.058	-0.002	-0.331**	0.135	-0.027
	(0.112)	(0.109)	(0.110)	(0.120)	(0.134)	(0.129)	(0.048)
Fifth quintile	0.107	0.203*	-0.029	-0.061	-0.253*	-0.049	-0.026
	(0.121)	(0.116)	(0.118)	(0.134)	(0.137)	(0.141)	(0.051)
Current living area (*baseline: Urban*)							
Semi-urban	0.341***	-0.053	-0.044	0.026	0.099	0.081	0.066*
	(0.087)	(0.082)	(0.104)	(0.088)	(0.098)	(0.091)	(0.037)
Country-side	-0.011	-0.046	-0.330*	-0.372***	-0.057	-0.060	-0.138***
	(0.135)	(0.094)	(0.178)	(0.108)	(0.141)	(0.143)	(0.050)
Constant	1.622***	1.446***	1.644***	1.811***	1.903***	1.939***	1.624***
	(0.183)	(0.166)	(0.167)	(0.295)	(0.219)	(0.176)	(0.166)
Regional fixed effects	Y	Y	Y	Y	Y	Y	Y
Country fixed effects	N	N	N	N	N	N	Y
Observations	999	1013	964	1042	1016	1055	6089
adj. R^2	0.086	0.083	0.030	0.087	0.088	0.105	0.092

$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: Standard errors in parentheses. All specifications include regional fixed effects for the place of residence of the respondent (relevant administrative level is the *province* in China and South Korea, the *region* in Japan, Italy and the United Kingdom, and the *state* in the United States). The outcome is a variable that counts how many of the following non-financial negative effects respondents report to be experiencing due to the pandemic. This includes (i) boredom, (ii) loneliness, (iii) trouble sleeping, (iv) general anxiety and stress, (v) increased conflicts with friends/relatives/neighbors.

Covid Economics 38, 16 July 2020: 196-217

Table A6: Ordinary least squares for index of belief of policy effectiveness (higher values denote belief in higher effectiveness)

	China	Japan	Korea	Italy	UK	US	6 countries
Female	0.209*	0.442***	0.435***	0.247*	0.561***	0.509***	0.387***
	(0.123)	(0.145)	(0.110)	(0.132)	(0.125)	(0.144)	(0.052)
Age group (*baseline: 18 to 25*)							
26 to 45	0.137	0.021	0.212	0.449**	0.504**	0.236	0.259***
	(0.186)	(0.257)	(0.173)	(0.226)	(0.216)	(0.240)	(0.087)
46 to 65	0.106	-0.000	0.355**	0.592**	0.731***	0.428*	0.333***
	(0.204)	(0.255)	(0.178)	(0.230)	(0.218)	(0.236)	(0.088)
Above 66	0.245	0.250	0.556***	0.781***	0.792***	0.597**	0.529***
	(0.253)	(0.274)	(0.209)	(0.256)	(0.248)	(0.247)	(0.099)
Income quintile (*baseline: First quintile*)							
Second quintile	-0.384*	-0.363	0.570***	0.345	0.174	-0.177	0.004
	(0.198)	(0.225)	(0.175)	(0.226)	(0.209)	(0.226)	(0.086)
Third quintile	-0.250	0.314	0.583***	0.332	0.266	0.251	0.240***
	(0.201)	(0.223)	(0.168)	(0.216)	(0.204)	(0.222)	(0.083)
Fourth quintile	0.238	-0.075	0.526***	0.209	0.062	0.404*	0.216***
	(0.203)	(0.233)	(0.169)	(0.214)	(0.199)	(0.218)	(0.083)
Fifth quintile	0.010	-0.263	0.596***	0.090	0.109	0.533**	0.182**
	(0.220)	(0.247)	(0.181)	(0.238)	(0.203)	(0.238)	(0.088)
Current living area (*baseline: Urban*)							
Semi-urban	0.085	0.114	-0.138	0.081	-0.053	-0.210	-0.019
	(0.157)	(0.176)	(0.159)	(0.156)	(0.145)	(0.155)	(0.063)
Country-side	0.233	0.011	0.017	0.066	0.049	-0.258	0.014
	(0.245)	(0.200)	(0.275)	(0.191)	(0.209)	(0.243)	(0.087)
Constant	0.283	-0.832**	-1.743***	-0.295	-1.425***	-0.806***	-0.101
	(0.331)	(0.354)	(0.258)	(0.524)	(0.325)	(0.298)	(0.287)
Regional fixed effects	Y	Y	Y	Y	Y	Y	Y
Country fixed effects	N	N	N	N	N	N	Y
Observations	999	1013	964	1042	1016	1055	6089
adj. R^2	0.005	0.013	0.030	-0.002	0.015	0.021	0.053

$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$
Notes: Standard errors in parentheses. All specifications include regional fixed effects for the place of residence of the respondent (relevant administrative level is the *province* in China and South Korea, the *region* in Japan, Italy and the United Kingdom, and the *state* in the United States).

CEPR PRESS

Covid Economics 38, 16 July 2020: 196-217

Table A7: Ordinary least squares for positive non-financial effects

	China	Japan	Korea	Italy	UK	US	6 countries
Female	-0.003	-0.018	0.101*	-0.010	0.075	0.082	0.025
	(0.062)	(0.051)	(0.059)	(0.063)	(0.069)	(0.069)	(0.025)
Age group (*baseline: 18 to 25*)							
26 to 45	-0.066	-0.140	-0.188**	0.005	-0.227*	0.064	-0.100**
	(0.094)	(0.090)	(0.093)	(0.108)	(0.118)	(0.114)	(0.042)
46 to 65	-0.430***	-0.153*	-0.273***	0.093	0.018	-0.099	-0.127***
	(0.103)	(0.089)	(0.095)	(0.110)	(0.120)	(0.112)	(0.042)
Above 66	-0.546***	-0.092	-0.371***	0.113	0.103	-0.244**	-0.162***
	(0.128)	(0.096)	(0.111)	(0.122)	(0.136)	(0.117)	(0.047)
Income quintile (*baseline: First quintile*)							
Second quintile	0.024	0.154**	0.210**	0.132	-0.136	0.152	0.101**
	(0.100)	(0.079)	(0.093)	(0.108)	(0.115)	(0.108)	(0.041)
Third quintile	0.076	0.188**	0.134	0.229**	-0.111	0.248**	0.144***
	(0.102)	(0.078)	(0.090)	(0.104)	(0.112)	(0.105)	(0.040)
Fourth quintile	-0.055	0.190**	0.191**	0.147	0.030	0.141	0.134***
	(0.103)	(0.082)	(0.090)	(0.102)	(0.109)	(0.104)	(0.040)
Fifth quintile	0.101	0.218**	0.023	0.259**	0.015	0.095	0.158***
	(0.111)	(0.086)	(0.097)	(0.114)	(0.111)	(0.113)	(0.042)
Current living area (*baseline: Urban*)							
Semi-urban	0.296***	-0.171***	0.014	0.149**	0.078	0.069	0.055*
	(0.080)	(0.061)	(0.085)	(0.075)	(0.079)	(0.073)	(0.030)
Country-side	0.163	-0.163**	0.242*	-0.083	0.051	0.016	0.016
	(0.124)	(0.070)	(0.147)	(0.092)	(0.115)	(0.115)	(0.042)
Constant	1.740***	0.518***	0.827***	1.541***	1.329***	1.220***	1.528***
	(0.167)	(0.124)	(0.137)	(0.251)	(0.178)	(0.142)	(0.137)
Regional fixed effects	Y	Y	Y	Y	Y	Y	Y
Country fixed effects	N	N	N	N	N	N	Y
Observations	999	1013	964	1042	1016	1055	6089
adj. R^2	0.060	0.025	0.019	0.016	0.006	0.018	0.115

$^*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: Standard errors in parentheses. All specifications include regional fixed effects for the place of residence of the respondent (relevant administrative level is the *province* in China and South Korea, the *region* in Japan, Italy and the United Kingdom, and the *state* in the United States). The outcome is a variable that counts how many of the following non-financial positive effects respondents report to be experiencing due to the pandemic. This includes (i) enjoying more free time, (ii) enjoying time with family, (iii) reduction of air pollution, (iv) reduction of noise pollution.

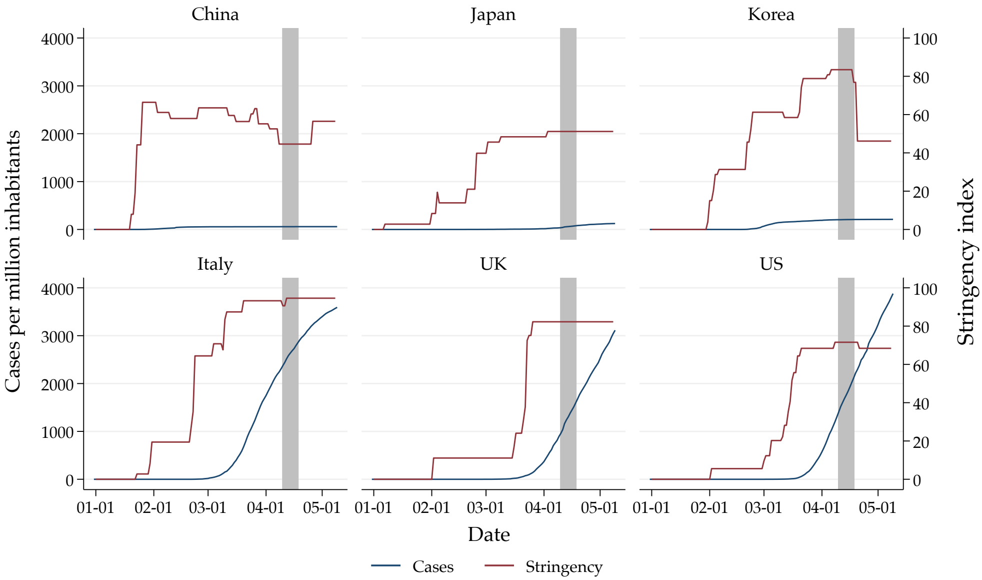
Table A8: Ordinary least squares for frequency of wearing masks "now" (i.e. around time of data collection)

	China	Japan	Korea	Italy	UK	US	6 countries
Female	0.133**	0.663***	0.212***	0.110	-0.174**	0.151	0.177***
	(0.064)	(0.076)	(0.062)	(0.078)	(0.086)	(0.099)	(0.032)
Age group (*baseline: 18 to 25*)							
26 to 45	-0.070	0.070	-0.058	0.474***	-0.058	0.140	0.087
	(0.097)	(0.134)	(0.098)	(0.135)	(0.148)	(0.163)	(0.054)
46 to 65	-0.089	0.029	0.039	0.418***	-0.627***	0.067	-0.049
	(0.107)	(0.133)	(0.100)	(0.137)	(0.149)	(0.161)	(0.054)
Above 66	-0.008	0.210	0.021	0.565***	-0.946***	0.237	0.004
	(0.132)	(0.143)	(0.118)	(0.152)	(0.171)	(0.169)	(0.061)
Income quintile (*baseline: First quintile*)							
Second quintile	0.081	-0.059	0.411***	0.074	0.125	0.199	0.117**
	(0.104)	(0.117)	(0.099)	(0.135)	(0.144)	(0.154)	(0.052)
Third quintile	0.133	-0.015	0.389***	0.042	0.246*	0.500***	0.199***
	(0.105)	(0.116)	(0.095)	(0.129)	(0.140)	(0.151)	(0.051)
Fourth quintile	0.082	0.186	0.328***	-0.000	-0.034	0.463***	0.157***
	(0.106)	(0.122)	(0.096)	(0.127)	(0.136)	(0.149)	(0.051)
Fifth quintile	0.214*	0.116	0.448***	0.026	-0.042	0.677***	0.249***
	(0.115)	(0.129)	(0.102)	(0.142)	(0.139)	(0.162)	(0.054)
Current living area (*baseline: Urban*)							
Semi-urban	0.123	-0.116	0.048	0.012	-0.280***	-0.422***	-0.150***
	(0.082)	(0.092)	(0.090)	(0.093)	(0.099)	(0.105)	(0.039)
Country-side	-0.090	-0.050	-0.217	0.025	-0.122	-0.641***	-0.178***
	(0.128)	(0.104)	(0.155)	(0.114)	(0.143)	(0.166)	(0.053)
Constant	4.330***	3.728***	4.040***	3.843***	2.471***	3.267***	4.280***
	(0.173)	(0.185)	(0.145)	(0.312)	(0.223)	(0.203)	(0.176)
Regional fixed effects	Y	Y	Y	Y	Y	Y	Y
Country fixed effects	N	N	N	N	N	N	Y
Observations	999	1013	964	1042	1016	1055	6089
adj. R^2	0.011	0.075	0.045	0.002	0.116	0.056	0.363

$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: Standard errors in parentheses. All specifications include regional fixed effects for the place of residence of the respondent (relevant administrative level is the *province* in China and South Korea, the *region* in Japan, Italy and the United Kingdom, and the *state* in the United States).

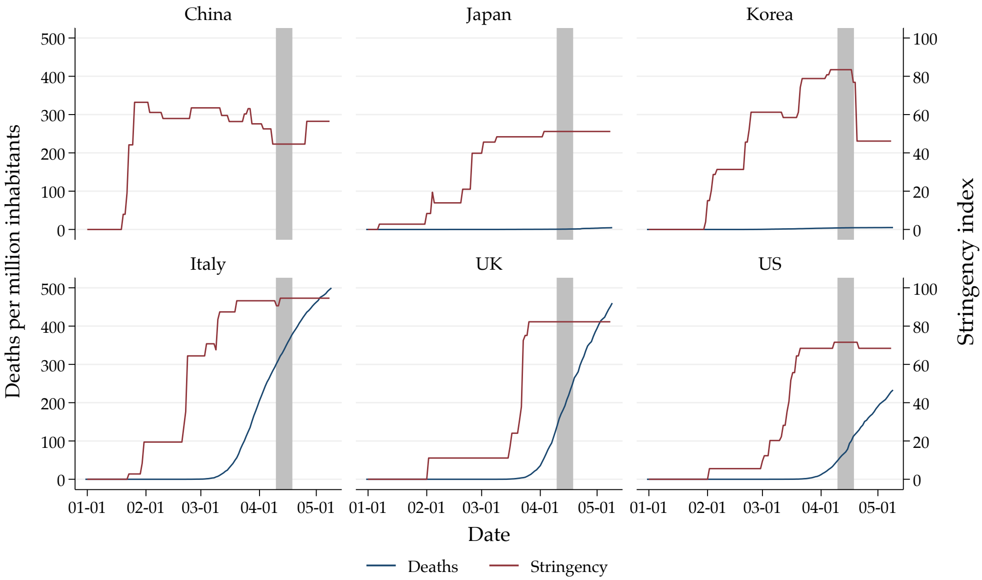
Covid Economics 38, 16 July 2020: 196-217



Note: The gray bar represents the third week of April 2020, in which the survey was conducted. Source: Hale et al. (2020).

Figure A1: Time series of the number of confirmed cases and stringency index of government responses

Covid Economics 38, 16 July 2020: 196-217



Note: The gray bar represents the third week of April 2020, in which the survey was conducted. Data on deaths from China is not available. Source: Hale et al. (2020).

Figure A2: Time series of the number of Covid-19 related deaths and stringency index of government responses

Consumption and geographic mobility in pandemic times: Evidence from Mexico[1]

Raymundo M. Campos-Vazquez[2] and Gerardo Esquivel[3]

Date submitted: 11 July 2020; Date accepted: 14 July 2020

We analyze the universe of point-of-sale (POS) transactions before and during the COVID-19 lockdown in Mexico. We find three key results. First, consumption in Mexico fell by 23 percent in the April-June quarter of 2020. Second, reductions in consumption were highly heterogeneous across sectors and states, with states and activities related to tourism the most affected. Third, using variation over time and states, we estimate the elasticity of POS expenditures with respect to geographic mobility (measured using cellphone location data) to be slightly less than 1. This estimate suggests that spending in developing countries may be more responsive to mobility than in developed countries, and that mobility indicators could be used as a real-time proxy for consumption in some economies.

Covid Economics 38, 16 July 2020: 218-252

1 We are grateful for superb research assistance from Raquel Yunoen Badillo and Kavik Rocha. The content of this study does not necessarily reflect the official opinion of the Banco de Mexico or its Board. Responsibility for the information, errors, omissions, and views expressed here lies entirely with the authors.
2 Professor, El Colegio de Mexico (on academic leave at the Banco de Mexico).
3 Banco de México.

1. Introduction

The economic consequences of COVID-19 are significant. Lockdown and health measures have substantially decreased geographic mobility, causing a drop in economic activity. The traditional economic indicators that measure these effects, like gross domestic product (GDP) and the industrial production index (IPI) are published by national statistical agencies with a lag: in Mexico, approximately two months after the fact. Researchers and policy-makers across the world are trying to overcome this delay by analyzing high-frequency data to quantify the magnitude of the shock and make prescriptions to avoid a more severe economic contraction (see, for example, the weekly economic index of Lewis, Mertens, and Stock 2020; the index of expenditures of Baker et al. 2020; and the labor market index of Kahn, Lange, and Wiczer 2020). Given the possibility of future waves of COVID-19, it is extremely important to measure the relationship of mobility and economic activity. In this paper, we use aggregated daily point-of-sale (POS) transaction data and cellphone location data in Mexico to quantify the magnitude of the shock and to estimate the effect of mobility patterns on POS expenditures.

It is now well known that a supply shock may cause a demand shock in the economy, thus amplifying the initial economic impact (Guerrieri et al. 2020). Sectors related to services, such as restaurants and tourism, are directly affected by a pandemic. One could then expect that the total shock should be proportional to the income losses of these sectors. However, income generated in other sectors may be affected as well, depending on the value of current versus future consumption and the value of goods and services not provided during the pandemic. If we have a high intertemporal elasticity of substitution (e.g., people can modify their consumption patterns relatively easily to spend more later rather than now) and a low intra-temporal elasticity of substitution (e.g., people

Covid Economics 38, 16 July 2020: 218-252

CEPR PRESS

Covid Economics 38, 16 July 2020: 218-252

prefer to buy the same goods and services, and there are no good substitutes for their consumption patterns), then a demand shock exacerbates the original shock, which can present an even greater problem in the presence of uncertainty and incomplete markets.

It is thus important to estimate how total expenditure is changing over time and which sectors are most affected, an estimate that requires high frequency data. POS expenditure data may meet the requirements for such analysis. Indeed, there are recent articles that make use of such information. In the United States, Baker et al. (2020) use de-identified non-random data from a Fintech company at the transaction and individual level. They find a spike in total spending when cases begin to increase (late February and early March) but a subsequent decrease of close to 50 percent with respect to January and early February. In Spain, Carvalho et al. (2020) use all POS transactions of customers of a commercial bank and transactions of others using the POS terminals of that bank. As in the U.S. study, they find a spike before the mid-March lockdown and then a sharp decline in total expenditure: 60 percent with respect to the same period in 2019. In Denmark, Andersen et al. (2020a) use data from the country's largest retail bank. They find a decrease in total spending of around 25 percent after lockdown starts. Similar results have been found in other countries: the United Kingdom shows a decline of 46 percent from April 2019 to April 2020 (Hacioglu, Känzig, and Surico 2020), France a decline of 60 percent (Bounie, Camara, and Galbraith 2020), Portugal a reduction of 55 percent in total purchases in April (Carvalho, Peralta, and Pereira 2020), and China a decline of 42 percent (Chen, Qian, and Wen 2020).

POS data is useful for shedding light on causes and potential solutions for the current crisis. Using U.S. data, Chetty et al. (2020) argue that the drop in POS expenditures is driven mainly by rich households due to health concerns. Expenditures in poor households generally returned to 2019 levels after their stimulus payments arrived.

COVID ECONOMICS
VETTED AND REAL-TIME PAPERS

Employment losses are greater in higher-income zip codes, especially in personal services like restaurants and barber shops. They conclude that economic recovery goes hand in hand with safety concerns.

Our paper makes important contributions to this literature. First, we show that the response in developing countries may be different than in developed countries. Although Mexico is an upper middle-income country, its financial sector is not as developed as in other countries. According to the World Bank (2020), domestic bank credit to the private sector accounts for only 27 percent of GDP, while in countries with similar consumption patterns, like China, Denmark, France, Spain, and the United Kingdom, it is close to or above 100 percent. Only in the United States is it less than that, and even there it is 52 percent of GDP. Also, the number of POS terminals in Mexico per 100,000 population is the lowest among similar countries (approximately 1000 in Mexico versus 2000 in China and 3000 in the other countries). Finally, internet penetration in Mexico (around 64 percent) is less than in the United States (76 percent) or similar European countries (all above 80 percent). Although this may mean that POS data are not as comprehensive for Mexico, our results indicate large negative effects of the pandemic, although not as large as in those in other countries.

Second, the data we analyze for Mexico includes all POS transactions in the country, in contrast to the data in previous studies, which is limited to selected banks or companies. The comprehensive nature of our data allows us to benchmark the effect of COVID-19 on POS expenditures to traditional measures like total consumption and GDP. Third, although we follow previous literature in calculating expenditure losses with respect to 2019, we also propose a simple model to calculate a counterfactual of what expenditure would have looked like in the absence of the pandemic. Finally, we estimate the elasticity of POS expenditures with respect to measures of geographic mobility using

Covid Economics 38, 16 July 2020: 218-252

variation over time within states in Mexico. This elasticity is important, as it could be used in theoretical models and simulation exercises to calculate expenditure losses for future waves of the pandemic. It is also an important consideration in the debate about the impact of lockdown measures on the level of expenditures.

We use the universe of point-of-sale (POS) transactions from January 1, 2019 to June 30, 2020, which is non-public data from the Banco de México (the Mexican central bank), consisting of aggregated daily information on total expenditures and certain other categories. This POS expenditure data provides important information about general consumption patterns. In 2019, there were 157 million debit and credit cards in Mexico, and the National Financial Inclusion Survey (INEGI, 2018) shows that more than two-thirds of the Mexican population (68 percent) aged 18-70 have at least one such financial product. In 2019, the average POS expenditure per transaction was $630 MXN (approximately $31 USD). Approximately 10 million transactions take place through POS terminals every day, 73 percent of which are with debit cards and the remaining 27 percent with credit cards. The average monthly total debit and credit card expenditure was almost $187 billion MXN during 2019 (approximately $9.2 billion USD). Annual total POS expenditure thus represents about 8 percent of GDP and 14 percent of consumption.

We are able to provide the first direct estimates of the elasticity of POS expenditures with respect to geographic mobility. Previous studies have provided only indirect or implicit estimates for this elasticity. For example, using the results in Andersen et al. (2020b), we can estimate an elasticity of 0.2 by exploiting the between-country variation in spending and mobility for Sweden and Denmark: consumption declined 29 percent in Denmark and by 25 percent in Sweden (Figure 3 in that study). Using mobility measures based on cellphone location data available from the Apple Corporation (2020)

Covid Economics 38, 16 July 2020: 218-252

for early April, we find that mobility decreased by only 12 percent in Sweden while it declined by 32 percent in Denmark. The implicit elasticity of POS expenditures with respect to mobility is thus around 0.2. In the current study, since we have daily data for expenditures in Mexico at the subnational level, we are able to estimate the elasticity of consumption with respect to mobility indicators by exploiting both the time and geographic variation in the data.

We find three key results. First, the percent loss in POS expenditures with respect to the estimate without the pandemic is 23 percent for April-June. This estimate is much lower than that for other countries. The estimate for Spain and France (for the last two weeks of March) is close to 50 percent (Bounie et al. 2020; Carvalho et al. 2020), for Portugal it is 55 percent (Carvalho, Peralta, and Pereira 2020), and for Denmark it is 30 percent (Andersen et al. 2020a). Although estimates for the U.S. vary, our result is similar to the live results from POS data in Chetty et al. (2020). In terms of GDP and consumption, for the April-June quarter it implies a loss of 2.6 percent of quarterly GDP and 3.9 percent of quarterly private consumption.

Second, losses vary significantly across sectors and regions. While some sectors were severely hit, like tourism, food services, and transportation, others, like insurance and telecommunications, were barely affected. This result is similar to that found in other studies. Mexican states that are highly dependent on tourism (beach resorts and other tourist destinations) are among the most affected.

Third, we estimate the elasticity of POS expenditures with respect to geographic mobility in Mexico, as measured using cellphone location data from Apple (2020) and Google (2020). Our estimates show that this elasticity is in most cases non-significantly different from one (0.93 using Apple's measure of mobility in one specification, and 0.91 for both Google's and Apple's measures of mobility in another). These estimates are

Covid Economics 38, 16 July 2020: 218-252

much larger and more precisely estimated than the estimate of 0.2 derived by comparing the effect of mobility on spending in Sweden and Denmark, as described above. This result suggests that POS expenditures in developing countries could be more responsive to mobility patterns than in developed countries, an interesting possibility that calls for further research. It may be possible, for example, that internet penetration and the strength of e-commerce affect the magnitude of this elasticity. This result is also important because it suggests that in economies like Mexico's, mobility indicators, which can be observed almost in real time, could serve as a good proxy for the behavior of expenditures.

2. Data and Methods

The data includes all point-of-sale (POS) transactions in Mexican territory, which is non-public information collected by the Banco de México under its mandate to assure a well-functioning payment system. The data is aggregated by type of card (debit or credit), at the state and national levels, and by type of expenditure, on a daily basis, from January 2019 to June 2020. We observe only aggregate information; we do not observe any individual transactions, any information about whether the credit or debit card is foreign or Mexican, or whether the transaction took place on the internet or in a physical location.

Most of the previous literature uses either a part of the universe of transactions or a sample of households. Our use of the full universe of transactions allows us to calculate total losses in the economy. However, one key challenge is how to construct a valid counterfactual for comparison. In general, previous studies calculate the percent change in 2020 with respect to 2019. This seems reasonable if the financial sector is stable. However, because transactions in Mexico were already growing before the pandemic arrived it seems more appropriate to construct a counterfactual scenario using data from

Covid Economics 38, 16 July 2020: 218-252

2019 and 2020. We propose a simple model that predicts the daily (t) outcome in 2020 (POS_t^{2020}) based on both the 2019 outcome (POS_t^{2019}) and pre-pandemic data observed for 2020.[1] We also include dummy variables related to paydays, Mondays, Fridays, and for the month of December.

$$POS_t^{2020} = \alpha + \beta POS_t^{2019} + dummies + \varepsilon_t \quad (1)$$

The regression is estimated for all days from January 1, 2020 to February 18, 2020. We select the final model minimizing the mean squared error for the prediction for February 19 to March 11, 2020, that is, during the pre-lockdown period. Then, we make a prediction for all the remaining days in 2020. All of the predictions are in constant pesos (MXN) of July 2018. The percent effect of the pandemic can then be calculated as:

$$\Delta\%\ Effect_t = \frac{POS_t^{2020} - \widehat{POS}_t^{2020}}{\widehat{POS}_t^{2020}} \quad (2)$$

The comparison with respect to 2019 replaces the predicted value $\widehat{POS}_t^{2020}$ with the value in 2019, POS_t^{2019}. As the daily expenditures are noisy, in some cases we smooth the lines in the figures by a simple moving average for the previous two weeks. We show below multiple estimates for total expenditures, for credit and debit cards, for type of expenditure, and at the state level.

We also calculate the elasticity of total expenditures with respect to indicators of geographic mobility, obtained from Google (2020) and Apple (2020). Google tracks mobility using the location history of the Google accounts on people's mobile devices; we use this data to calculate the percent change compared with the median value for baseline days in the five-week period January 3 to February 6, 2020. We focus on the mobility trends for workplaces. Apple mobility is an index with a baseline set at January

[1] We compare this model with other ARIMA models with the form $POS_t^{2020} = \alpha + \beta POS_t^{2019} + \theta ARIMA + dummies + \varepsilon_t$ in terms of the root mean squared predicted error (RMSPE) for February 19 to March 11. This comparison is for total, credit, and debit expenditures, varying the introduction of dummies. For a complete table of the evaluated models see supplementary material Table S1. For simplicity, we choose the model with dummies to make the predictions.

Covid Economics 38, 16 July 2020: 218-252

Covid Economics 38, 16 July 2020: 218-252

13, 2020. Apple also uses people's mobile devices to track their location (monitoring the requests made to the Apple map application). For Apple, we use the mobility measure based on driving. Data is available for the period January 13 to June 30.[2] For purposes of comparability between the Google and Apple datasets, we change the baseline to February 17. We thus obtain a dataset for the period February 15 to June 30, with each row including two columns: the percent change of total POS expenditure from state *s* in week *w*, and the mean percent change in mobility from each source from state *s* in week *w*. The percent change is with respect to February 17 in all columns.

$$\Delta\%POS_{s,w}^{2020} = \beta\Delta\%Mobility_{s,w}^{2020} + \delta_s + \delta_w + \varepsilon_{s,w} \qquad (3)$$

The regression controls for fixed effects of week and state. The first control is for shocks that affect all states at the same time, and the second is for permanent differences across states. For example, some states may specialize in occupations or industries that make them either more resilient or more susceptible to an economic shock, and this specialization may at the same time be correlated with geographic mobility.[3]

3. Descriptive Results

The first case of COVID-19 in Mexico was diagnosed on February 27, later than in European countries. On March 14, the government announced the suspension of non-essential activities and rescheduled mass events. A soft lockdown began on March 23. The government has taken different steps to address the health and economic shocks. First, it implemented recommendations for social distancing, travel restrictions, and the suspension of non-essential activities to prevent the spread of the virus. Second, the

[2] Two days, May 11 and May 12, were not available. We impute values for these days with the mean values for May 10 and May 13.

[3] In particular, some states may be more prepared for telecommuting than others, making them more resistant to employment losses. If the latter states show greater mobility and expenditure, that could bias the elasticity estimate.

government and the Banco de México have taken action to mitigate the effects of the pandemic. Like other central banks across the world, the Banco de México has implemented measures to provide liquidity to the market, injecting the equivalent of 3.3 percent of GDP into the economy.[4] The fiscal policy response has been more limited: it has offered access to microcredits and has implemented a frontloaded payment of some social programs (close to 1 percent of GDP). The government has also announced an austerity program and the continuation of some public works.

Table 1 shows the main descriptive statistics for May 2019 and May 2020, including the total amount spent in POS terminals, the average amount of each transaction, and the share of expenditures in each group. For simplicity the data is grouped into 12 categories: tourism (travel agencies and hotels), education (universities, colleges, basic education, and daycare), health care (pharmacies, hospitals, physicians, and dentists), food services (restaurants and fast food), trade (wholesale and retail), transportation (air transportation, ground transportation, tolls, parking lots, and car rental), insurance, telecommunications, supermarkets, big-box stores, and others.

The average transaction amount did not change substantially. It was $601 in May 2019 and $589 in May 2020 (in constant MXN pesos of July 2018). However, there is an overall decline of approximately $34 billion, or 16 percent, representing an average monthly decline in total private consumption of 2.5 percent a month. The sectors with the largest expenditures in 2019 (a combined total of 80 percent) were big-box stores, trade, gasoline, food services, and other. In May 2020, most sectors showed reduced total POS transactions. Services related to tourism, food services, and transportation were hit

[4] These measures have included bond swaps, loosened rules for minimum deposits from commercial banks, and facilities to swap assets with the central bank in order to obtain credit. These measures have the goal of directing credit to small and medium-sized business.

Covid Economics 38, 16 July 2020: 218-252

especially hard. However, insurance, telecommunications, big-box stores, and other maintained or increased sales.

Table 1. Descriptive Statistics

	May 2019			May 2020		
	Total Amount (millions of pesos)	Avg. Transaction amount (pesos)	Share (%)	Total Amount (millions of pesos)	Avg. Transaction amount (pesos)	Share (%)
Total	$206,669	$ 601		$172,800	$589	
Tourism	$5,333	$ 2,580	2.6	$786	$1,988	0.5
Education	$6,851	$ 4,026	3.3	$4,607	$4,639	2.7
Health Care	$8,239	$ 524	4.0	$7,757	$485	4.5
Food Services	$11,747	$ 383	5.7	$2,681	$257	1.6
Trade	$42,312	$ 473	20.5	$32,978	$462	19.1
Transportation	$8,961	$ 589	4.3	$1,833	$286	1.1
Insurance	$5,153	$1,811	2.5	$5,445	$2,207	3.2
Telecomm. Services	$6,394	$ 701	3.1	$6,779	$525	3.9
Gasoline	$18,400	$616	8.9	$ 10,979	$538	6.4
Other	$35,244	$641	17.1	$ 41,564	$601	24.1
Supermarkets	$28	$348	0.0	$19	$314	0.0
Big-Box Stores	$58,007	$630	28.1	$57,373	$692	33.2

Notes: Authors' calculations. Amounts are in constant MXN for July 2018.

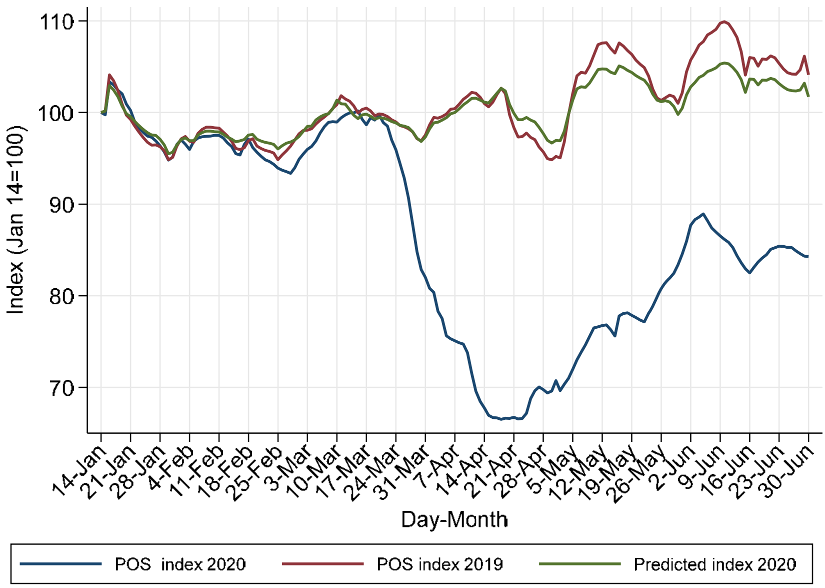
3.1 Aggregate Results

Figure 1 shows smoothed lines of daily expenditure in POS terminals for 2019 and 2020. For comparison purposes, the series are in relative terms with respect to January 14 of each year. The red line is the index for 2019 and the blue line for 2020. Using the method described above, we obtain a prediction for 2020 using data for the early part of the year. The green line is the prediction for 2020. Before the lockdown, the patterns for 2019 and 2020 are similar. When the lockdown starts, POS expenditures fall drastically. The worst days were in mid-April, with expenditures about 35 percent lower than in 2019

Covid Economics 38, 16 July 2020: 218-252

or in the prediction for 2020. After that point, expenditures slowly started to recover. By late May and early June, the shortfall was only about 15 percent lower than the prediction.

Figure 1. Effect of COVID-19 on expenditures in POS terminals



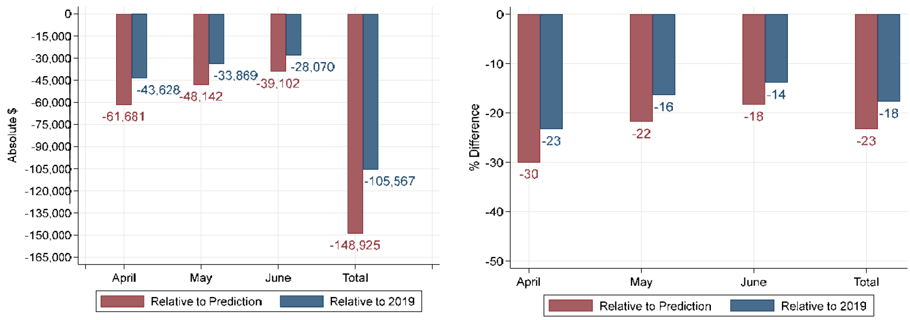
Notes: Authors' calculations. Lines are smoothed using a moving average for the previous two weeks. Predicted line is obtained with equation (3), an OLS of the amount in 2020 with the amount in 2019. Expenditures are in constant pesos (MXN) of July 2018. Expenditures in January 2020 are 9 percent larger than in January 2019.

Figure 2 shows the decline in POS expenditures by month (constant pesos of July 2018), with comparisons to 2019 and the predicted expenditures for 2020. The greatest decline is in April, with expenditures 30 percent lower than predicted and 23 percent lower than in the corresponding period in 2019. Subsequent months show lesser declines: May is 22 percent and June is 18 percent below the prediction. The total decline in POS expenditures from the predicted figure for April through June is around $149 billion MXN, a loss of 3.9 percent of an average quarter of private consumption in 2019, and a loss of 2.6 percent of an average quarter of GDP.

Covid Economics 38, 16 July 2020: 218-252

Figure 2. Decline in POS expenditures

A. Absolute decline **B. Percent decline**



Note: Authors' calculations. This graph shows the difference between actual and predicted values, and the difference between actual 2020 and 2019 values (in constant pesos of July 2018).

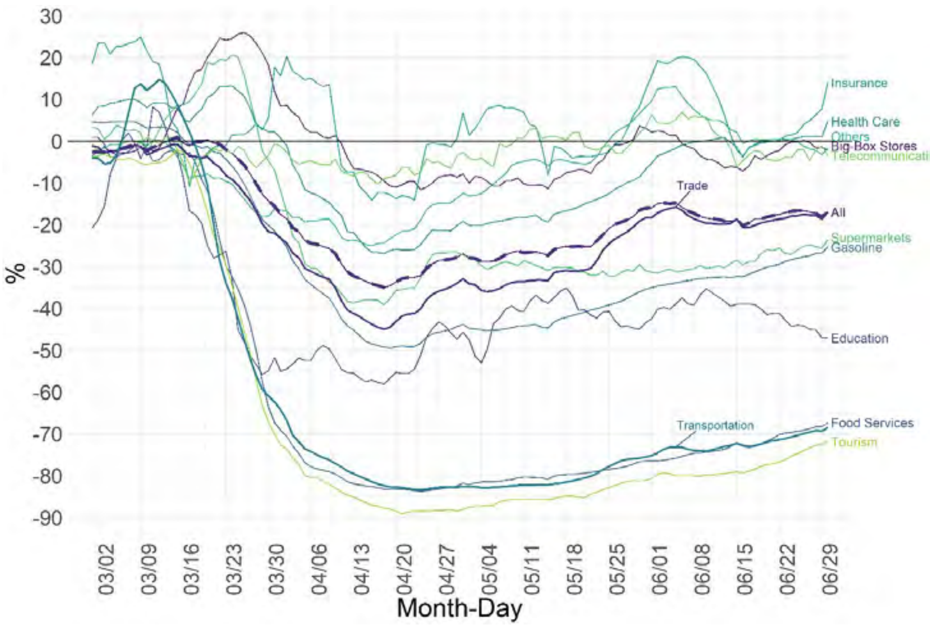
3.2 Results by Sector

Figure 3 shows the change in consumption patterns by sector. The lines are smoothed using a moving average of the previous two weeks (Leatherby and Gelles 2020), and the comparison is to the predicted sales in each sector. The comparison with respect to 2019 can be found in the Supplementary Materials. After the beginning of the lockdown, there is a sharp decline in education, tourism, food services, and transportation. Only education recovers, but at the end of May it is still about 40 percent below the prediction. Tourism, food services, and transportation fall from 80 to 90 percent by mid-April. Because of the decline in mobility and in domestic prices, POS expenditures for gasoline decrease almost 50 percent in mid-April, and by the end of May they were still approximately 35 percent below the prediction. Expenditures in June have been relatively stable, with a slight recovery in most cases.

Similar to the experience of other countries, POS expenditures in big-box stores increased in the last two weeks of March, an effect of panic buying to stockpile goods.

Other sectors, like insurance, health care, and telecommunications, were not affected by the mobility restrictions. At least with insurance and telecommunications, this is likely related to direct billing options as well as the inelasticity of demand for this type of goods. While in the U.S. there was a large decline in health expenditures in April (Chetty et al. 2020), in Mexico the decline was smaller and it quickly recovered, by the end of May.

Figure 3. Changes in consumption patterns by sector. Smoothed lines.



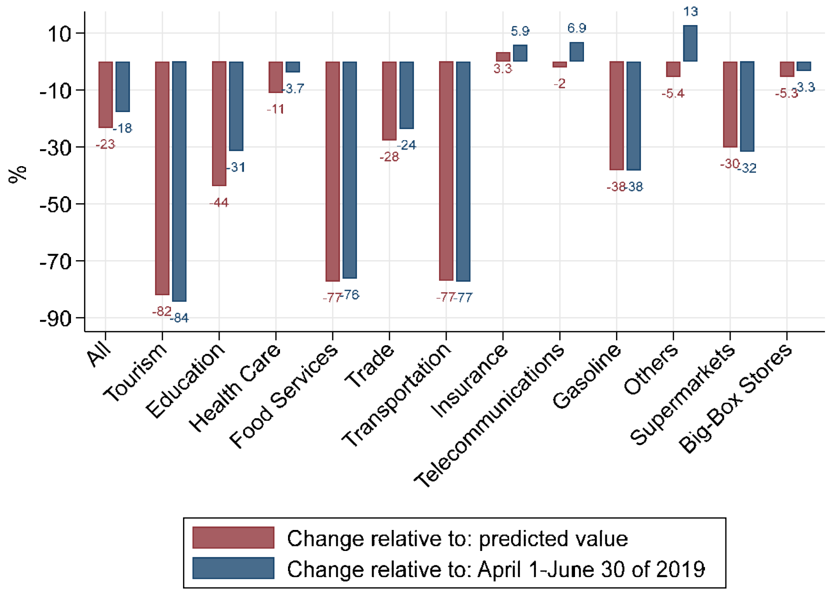
Notes: Authors' calculations. Comparison is to predicted sales in each sector. Constant pesos (MXN) of July 2018. Smoothed with moving average of the previous two weeks.

Figure 4 summarizes previous estimates. It indicates the percent difference of POS expenditures in 2020 with respect to the predicted expenditures and with respect to the same period in 2019 (in constant pesos). The losses total 23 percent of predicted expenditures: one quarter of expected POS sales did not take place. Total expenditures were 18 percent lower than in the same period in 2019. Comparisons are difficult because

Covid Economics 38, 16 July 2020: 218-252

lockdowns were implemented at different times in different countries, but the Mexican loss estimate is among the lowest. In the last two weeks of March, France and Spain had expenditure losses of 50 percent (Bounie et al. 2020; Carvalho et al. 2020); in April, Portugal had losses of 55 percent (Carvalho, Peralta, and Pereira, 2020) and Denmark had more moderate losses of approximately 30 percent.

Figure 4. Summary of expenditure losses. April-June 2020.



Note: Authors' calculations. This graph shows the change in expenditures relative to 2019 values and to predicted values for 2020. Constant pesos (MXN) of July 2018.

The comparison with the U.S. depends on the source. The estimates of Baker et al. (2020) imply a decline of 50 percent, while Chetty et al. (2020) find a decline of 30 percent in the last two weeks of March. In fact, the change in All Expenditures in Figure 3 is very close to that found in Chetty et al. (2020) (see Figure S3 in Supplementary Materials). The decline in expenditures in the U.S. was larger before mid-April. Stimulus payments began on April 15 in the U.S., and POS expenditures recovered faster around that date. The decline in POS expenditures from January to mid-June was 10 percent in the U.S., while in Mexico it was still 20 percent. There is significant heterogeneity across sectors, however. Those affected most severely in Mexico were tourism, food services, and transportation, where expenditures declined approximately 80 percent. This is similar to what previous studies have found in Denmark, Spain, the United States, and other countries (Andersen et al. 2020a; Baker et al. 2020; Carvalho et al. 2020; Chetty et al. 2020; Leatherby and Gelles 2020).

Some sectors in Mexico even had gains or only small losses. Expenditures on insurance increased slightly in the period, and expenditures on telecommunications decreased slightly. We interpret these sectors as supplying highly inelastic necessities. Expenditures in big-box stores decreased by 5.4 percent. The pattern for these stores is mixed: in mid-March their expenditures increased, in mid-April they declined, and by the end of May they recovered. This group includes large supermarkets (such as Walmart and Soriana) as well as department stores (such as Liverpool, Palacio de Hierro, and Sears). It is likely that sales increased in large supermarkets and decreased in department stores.

There were decreased sales in health care, gasoline, trade, small supermarkets, and other, which accounted for close to 50 percent of all expenditures in 2019 (Table 1). The decline in trade and gasoline (28 and 38 percent, respectively) is directly related to restrictions in mobility.

3.3 Results by State

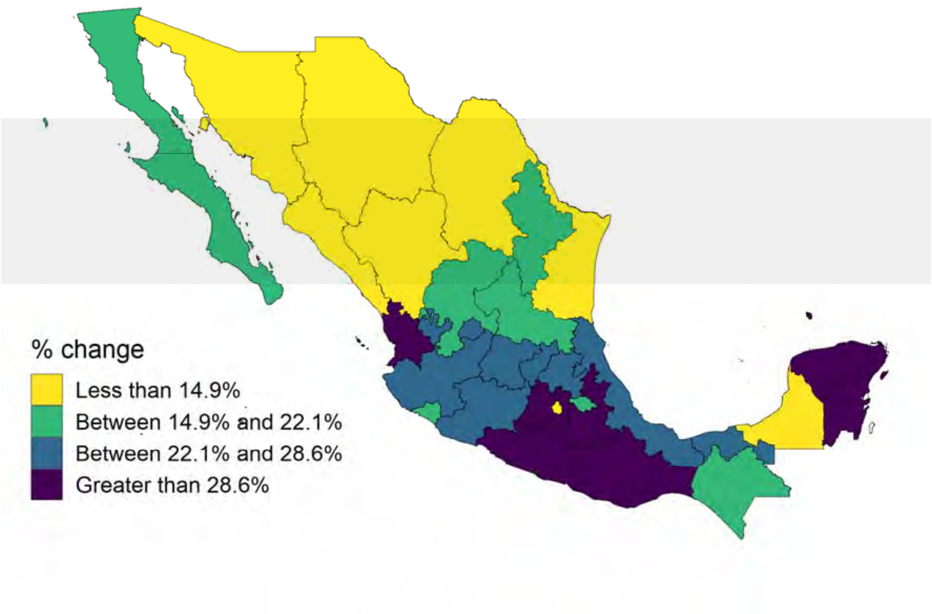
We estimate the model in equation (1) for each state in Mexico. Figure 5 shows percent losses by state with respect to the predictions of the model. States shown in purple are the hardest hit and those in yellow are the least affected. The hardest hit regions depend on international tourism: Quintana Roo and Yucatan in the south, as well as

Covid Economics 38, 16 July 2020: 218-252

Guerrero and Nayarit. These states lost all expected revenue from the spring vacation season. Other states closer to Mexico City are also greatly affected: Michoacán, Estado de México, Puebla, and Morelos, probably related to the loss of domestic tourism around Easter. Mexico City is not as affected as other states. We suspect that here the effects of the pandemic were partially compensated by online sales, but our data unfortunately does not distinguish online from other sales. Finally, states in the north are not as affected as the rest, an effect of greater mobility than in the rest of Mexico, as explained in the next section.

Covid Economics 38, 16 July 2020: 218-252

Figure 5. Losses by state in total POS expenditure (percent).



Notes: Authors' calculation. The map shows the percent change of POS expenditures from April to June with respect to the predicted sales for each state. Constant pesos of July 2018

Covid Economics 38, 16 July 2020: 218-252

4. Consumption and Mobility

We use geographic mobility data from Google (2020) and Apple (2020) through June 2020. There is an ongoing debate about the relationship between mobility and POS expenditures. The case and evidence from Sweden are relevant.[5] Unlike other European countries, Sweden did not impose a lockdown in response to the COVID-19 pandemic, which was responsible for a higher mortality rate than in similar Nordic countries. One might expect that the lack of restrictions on mobility at least lessened the economic effects of the pandemic. However, Andersen et al. (2020b) found that this was not the case. Sweden experienced a 25 percent reduction in POS expenditures from March 11 to April 12; the corresponding figure for Denmark was 29 percent. Apple's measure of driving mobility for early April shows a reduction in Sweden of 12 percent and a reduction in Denmark of 32.4 percent. The between-country variation suggests that the elasticity of mobility is around 0.20. However, elasticity may depend on the relative importance of internet sales, which depends in turn on the depth of the financial sector. In a less developed country like Mexico, in-person sales and therefore mobility may matter much more than in developed economies.

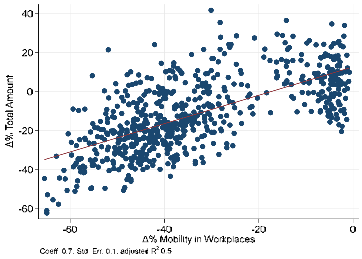
To show how mobility and expenditures are related, we use Google's measure of workplace mobility and Apple's measure of driving mobility. We calculate mobility and expenditure patterns for each of the 20 weeks and 32 states under study. We thus have a panel dataset with 640 observations. The patterns are shown in Figure 6. The variation in the mobility measures is positively correlated with the variation in the total amount spent. Panel A uses Google's workplace mobility and it finds a coefficient of 0.7 using a simple OLS regression. Panel B uses Apple's driving mobility and it finds a coefficient of 0.9

[5] https://www.politico.eu/article/swedens-cant-escape-economic-hit-with-covid-19-light-touch/, https://www.ft.com/content/93105160-dcb4-4721-9e58-a7b262cd4b6e.

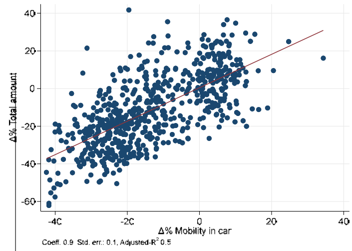
with the same type of regression. States with the largest declines in mobility are related to the largest declines in expenditures at the weekly level.

Figure 6. Relationship between mobility (Google and Apple) and POS expenditures.

A. Google: Workplace Mobility



B. Apple: Driving Mobility



Notes: Authors' calculations. Each dot is the percent change of mobility or POS expenditures (constant pesos of July 2018) in week *w* with respect to February 17 for each of the 32 states in Mexico. Period of estimation is February 15 to June 30.

In order to analyze this claim more carefully, we estimate different versions of equation (3). Panel A in Table 2 estimates the relationship between changes in POS expenditures and mobility in Mexico including week and state fixed effects. These effects control for permanent differences across states (for example, density or geographic characteristics) as well as for temporal shocks that affect all states at the same time. Table 2 shows the results for all expenditures as well as for expenditures differentiated by credit versus debit card. The elasticity coefficient for total expenditure using Google's mobility is 0.73; for Apple's mobility it is 0.93. These estimates, which exploit the within-state variation, are very similar to those obtained simply by pooling the spending and mobility information (Figure 6). All of the estimates in Panel A are very precisely estimated and they are all statistically significant. The elasticity using Apple's mobility information is not statistically different from 1. The elasticity for credit card spending is greater than for debit cards, regardless of the mobility indicator used.

Panel B in Table 2 estimates the same spending-mobility relationship but instead of including fixed effects, it includes as an additional control variable the proportion of work that can be performed by telecommuting from home in each Mexican state, as estimated by Monroy-Gómez-Franco (2020). This specification exploits both the between and the within variation across states in Mexico to estimate the elasticity of POS expenditures with respect to mobility. The elasticity results obtained with this specification have larger standard errors, but they are similar for both mobility indicators (0.91). In both cases, they are statistically different from 0 but not from 1. As before, credit cards are more elastic with respect to mobility than debit cards. These estimated elasticities are also much larger than that implied by Andersen et al. (2020b) for the case of Sweden (0.2).

Table 2. Elasticity Estimates: Change in % POS Expenditures with Respect to Change in % Mobility

	Google: Workplace Mobility			Apple: Driving Mobility		
	Total	Credit	Debit	Total	Credit	Debit
A. Including state fixed effects						
Coefficient	0.73	1.09	0.56	0.93	1.33	0.74
Standard Error	[0.05]	[0.05]	[0.04]	[0.06]	[0.07]	[0.05]
R^2	0.45	0.54	0.35	0.46	0.44	0.41
Total Obs.	640	640	640	640	640	640
B. Controlling for telecommuting (without state fixed effects)						
Coefficient	0.91	1.34	0.67	0.91	1.03	0.85
Standard Error	[0.45]	[0.51]	[0.42]	[0.20]	[0.29]	[0.18]
R^2	0.57	0.63	0.51	0.65	0.66	0.62
Total Obs.	640	640	640	640	640	640

Notes: Authors' calculations. The dependent variable is the percent change in POS expenditures in week *w* with respect to February 17 for each state in Mexico, and the independent variable is the percent change in mobility for the same period. The regression in Panel A includes fixed effects for state and week. Estimation period is February 15 to June 27. Panel B includes dummies for weeks and proportion of telecommuting (defined as in Monroy-Gómez-Franco 2020). Standard errors clustered at the state level in brackets.

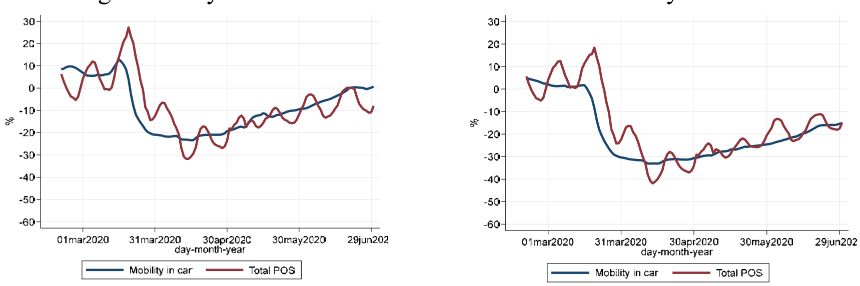
To further analyze these results, in Figure 7 we show Apple's mobility measure (blue line) and POS expenditures (red line) in high- and low-mobility states in Mexico.

Covid Economics 38, 16 July 2020: 218-252

In high-mobility states, mobility and POS expenditures declined close to 10 percent from early in the year to the end of May. By mid-June, mobility and expenditures in these states were similar to pre-pandemic levels. In low-mobility states the decline was close to 25 percent in mid-May and by mid-June it was still around 15 percent below pre-pandemic levels. The high correlation between spending and mobility in both types of states is evident. As expected, the estimates of the elasticity of spending with respect to mobility are also very high for each of these groups of states: around 0.80 for high-mobility states and 1.04 for low-mobility states, and in both groups the elasticity for credit cards is larger than for debit cards. These elasticity estimates are between four and five times the implied elasticity estimated by Andersen et al. (2020b).

Figure 7. Apple's mobility and POS expenditures in high- versus low-mobility states.

A. High mobility B. Low mobility.



Notes: Authors' calculations. High-mobility states include Aguascalientes, Campeche, Chihuahua, Coahuila, Colima, Durango, Guerrero, Michoacán, Morelos, San Luis Potosí, Sinaloa, Sonora, Tamaulipas, Tlaxcala, Veracruz, and Zacatecas. Low-mobility states include Baja California, Baja California Sur, Chiapas, Mexico City, Estado de México, Guanajuato, Hidalgo, Jalisco, Nayarit, Nuevo León, Oaxaca, Puebla, Querétaro, Quintana Roo, Tabasco, and Yucatán. Mobility refers to driving mobility measured by Apple.

Why is the elasticity of POS expenditures to mobility larger in Mexico? We conjecture that this difference is driven mainly by the strength (or lack thereof) of e-commerce, financial inclusion, and internet penetration. As mentioned in the introduction,

Covid Economics 38, 16 July 2020: 218-252

financial inclusion is lower in Mexico than in China, the United States, and European countries. In 2014, the proportion of individuals in Mexico with an account at a financial institution was 40 percent, while it is 80 percent in China and close to 100 percent in developed countries. If we consider that internet penetration is lower as well, then we have a weaker market for e-commerce in Mexico than elsewhere. Indeed, results from the United Nations Conference on Trade and Development (2016) show that Mexico has an e-commerce readiness index much lower than other countries (Mexico's index is 49.1 while the U.S.'s is 82.6). POS transactions thus depend much more on mobility in Mexico than in other countries.

Finally, we cannot test Chetty's et al. (2020) claim that the channels of the decline in POS expenditures are mainly rich individuals in fear of contagion. We attempt to compare our results, at least at the aggregate level, by computing POS expenditures by credit versus debit cards, which are highly segregated in Mexico. We calculate that approximately 77 percent of credit cards and 62 percent of debit cards are held by individuals in the top 30 percent of the wealth distribution (see figures in Supplementary Materials). The decline in POS expenditures is larger for credit cards (28.6 percent) than for debit cards (10 percent). The elasticity of POS expenditures with respect to mobility is also much larger for credit cards than for debit cards. We thus conjecture that the decline in POS expenditures is partially driven by richer individuals concerned for their health, as in Chetty et al. (2020).

5. Summary

This paper analyzes consumption patterns in Mexico using the universe of POS transactions for the period from January 2019 to June 2020. Unlike some other countries, Mexico implemented a soft lockdown as well as a moderate countercyclical fiscal policy.

Covid Economics 38, 16 July 2020: 218-252

We find that POS expenditures for the April-June quarter are 23 percent less than they would have been in the absence of the pandemic. This difference is less than that calculated for European countries using similar data, and comparable to that reported for the U.S. by Chetty et al. (2020), also using results based on live POS data.

The losses we find for Mexico are heterogeneous across economic sectors and region. As in other studies, the more severely affected sectors are those related to tourism (travel agencies and hotels), food services (such as restaurants), and transportation. States that benefit more directly from tourism (beach resorts and other tourist destinations) were also more affected.

There is a debate about whether mobility patterns affect POS expenditures and thus economic activity. We find that the elasticity of POS expenditures with respect to mobility is close to 1 (0.93 using Apple's measure of mobility in one specification and 0.91 for both Google's and Apple's measures of mobility in another). These estimates are much larger than the implied elasticity estimated by Andersen et al. (2020b) for Sweden. Our estimate likely indicates that POS expenditures in developing countries with shallower financial sectors are more responsive to mobility patterns than developed countries. It also suggests that mobility indicators, which can be observed almost in real time, could serve as a good proxy for the behavior of expenditures in some economies.

Covid Economics 38, 16 July 2020: 218-252

Covid Economics 38, 16 July 2020: 218-252

REFERENCES

Andersen, Asger, Emil Toft Hansen, Niels Johannesen and Adam Sheridan (2020a). "Consumer Responses to the COVID-19 Crisis: Evidence from Bank Account Transaction Data". *Centre for Economic Policy Research.* http://www.cepr.org/active/publications/discussion_papers/dp.php?dpno=14809

Andersen, Asger, Emil Toft Hansen, Niels Johannesen, and Adam Sheridan (2020b). "Pandemic, Shutdown and Consumer Spending: Lessons from Scandinavian Policy Responses to COVID-19". Working Paper arXiv:2005.04630v1.

Apple Inc. (2020). "Apple Mobility Trends reports". https://www.apple.com/covid19/mobility/ (Accessed: June 5th, 2020)

Baker, Scott R., R.A. Farrokhnia, Steffen Meyer, Michaela Pagel and Constantine Yannelis (2020). "How Does Household Spending Respond to an Epidemic? Consumption During the 2020 COVID-19 Pandemic". *Covid Economics* 18, 15 May 2020: 73-108.

Bounie, David, Youssouf Camara and John W. Galbraith (2020). "Consumers' Mobility, Expenditure and Online-Offline Substitution Response to COVID-19: Evidence from French Transaction Data". CIRANO Working Papers 2020-28.

Carvalho, Bruno P., Susana Peralta and Joao Pereira (2020). "What and How did People Buy during the Great Lockdown? Evidence from Electronic Payments". *Covid Economics* 28, 12 June 2020: 119-158.

Carvalho, Vasco, M., Stephen Hansen, Álvaro Ortiz, Juan Ramón García, Tomasa Rodrigo, Sevi Rodríguez Mora, and José Ruiz (2020). "Tracking the COVID-19 Crisis with High-Resolution Transaction Data". *Centre for Economic Policy Research*. https://cepr.org/active/publications/discussion_papers/dp.php?dpno=14642

Chen, Haiqiang, Wenlan Qian and Qiang Wen (2020). "The Impact of the COVID-19 Pandemic on Consumption: Learning from High Frequency Transaction Data". *SSRN*. http://dx.doi.org/10.2139/ssrn.3568574

Chetty, Raj, John N. Friedman, Nathaniel Hendren, Michael Stepner, and the Opportunity Insights Team (2020). **"**How Did COVID-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data." National Bureau of Economic Research, NBER Working Paper 27431, https://doi.org/10.3386/w27431.

Google LLC. (2020). "Google COVID-19 Community Mobility Reports". https://www.google.com/covid19/mobility/ (Accessed: June 5th, 2020)

Guerrieri, Veronica, Guido Lorenzoni, Ludwig Straub, and Ivan Werning (2020). "Macroeconomic implications of COVID-19: Can negative supply shocks cause demand shortages?" National Bureau of Economic Research, NBER Working Paper 26918, https://doi.org/10.3386/w26918.

Hacioglu, Sinem, Diego Känzig and Paolo Surico (2020). "Consumption in the time of Covid 19: Evidence from UK transaction data". Centre for Economic Policy Research, CEPR Discussion Paper 14733. http://www.cepr.org/active/publications/discussion_papers/dp.php?dpno=14733

Instituto Nacional de Estadística y Geografía. (2018). "Encuesta Nacional de Inclusión Financiera (ENIF) 2018". https://www.inegi.org.mx/programas/enif/2018/ (Accessed: June 5th, 2020)

Kahn, Lisa B., Fabian Lange, and David G. Wiczer (2020). "Labor Demand in the Time of COVID-19: Evidence from Vacancy Postings and UI Claims." National Bureau of Economic Research, NBER Working Paper 27061, https://doi.org/10.1017/CBO9781107415324.004.

Lewis, Daniel, Karel Mertens, and James H. Stock (2020). "US economic activity during the early weeks of the SARS-Cov-2 outbreak". National Bureau of Economic Research, NBER Working Paper 26954, https://doi.org/10.3386/w26954

Leatherby, Lauren and David Gelles (2020). "How the Virus Transformed the Way Americans Spend Their Money" *New York Times*. April 11. https://www.nytimes.com/interactive/2020/04/11/business/economy/coronavirus-us-economy-spending.html?referringSource=articleShare

Monroy-Gómez-Franco, Luis (2020). "¿Quién puede trabajar desde casa? Evidencia desde México.", Documento de Trabajo no. 06/2020. Centro de Estudios Espinosa Yglesias. https://ceey.org.mx/wp-content/uploads/2020/05/06-Monroy-G%C3%B3mez-Franco-2020.pdf

United Nations Conference on Trade and Development (2016). *UNCTAD B2B E-Commerce Index 2016* (Vol. 7). https://unctad.org/en/PublicationsLibrary/tn_unctad_ict4d07_en.pdf

World Bank (2020). Data indicators. Data.worldbank.org. Series: https://data.worldbank.org/indicator/FD.AST.PRVT.GD.ZS?locations=US-ES-GB-CN-DK-FR&name_desc=true, https://databank.worldbank.org/reports.aspx?source=1251&series=GPSS_4

Covid Economics 38, 16 July 2020: 218-252

Covid Economics 38, 16 July 2020: 218-252

Supplementary Materials

Table S1. RMSPE Tested Models

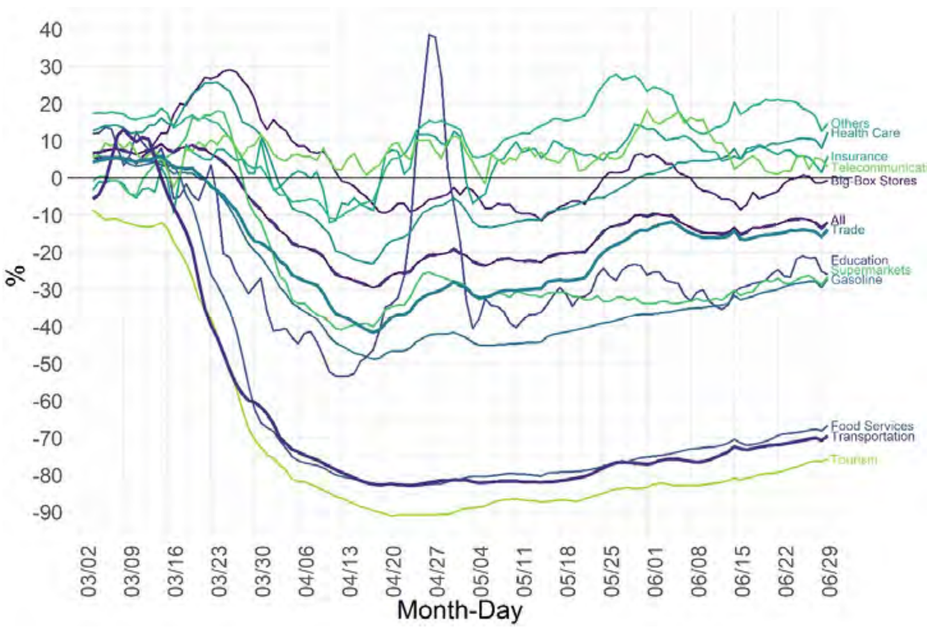
		RMPSE February 19-March 11		
Model #	**Model description**	**Total**	**Credit**	**Debit**
1	OLS	547.83	267.93	414.86
2	OLS dummies	425.40	197.64	299.30
3	ARIMA(0, 0 ,0)	547.83	267.93	414.86
4	ARIMA(0, 1 ,0)	598.36	268.81	479.59
5	ARIMA(0, 2 ,0)	6153.22	2450.88	1082.39
6	ARIMA(0, 0 ,1)	551.20	259.99	413.24
7	ARIMA(0, 1 ,1)	538.81	281.95	396.49
8	ARIMA(0, 2 ,1)	695.04	450.37	434.59
9	ARIMA(0, 0 ,2)	559.96	278.89	420.63
10	ARIMA(0, 1 ,2)	544.82	271.71	391.56
11	ARIMA(0, 2 ,2)	549.44	277.89	412.70
12	ARIMA(1, 0 ,0)	547.46	267.40	413.63
13	ARIMA(1, 1 ,0)	519.04	296.20	455.55
14	ARIMA(1, 2 ,0)	2892.21	645.85	817.03
15	ARIMA(1, 0 ,1)	561.43	282.26	414.88
16	ARIMA(1, 1 ,1)	539.55	280.94	394.15
17	ARIMA(1, 2 ,1)	586.35	386.94	451.30
18	ARIMA(1, 0 ,2)	556.11	276.08	418.16
19	ARIMA(1, 1 ,2)	531.36	272.46	394.61
20	ARIMA(1, 2 ,2)	548.14	276.24	433.78
21	ARIMA(2, 0 ,0)	543.16	276.84	420.12
22	ARIMA(2, 1 ,0)	509.93	253.88	440.78
23	ARIMA(2, 2 ,0)	1937.40	963.75	2354.00
24	ARIMA(2, 0 ,1)	545.08	265.54	419.62
25	ARIMA(2, 1 ,1)	537.28	285.54	403.67
26	ARIMA(2, 2 ,1)	512.18	246.57	436.00
27	ARIMA(2, 0 ,2)	436.47	277.93	422.59
28	ARIMA(2, 1 ,2)	536.05	273.96	399.86
29	ARIMA(2, 2 ,2)	548.95	288.44	424.83
30	ARIMA(0, 0 ,0) dummies	425.40	197.64	299.30
31	ARIMA(0, 1 ,0) dummies	476.93	225.70	276.60
32	ARIMA(0, 2 ,0) dummies	7886.29	2126.09	1792.76
33	ARIMA(0, 0 ,1) dummies	456.05	218.66	296.66
34	ARIMA(0, 1 ,1) dummies	415.38	202.44	284.85
35	ARIMA(0, 2 ,1) dummies	798.20	557.02	318.68
36	ARIMA(0, 0 ,2) dummies	477.72	241.28	306.00
37	ARIMA(0, 1 ,2) dummies	425.18	219.22	282.93
38	ARIMA(0, 2 ,2) dummies	405.73	650.48	339.80
39	ARIMA(1, 0 ,0) dummies	425.47	196.96	297.46
40	ARIMA(1, 1 ,0) dummies	417.76	225.53	257.01
41	ARIMA(1, 2 ,0) dummies	4905.65	980.97	2054.57
42	ARIMA(1, 0 ,1) dummies	469.24	212.66	310.32
43	ARIMA(1, 1 ,1) dummies	415.31	199.39	283.38
44	ARIMA(1, 2 ,1) dummies	719.47	593.73	298.22
45	ARIMA(1, 0 ,2) dummies	476.83	211.87	308.88
46	ARIMA(1, 1 ,2) dummies	429.18	214.44	286.64
47	ARIMA(1, 2 ,2) dummies	714.69	611.71	265.89
48	ARIMA(2, 0 ,0) dummies	438.86	269.24	303.32
49	ARIMA(2, 1 ,0) dummies	456.70	248.57	259.42

Covid Economics 38, 16 July 2020: 218-252

50	ARIMA(2, 2 ,0) dummies	827.67	1036.05	1455.96
51	ARIMA(2, 0 ,1) dummies	455.71	265.13	313.65
52	ARIMA(2, 1 ,1) dummies	420.19	251.79	288.18
53	ARIMA(2, 2 ,1) dummies	438.43	250.41	279.31
54	ARIMA(2, 0 ,2) dummies	No convergence	260.20	311.00
55	ARIMA(2, 1 ,2) dummies	415.41	268.92	288.83
56	ARIMA(2, 2 ,2) dummies	381.85	278.92	270.50

Notes: Models estimated for first 7 weeks of 2020. All the models are controlled for the amount in the corresponding weeks of 2019. Dummies refers to paydays, Mondays, Fridays, and December. RMPSE calculated for February 19 to March 11.

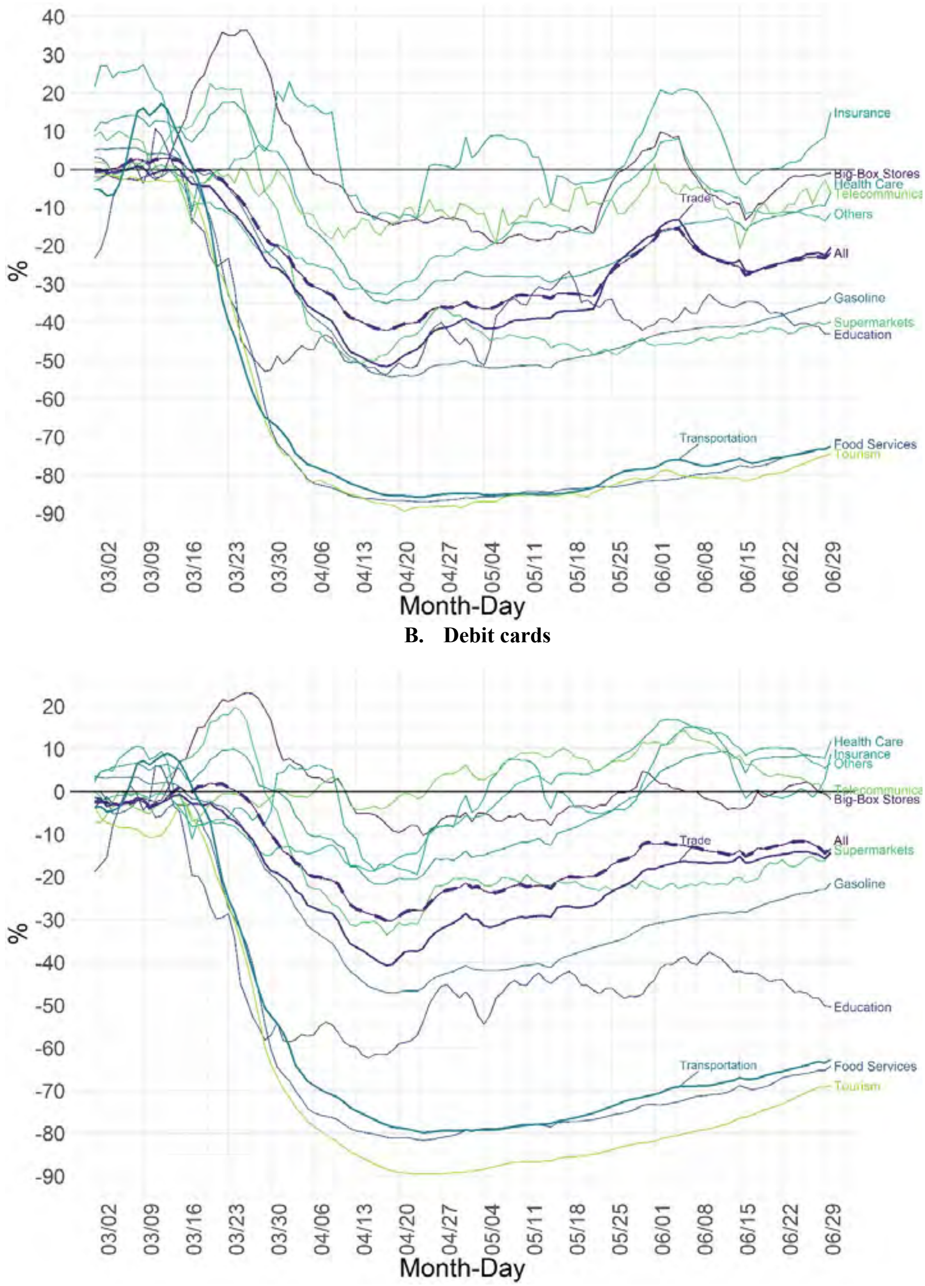
Figure S1. Changes in consumption patterns by sector with respect to corresponding period in 2019.



Notes: Authors' calculations. Comparison is to corresponding period in 2019 (in constant pesos of July 2018).

Figure S2. Changes in consumption patterns by sector with respect to corresponding period in 2019.

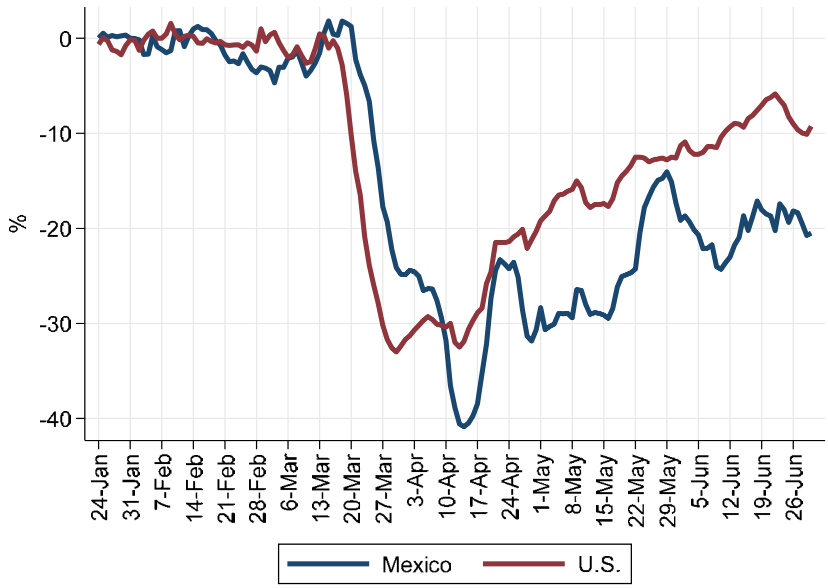
A. Credit cards



Notes: Authors' calculations. Comparison is to corresponding period in 2019 (in constant pesos of July 2018).

Covid Economics 38, 16 July 2020: 218-252

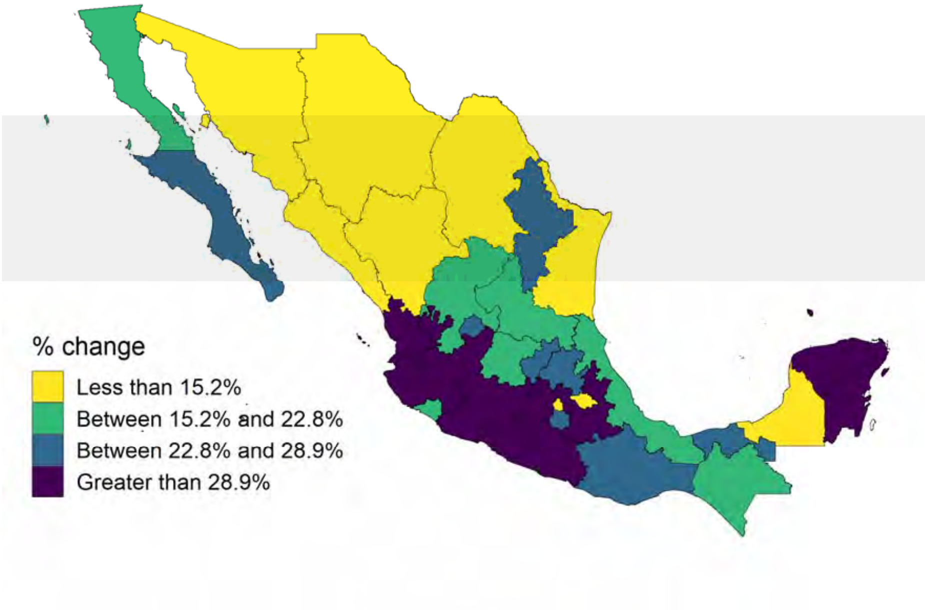
Figure S3. Comparison of expenditures in Mexico and U.S. Smoothed lines.



Note: Authors' calculations for the Mexican series. For the U.S., we used the calculation published by Chetty et al. (2020) and the https://tracktherecovery.org/ webpage. Both series use exactly the same construction. We first take a seven-day moving average, then we divide the 2020 series by the 2019 calendar day-month values. Finally, we divide the series by its average value for January 4-31.

Covid Economics 38, 16 July 2020: 218-252

Figure S4. Losses by state in total POS expenditure.



Notes: Authors' calculation. The map shows the percent change in POS expenditures from April to June with respect to sales for each state in the same period in 2019. Constant pesos of July 2018.

Table S3. Elasticity Estimates: Change in Percent Sales with Respect to Change in Percent Mobility, Panel by Day

	Google: Workplace Mobility			Apple: Driving Mobility		
	Total	Credit	Debit	Total	Credit	Debit
Coefficient	0.59	0.83	0.49	0.84	1.17	0.68
Standard Error	[0.04]	[0.05]	[0.03]	[0.05]	[0.08]	[0.04]
R^2	0.27	0.31	0.21	0.31	0.32	0.26
Total Obs.	4384	4384	4384	4384	4384	4384

Notes: Authors' calculations. The dependent variable is the percent change in POS expenditures on day *d* with respect to February 17 for each state in Mexico; the independent variable is the percent change in mobility. The regression includes fixed effects for state and day. The estimation period is February 15 to June 30. Clustered standard errors at state level.

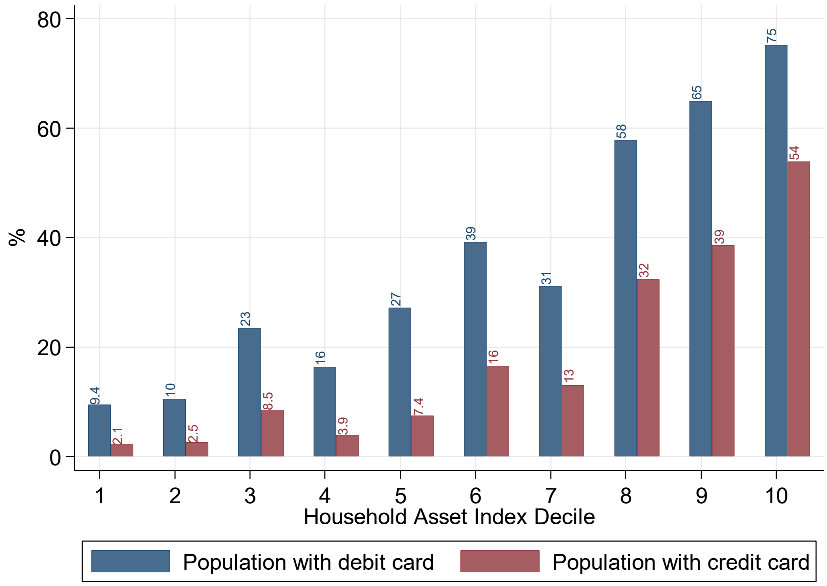
Table S4. Elasticity estimates: Change in % POS expenditures with respect to change in % Mobility

	Apple: High Driving Mobility			Apple: Low Driving Mobility		
	Total	Credit	Debit	Total	Credit	Debit
Coefficient	0.80	1.21	0.63	1.04	1.43	0.83
Standard Error	[0.05]	[0.08]	[0.04]	[0.09]	[0.11]	[0.07]
R^2	0.31	0.35	0.24	0.57	0.56	0.52
Total Obs.	320	320	320	320	320	320

Notes: Authors' calculations. The dependent variable is the percent change in POS expenditures in week *w* with respect to February 17 for each state in Mexico; the independent variable is the percent change in mobility for the same period. The regression includes fixed effects for state and week. Estimation period is February 15 to June 27. Clustered standard errors at the state level in brackets. High mobility states include Aguascalientes, Campeche, Chiapas, Chihuahua, Coahuila, Colima, Durango, Michoacán, Morelos, San Luis Potosi, Sinaloa, Sonora, Tamaulipas, Tlaxcala, Veracruz, and Zacatecas; low mobility states include Baja California, Baja California Sur, Mexico City, Estado de México, Guanajuato, Guerrero, Hidalgo, Jalisco, Nayarit, Nuevo León, Oaxaca, Puebla, Querétaro, Quintana Roo, Tabasco, and Yucatán. High mobility refers to driving mobility measured by Apple.

Covid Economics 38, 16 July 2020: 218-252

Figure S5. Cardholding by household asset index decile

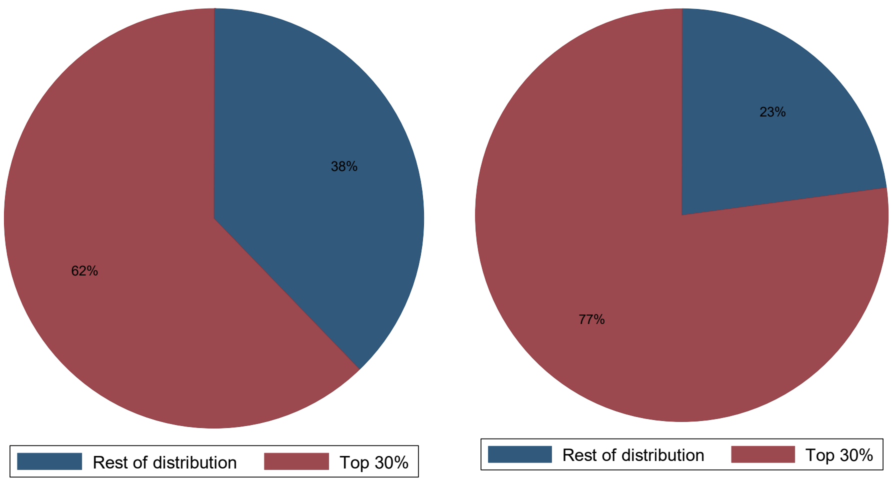


Notes: Authors' calculation with data from the INEGI intergenerational social mobility module (2016). Household asset index was constructed with PCA of the covariance matrix using asset holding (television, vehicles, home ownership, telephone, internet access, radio, DVD, blender, toaster, microwave, refrigerator, stove, washing machine, iron, sewing machine, fan, tablet computer, videogame console, computer, printer, and livestock) and years of schooling.

Covid Economics 38, 16 July 2020: 218-252

Figure S6. Cardholding by household asset index position

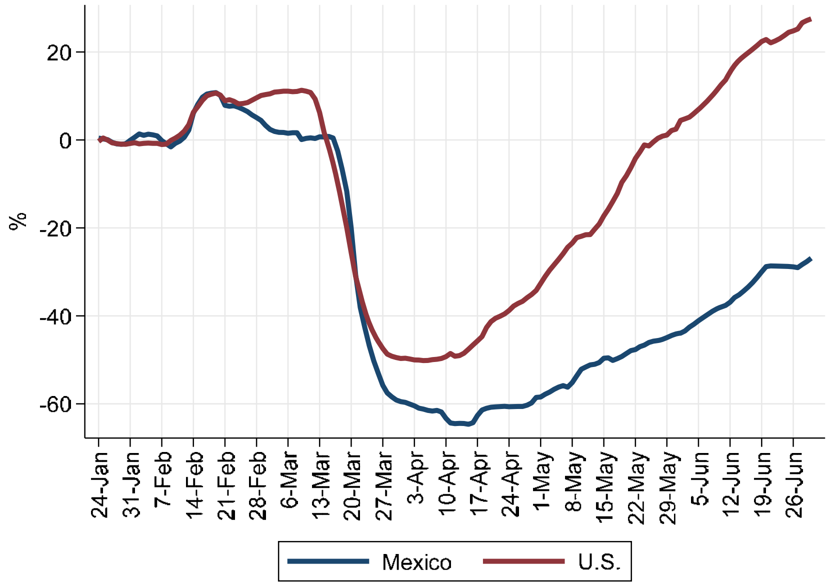
A. Debit card **B. Credit card**



Notes: Authors' calculation with data from INEGI intergenerational social mobility module (2016). The graph shows the percent of credit (debit) cardholding by household asset index distribution. Household asset index was constructed with PCA of the covariance matrix using asset holding (television, vehicles, home ownership, telephone, internet access, radio, DVD, blender, toaster, microwave, refrigerator, stove, washing machine, iron, sewing machine, fan, tablet computer, videogame console, computer, printer, and livestock) and years of schooling.

Covid Economics 38, 16 July 2020: 218-252

Figure S7. Comparison of mobility between United States and Mexico



Notes: Authors' calculation. Mobility uses Apple (2020) driving mobility. Base index is January 2020.

Covid Economics 38, 16 July 2020: 218-252