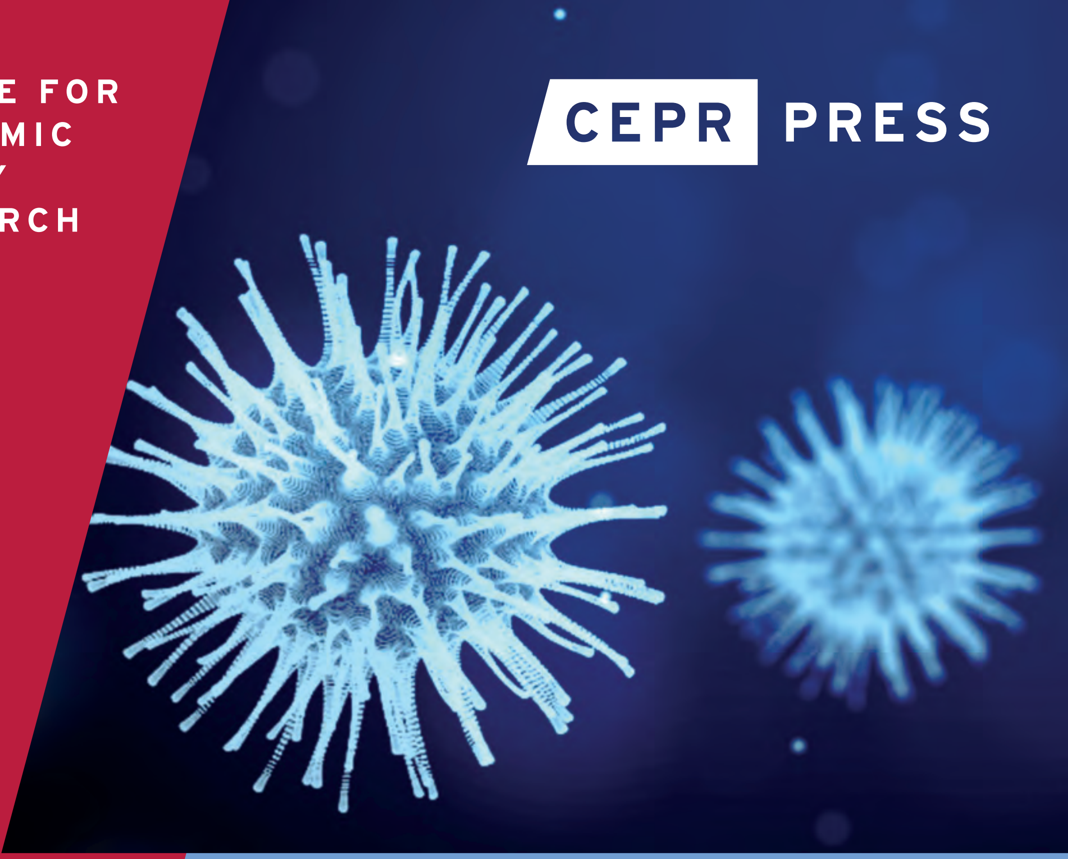
CENTRE FOR ECONOMIC POLICY RESEARCH

CEPR PRESS



COVID ECONOMICS

VETTED AND REAL-TIME PAPERS

ISSUE 51
7 OCTOBER 2020

RACIAL DISPARITIES
Milena Almagro, Joshua Coven, Arpit Gupta and Angelo Orane-Hutchinson

TRACKING AND FORECASTING
Andrew Harvey and Paul Kattuman

LIQUIDITY: BONDS BEAT BANKS
Olivier Darmouni and Kerry Y. Siani

LOCKDOWN IN INDIA
Kenneth Lee, Harshil Sahai, Patrick Baylis and Michael Greenstone

FAMILY BUSINESS IN NIGERIA
Elvis Korku Avenyo and Gideon Ndubuisi

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 Public Finance and Public Choice
Journal of Political Economy
Journal of Population Economics
Quarterly Journal of Economics
*Review of Corporate Finance Studies**
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 51, 7 October 2020

Contents

Racial disparities in frontline workers and housing crowding during COVID-19: Evidence from geolocation data 1
Milena Almagro, Joshua Coven, Arpit Gupta and Angelo Orane-Hutchinson

A farewell to R: Time series models for tracking and forecasting epidemics 36
Andrew Harvey and Paul Kattuman

Crowding out bank loans: Liquidity-driven bond issuance 74
Olivier Darmouni and Kerry Y. Siani

Job loss and behavioral change: The unprecedented effects of the India lockdown in Delhi 134
Kenneth Lee, Harshil Sahai, Patrick Baylis and Michael Greenstone

Coping during COVID-19: Family businesses and social assistance in Nigeria 159
Elvis Korku Avenyo and Gideon Ndubuisi

CEPR PRESS

Covid Economics **Issue 51, 7 October 2020**

Racial disparities in frontline workers and housing crowding during COVID-19: Evidence from geolocation data[1]

Milena Almagro,[2] Joshua Coven,[3] Arpit Gupta[4] and Angelo Orane-Hutchinson[5]

Date submitted: 25 September 2020; Date accepted: 29 September 2020

We document that racial disparities in COVID-19 in New York City stem from patterns of commuting and housing crowding. During the initial wave of the pandemic, out-of-home activity related to commuting is strongly associated with COVID-19 cases at the ZIP code level and hospitalization at an individual level. After layoffs of essential workers decreased commuting, case growth continued through household crowding. A larger share of individuals in crowded housing or commuting to essential work are Black, Hispanic, and lower-income. As a result, structural inequalities, rather than population density, help determine the cross-section of COVID-19 risk exposure in urban areas.

Covid Economics 51, 7 October 2020: 1-35

1 The views expressed in this paper are those of the authors and don't necessarily reflect the position of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

2 Visiting Scholar, Federal Reserve Bank of Minneapolis and Assistant Professor, Booth School of Business, University of Chicago.
3 Graduate Student, Department of Finance, Stern School of Business, New York University.
4 Assistant Professor, Department of Finance, Stern School of Business, New York University.
5 Graduate Student, Department of Economics, New York University.

I INTRODUCTION

The novel coronavirus disease 2019 (COVID-19) has disproportionately and negatively impacted disadvantaged populations. The hardest-hit regions of New York City include parts of the Bronx, Brooklyn, and Queens with high fractions of Black, Hispanic, and low-income populations as has been noted by Borjas (2020) and Schmitt-Grohé et al. (2020).[1] Nationwide, infections are three times as likely among Latinos and African-Americans compared to infections among whites.[2] While racial disparities in COVID-19 disease burdens have been widely recognized, the ultimate drivers of these inequities remain unclear.

This paper connects racial disparities in COVID-19 cases with mobility patterns for vulnerable populations by using anonymized mobile phone Global Positioning System (GPS) data. We focus on New York City, the global epicenter for the pandemic in Spring 2020. Our work splits the COVID-19 pandemic in New York City into two periods. In the initial stage of the crisis—lasting from March until early April—we document that the commuting behavior of essential workers placed them at greater risk of infection. We find these commuting patterns changed after early April, when many of these workers were laid off. At this point in the crisis, disease spread continued through a household crowding channel. The relative importance of household crowding compared to mobility patterns doubled during April. We find that racial minorities are over-represented in both mobility-based risk factors—occupational specialization in essential work professions and household overcrowding.

Our analysis is conducted at both the neighborhood and individual levels, allowing us to finely measure the nature of mobility responses in the wake of this pandemic. To do so, we link individual mobile phone data with ZIP Code-level data on daily COVID-19 infection and hospitalization rates, as well as Census data on occupation and household occupancy. A key advantage of our identification approach is that we are able to measure hospitalizations at aggregate levels as well as the *individual* level by isolating mobile phone users who appear in local hospitals. Doing so allows us to control for important local characteristics and use within-tract variation in commuting and housing characteristics.

[1]We document the demographic associations of COVID-19 in Section III.A. New York City official data suggest that African-Americans were 59% more likely to be diagnosed with COVID-19 relative to whites, while Hispanics were 64% more likely. See: https://www1.nyc.gov/site/doh/covid/covid-19-data.page.

[2]See data from the C.D.C. https://www.nytimes.com/interactive/2020/07/05/us/coronavirus-latinos-african-americans-cdc-data.html.

Covid Economics 51, 7 October 2020: 1-35

Our results suggest sizeable effects of mobility on disease exposure: increasing time outside of an individual's home census tract from the 10^{th} to the 90^{th} percentile is associated with a 30% increase in the hazard rate of hospitalizations. Similarly, individuals at the 90^{th} percentile of housing crowding are 7% more likely to be hospitalized than those at the 10^{th} percentile. While both measures suggest large and statistically significant impacts of our mobility measures on hospitalization outcomes, we also find that including the housing crowding measure lowers the measured impact of the commuting measure to 25%. This suggests a substantial correlation between mobility patterns and housing density, and it points to possible implications of policies that target these measures separately. For example, shutting down workplaces through lockdowns may lower infectious spread through a commuting channel but may instead result in individuals interacting more in crowded home settings.

Our analysis has implications for ongoing debates on the role of density and urban form on disease exposure. In contrast to research which emphasizes the role of static characteristics of urban design such as density (Duranton and Puga, 2020; Carozzi et al., 2020) or subways (Harris, 2020), we highlight the dynamic responses of individuals and groups which depend on access to preexisting resources. Notably, Manhattan—the densest and wealthiest borough—saw many fewer infections than the other boroughs.

Instead, our results suggest that the types of density that matter the most are the experienced density of front-line workers exposed to contact through direct physical proximity, as well as the crowding of individuals in multi-family households. We document that these mobility-induced densities are disproportionately experienced by vulnerable populations. In turn, structural inequalities lead disadvantaged groups to disproportionately live in crowded housing and specialize in jobs that require physical presence. These underlying inequalities in job presence and housing create temporary pockets of density through which SARS-CoV-2 virus propagates.

We contribute to a growing literature on COVID-19 by providing direct evidence on the role of specific mobility factors in contributing to the spread of the disease, as well as on the racial dimensions of these factors. Many papers have used geolocation data in the context of COVID-19 (Chen et al., 2020; Couture et al., 2020; García-Lopez and Puga, 2020). Our work is most closely related to Glaeser et al. (2020). We differ in four key ways. First, we consider both aggregated data and individual-level mobility data, allowing us to highlight individual risk factors for hospitalization. Second, our central focus is examining racial disparities. Third, we consider an additional housing crowding

Covid Economics 51, 7 October 2020: 1-35

Covid Economics 51, 7 October 2020: 1-35

dimension which was crucial at the stage in the pandemic when many workers stopped commuting due to unemployment. Finally, we complement the approach in Glaeser et al. (2020) by proposing a new identification strategy that leverages the granularity of our data. To do so, we construct a panel of buildings where our main outcome variable is the hospitalization of a building's resident. Our main identification assumption is that daily unobservables are common across all individuals who live in buildings in the same census tract. Chiou and Tucker (2020) also examine income mobility responses and focus on variation in the ability to work from home and internet access. We emphasize both income and racial disparities within urban areas and highlight both commuting- and housing-related disparities.

We build on methods used in prior works such as Athey et al. (2019), Chen et al. (2019), and Chen and Rohla (2018), which used mobile phone geolocation data to examine segregation, racial disparities in voting waiting times, and partisanship.

A growing literature also examines racial disparities specifically in the context of COVID-19 (Borjas, 2020; McLaren, 2020; McCormack et al., 2020; Almagro and Orane-Hutchinson, 2020; Sá, 2020; Karaca-Mandic et al., 2020). Our work adds to this literature by linking important mobility components of this racial disparity, and further directly connecting them to case exposure. Notably, the emerging medical literature, such as Rentsch et al. (2020) and Price-Haywood et al. (2020), finds evidence of racial disparities in exposure to COVID-19—but does not find evidence of racial differences in mortality. This highlights the importance of understanding why different racial groups are potentially exposed to infection at different rates, which our analysis does through considering commuting and housing crowding channels. Finally, our findings complement other well-established health disparities which may impact the severity of COVID-19 for different populations, such as those presented in Wong et al. (2002) and Trivedi et al. (2005).

II Data

II.A Geolocation Data

Mobile location data were sourced from VenPath, a holistic global provider of compliant smartphone data. Our data provider aggregates information from approximately 120 million smart phone users across the United States. GPS data were combined across applications for a given user to produce "pings" corresponding to time stamp–location pairs.

The provider anonymizes information on individual users. Ping data include both background pings (location data provided while the application is running in the background) and foreground pings (activated while users are actively using the application). Our sample period covers the period February 1st–July 12th, 2020.

II.B Estimating Time Outside of Home Tract

To isolate the mobility behavior of New York City residents, we employ multiple screens to filter out commuters, visitors, and those who leave the city either temporarily or permanently.

First we separate those who spend the night in New York City from those who commute into or visit New York City during the day in order to measure the behavioral responses of local residents. We select from the anonymous users those who have the majority of their pings between 6pm and 8am (night hours) in New York City (as opposed to any non-New York City county in the US) on at least three different days in a specific month. We then enforce a minimum required data density and keep only those with at least three pings on at least five nights in the data in New York City, with the same requirements during work hours.

Second, we filter out those who commute to, visit, or leave NYC by joining our remaining users' pings to census tracts from New York City Open Data. On every date, we identify each user's modal tract during night hours. If they ping in the modal tract at least twice a night, on at least five different nights in a month, we assign their most frequent remaining modal night tract as their "home census tract" (HCT). If a user spends an equal amount of nights in more than one HCT, we choose the one that they ping in the most throughout the data.

We repeat this process each month from February to June and exclude those who have been identified as residents in previous months. We use only one month of data at a time to identify residents' home tracts. We then analyze their data in the months after the month that was used to identify their home locations. This gives us a sample population of 647,068 unique users for our base analysis.[3]

Next, we exclude those who have left New York City from our analysis by requiring the county of their HCT to match each day with the county they spent the night in. The resulting dataset allows us to measure the mobility responses among NYC residents.

[3] We find that our estimated mobile phone population correlates with Census population at 0.89, suggesting representative sample coverage.

Covid Economics 51, 7 October 2020: 1-35

Covid Economics 51, 7 October 2020: 1-35

Finally, to measure out-of home-behavior, we estimate the fraction of pings that occur each hour within that user's home tract, and we then estimate the number of hours a user spends entirely outside of their home tract during the range of 8am to 10pm (the "number of hours entirely outside").

II.C Household Crowding

To construct our housing crowding metrics, we connect the ping data with the geographic data for all building footprints in NYC, which are created by Microsoft using satellite images.[4] We join these building shapes to land use data from the NYC Department of Planning at the lot level.[5] Multiple lots correspond to each Microsoft building. We do a geospatial join of the building lots to the Microsoft building shapes and then aggregate to arrive at the total number of residential units and residential square footage for each building.

To define a metric of housing crowding, we identify each user's modal home building each night and then their modal building across nights. We aggregate to the building level and count unique users for whom that building is their home building. For each building, we calculate the people per housing unit on each date. While our measures of mobility are measured at the individual level, our housing crowding measure is estimated at the building level. We then take the average of buildings for each tract and ZIP Code to get our "mean people per unit" measure of housing crowding. Appendix Figure B2 shows the spatial variation of the housing crowding measure, averaged across our sample. We tend to observe greater housing crowding in the outer boroughs of the city.

II.D Measuring Hospitalizations

To construct our hospitalization measure, we first determine which of the Microsoft building shapes correspond to hospitals by joining them to latitudes and longitudes of hospitals provided by HIFLD Open Data.[6] We include only hospitals within NYC that are not long-term care facilities or psychiatric hospitals. We can then see which pings correspond to which designated hospitals within NYC.

[4] This dataset can be found at: https://github.com/microsoft/USBuildingFootprints.

[5] See: https://www1.nyc.gov/site/planning/data-maps/open-data/dwn-pluto-mappluto.page.

[6] See: https://hifld-geoplatform.opendata.arcgis.com/datasets/6ac5e325468c4cb9b905f1728d6fbf0f_0.

Covid Economics 51, 7 October 2020: 1-35

We identify whether a user pings within a hospital from our hospital building shapefiles. We observe the first date when a user pings within a hospital as their hospitalization date. To verify that our measured hospitalizations line up with other data sources, we compare against actual hospitalization data for COVID-19 from the NYC Department of Health in Appendix Figure A2. Across the period from March 25th–April 22nd, we find a correlation of 0.78 between our mobility measure of hospitalization and actual hospitalizations, suggesting that we are able to accurately estimate individual hospitalized COVID-19 cases.

II.E NYC COVID-19 Data

Our source of incidence rates of COVID-19 and number of tests performed is the NYC Department of Health (DOH) data release.[7] The DOH releases (almost) daily data on the cumulative count of COVID-19 cases and the total number of residents who have been tested, divided by the ZIP code of residence. We have collected data covering the months of April and May.[8] In our analysis, we drop the first week of April due to these missing dates, and also because the first few days in our sample appear very noisy.

II.F Census and Occupation Shares

We obtain demographic and occupation data at the ZIP code and Census tract level from the American Community Survey (ACS). The demographic characteristics we include are ZIP Code median income, average age, racial breakdown, and health insurance status. We also include commuting-related variables: average commute time to work as well as means of transportation.

We also construct the shares of the working-age population employed in different occupation categories, similar to Almagro and Orane-Hutchinson (2020). The ACS provides the number of workers employed in each occupation by ZIP code of residence. We first divide occupation between flexible and non-flexible occupations. Then, we categorize non-flexible occupations according to their essential definition and similarity in work environments and social exposure. For our occupational variables, we count the number for workers in each of these occupations and normalize by working-age population, which includes everyone with age ranging from 18 to 65 years. The summary statistics of demographics and occupations can be found in Appendix A, which also breaks out the occupational groups in the non-flexible category separately.

[7] See: https://github.com/nychealth/coronavirus-data.

[8] Unfortunately April 2nd and April 6th are missing from our sample as these data have never been made publicly available.

II.G Aggregating Mobility and Crowding Measures

To aggregate our mobility and crowding metrics to the ZIP Code level, we use the geospatial shapes of NYC's census tracts provided by NYC Open Data, and we link to ZIP Codes using a crosswalk provided by the Department of Housing and Urban Development.[9] We select the ZIP and tract mapping that has the highest number of residents residing in the ZIP for a given tract to get a 1:1 mapping of tracts to ZIP Codes.

We can aggregate housing crowding data and individual mobility data to the building level. Panel A of Appendix Table A1 presents the summary statistics of the building-level dataset used for our housing crowding analysis. We identified 333,132 buildings in NYC with residential space, from March 1st to July 12th. We were able to link 538 hospitalizations to residents of these buildings, where the hospitalizations occurred between March 18th and April 22nd.

For the individual analysis, we aggregate these variables to the ZIP and date level. We winsorize these variables at the 1% level. All of the daily values are constructed using a seven-day moving average to reduce the noise of daily raw values and to account for weekly seasonality. Panel B of Appendix Table A1 presents the summary statistics of the individual-level dataset used for our survival analysis. We identified 286,367 individual residents in NYC from March 1st to July 12th. We were able to link 597 hospitalizations to these residents, where the hospitalizations occurred between March 18th and April 22nd.

III Results

III.A Descriptive Analysis

We begin with a descriptive analysis of our sample to highlight the key features of the COVID-19 pandemic in New York City. Appendix Figure A1 shows how our mobility metrics, housing crowding, and the daily share of positive tests evolve over time. We observe that the share of positive tests steadily decreases over time. We also observe a large decrease in mobility in early March that started prior to the stay-at-home order issued by Governor Cuomo on March 20. Our finding that mobility responds primarily to the pandemic, rather than the state-imposed order, is consistent with similar nationwide findings in Goolsbee and Syverson (2020). Mobility in our sample hits a low in

[9]See: `https://data.cityofnewyork.us/City-Government/2010-Census-Tracts/fxpq-c8ku` for the list of tracts and `https://www.huduser.gov/portal/datasets/usps_crosswalk.html` for the crosswalk.

Covid Economics 51, 7 October 2020: 1-35

Covid Economics 51, 7 October 2020: 1-35

early April before steadily recovering towards pre-pandemic levels later in our sample. On the other hand, we do not see any stark trend for the average number of people per housing unit.

We also contrast the time series of mobility measures in Appendix Figure B1 across different boroughs of NYC. In the early period of our sample, we observe the greatest sheltering responses in Queens and Staten Island (Richmond County). However, after May, we observe the highest sheltering responses in Manhattan. These differential patterns across boroughs may reflect the ability of different populations to shelter effectively given the tendency for these jobs to be precarious and non-local. We compare across both the time series and the cross-section in Appendix Figure B3. In the key months of the pandemic, in March and April, measured mobility patterns shows sheltering in certain high-income neighborhoods of Manhattan and Brooklyn—while residents in other boroughs were much more likely to spend time outside of their home tract.

We then plot some basic correlations between our mobility and housing density measures with demographics and occupations. Panel A of Figure 1 shows correlations of mobility with certain neighborhood demographics: the fraction of tract residents who are Black, Hispanic, and/or low-income.

We find substantial positive correlations of increased out-of-home mobility in areas with more low-income, Black, and Hispanic individuals. We also observe a positive correlation between crowded spaces and neighborhoods with a higher share of minorities and lower average income in Panel B of Figure 1.

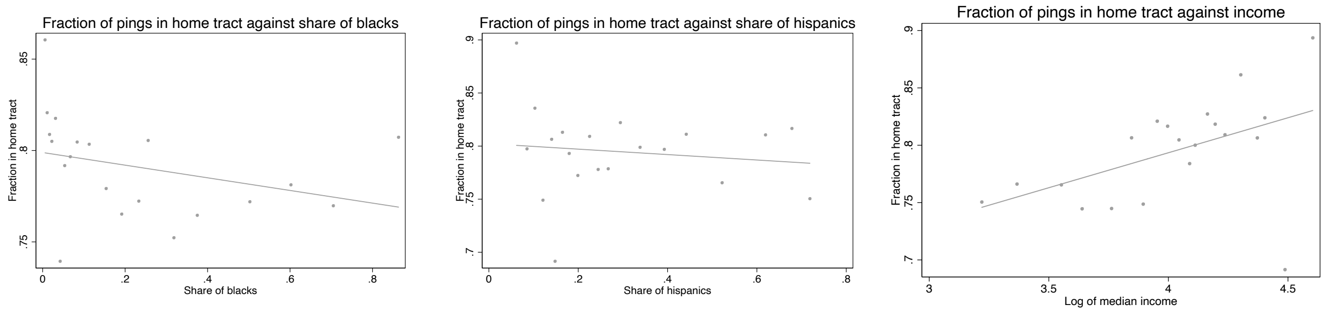
Given that many workers were laid off during April and May due to the pandemic, we turn to a more dynamic analysis to illustrate correlations between our two mobility measures (commuting and housing crowding) across occupations. In Panel C of Figure 1, we show the coefficients obtained when we regress daily mobility patterns and housing crowding on the share of flexible occupations after controlling for time trends. We find that ZIP Codes with higher shares of flexible occupations are consistently positively correlated with less mobility throughout our sample, as shown by the positive daily coefficients plotted on the left graph. This suggests that a key driver of our mobility results is related to commuting to essential work.[10] We conduct the same exercise for our measure of housing crowding and find a negative, although not significant, coefficient for most of our sample.

[10]The lack of significance in most of these regressions is mainly driven by our having a small number of observations for each daily regression.

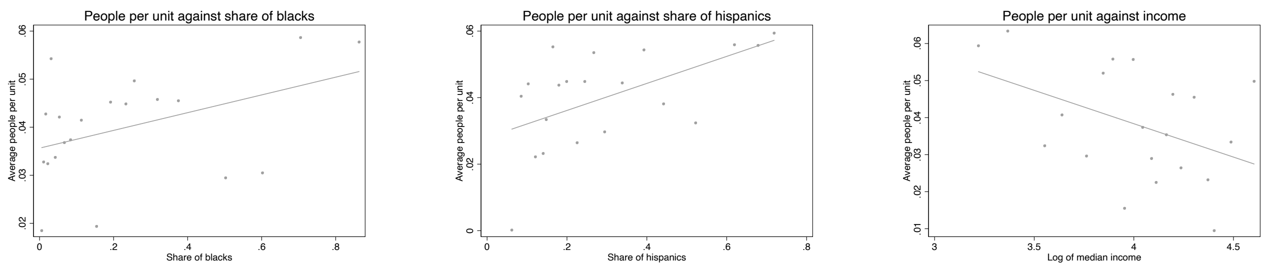
Covid Economics 51, 7 October 2020: 1-35

Figure 1: Demography and Mobility Measures

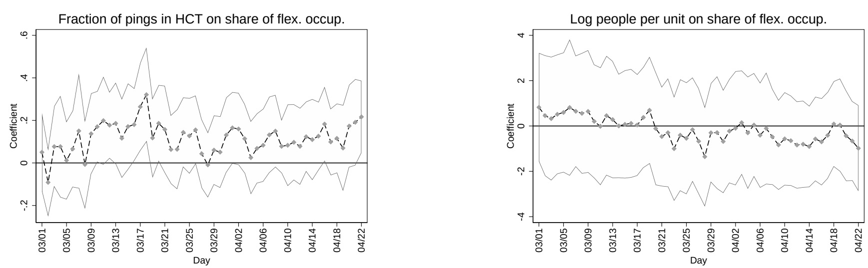
Panel A: Binscatter plots of share of mobile phone pings in home tract and demographics



Panel B: Binscatter plots of crowded spaces and demographics



Panel C: Daily correlations of mobility patterns and housing crowding, and share of flexible occupations



III.B ZIP Code-Level Analysis

Having established our basic variables, we turn next to a deeper analysis of the relationship between structural inequalities and the incidence of COVID-19. We start by constructing a panel of the daily share of positive tests across NYC ZIP codes from April 8$^{\text{th}}$ to May 26$^{\text{th}}$. For this specification, we estimate the following equation:

$$\text{share of positive tests}_{jt} = \beta_1 \text{mobility}_{jt} + \beta_2 \text{housing density}_{jt} + \gamma X_j + \mu_t + \varepsilon_{jt}$$

where X_j contains demographic and occupational characteristics at the ZIP Code level and μ_t is a day fixed effect that controls for the aggregate evolution of the pandemic in NYC.

Table 1 shows the estimation results of regressing the daily share of positive tests across ZIP Codes on mobility and housing crowding measures for several specifications that vary in their set of neighborhood controls. The first specification, column (1), includes only basic demographics such as race and income, while column (2) includes only our mobility and housing density measures. Column (3) includes all of these covariates together. A comparison of column (1) with column (3) shows that the initial racial disparities are partially explained by differences in mobility patterns and housing density, as all coefficients for racial groups shrink towards zero. Similarly, a comparison of column (2) with column (3) shows that much of the correlation in mobility and housing density is mediated through race and income differences: their coefficient sizes decrease by 10% and 52%, respectively, when including both set of controls.

For this basic specification, column (3), which presents the interpretation of the magnitudes for our mobility measure, is as follows: if the fraction of pings inside a home census tract (HCT) increases by 10%, an average level increase of 0.06, the daily share of positives decreases by 0.00396 points, corresponding to a 1.8% lower share of daily positive tests. On the other hand, a 10% increase in the number of people per unit corresponds to a 0.0019-point increase, or equivalently a 0.8% increase in the share of positive tests.[11] When we include other controls, we find that the size of the coefficient on the out-of-tract mobility decreases, becoming non-significant with the inclusion of the occupation controls. This result indicates that most of the housing density variation can be captured by the occupational composition of neighborhoods.

[11] The average rate of daily tests that came out positive between April 8$^{\text{th}}$ and May 26$^{\text{th}}$ is 22%.

Covid Economics 51, 7 October 2020: 1-35

Covid Economics 51, 7 October 2020: 1-35

Examining only demographic variables shows evidence of strong racial disparities in test positivity rates. Incorporating mobility, demographic, and occupational controls lowers the coefficient on fraction Black by 45%, and lowers the coefficient on the fraction of population members who are Asian to zero. This suggests that racial disparities in test positivity rates, at least for these groups, can be accounted for by variation in background variables related to mobility and occupation—though we observe larger residual disparities for Hispanic populations.

III.B.1 Weekly Analysis

In principle, the effects of mobility measures may vary over time given the dynamic behavior of other factors such as the natural evolution of the pandemic or its effects on the economy. Motivated by this hypothesis, we estimate the following equation:

$$\text{daily share of positives}_{jt} = \alpha_{1,t}\text{mobility}_{jt} + \alpha_{2,t}\text{housing density}_{jt} + \gamma_t X_j + \mu_t + \varepsilon_{jt}.$$

Panels A and B of Appendix Figure A3 plot the evolution of the daily coefficients for our mobility measure as well as for housing density. These patterns suggest that mobility had a larger impact for the first three weeks of April, while the magnitude of the coefficient of housing crowding was larger for the last week of April and the first week of May.

As pointed out by Glaeser et al. (2020), the estimates of these regressions are likely masking correlations with unobservables that bias the coefficients towards zero. For example, a change in behaviour following the evolution of the pandemic correlates with mobility measures as well as with the number of new infections. For this reason, we turn to our next analysis, where we leverage the granularity of our data to overcome these endogeneity concerns.

Covid Economics 51, 7 October 2020: 1-35

Table 1: Neighborhood Associations of Positive Tests

Dependent Variable:	Daily Share of Tests that are Positive									
	(1) Race & Income		(2) Mobility		(3) Mobility & Race, Income		(4) Mobility & Demographics		(5) Mobility, Dem. & Occupations	
Fraction of Pings in Home Tract			-0.076***	(0.024)	-0.066**	(0.026)	-0.144***	(0.026)	-0.065***	(0.025)
Log People per Unit			0.044***	(0.002)	0.019***	(0.002)	0.011***	(0.003)	0.000	(0.003)
Log Income	-0.003	(0.003)			-0.001	(0.003)	-0.015**	(0.006)	-0.000	(0.011)
% Black	0.113***	(0.004)			0.095***	(0.005)	0.061***	(0.008)	0.051***	(0.008)
% Hispanic	0.154***	(0.006)			0.145***	(0.006)	0.125***	(0.008)	0.157***	(0.010)
% Asian	0.170***	(0.007)			0.152***	(0.008)	0.006	(0.013)	0.002	(0.013)
Share $\geq 20, \leq 40$							0.066	(0.059)	0.040	(0.054)
Share $\geq 40, \leq 60$							0.016	(0.090)	0.027	(0.108)
Share ≥ 60							0.456***	(0.058)	0.371***	(0.072)
Share Male							0.192***	(0.050)	0.362***	(0.055)
Log Household Size							0.124***	(0.020)	0.086***	(0.025)
Log Density							0.007***	(0.002)	0.004**	(0.002)
% Public Transport							0.020	(0.014)	-0.016	(0.018)
Log Commute Time							0.016	(0.011)	0.017	(0.012)
% Uninsured							0.203***	(0.038)	0.351***	(0.047)
Bronx							-0.010	(0.007)	-0.017**	(0.007)
Brooklyn							0.026***	(0.008)	0.028***	(0.007)
Queens							0.022***	(0.007)	0.030***	(0.008)
Staten Island							-0.019**	(0.008)	-0.003	(0.009)
% Flexible Occupations									-0.034	(0.058)
% Health Practitioners									-0.121	(0.138)
% Other Health									0.618***	(0.107)
% Firefighting									-0.323	(0.305)
% Law Enforcement									-1.705***	(0.397)
% Essential: Service									-0.287***	(0.068)
% Non Ess.: Service									-0.017	(0.144)
% Ind. and Construction									-0.240**	(0.096)
% Essential: Technical									-0.683**	(0.294)
% Transportation									0.312***	(0.106)
Constant	0.280***	(0.014)	0.570***	(0.019)	0.395***	(0.023)	0.014	(0.059)	-0.105	(0.066)
Day FE	✓		✓		✓		✓		✓	
N	6401		6401		6401		6401		6401	
adj. R^2	0.89		0.88		0.89		0.91		0.91	

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

III.C Building-Level Analysis

In this section we exploit the panel structure of our sample including both mobility and housing crowding measures in order to better address identification concerns. Unfortunately, test data at more granular geographical levels are unavailable. To overcome this challenge, we instead measure hospitalizations at the individual level by measuring direct individual visits to hospitals. We classify individuals as being hospitalized if they spend more than 24 hours at a hospital. We focus on the first month after the issuance of the stay-at-home order to maximize the probability that new hospitalizations that we observe in our data are due to COVID-19 and not due to something else.[12] Our main outcome variable is defined as the event of a building's resident being hospitalized per building occupancy. Because our unit of observation is the crowding at the building level, we analyze as our key outcome the hospitalization rate per building.

Table 2 shows the result of an analysis in which our dependent variable is the event of a new hospitalization within a building and different specifications include different sets of neighborhood controls. Our main regression equation is

$$\text{hospitalization}_{bt} = \alpha_1 \text{mobility}_{bt} + \alpha_2 \text{housing density}_{bt} + \gamma X_{j(b)} + \mu_t + \varepsilon_{bt},$$

where mobility_{bt} and $\text{housing density}_{bt}$ are respectively the average mobility and housing density measures for date t, $X_{j(b)}$ are demographic and occupational controls for the Census tract where the building b is located, and μ_t are date fixed effects.

Specifications (1)–(4) are similar to those of Table 1 but are done at a different level of aggregation. Given that we have variation across buildings within the same census tract for any given date, we can include a fixed effect at the census tract and date level, $\delta_{j(b)t}$, to control for daily factors common to all individuals within a census tract. That is, our regression equation in this case is:

$$\text{hospitalization}_{bt} = \alpha_1 \text{mobility}_{bt} + \alpha_2 \text{housing density}_{bt} + \delta_{j(b)t} + \varepsilon_{bt}.$$

Our identifying assumption for this specification is based on the hypothesis that unobservables with temporal variation that correlate with mobility and housing density measures are common to all buildings inside the same census tract.

[12]Using data available at `https://github.com/thecityny/covid-19-nyc-data` that constructed total hospitalizations from reports of Governor Cuomo's office, we observe that more than 50% of all hospitalizations for the first three weeks of April were related to COVID-19. This measure includes new hospitalizations as well as as patients with more long-term diseases or patients in palliative care. For the time-series correlation of our measure of hospitalizations and the official numbers, see Appendix A2.

Covid Economics 51, 7 October 2020: 1-35

Table 2: Impact of Mobility on Hospitalizations

Dependent Variable	Hospitalizations per Occupancy				
	(1) Mobility	(2) Mobility & Race, Income	(3) Mobility & Demographics	(4) Mobility & Dem., Occ.	(5) Census Tract × Day
Fraction of Pings in Home Tract	-1.815e-04***	-1.780e-04***	-1.766e-04***	-1.770e-04***	-1.837e-04***
	(2.624e-05)	(2.618e-05)	(2.620e-05)	(2.626e-05)	(2.766e-05)
Log People per Unit	1.226e-05***	1.217e-05***	1.283e-05***	1.281e-05***	1.347e-05***
	(2.313e-06)	(2.340e-06)	(2.329e-06)	(2.333e-06)	(2.482e-06)
Constant	3.135e-04***	3.496e-04*	2.832e-04	1.373e-04	3.191e-04***
	(2.873e-05)	(1.425e-04)	(3.389e-04)	(3.547e-04)	(2.979e-05)
Day FE	✓	✓	✓	✓	
Demographic Controls			✓	✓	
Occupation Controls				✓	
Census Tract × Day FE					✓
N	1,355,459	1,354,504	1,354,504	1,354,504	1,355,459

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The interpretation of the coefficients is as follows. For our preferred specification, column (5), if a building's residents increase their fraction of pings within their HCT by 10%, the number of hospitalizations per occupant in that building decreases by 7.8%.[13] Similarly, if a building's number of people per housing unit increases by 10%, we expect to see hospitalizations per occupant increase by 0.96%. These results are similar to those presented in Table 1 where the effect sizes of mobility patterns are larger compared to housing crowding.

Also, we observe that coefficients on mobility patterns are very stable across specifications, showing small differences that are no larger than 5%. This result implies that for this particular specification, unobservables are not producing meaningful biases in the coefficients. On the other hand, we see more pronounced changes in housing density with a bias of 10% when we compare the results in column (1) to those in column (5). In both cases, we see a bias that shrinks the coefficients towards zero, which is consistent with the fact that people adjust their behaviour counter-cyclically with the evolution of the pandemic.

[13] The average probability of hospitalization per occupant is 1.396e-4, and the average number of pings per building is 0.6.

Covid Economics 51, 7 October 2020: 1-35

Covid Economics 51, 7 October 2020: 1-35

III.C.1 Building-level weekly analysis

Similarly, motivated by the dynamic evolution of different channels of transmission, we estimate the following equation:

$$\text{hospitalization}_{bt} = \alpha_{1,w(t)}\text{mobility}_{bt} + \alpha_{2,w(t)}\text{housing density}_{bt} + \gamma_{w(t)}X_{j(b)} + \mu_t + \varepsilon_{bt},$$

where coefficients are allowed to change week by week as denoted by subindex $w(t)$.

Appendix Table C2 shows that the coefficients for mobility patterns had a larger impact at early stages of the pandemic and that they decrease in magnitude over time, similar to Glaeser et al. (2020). We also find a similar pattern for housing density. Between week 1 (March 25^{th} to March 31^{st}) and week 4 (April 15^{th} to April 21^{st}), the coefficient for mobility patterns declines by 48% while decreasing by 41% for housing density—suggesting that housing density gained more importance over time with the progression of the pandemic, the issuance of the state-at-home order, and the large economic shock that led to high unemployment levels.

III.D Individual-Level Analysis

In this section we present results obtained using anonymized individual-level data. Given that we can track individual pings over time, we associate mobility measures to individual phones. While we cannot see whether an individual has been tested, we can observe whether an individual pings inside a hospital (in which case we assign them as a hospitalized individual). Our measurement of individual hospitalizations is an important contribution to the literature, which has generally focused on cases measured at more aggregate levels—and hence has been unable to control for possibly important local covariates. However, we face an important challenge of censoring. The event of being hospitalized due to COVID-19 generally happens only once with a probability that increases over time. For this reason, we borrow tools from the survival analysis literature to appropriately account for this censoring issue as well as for the fact that the probability of this event is not independent of what happened in the past. In doing so, we construct a panel of individuals with the hazard rate of being hospitalized as the outcome variable whose covariates are seven-day moving averages of mobility measures and housing density lagged by two weeks.

We start by plotting Kaplan-Meier graphs with the cumulative probability of failure on a daily time scale. We observe that the probability of being hospitalized increases

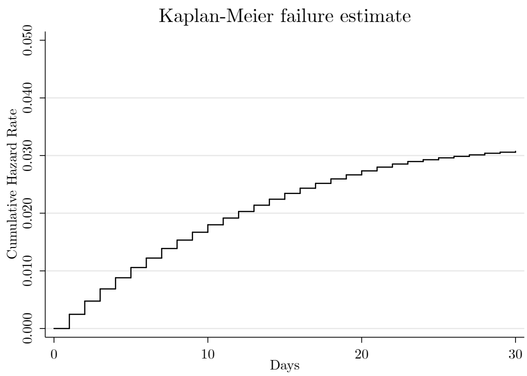
Covid Economics 51, 7 October 2020: 1-35

over time. For the following graphs, we have divided the population into two bins corresponding to above and below median. We do so for two variables: share of pings in HCT and average number of people in the same housing unit.

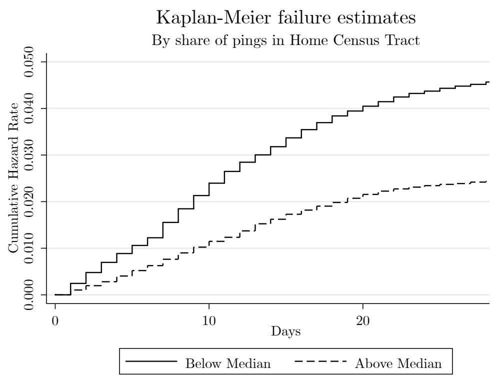
Panel B of Figure 2 shows that spending more time outside of the HCT is associated with a greater cumulative probability of being hospitalized. Similarly, individuals with more people in their housing unit also experience a higher cumulative hazard rate of being hospitalized, as shown in Panel C of Figure 2.

Figure 2: Kaplan-Meier graphs of survival probability

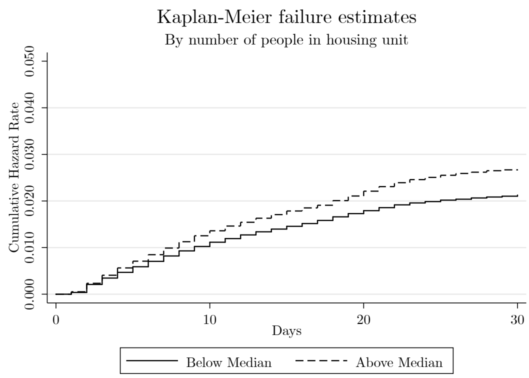
Panel A: Whole Sample



Panel B: By Share of Pings Outside HCT



Panel C: By Average People per Housing Unit



Covid Economics 51, 7 October 2020: 1-35

III.E Survival Analysis Regressions

In this section we present estimation results obtained from semiparametric Cox regressions where a failure in our sample is the event of being hospitalized in a panel of individuals. This type of estimation constructs hazard rates of being hospitalized nonparametrically and then uses such hazard rates as the outcome variable in a regression where covariates can be similarly defined as in any standard linear regression.

We first start by looking at all individuals.[14] To understand how mobility and housing crowding measures correlate with each other, we perform our analysis in two steps: (1) including only mobility measures, and (2) including mobility as well as housing crowing measures.

Our estimation results from the Cox regression are presented in Table 3, and they highlight the central role of out-of-home mobility and housing crowding in determining individual hospitalization rates. First, we observe a similar pattern as in Table 2 for mobility patterns in both Panel A and B of Table 3: including demographics and occupational controls decreases the magnitude of the coefficient, suggesting that part of the mobility patterns can actually be mediated by occupations and demographics.

Moreover, we can employ a similar identification strategy as in our aggregate analysis at the Census tract level. Unfortunately, Cox regressions do not allow fixed effects at the same level as the temporal unit level, which in this case is days. Hence, we cannot include day fixed effects. To overcome this problem, we pool days in the same week. We then interact week with ZIP Code to construct fixed effects that control for common factors in a given week and inside a given ZIP Code. The identifying variation comes from differences in individual patterns for those who live in the same ZIP Code in any given week. Our exclusion restriction is that unobservables that correlate with mobility and housing crowding patterns are common to all individuals living in the same ZIP Code in a given week. Reassuringly, we obtain similar results as in Table 2, where we find that the coefficient on mobility increases in magnitude after controlling for such types of unobservables, which possibly indicates that individuals countercyclically adjust their mobility patterns with the evolution of the pandemic.[15]

[14]For 34% of individuals in our sample, we cannot identify a modal building and thus we cannot construct a housing density measure for them.

[15]For full results, see Appendix D.

Table 3: Cox Regression of Mobility on Hospitalization

Panel A: Personal Mobility

Dependent Variable:	Hazard Rate of Hospitalization				
	(1) Baseline	(2) Demographics	(3) Dem. & Occupations	(4) ZIP Code Fixed Effect	(5) ZIP × Week Fixed Effect
Share of Pings in HCT	-0.224*** (0.060)	-0.216*** (0.060)	-0.215*** (0.060)	-0.211*** (0.060)	-0.284*** (0.063)
Demographic Controls		✓	✓		
Occupation Controls			✓		
ZIP Fixed Effects				✓	
ZIP × Week Fixed Effects					✓
Number of Observations	1,798,442	1,795,810	1,795,810	1,798,442	1,795,810

Standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel B: Personal Mobility and Housing Crowding

Dependent Variable:	Hazard Rate of Hospitalization				
	(1) Baseline	(2) Demographics	(3) Dem. & Occupations	(4) ZIP Code Fixed Effect	(5) ZIP × Week Fixed Effect
Share of Pings in HCT	-0.196*** (0.064)	-0.186*** (0.064)	-0.186*** (0.064)	-0.187*** (0.065)	-0.243*** (0.067)
Log People per Unit	0.023*** (0.006)	0.022*** (0.006)	0.021*** (0.006)	0.025*** (0.006)	0.015*** (0.006)
Demographic controls		✓	✓		
Occupation controls			✓		
ZIP Fixed Effects				✓	
ZIP × Week Fixed Effects					✓
Number of Observations	1,677,868	1,675,466	1,675,466	1,677,868	1,675,466

Standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The magnitude of mobility patterns decreases when including housing crowding measures when we compare Panel A to Panel B for all specifications. For example, for column (5), the magnitude of mobility patterns decreases by 14%. This pattern indicates that mobility patterns and housing crowding measures share a component that affects the individual level of exposure to the pandemic: if we ignore the housing crowding channel we will overestimate the effect of mobility patterns. However, it is important to disentangle these two channels in order to understand different effects of containment policies. For example, the most common non-pharmaceutical intervention, that is a quarantine on the population, can mitigate contagion by minimizing mobility but will undoubtedly fuel

Covid Economics 51, 7 October 2020: 1-35

Covid Economics 51, 7 October 2020: 1-35

exposure within the household.

In our preferred specification, column (5) of Panel B, a 10% increase in the share of pings inside HCT translates into a hazard rate of being hospitalized that is 1.46%. Similarly, a 10% increase in the number of people per unit leads to a hazard rate of being hospitalized that is 0.15% higher. These estimates are quite large, and they point to economically and statistically important impacts of our estimated risk factors on hospitalization risk.

Following the same specification, column (5) of Panel A in Table 3, we find that individuals at the 10^{th} percentile of share of pings inside HCT have a hazard rate of being hospitalized that is 30% larger than individuals at the 90^{th} percentile when only mobility measures are taken into account. This decreases to 25% when we control for the housing crowding channel, in Panel B . When we do the same comparison for the housing crowding distribution we find a hazard rate of being hospitalized that is 7% higher for individuals at the 90^{th} percentile compared to individuals at the 10^{th} percentile.

IV Conclusion

We document that important inequities in occupations and housing lead to racially disparate outcomes in exposure to COVID-19. We focus on the epicenter of the global pandemic in New York City, showing that infections spread in two waves. First, infections spread through essential workers, who continued to commute to establishments. Next, even after many of these essential workers were laid off, infections continued to spread within more crowded households, and the relative importance of this channel grew.

Racial disparities in infections reflect inequalities in access to both jobs and housing. Black, Hispanic, and low-income workers are more likely to be employed in an essential work occupation and hence exhibit mobility patterns which put them at greater risk of infection in the initial phase of the pandemic. We use novel data drawn from cell phones to measure these mobility patterns, which we use to establish a direct link between outside mobility at both neighborhood and individual levels. Our individual-level analysis advances on prior research using geolocation data by directly linking greater mobility for individual workers and presence in hospitals, controlling for other unobserved local factors.

We also connect both cell phone mobility and Census data on housing occupancy, and we see case increases in the second phase of the pandemic. We find that housing overcrowding predicts a greater caseload, and we also document more Black, Hispanic, and low-income households present in overcrowded buildings. As a result, vulnerable populations are disproportionately burdened by disease exposure through this housing crowding channel.

Our results present a stark contrast to some existing work on the COVID-19 pandemic which highlights the role of static factors such as population density or public transportation. We find that population density per se is not the dominant factor in explaining the cross-section in infections seen throughout this crisis: the densest borough, Manhattan, was relatively less affected. Instead, we find that underlying inequalities in access to jobs and housing explain the racial disparities in outcomes. Crowding and exposure at work, rather than density, best explains the pattern of exposure through the pandemic.

Covid Economics 51, 7 October 2020: 1-35

References

Almagro, Milena and Angelo Orane-Hutchinson, "The Determinants of the Differential Exposure to COVID-19 in New York City and Their Evolution Over Time," *Covid Economics, Vetted and Real-Time Papers*, 2020.

Athey, Susan, Billy Ferguson, Matthew Gentzkow, and Tobias Schmidt, "Experienced Segregation," Working Paper 2019.

Borjas, George J., "Demographic determinants of testing incidence and Covid-19 infections in New York City neighbourhoods," *Covid Economics, Vetted and Real-Time Papers*, 2020.

Carozzi, Felipe, Sandro Provenzano, and Sefi Roth, "Urban Density and Covid-19," *CEP Discussion Paper No 1711*, 2020.

Chen, M. Keith and Ryne Rohla, "The effect of partisanship and political advertising on close family ties," *Science*, 2018, *360* (6392), 1020–1024.

_, Judith A. Chevalier, and Elisa F. Long, "Nursing Home Staff Networks and COVID-19," Working Paper 27608, National Bureau of Economic Research July 2020.

_, Kareem Haggag, Devin G. Pope, and Ryne Rohla, "Racial Disparities in Voting Wait Times: Evidence from Smartphone Data," Working Paper 26487, National Bureau of Economic Research November 2019.

Chiou, Lesley and Catherine Tucker, "Social Distancing, Internet Access and Inequality," Working Paper 26982, National Bureau of Economic Research April 2020.

Couture, Victor, Jonathan Dingel, Allison Green, Jessie Handbury, and Kevin Williams, "Measuring movement and social contact with smartphone data: A real-time application to COVID-19," 2020.

Duranton, Gilles and Diego Puga, "The economics of urban density," Technical Report 2020.

García-Lopez, Miquel-Angel and Diego Puga, "Cities reshaped: Mobility patterns and COVID-19's impact on cities," 2020.

Covid Economics 51, 7 October 2020: 1-35

Glaeser, Edward L., Caitlin S. Gorback, and Stephen J. Redding, "How Much does COVID-19 Increase with Mobility? Evidence from New York and Four Other U.S. Cities," Working Paper 27519, National Bureau of Economic Research July 2020.

Goolsbee, Austan and Chad Syverson, "Fear, Lockdown, and Diversion: Comparing Drivers of Pandemic Economic Decline 2020," Working Paper 27432, National Bureau of Economic Research June 2020.

Harris, Jeffrey E., "The Subways Seeded the Massive Coronavirus Epidemic in New York City," Working Paper 27021, National Bureau of Economic Research April 2020.

Karaca-Mandic, Pinar, Archelle Georgiou, and Soumya Sen, "Assessment of COVID-19 Hospitalizations by Race/Ethnicity in 12 States," *JAMA Internal Medicine*, 08 2020.

McCormack, Grace, Christopher Avery, Ariella Kahn-Lang Spitzer, and Amitabh Chandra, "Economic Vulnerability of Households With Essential Workers," *JAMA*, 07 2020, *324* (4), 388–390.

McLaren, John, "Racial Disparity in COVID-19 Deaths: Seeking Economic Roots with Census data," Working Paper 27407, National Bureau of Economic Research June 2020.

Price-Haywood, Eboni G., Jeffrey Burton, Daniel Fort, and Leonardo Seoane, "Hospitalization and mortality among black patients and white patients with Covid-19," *New England Journal of Medicine*, 2020.

Rentsch, Christopher T., Farah Kidwai-Khan, Janet P. Tate, Lesley S. Park, Joseph T. King Jr., Melissa Skanderson, Ronald G. Hauser, Anna Schultze, Christopher I. Jarvis, Mark Holodniy et al., "Covid-19 by Race and Ethnicity: A National Cohort Study of 6 Million United States Veterans," *medRxiv*, 2020.

Sá, Filipa, "Socioeconomic Determinants of Covid-19 Infections and Mortality: Evidence from England and Wales," *Covid Economics, Vetted and Real-Time Papers*, 2020.

Schmitt-Grohé, Stephanie, Ken Teoh, and Martín Uribe, "COVID-19: Testing inequality in New York City," Working Paper 27019, National Bureau of Economic Research 2020.

Trivedi, Amal N., Alan M. Zaslavsky, Eric C. Schneider, and John Z. Ayanian, "Trends in the quality of care and racial disparities in Medicare managed care," *New England Journal of Medicine*, 2005, *353* (7), 692–700.

Covid Economics 51, 7 October 2020: 1-35

Wong, Mitchell D., Martin F. Shapiro, W. John Boscardin, and Susan L. Ettner, "Contribution of major diseases to disparities in mortality," *New England Journal of Medicine*, 2002, *347* (20), 1585–1592.

CEPR PRESS

Covid Economics 51, 7 October 2020: 1-35

Online Appendix

A Data Appendix

Table A1: Summary Statistics

Variable	Mean	Std. Dev.	p10	Median	p90
Panel A: Housing Crowding Measures					
People per Unit	0.043	0.228	0.000	0.000	0.031
Residential Units per Building	8.306	4.420	1.000	2.000	9.000
Residential Area (sqft)	8071	38,034	1140	2160	9011
Panel B: Individual Level Mobility					
Average Number of Pings in Home Tract	0.62	0.41	0.00	0.80	1.00
Average Distance From Home Tract (km)	2.42	4.50	0.00	0.22	8.13
Number of Hours Entirely Outside of Home Tract	1.46	2.37	0.00	0.00	5.00
Panel C: Other Variables and Local Controls					
Share of Positive Tests	0.563	0.085	0.438	0.583	0.645
Tests per Capita	0.018	0.006	0.012	0.017	0.026
Median Income (in $1000ss)	68.604	31.878	34.122	62.202	115.084
Share $\geq$ 20, $\leq$ 40	0.323	0.084	0.246	0.308	0.433
Share $\geq$ 40, $\leq$ 60	0.258	0.033	0.220	0.261	0.296
Share $\geq$ 60	0.200	0.079	0.132	0.190	0.276
Share Male	0.477	0.029	0.446	0.479	0.508
Household Size	2.683	0.537	1.930	2.750	3.300
% Black	0.200	0.240	0.010	0.076	0.600
% Hispanic	0.263	0.195	0.078	0.189	0.634
% Asian	0.144	0.139	0.017	0.094	0.335
Density (in 1000s of people per unit)	43.380	31.045	10.784	36.639	90.075
% Public Transport	0.532	0.150	0.312	0.543	0.712
Commuting Time (in mins)	40.647	7.054	27.200	42.100	48.100
% Uninsured	0.089	0.043	0.042	0.084	0.143
% Essential: Professional	0.126	0.089	0.046	0.092	0.285
% Essential: Service	0.065	0.033	0.035	0.060	0.107
% Essential: Technical	0.014	0.009	0.004	0.013	0.022
Non-Flexible Occupations:					
- % Health Practitioners	0.029	0.018	0.009	0.026	0.050
- % Other Health	0.038	0.024	0.010	0.035	0.073
- % Firefighting	0.012	0.009	0.003	0.012	0.023
- % Law Enforcement	0.007	0.007	0.001	0.006	0.014
- % Ind. and Construction	0.054	0.027	0.014	0.056	0.090
- % Transportation	0.029	0.016	0.004	0.032	0.048
- % Non Ess.: Professional	0.279	0.075	0.195	0.271	0.359
- % Science Fields	0.006	0.007	0.001	0.004	0.015
- % Law and Related	0.018	0.026	0.003	0.008	0.049
- % Non Ess.: Service	0.032	0.013	0.016	0.032	0.047

Covid Economics 51, 7 October 2020: 1-35

Figure A1: Time series of share of positive tests, mobility patterns, and housing density

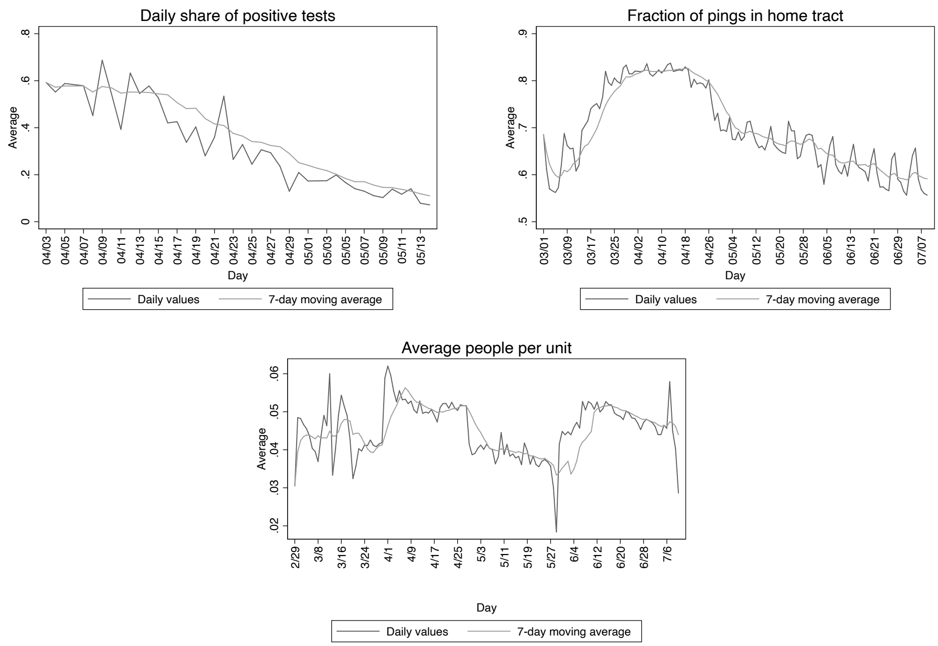
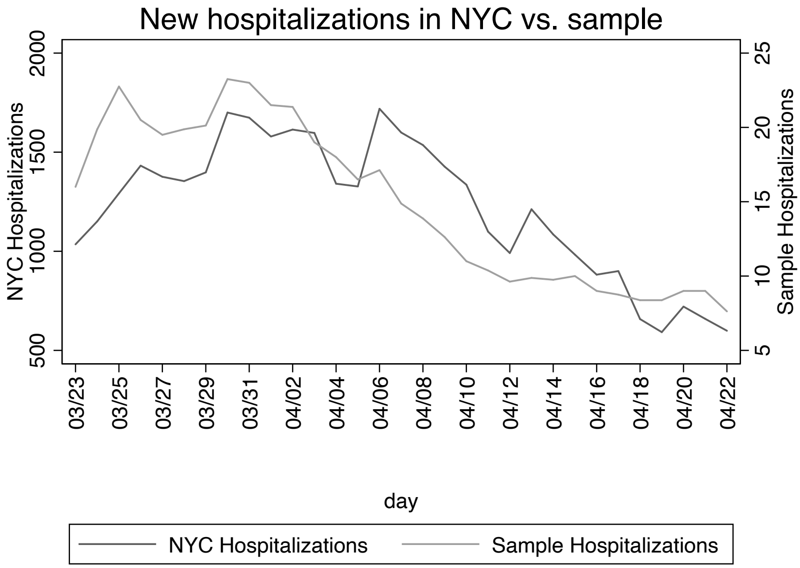


Figure A2: Comparing Hospitalizations in Mobile Phone Sample

These graphs plots the time series for our individual-level mobility-derived measure of hospitalization in comparison with the official figures provided by the DOH. The correlation between the two is 0.79



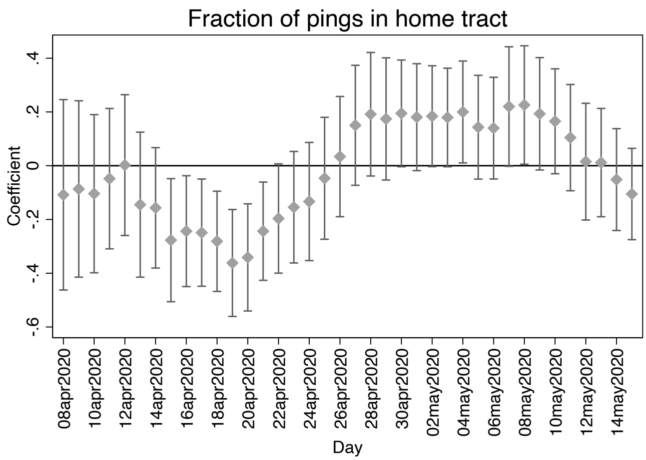
Covid Economics 51, 7 October 2020: 1-35

CEPR PRESS

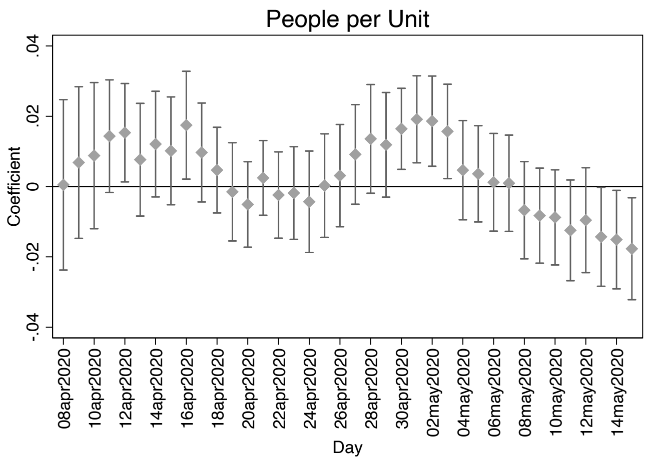
Covid Economics 51, 7 October 2020: 1-35

Figure A3: Time series of mobility and crowding coefficients

Panel A: Coefficient of share of pings in HCT on daily positive share



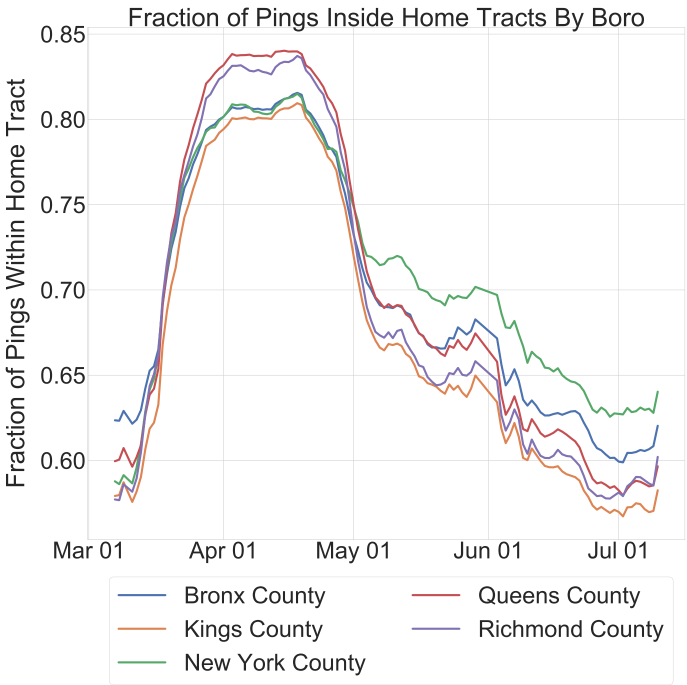
Panel B: Coefficient of log of people per unit on daily positive share



CEPR PRESS

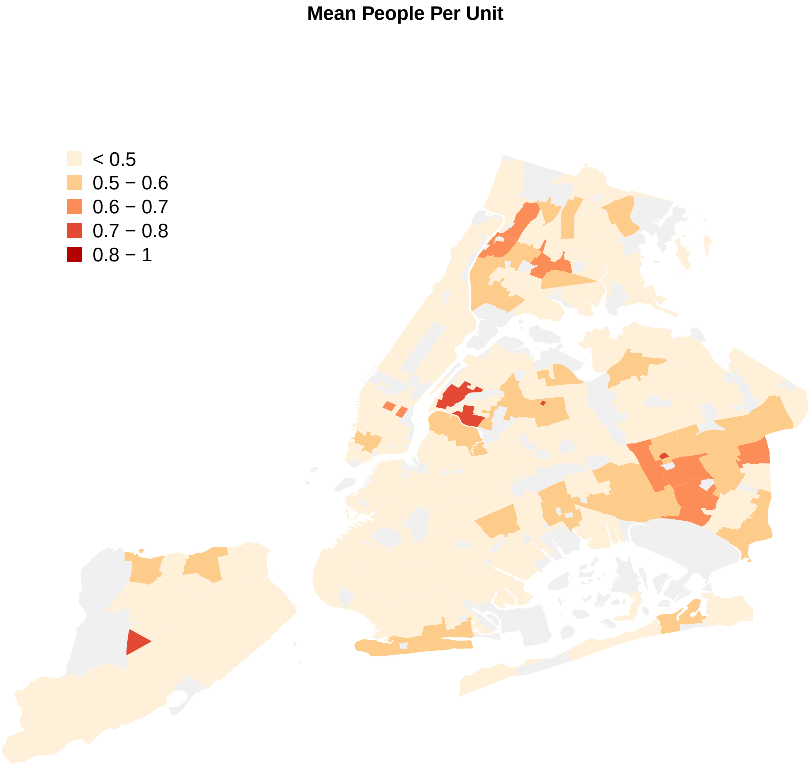
B Time Series and Cross-Section of Mobility measures

Figure B1: Time Series of Mobility Measures



Covid Economics 51, 7 October 2020: 1-35

Figure B2: Cross-Section of Housing Crowding Measure

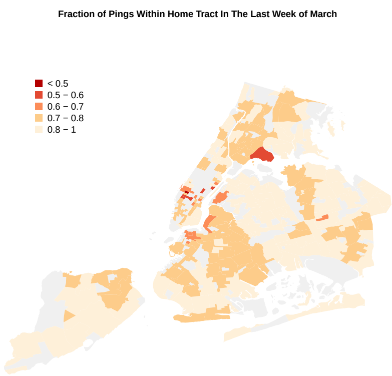


Covid Economics 51, 7 October 2020: 1-35

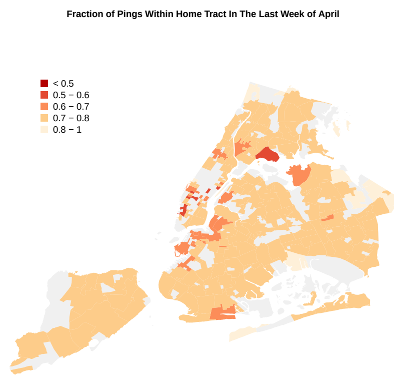
Covid Economics 51, 7 October 2020: 1-35

Figure B3: Cross-Section of Outside of HCT Mobility Measure

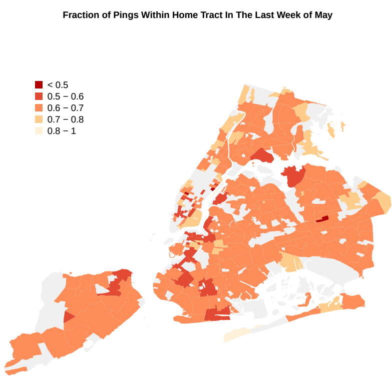
Panel A: March



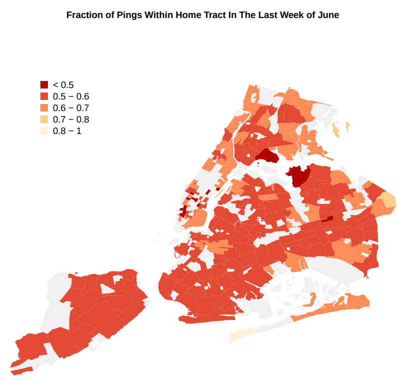
Panel B: April



Panel C: May



Panel D: June



C Building Level Analysis

Table C1: Dependent Variable: Hospitalizations per Occupancy

	(1) Mobility	(2) Demographics	(3) Demographics & Occupations	(4) Mobility & Demographics	(5) Mobility, Dem. & Occupations	(6) Mobility
Fraction of Pings in Home Tract	-1.815e-04***			-1.766e-04***	-1.770e-04***	-1.837e-04***
	(2.624e-05)			(2.620e-05)	(2.626e-05)	(2.766e-05)
Log People per Unit	1.226e-05***			1.283e-05***	1.281e-05***	1.347e-05***
	(2.313e-06)			(2.329e-06)	(2.333e-06)	(2.482e-06)
Log Income		1.669e-05	3.709e-05	1.977e-05	3.564e-05	
		(3.674e-05)	(4.768e-05)	(3.673e-05)	(4.768e-05)	
Share $\geq 20, \leq 40$		8.318e-05	1.696e-04	7.534e-05	1.611e-04	
		(2.365e-04)	(2.459e-04)	(2.365e-04)	(2.458e-04)	
Share $\geq 40, \leq 60$		2.038e-04	3.018e-04	1.846e-04	2.891e-04	
		(2.898e-04)	(2.924e-04)	(2.894e-04)	(2.922e-04)	
Share ≥ 60		2.264e-05	2.130e-04	2.546e-05	2.091e-04	
		(2.284e-04)	(2.477e-04)	(2.284e-04)	(2.476e-04)	
Share Male		-4.539e-04	-5.037e-04	-4.554e-04	-5.023e-04	
		(2.839e-04)	(2.915e-04)	(2.841e-04)	(2.915e-04)	
Log Household Size		-1.076e-06	1.277e-05	-1.182e-05	7.429e-06	
		(7.604e-05)	(8.391e-05)	(7.582e-05)	(8.380e-05)	
% Black		1.478e-05	2.735e-05	3.124e-06	1.856e-05	
		(5.735e-05)	(6.633e-05)	(5.751e-05)	(6.643e-05)	
% Hispanic		1.374e-05	5.078e-07	1.355e-05	2.703e-07	
		(8.876e-05)	(9.829e-05)	(8.877e-05)	(9.829e-05)	
% Asian		-1.067e-04	-1.388e-04	-1.018e-04	-1.345e-04	
		(8.472e-05)	(8.754e-05)	(8.462e-05)	(8.746e-05)	
% Public Transport		-7.439e-05	-5.389e-05	-6.890e-05	-4.948e-05	
		(1.031e-04)	(1.074e-04)	(1.030e-04)	(1.073e-04)	
Log Commute Time		3.133e-05	7.468e-05	3.125e-05	7.561e-05	
		(7.924e-05)	(8.154e-05)	(7.970e-05)	(8.192e-05)	
% Uninsured		7.045e-04*	7.817e-04*	7.054e-04*	7.798e-04*	
		(2.889e-04)	(3.044e-04)	(2.889e-04)	(3.044e-04)	
Bronx		-4.522e-05	-4.126e-05	-4.450e-05	-4.026e-05	
		(5.039e-05)	(5.096e-05)	(5.021e-05)	(5.078e-05)	
Brooklyn		7.926e-06	5.342e-06	9.103e-06	6.997e-06	
		(4.639e-05)	(4.764e-05)	(4.622e-05)	(4.750e-05)	
Queens		-5.847e-05	-6.588e-05	-5.814e-05	-6.561e-05	
		(4.749e-05)	(4.954e-05)	(4.735e-05)	(4.941e-05)	
Staten Island		-6.500e-05	-8.138e-05	-6.649e-05	-8.162e-05	
		(5.649e-05)	(5.927e-05)	(5.641e-05)	(5.919e-05)	
% Flexible Occupations			-2.138e-04		-1.953e-04	
			(1.781e-04)		(1.777e-04)	
% Health Practitioners			-3.211e-06		3.794e-05	
			(4.518e-04)		(4.522e-04)	
% Other Health			-6.136e-04		-6.307e-04	
			(3.278e-04)		(3.279e-04)	
% Firefighting			-1.482e-03*		-1.517e-03*	
			(7.495e-04)		(7.502e-04)	
% Law Enforcement			9.929e-04		1.048e-03	
			(9.474e-04)		(9.484e-04)	
% Essential: Service			-2.094e-04		-1.834e-04	
			(3.328e-04)		(3.330e-04)	
% Non Ess.: Service			-3.793e-04		-3.278e-04	
			(5.994e-04)		(5.995e-04)	
% Ind. and Construction			-3.336e-04		-3.357e-04	
			(3.358e-04)		(3.359e-04)	
% Essential: Technical			-1.397e-03		-1.365e-03	
			(7.253e-04)		(7.255e-04)	
% Transportation			4.253e-04		4.248e-04	
			(4.789e-04)		(4.791e-04)	
Constant	3.135e-04***	1.054e-04	-3.692e-05	2.832e-04	1.373e-04	3.191e-04***
	(2.873e-05)	(3.367e-04)	(3.527e-04)	(3.389e-04)	(3.547e-04)	(2.979e-05)
Day FE	✓	✓	✓	✓	✓	
Demographic Controls		✓	✓	✓	✓	
Occupation Controls			✓		✓	
Census Tract × Day FE						✓
N	1354631	1354631	1354631	1354631	1354631	1354631

Covid Economics 51, 7 October 2020: 1-35

Table C2: Weekly Analysis of Mobility Exposures and Hospitalization

Dependent Variable:	Hospitalizations per Occupancy			
	(1) Mar 25–31	(2) Apr 1–7	(3) Apr 8–14	(4) Apr 15–21
Fraction of Pings in Home Tract	-1.962e-04***	-2.715e-04***	-1.890e-04***	-1.020e-04**
	(6.609e-0.5)	(6.700-06)	(5.070e-0.5)	(4.990-0.5)
Log People per Unit	1.220e-05***	2.370e-05***	1.320e-05***	7.210e-06***
	(5.860e-06)	(5.460e-06)	(3.760e-06)	(4.810e-06)
Census Tract × Day FE	✓	✓	✓	✓
N	336,455	312,529	288,467	280,810

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Covid Economics 51, 7 October 2020: 1-35

D Survival Analysis

Table D1: Cox Regression of Mobility on Hospitalization

Dependent Variable:	Hazard Rate of Hospitalization									
	(1) Baseline		(2) + Demographics		(3) + Demographics & Occupations		(4) ZIP Code Fixed Effects		(5) ZIP Code × week Fixed Effects	
Fraction of Pings in Home Tract	-0.224***	(0.060)	-0.215***	(0.060)	-0.216***	(0.060)	-0.211***	(0.060)	-0.284***	(0.063)
Log Income			0.396***	(0.079)	0.661***	(0.104)				
Share $\geq 20, \leq 40$			-1.247*	(0.562)	-1.058	(0.571)				
Share $\geq 40, \leq 60$			-3.949***	(0.703)	-4.027***	(0.716)				
Share ≥ 60			-1.123	(0.580)	-0.501	(0.623)				
Share Male			-3.166***	(0.642)	-3.370***	(0.666)				
Log Household Size			1.113***	(0.200)	0.852***	(0.215)				
% Black			0.695***	(0.122)	0.450**	(0.141)				
% Hispanic			-0.689***	(0.201)	-1.052***	(0.217)				
% Asian			0.535*	(0.212)	0.299	(0.219)				
% Public Transport			0.051	(0.215)	0.260	(0.236)				
Log Commute Time			-0.092	(0.183)	-0.299	(0.198)				
% Uninsured			2.903***	(0.527)	2.196***	(0.600)				
Bronx			-0.192*	(0.084)	-0.124	(0.087)				
Brooklyn			-0.460***	(0.074)	-0.422***	(0.077)				
Queens			-0.807***	(0.081)	-0.762***	(0.083)				
% Flexible Occupations					-1.457***	(0.385)				
% Health Practitioners					-0.440	(1.128)				
% Other Health					-0.180	(0.786)				
% Firefighting					2.174	(1.712)				
% Law Enforcement					2.775	(2.333)				
% Essential: Service					1.587*	(0.742)				
% Non Ess.: Service					-1.167	(1.279)				
% Ind. and Construction					-0.057	(0.843)				
% Essential: Technical					-3.300	(1.962)				
% Transportation					0.168	(1.050)				
Demographic Controls			✓		✓					
Occupation Controls					✓					
ZIP Code FE							✓			
ZIP Code × week FE									✓	
Number of Observations	1798442		1795810		1795810		1798442		1795810	

Standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Covid Economics 51, 7 October 2020: 1-35

Table D2: Cox Regression of Mobility and Housing Crowding on Hospitalization

Dependent Variable:	Hazard Rate of Hospitalization									
	(1) Baseline		(2) + Demographics		(3) + Demographics & Occupations		(4) ZIP Code Fixed Effects		(5) ZIP Code × week Fixed Effects	
Fraction of Pings in Home Tract	-0.196**	(0.064)	-0.186**	(0.064)	-0.186**	(0.064)	-0.187**	(0.065)	-0.243***	(0.067)
Log People per Unit	0.023***	(0.006)	0.022***	(0.006)	0.021***	(0.006)	0.025***	(0.006)	0.015*	(0.006)
Log Income			0.450***	(0.084)	0.668***	(0.110)				
Share $\geq 20, \leq 40$			-1.592**	(0.595)	-1.337*	(0.605)				
Share $\geq 40, \leq 60$			-4.615***	(0.745)	-4.671***	(0.760)				
Share ≥ 60			-1.515*	(0.615)	-0.909	(0.662)				
Share Male			-2.978***	(0.679)	-3.188***	(0.705)				
Log Household Size			0.996***	(0.212)	0.767***	(0.228)				
% Black			0.748***	(0.130)	0.544***	(0.151)				
% Hispanic			-0.564**	(0.213)	-0.895***	(0.230)				
% Asian			0.672**	(0.225)	0.441	(0.234)				
% Public Transport			0.084	(0.228)	0.275	(0.250)				
Log Commute Time			0.043	(0.195)	-0.130	(0.211)				
% Uninsured			2.587***	(0.562)	1.830**	(0.640)				
Bronx			-0.266**	(0.089)	-0.199*	(0.092)				
Brooklyn			-0.492***	(0.079)	-0.457***	(0.081)				
Queens			-0.831***	(0.085)	-0.786***	(0.088)				
% Flexible Occupations					-1.323**	(0.408)				
% Health Practitioners					0.468	(1.179)				
% Other Health					-0.526	(0.840)				
% Firefighting					2.436	(1.819)				
% Law Enforcement					1.719	(2.497)				
% Essential: Service					1.289	(0.788)				
% Non Ess.: Service					-0.867	(1.359)				
% Ind. and Construction					0.450	(0.895)				
% Essential: Technical					-4.681*	(2.099)				
% Transportation					0.685	(1.115)				
Demographic Controls			✓		✓					
Occupation Controls					✓					
ZIP Code FE							✓			
ZIP Code × week FE									✓	
Number of Observations	1677868		1675466		1675466		1677868		1675466	

Standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Covid Economics Issue 51, 7 October 2020

A farewell to R: Time series models for tracking and forecasting epidemics

Andrew Harvey[1] and Paul Kattuman[2]

Date submitted: 30 September 2020; Date accepted: 1 October 2020

The time-dependent reproduction number, R_t, is a key metric used by epidemiologists to assess the current state of an outbreak of an infectious disease. This quantity is usually estimated using time series observations on new cases, or deaths, combining this information with the distribution of the serial interval of transmissions. For a new epidemic, such as COVID-19, the available information on the serial interval is limited. Bayesian methods are often used to combine this limited information with the new cases, with the new cases usually being smoothed by a simple, but to some extent arbitrary, moving average. This paper describes a new class of time series models for tracking and forecasting new cases. The viability of these models and their ability to deal with spikes and second waves is illustrated with data from Germany and Florida. As a by-product, estimates of R_t, together with their standard deviations, can be obtained from the growth rate of new cases. Very few assumptions are needed and those that are made can be checked. This leads us to the conclusion that tracking an epidemic by trying to estimate R_t may be neither necessary nor desirable.

Covid Economics 51, 7 October 2020: 36-73

1 Professor of Econometrics, Faculty of Economics, Cambridge University.
2 Reader in Economics, Cambridge Judge Business School, Cambridge University.

Covid Economics 51, 7 October 2020: 36-73

1 Introduction

The degree of infectiousness of a disease is given by the basic reproduction number, R_0, defined as the number of infections that are expected to result from a single infectious individual in a completely susceptible population. As an infection spreads, immunity starts to develop and for serious diseases, such as COVID-19, social behaviour may change endogenously, or may be modified, perhaps by the imposition of lockdown and social distancing measures. The progress of an epidemic is then usually tracked by the effective, or instantaneous, reproduction number, R_t, which is the number of people in a population who get infected by an individual at any specific time; see, for example, Wallinga and Lipsitch (2006), Gostic et al (2020) and Birrell et al (2020). Such tracking is of considerable importance for planning, but it raises the question whether estimating R_t is to be regarded as an end in itself or as a means to an end, namely tracking and forecasting the number of new cases, hospital admissions and deaths.

Harvey and Kattuman (2020) - hereafter HK - develop a class of generalized logistic time series models for predicting future values of a variable which when cumulated is subject to an unknown saturation level. These models are relevant for many disciplines, but attention in HK was focussed on applications for coronavirus. Observations on the cumulative series are transformed to growth rates and the logarithms of these growth rates are modelled with a time trend. Allowing this trend to be time-varying introduces flexibility and enables the effects of changes in policy and the environment to be tracked by filters for the level and slope. The filters are functions of current and past observations implied by the model. They can produce nowcasts of the current level of the incidence curve, together with forecasts of its future direction. Estimation is by maximum likelihood (ML) and goodness of fit can be assessed by standard statistical test procedures.

The methods used by epidemiologists to assess the current state of an infectious disease use time series observations on new cases, or deaths, and combine this information with information on the distribution of the serial interval of transmissions, sometimes called the infection profile; see, for example, Cori et al (2013), Thompson et al (2019), Chowell et al (2006) and Bettencourt and Ribeiro (2008). The serial interval is the gap between consecutive symptom onsets and for a new epidemic, such as COVID-19, information on it is limited. Bayesian methods are often used to combine the information on the serial interval with the observations on new cases, often

Covid Economics 51, 7 October 2020: 36-73

smoothed by a simple, but to some extent arbitrary, moving average. More practical and transparent methods, such as those used by the Robert Koch Institute (RKI) in Germany, adopt measures based solely on moving averages. These formulae effectively link estimates of R_t to the growth rate in new cases, or deaths. This growth rate is produced by the time series model in HK, and so it is possible to make nowcasts and forecasts of R_t under the assumptions made by RKI when a particular version of their moving average formula is adopted. The more general formulae given in Wallinga and Lipsitch (2006) may also be used. The nowcasts and forecasts of R_t therefore emerge as a by-product of the time series model and only require observations on the numbers of new cases or deaths. The underlying assumptions are clear and are subject to diagnostic tests, so estimates of R_t are implicitly validated. In contrast to R_t, which is not observed directly, the accuracy of forecasts of future observations can be assessed ex post, providing further testing of the effectiveness of the model.

The HK model is reviewed in Section 2 and in Section 3 it is shown that it can be used to estimate R_t. The conditional distributions of R_t implied by these estimators are given and it is shown how predictions of future values of R_t can be made. The implicit weights in the model-based filter are compared with the weights in the simple moving average ratio estimators used by RKI and the similarities and differences are highlighted. In Section 4 we briefly review and comment on the methods currently used in epidemiology. Modelling issues that arise at the beginning of an epidemic are discussed in Section 5 and illustrated by the analysis of the Spanish flu outbreak of 1919. Section 6 examines how the model can be used to assess the importance of spikes in new cases and to track second waves. Data from Germany and Florida are used to test the viability of the new approach. Section 7 concludes by suggesting that tracking an epidemic by methods dependent on R_t may be neither necessary nor desirable: the focus should be on the growth rate of new cases and deaths.

2 The time series model and its implementation

This section sets out the basic model in which the logarithm of the growth rate of the cumulative series depends on time-varying, or stochastic, trend.

CEPR PRESS

Covid Economics 51, 7 October 2020: 36-73

Other components, such as day of the week effects, may also be included.

2.1 Statistical modeling

The observational model uses data on the time series of the cumulative total of confirmed cases or deaths, Y_t, and the daily change, $y_t = \Delta Y_t = Y_t - Y_{t-1}$. HK show how the theory of generalized logistic (GL) growth curves, described briefly in Appendix A, suggests models of the form

$$\ln y_t = \rho \ln Y_{t-1} + \delta - \gamma t + \varepsilon_t, \quad \rho \geq 1,\ \gamma > 0, \qquad t = 2, ..., T, \tag{1}$$

where ε_t is a disturbance term. Subtracting $\ln Y_{t-1}$ from both sides gives

$$\ln g_t = (\rho - 1) \ln Y_{t-1} + \delta - \gamma t + \varepsilon_t, \tag{2}$$

where $g_t = y_t / Y_{t-1}$, although it may also be defined as $\Delta \ln Y_t$.

The observational model for the Gompertz curve, obtained by setting $\rho = 1$, is

$$\ln y_t = \ln Y_{t-1} + \delta - \gamma t + \varepsilon_t, \quad t = 2, ..., T, \tag{3}$$

or the simple time trend regression

$$\ln g_t = \delta - \gamma t + \varepsilon_t, \quad t = 2, ..., T. \tag{4}$$

Remark 1 *Viboud et al (2016) propose a growth curve for the ascending phase of an epidemic that implies an observational equation of the form (1) with* $\gamma = 0$ *and with* ρ *a deceleration parameter in the range* $0 \leq \rho \leq 1$*; see also Chowell et al (2016). When* $\rho = 1$ *there is exponential growth. The introduction of the time trend gives sub-exponential growth when* $\gamma > 0$.

Remark 2 *Some researchers, such as Levitt et al (2020), have exploited the relationship in Equation (4) to make forecasts. However they do not allow the trend to be stochastic. This is a fundamental part of our modelling approach.*

2.2 Dynamic trend models

Deterministic trends are too inflexible for most practical time series modeling. A stochastic, or time-varying, trend may be introduced into equation (4) to give the dynamic trend model

$$\ln g_t = \delta_t + \varepsilon_t, \quad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2), \qquad t = 2, ..., T, \tag{5}$$

Covid Economics 51, 7 October 2020: 36-73

where

$$\begin{array}{rcll} \delta_t & = & \delta_{t-1} - \gamma_{t-1} + \eta_t, & \eta_t \sim NID(0, \sigma_\eta^2), \\ \gamma_t & = & \gamma_{t-1} + \zeta_t, & \zeta_t \sim NID(0, \sigma_\zeta^2), \end{array} \tag{6}$$

and the normally distributed irregular, level and slope disturbances, ε_t, η_t and ζ_t respectively, are mutually independent. When σ_ζ^2 is positive, but $\sigma_\eta^2 = 0$ the trend is an integrated random walk (IRW). It is this form of the stochastic trend that turns out to be most useful for tracking an epidemic because it is the movements in γ_t which are crucial for that purpose. The key parameter is then the signal-noise ratio, $q = \sigma_\zeta^2/\sigma_\varepsilon^2$. A deterministic trend is obtained when q is zero.

Stochastic trend models can be estimated using techniques based on state space models and the Kalman filter (KF). Here the computations were performed using the STAMP package of Koopman et al. (2020). The Kalman filter outputs the estimates of the state vector $(\delta_t, \gamma_t)'$. Estimates of the state at time t conditional on information up to and including time t are denoted $(\delta_{t|t}, \gamma_{t|t})'$ and given by the contemporaneous filter; the predictive filter, which outputs $(\delta_{t+1|t}, \gamma_{t+1|t})'$, estimates the state at time $t+1$ from the same information set. It may sometimes be useful to review past movements by the smoother, which is the estimate of the state at time t based on all T observations in the series. Estimation of the unknown variance parameters is by maximum likelihood. Tests for normality and residual serial correlation are based on the standardized innovations, that is one-step ahead prediction errors, $v_t = y_t - \delta_{t|t-1}$, $t = 3, ..., T$.

Remark 3 *A stochastic trend can be introduced into the more general GL model. However, unless ρ is fixed, in which case g_t is replaced by y_t/Y_{t-1}^ρ, it may be hard to estimate in small samples. There is a strong case for concentrating on Equation (5) since it simply sets up a flexible model for the logarithm of the growth rate; if a more general GL model were more appropriate this model would adapt to it through changes in the slope.*

Remark 4 *The Kalman filter can be by-passed by adopting the reduced form, which comes from the innovations form (IF) of the Kalman filter; see HK.*

Remark 5 *When y_t is small, it may be better to specify its distribution, conditional on past values, as discrete. The usual choice is the negative binomial. The way in which a dynamic model may be constructed is set out in HK and in the Appendix where it is shown how nowcasts may be obtained. However,* *it should be borne in mind that estimates of R_t based on small numbers are likely to be unreliable and not suited to policy decisions.*

2.3 Forecasts

Recursions for making forecasts of future observations and constructing an estimate of the saturation level in the GL case are given in HK. For the dynamic Gompertz model it is convenient to write

$$\begin{aligned}\widehat{g}_{T+\ell|T} &= \exp(\delta_{T+\ell|T}), \quad \ell = 1, 2, .. &(7)\\ \widehat{\mu}_{T+\ell|T} &= \widehat{\mu}_{T+\ell-1|T}(1 + \widehat{g}_{T+\ell|T}), &(8)\end{aligned}$$

so that $\widehat{y}_{T+\ell|T} = \widehat{g}_{T+\ell|T}\widehat{\mu}_{T+\ell-1|T}$ and $\widehat{Y}_{T+\ell|T} = \widehat{\mu}_{T+\ell|T}$; the initial value is $\widehat{\mu}_{T|T} = Y_T$. The prediction of $\delta_{T+\ell}$ is simply $\delta_{T+\ell|T} = \delta_{T|T} - \gamma_{T|T}\ell$. Other components, such as daily effects, may be added to it and in more general cases the predictions are given by the KF predictive recursions. Combining equations (7) and (8) gives

$$\widehat{y}_{T+\ell+1|T} = Y_T \exp(\delta_{T+\ell|T}) \prod_{j=1}^{\ell}(1 + \exp \delta_{T+j|T}), \quad \ell = 1, 2, .. \tag{9}$$

and $\widehat{y}_{T+1|T} = Y_T \exp(\delta_{T+1|T})$.

3 Tracking R

Definitions of the instantaneous reproduction number, R_t, vary. In Germany the national figure used by the Robert Koch Institute[1] is the average of new cases in the past four days divided by the average of the preceding four days. More generally

$$\widehat{R}_{t,k,\tau} = \frac{\sum_{j=0}^{k-1} y_{t-j}}{\sum_{j=\tau}^{k+\tau-1} y_{t-j}} = \frac{\sum_{j=0}^{k-1} y_{t-j}}{\sum_{j=0}^{k-1} y_{t-\tau-j}}, \tag{10}$$

where the sum in the denominator starts at a lag of four and the sums in the numerator and denominator may overlap. The RKI estimator is a special case of this in which $\tau = k = 4$. The lag of τ reflects the generation interval,

[1] Erläuterung der Schätzung der zeitlich variierenden Reproduktionszahl R. Robert Koch-Institut, 15. Mai 2020.

Covid Economics 51, 7 October 2020: 36-73

Covid Economics 51, 7 October 2020: 36-73

which is number of days that must elapse before an infected person can transmit the disease; for COVID-19 a reasonable choice seems to be $\tau = 4$. The rationale for $\widehat{R}_{t,k,\tau}$ comes from Cori et al (2013) as outlined in sub-section 4.1: see equation (23).

A little algebraic manipulation shows that $\widehat{R}_{t,k,\tau} = 1 + \tau \widehat{g}_{yt}$, where $\widehat{g}_{yt}$ is an implicit estimator of the growth rate in y_t. It is given by

$$\widehat{g}_{yt} = \frac{1}{\tau} \sum_{j=0}^{\tau-1} \frac{\sum_{j=0}^{k-1} \Delta y_{t-j}}{\sum_{j=\tau}^{k+\tau-1} y_{t-j}}$$

where $\Delta y_t = y_t - y_{t-1}$. When $\widehat{g}_{yt}$ is small,

$$\widehat{R}_{t,k,\tau} = 1 + \tau \widehat{g}_{y,t} \simeq \exp(\tau \widehat{g}_{yt}). \tag{11}$$

Dynamic Gompertz models can monitor $g_{y,t}$. Writing the growth rate of the total in continuous time, differentiating its logarithm and re-arranging gives

$$g_y(t) = g(t) + g_g(t), \tag{12}$$

where $g(t)$ is the growth rate of $\mu(t)$ and $g_g(t)$ is the growth rate of the growth rate. The negative of the growth rate of the growth rate is tracked by the filtered estimates of the slope, that is $\gamma_{t|t}$, while the growth rate itself can be tracked by the exponent of the filtered level, that is $g_{t|t} = \exp \delta_{t|t}$. Thus

$$g_{y,t|t} = g_{t|t} - \gamma_{t|t}, \quad t = t', ..., T, \tag{13}$$

where t' is the time at which the estimates are deemed to be reasonably reliable. The same formula applies to the Negbin model.

The nowcast of R_t suggested by equation (11) with $k = \tau$ is

$$\widetilde{R}_{t,\tau} = 1 + \tau g_{y,t|t} \quad \text{or} \quad \widetilde{R}^e_{t,\tau} = \exp(\tau g_{y,t|t}). \tag{14}$$

The exponential form ensures that the estimator is always positive. The nowcasts of number of cases or deaths peaks when $g_{t|t} - \gamma_{t|t} = 0$ which gives $\widetilde{R}_{t,\tau} = \widetilde{R}^e_{t,\tau} = 1$. The RKI estimator implies $\tau = 4$ in equation (14).

Wallinga and Lipsitch (2006) give a general formula linking R_t to $g_{y,t}$ for a given serial interval distribution. Although stated for the initial, exponential phase of an epidemic it can also be used to estimate R_t from $g_{y,t|t}$. The expression is

$$R_t^M = 1/M(-g_{y,t|t}), \tag{15}$$

Covid Economics 51, 7 October 2020: 36-73

where $M(.)$ is the moment generating function (MGF) of the serial interval distribution, which is the time between the onset of symptoms in a primary case and the onset of symptoms of secondary cases. When the distribution is degenerate, so that all secondary infections occur with lead exactly equal to τ, the mean generation interval, $R_t^M = \exp(\tau g_{y,t|t})$ which is the same as $\widetilde{R}^e_{t,\tau}$. When the serial interval has a gamma distribution with parameters a and b, implying a mean of ab and a variance of ab^2, equation (15) is simply

$$R_t^M = (1 + b g_{y,t|t})^a. \tag{16}$$

Keeping the mean constant and letting $b \to 0$ confirms that $R_t^M = \exp(\tau g_y)$. At the same time, if $a = b = 2$, which is not inconsistent with some of the estimates of the mean and variance obtained for COVID-19,

$$R_t^M = (1 + 2g_{y,t|t})^2 = 1 + 4g_{y,t|t} + 4g^2_{y,t|t}.$$

and, so when $g_{y,t|t}$ is small, $R_t^M \simeq 1 + 4g_{y,t|t}$, just as for the RKI estimator, $\widehat{R}_{t,4,4}$. Finally we note that Wallinga and Lipsitch (2006, p 602) observe that $\exp(\tau g_{y,t|t})$ is an upper bound for R_t^M in equation (15). Overall it seems that if a single formula is to be adopted for calculating estimates of R_t from $g_{y,t|t}$ for COVID-19, $\widetilde{R}_{t,4}$ or $\widetilde{R}^e_{t,4}$ is not a bad choice.

Remark 6 *As an epidemic dies down, $g_{t|t}$ tends towards zero so if $\gamma_{t|t}$ is constant, the estimators of R_t will tend towards a positive lower bound; see also the growth curve analysis in Appendix B.*

3.1 Example: new cases in Germany

Figure 1 shows $\gamma_{t|t}$ and $g_{t|t}$ estimates for Germany from a dynamic Gompertz model, together with the daily number of new cases[2]. A similar graph can be found in HK but the revised data used here results in a peak in mid-March rather than early April. However, the message is the same: initially $g_{t|t}$ dominates but after the peak it starts to become small relative to $\gamma_{t|t}$.

Figure 2 shows estimates of R_t with $\tau = 4$. The daily effect is included in the time series model and the the signal-noise ratio, $q = 0.01$. As can be

[2]The data for new cases in Germany is from the Robert Koch Institut and are the confirmed cases of COVID-19 in all national hospitals and testing centres. https://www.rki.de/DE/Home/homepage_node.htm

Covid Economics 51, 7 October 2020: 36-73

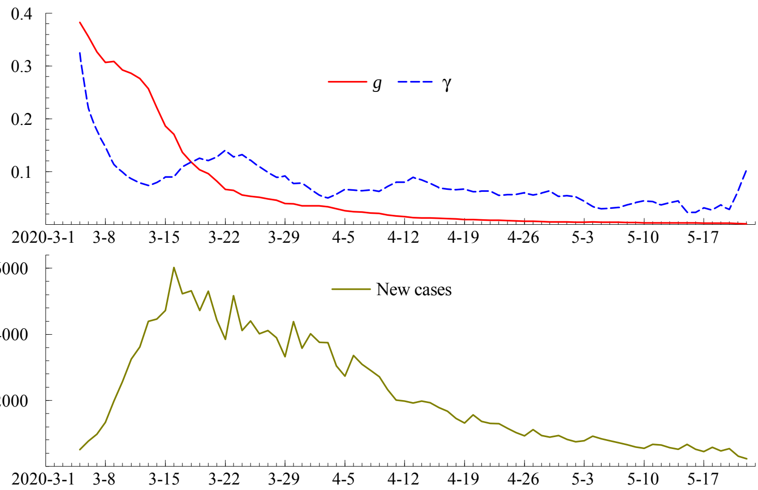


Figure 1: New cases in Germany from March 1st to May 22nd 2020 together with filtered growth rate and its rate of change from a daily model with $q = 0.01$.

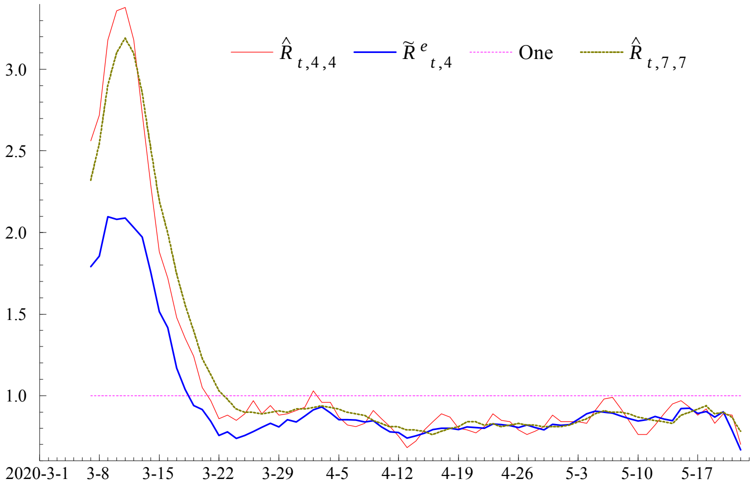


Figure 2: Estimates of R from 8th March computed from new cases in Germany, using data starting on March 1st : $\widetilde{R}^e_{t,4}$ is from a model with $q = 0.01$.

seen, $\widetilde{R}^e_{t,4}$ goes below one when $g_{t|t} < \gamma_{t|t}$. The movements in the RKI moving average estimates, $\widehat{R}_{t,4,4}$, are similar but the message is less clear because of the day of the week effect. Indeed $\widehat{R}_{t,4,4}$ periodically goes moves close to one for precisely this reason. The longer moving average $\widehat{R}_{t,7,7}$ is also included because it irons out the daily effects and so is more stable.[3] The last two observations on new cases are much smaller than those immediately before, so there is a significant drop in $\widehat{R}_{t,4,4}$ and $\widetilde{R}^e_{t,4}$ both of which lead $\widehat{R}_{t,7,7}$.

3.2 Sampling variability

When γ_t changes, some idea of the sampling variability in the estimators of $g_{y,t}$ and R_t is needed. When q, the signal-noise ratio in the Gaussian IRW model, is treated as known, the unknown state γ_t is normally distributed

[3]In the RKI program it is dated at time $t-1$ so it lags behind $\widehat{R}_{t,4,4}$ and uses a future observation (at t). The estimator $\widehat{R}_{t,7,4}$ could also be used.

Covid Economics 51, 7 October 2020: 36-73

Covid Economics 51, 7 October 2020: 36-73

with mean $\gamma_{t|t}$ and a variance $p_{t|t}$ that is produced by the KF. (Note that $p_{t|t}$ does not depend on the value of the estimated state, only on q and σ^2). The growth rate of the incidence curve, $g_{y,t}$, depends on g_t as well as γ_t but as argued below its contribution to the variability of $g_{y,t}$ is dominated by that of γ_t. When the variability in g_t is ignored, the probability that R_t exceeds one, that is $\Pr(\gamma_t < g_{t|t})$ where $g_{t|t}$ is fixed, can be obtained directly from the (posterior) distribution of γ_t. *This probability does not depend on which equation is chosen to estimate R_t from the estimate of $g_{y,t}$ and it does not depend on the choice of τ.*

When $\widetilde{R}_t$ is defined as $1 + \tau g_{y,t}$, its distribution, conditional on current and past observations, is normal with mean $1 + \tau\widehat{g}_{y,t}$ and standard deviation $\tau p_{t|t}^{1/2}$. If, on the other hand, $\widetilde{R}_t = \exp(\tau g_{y,t})$, its conditional distribution is lognormal with mean

$$E_t(\widetilde{R}^e_{t,\tau}) = \exp(\tau(g_{t|t} - \gamma_{t|t} + (\tau/2)p_{t|t})) \tag{17}$$

and standard deviation

$$SD_t(\widetilde{R}^e_{t,\tau}) = E_t(R_{t,\tau})\sqrt{(\exp\tau^2 p_{t|t} - 1)}. \tag{18}$$

Why is the variability in $g_{t|t}$ ignored? From equation (5), $g_t = \exp\delta_t$ and because δ_t is normal, g_t is lognormal with mean $\mu_{g,t|t} = \exp(\delta_{t|t} + 0.5p_{\delta,t|t})$ and variance $Var(g_t) = \mu^2_{g,t|t}(\exp p_{\delta,t|t} - 1)$, where $p_{\delta,t|t}$ is the variance of δ_t. However, $p_{\delta,t|t}$ is typically small so $\mu_{g,t|t} \simeq \exp\delta_{t|t} = g_{t|t}$ and $Var(g_t) \simeq \mu^2_{g,t|t}p_{\delta,t|t} \simeq g^2_{t|t}p_{\delta,t|t}$. Now

$$Var(g_{y,t}) = Var(g_t) + Var(\gamma_t) - 2Cov(g_t, \gamma_t)$$

Although $p_{\delta,t|t}$ is usually larger than $p_{\gamma,t|t}$ the former is multiplied by $g^2_{t|t}$; note that $p_{\delta,t|t}$ itself does not depend on the value of $g_{t|t}$. While $g_{t|t}$ can be high near the beginning of an epidemic, it tends to fall quite rapidly and once the epidemic is underway it rarely exceeds 0.05; see Figure 1. The example of Florida, where the second wave increases $g_{t|t}$, shows that, even in this case, $Var(g_t)$ remains negligible compared with $Var(\gamma_t)$.

3.3 Predictions of R

Predictions of R_t in the dynamic Gompertz model can be made from predictions of $g_{y,t}$, that is

$$\widetilde{g}_{y,T+\ell|T} = \exp\delta_{T+\ell|T} - \gamma_{T+\ell|T}\exp(\delta_{T|T} - \gamma_{T|T}\ell) - \gamma_{T|T}, \qquad \ell = 1, 2, . \tag{19}$$

Covid Economics 51, 7 October 2020: 36-73

As $T \to \infty$, $\widetilde{R}^{e}_{T+\ell|T} \to \exp(-\tau\gamma_T))$.

If, as in the previous sub-section, it is assumed that g_t is relatively small, the predictive distribution of $g_{y,T+\ell}$, and hence of $R_{T+\ell}$, is available: the conditional distribution of $\gamma_{T+\ell}$ given observations up to and including time T is Gaussian with mean $\gamma_{T|T}$ and variance $p_{T+\ell|T}$, which is produced by the predictive equations of the KF.

The basic forecasts are made with the estimates of δ_T and γ_T. However, alternative scenarios in which γ_t is assumed to evolve in a certain way, perhaps to reflect changing behaviour and policies, may also be envisaged. If a future scenario arises in terms of a time path for $R_{T+\ell|T}$, it can easily be translated into one for $\gamma_{T+\ell|T}$. The time path for $\gamma_{T+\ell|T}$ leads directly to the forecasting equations of (9) and so no simulations are not needed for the predictions of $y_{T+\ell}$.

Finally the ability to make predictions offers insight into how to deal with reporting delay, as described in Abbott et al (2020, p3-4). If the observation at time t actually relates to an event ℓ days earlier, the current R_t is better estimated by $R_{t+\ell|t}$. However, when $R_{T+\ell}$ is estimated from $\gamma_{T+\ell|T}$ the lag makes no difference when $\gamma_{T+\ell|T} = \gamma_{T|T}$.

3.4 Weights

The filtered estimates of g_t and γ_t in the dynamic Gompertz model, equation (5), are obtained by discounting past observations, with the rate of discounting depending on the signal-noise ratio, q. Weights implied by the Kalman filter and smoother for estimated states in a linear model can be obtained as output from the STAMP package, using a method described in Koopman and Harvey (2003). The forcing variable in the filter is $\ln y_t$ and the weights assigned to it in the contemporaneous filter are the weights for $-\gamma_{t|t}$ plus the weights for $g_{t|t}$. If the weights for the slope, $-\gamma_{t|t}$, are denoted w_j, $j = 0, 1, 2, ...$, and the weights for $g_{t|t}$ are v_j, $j = 0, 1, 2..$, with $\sum v_j = 1$, equation (13) gives

$$g_{y,t|t} = g_{t|t} - \gamma_{t|t} = \prod_{j=0}^{t} (y_{t-j}/Y_{t-j-1})^{v_j} + \sum_{j=0}^{t} \ln(y_{t-j}/Y_{t-j-1})^{w_j}, \qquad (20)$$

When Y_t is much larger than y_t, as will be the case when an epidemic has been underway for some time and new cases are falling, $g_{t|t}$ will be relatively

Covid Economics 51, 7 October 2020: 36-73

small and attention can be focussed on $\gamma_{t|t}$. Then

$$g_{y,t|t} \simeq -\gamma_{t|t} = \sum_{j=0}^{t} w_j(\ln y_{t-j} - \ln Y_{t-j-1}) \simeq \sum_{j=0}^{t} w_j \ln y_{t-j} \tag{21}$$

where the last approximation follows because $\ln Y_{t-j-1}$ is assumed to be changing very slowly and $\sum_{j=0}^{t} w_j = 0$.

When multiplied by τ, the weights in equation (21) feed directly into the estimators of R_t implied by equation (14). It is interesting to compare the weights for $\ln \widetilde{R}^{e}_{t,\tau}$ given by the exponential estimator, that is $\tau g_{y,t|t}$ in (21), with the weighting structure for $\widehat{R}_{t,k,\tau}$ in equation (11) where

$$\ln \widehat{R}_{t,k,\tau} = \ln \sum_{j=0}^{k-1} y_{t-j} - \ln \sum_{j=\tau}^{k+\tau-1} y_{t-j}; \tag{22}$$

there is no overlap when $k = \tau$. The contrast is between the logarithms of weighted sums of past values and a weighted sum of past logarithms. Only for $k = \tau = 1$, are the weights in (22) the same as those in (21), obtained by setting $w_0 = 1$ and $w_1 = -1$, but this estimator is just the first difference of $\ln y_t$ which is usually too noisy to be of any practical value.

Figure 3 shows the weights for the slope produced when $q = 0.0015$, a number reported in Table 2 of HK for Germany. The initial six positive weights are declining. They are contrasted with the previous 21 weights which are smaller and negative. The sum of the weights over all past observations is zero and the negative weights provide a stable base for the contrast. Sensitivity to recent change is therefore combined with stability. When a daily effect is included in the model, the weights are adjusted accordingly but the impications for the slope remain much the same.

Increasing the signal-noise ratio to $q = 0.01$ puts more weight on the most recent observations and so gives the faster response shown in Figure 4. However this comes at the price of a less stable measure of R_t. It could be argued that an increase in q is appropriate when there is a sharp change in the environment, perhaps due to a change in policy. However, if the level of daily infections rises, the variance will tend to rise as well when the conditional distribution of y_t is lognormal or negative binomial. Thus keeping q the same may not be unreasonable.

Covid Economics 51, 7 October 2020: 36-73

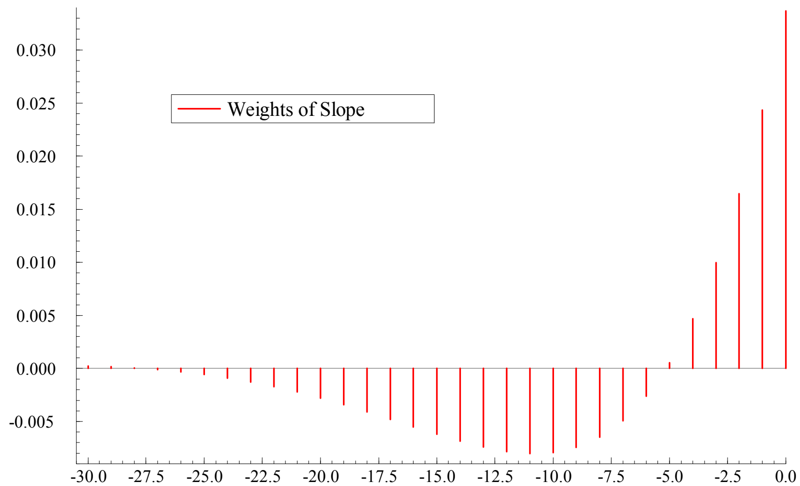


Figure 3: Weights for slope with $q = 0.0015$ (no daily component)

Covid Economics 51, 7 October 2020: 36-73

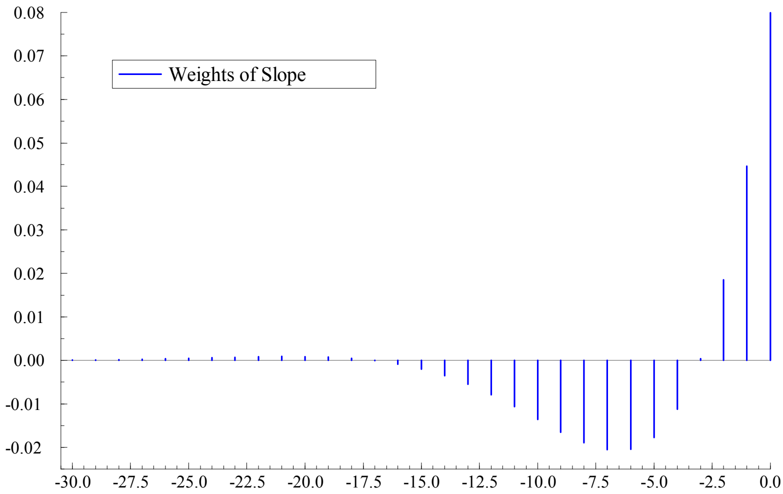


Figure 4: Weights for slope with $q = 0.01$

Covid Economics 51, 7 October 2020: 36-73

4 Current methods for estimating R

In epidemiology, models for estimating R_t have tended to use Bayesian methods to deal with updating of estimates. These methods contrast with updating based on a state space time series model.

4.1 Cori et al (2013)

Following Cori et al (2013), Thompson et al (2019) generate an estimate of the current level of new cases, y_t, that combines the estimate of R_t with an estimate, $\Lambda_{t|t-1}$, of the previous level based on the sum of new cases in the previous time period weighted by the infectivity function, or infectious profile after infection, $f_j, j = 0, 1, 2...$ This estimate can be written $\Lambda_{t|t-1} = \sum_{j=1}^{t} f_j y_{t-j}$, where $\sum_{j=\tau}^{t} f_j = 1$ with f_j describing the serial distribution. This is based on prior knowledge, such as data collected from household studies in the early phase of an infection. The estimate of R_t is obtained by Bayesian methods. Cori et al (2019, Appendix 1) assume a Poisson distribution for y_t and a (conjugate) gamma prior for R_{t-1} with parameters a and b, giving a mean of ab and a variance of ab^2. The posterior mean of R_t - its nowcast - is then

$$\widehat{R}_{t,k}^{*} = \frac{a + \sum_{j=0}^{k-1} y_{t-j}}{b^{-1} + \sum_{j=1}^{t-1} \Lambda_{t-j|t-j-1}} = \frac{a + \sum_{j=0}^{k-1} y_{t-j}}{b^{-1} + \sum_{j=1}^{t-1} \sum_{i=1}^{j} f_j y_{t-i}} = \frac{a + \sum_{j=0}^{k-1} y_{t-j}}{b^{-1} + \sum_{j=1}^{t} w_j y_{t-j}} \tag{23}$$

while the associated nowcast for the mean of new cases is

$$\widehat{\mu}_{t|t} = \widehat{R}_{t|t,k}^{*} \Lambda_{t|t-1} = \frac{a + \sum_{j=0}^{k-1} y_{t-j}}{b^{-1} + \sum_{j=1}^{t} w_j y_{t-j}} \sum_{j=1}^{t} f_j y_{t-j}.$$

Remark 7 *Thompson et al (2019) propose $a = 1$ and $b = 5$ at the outset so that both mean and standard deviation are set to five. With no prior information $a = 1/b = 0$.*

The numerator in $\widehat{R}_{t,k}^{*}$ provides an estimate of the level at time t based on the last k observations. The value of k reflects a trade off between response and stability. The choice of equal weights seems to be arbitrary. The weights in the second term reflect the structure implied by the sometimes imperfect knowledge of the distribution of serial intervals.

Covid Economics 51, 7 October 2020: 36-73

Cori et al (2013, Appendix 2) suggest letting the summation in the denominator of equation (23) start at $j = \tau$, where τ is the generation interval or threshold value for transmission. Approximating the weights by a simple moving average of length k and assuming no prior information then gives equation (10), which is the formula for the instantaneous reproduction number used by RKI. The value of k may be set equal to the length of the serial interval.

In our time series approach, the lag structure depends solely on the observations, y_t, and the properties of the fitted model. The weights in Figure 3 are comparable with those used to construct $\widehat{R}^*_{t,k}$ in equation (23), except that, as is apparent from equation (14), they are multiplicative in the implied estimator of R_t. The negative weights in Figure 3 correspond to the weights in the denominator of $\widehat{R}^*_{t,k}$.

4.2 Bettencourt and Ribeiro (2008)

Bettencourt and Ribeiro (2008) propose an autoregressive model in which

$$E_{t-\tau}y_t = \phi(R_t)y_{t-\tau} \tag{24}$$

where $\phi(R_t) \simeq \exp((\tau/k)(R_t - 1))$ where k is the infectious period or 'average residence time'. A more elaborate version allows for infections from non-human sources. They use Bayesian methods to update the estimates of R_t as new observations on daily cases arrive. The model is based on assumptions about the mechanism by which an epidemic spreads and the values of key parameters. The method has been adopted somewhat uncritically by many agencies, particularly in the United States, for tracking COVID-19. However, the recent article by Gostic et al (2020) cautions against the use of this approach and other methods discussed by Bettencourt and Ribeiro (2008) because of their reliance on underlying structural assumptions that may not be met.

Bettencourt and Ribeiro (2008) assume that y_t is Poisson distributed but note the possibility of extension to negative binomial. Some of the examples they consider involve small numbers but others do not; Figure 1 in their paper shows both. The example of H3N2 flu in the US has numbers in the hundreds so the validity of a Poisson distribution, where mean is equal to variance, is questionable. In the Spanish flu example of the next section it was shown that the Poisson distribution is easily rejected. For COVID-19 where

Covid Economics 51, 7 October 2020: 36-73

the daily numbers can be very large the adoption of a Poisson distribution could give poor estimates an incredible Bayesian 'credible intervals'. In the Gaussian model the distribution is subject to diagnostic checking and so the implied CIs for R_t are constructed on firmer ground.

5 The early phase of an epidemic and the role of prior information

5.1 Prior information

Any modelling is very difficult at start of an epidemic because of the lack of data; see, for example, the remarks in Appendix 3 of Cori et al (2013). Thus any prior information is potentially valuable. In the dynamic trend model for the logarithm of the growth rate, starting the Kalman filter in the absence of prior information is generally done with a non-informative (diffuse) prior on the level and slope, as in the STAMP package, or they are estimated as unknown parameters. An informative prior requires a mean and variance; the bigger the variance, the smaller the weight put on the initial mean. When elements of the state vector in a state space model are stochastic, assigning a prior distribution to them does not amount to a fully fledged Bayesian treatment because this would require a prior distribution for parameters, such as the signal-noise ratio q, that we regard as fixed. In our treatment q is initially given a fixed value, chosen according to the criteria outlined in the sub-section on weights, but it may later be estimated by ML when this becomes viable.

5.2 Exponential growth

In the early part of an epidemic the growth is exponential or very close to it. Thus, following the discussion in sub-section 2.1, we could set γ equal to zero or a very small positive number; see also Appendix B. The filter for the level of $\ln g_t$ can be given a prior distribution informed by information about the basic reproduction number, R_0. When $g_g(t) = \gamma = 0$, it follows from equation (12) that $g_y(t) = g(t)$. Thus given R_0 and τ, a rough estimate of $\ln g_0$ is given from $\widehat{g}_0 = (1/\widehat{\tau}) \ln \widehat{R}_0$ or $\widehat{g}_0 = (\widehat{R}_0 - 1)/\widehat{\tau}$. Choosing a suitable variance is somewhat more problematic. For a negative binomial distribution, initial values of $\ln g_0$ and γ can be constructed in a similar way.

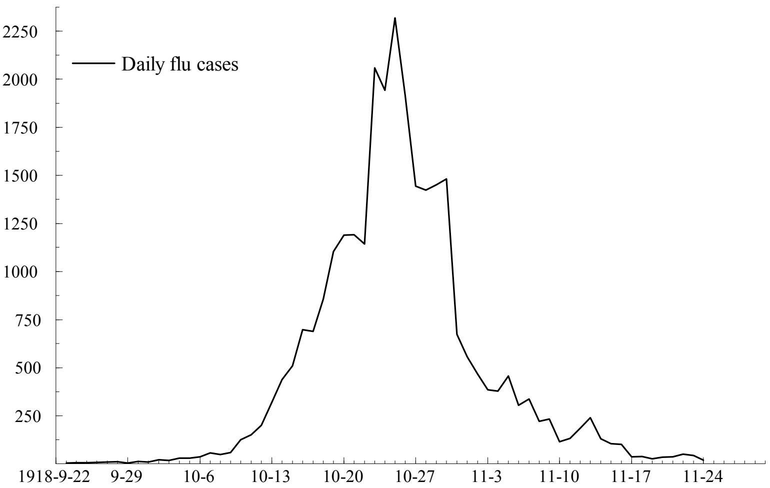


Figure 5: Daily flu cases in San Francisco in 1918

5.3 Spanish flu

Figure 5 shows daily cases of flu[4] in San Francisco during the worldwide outbreak of Spanish flu in 1918. For the upward phase Chowell et al (2007) assumed the growth to be exponential and suggested estimating the growth rate by regressing $\ln Y_t$ on a time trend. They proposed estimating R in this initial phase with the formula

$$R_\tau = 1 + \tau g + f(1-f)(\tau g)^2 + 0[(\tau g)^3]$$

where τ is the mean serial interval and f is the ratio of the mean infectious period to the mean serial interval. For Spanish flu, τ is 3-6 days. If τg is small, $R_\tau \simeq 1 + \tau g \simeq \exp(\tau g)$; so there is a connection with the formulae in equation (14) because with exponential growth $g_{y,t|t} = g_{t|t} = g$. Indeed Wallinga and Lipsitch (2006, p 599) give the formula $R_\tau = 1 + \tau g$.

[4]The data are supplementary material to the article by Chowell (2007) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2358966/

Covid Economics 51, 7 October 2020: 36-73

Covid Economics 51, 7 October 2020: 36-73

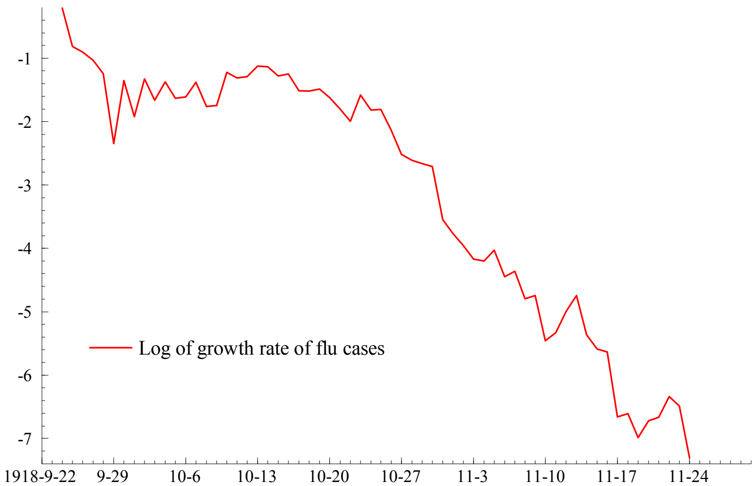


Figure 6: Logarithm of growth rate of flu cases in San Francisco in 1918.

The regression of $\ln Y_t$ on a time trend is over a 'best' exponential phase, which is determined by goodness of fit tests. For the San Francisco data it is 17 days. An alternative estimate could be obtained as a simple average of $\ln g_t$'s over the same period. Diagnostics show the fit is better with residuals that are approximately random, in contrast to those from the regression which, judging from the high residual serial correlation, are not. If $\ln g$ is assumed to be normally distributed, the exponential of its sample mean, $\widehat{g}$, is lognormal and so is $\widehat{R}_\tau = 1 + \tau\widehat{g}$. Using observations from September 30th to October 16th inclusive, the mean and variance of $\ln g_t$ are estimated to be -1.434 and 0.057 respectively, so $E\widetilde{R}_\tau = 1 + 4\exp(-1.434 + 0.056) = 2.01$ and $SD(\widetilde{R}_\tau) = 0.23$.

Once beyond the initial exponential growth phase R_t is no longer constant. Chowell et al (2007) discuss two approaches to estimating R_t based on SEIR models, the more complex one having eight nonlinear differential equations. They also use the Bayesian method of Bettencourt and Ribeiro (2008).

Fitting the Gaussian dynamic Gompertz to the whole series gives $q = 0.049$. The slope in the logarithm of the growth rate adapts so it is close to

Covid Economics 51, 7 October 2020: 36-73

zero in early October and then falls so as to capture the downward phase; see Figure 6. The fit is quite good as shown in diagnostics and Figure 7. The residuals become slightly bigger towards the end as numbers decrease but the heteroscedasticity statistic is not unduly large. The corresponding estimates of R_t with $\tau = 4$ are in Figure 8. The estimates of g_t and γ_t take time to settle down and $g_{y,t|t}$ peaks on October 14. If $\gamma_{t|t}$ were set to zero for this early period, the estimates of R_t would be similar to those obtained from R_τ in Method 1 of Chowell et al (2007) and from the initial estimates reported above. The estimates of R_t after mid-October may be compared with those in Figure 6 of Chowell et al (2007) where there is more smoothing so $R_t \simeq 1$ towards the end. The SD of $\gamma_{t|t}$ is 0.106 so when $g_{y,t|t} = 0$, it follows from (18) that $SD(R^e_{t,4}) = 0.55$ but if $E_t(R^e_{t,4}) = 1$ it is 0.44. The SD of the estimate of R_t given by $1 + 4\gamma_{t|t}$ is somewhat smaller at 0.42. The credible interval in Chowell et al (2007) is very small which is partly because there is more smoothing but it may also be a reflection of imposing the Poisson distribution[5]. A smaller q in the time series model would reduce the SD. (But when the sample was broken into two parts, with the second beginning in the downward phase, the estimate of q was very similar.)

In summary the time series model does remarkably well, adapting to the downward slope in the logarithm of the growth rate after the initial exponential growth. The estimates of R_t are not out of line with those reported by Chowell et all (2007) for methods which can be quite complex. Finally the fact that the classical approach allows diagnostic checks on the statistical assumptions raises issues about some of the assumptions made by other methods.

6 Waves and spikes

After an epidemic has peaked, daily cases start to fall and the concern shifts to the possibility of a second wave and the need to deal with outbreaks indicated by spikes in the data so that they do not morph into waves. The monitoring of waves and spikes raises different issues, primarily because a wave applies to a whole nation or a relatively large geographical unit, whereas

[5]If the predicted level of y_t is taken to be 1000, the implied SD of the predictive lognormal distribution is 92 whereas that of the Poisson distribution is 32. For a level of 100, the implied SDs are very similar: for the fitted lognormal model it is 9.5 as opposed to 10 for the Poisson.

Covid Economics 51, 7 October 2020: 36-73

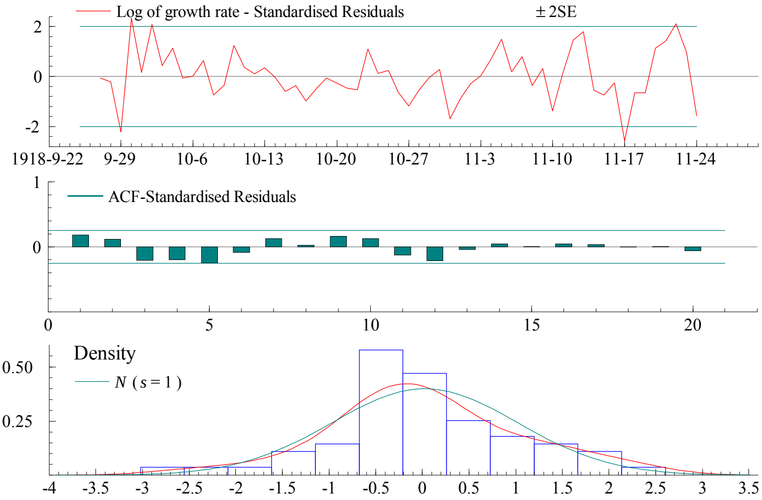


Figure 7: Diagnostics for dynamic Gompertz model fitted to flu cases in San Francisco in 1918.

Covid Economics 51, 7 October 2020: 36-73

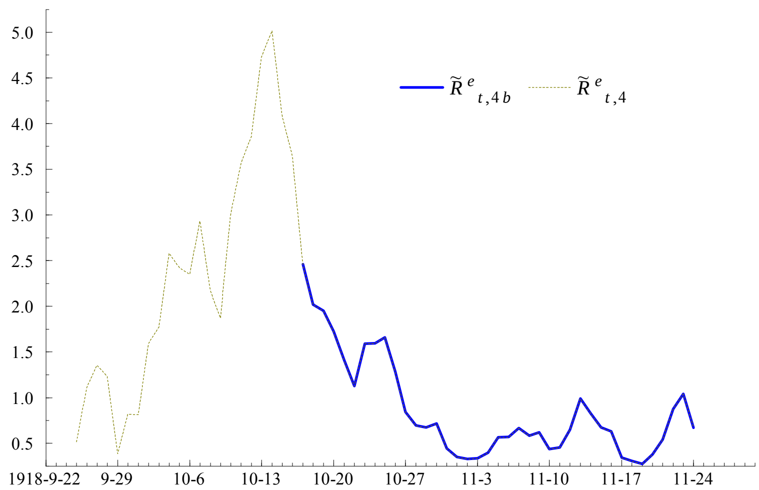


Figure 8: R_4 for San Francisco flu from time series model. Bold line (R_{4b}) is after October 15th.

Covid Economics 51, 7 October 2020: 36-73

a spike is localized.

6.1 Spikes

As the epidemic subsides, the daily numbers of infected start to become small and so, as at the beginning, the negative binomial distribution may be more appropriate. At the same time attention switches to localized outbreaks. The value of R_t in a small area where there is an outbreak typically rises very quickly because it is based on an increase in the number of new cases when the current number is small. This may be compounded by the fact that when an outbreak occurs there is an associated increase in testing. The importance of such an increase in a local R_t needs to be put in a wider context by considering it together with the number of cases per 100,000 people.

When national numbers are low, a localized outbreak can also result in a jump in the national estimate of R_t. However, such a jump does not indicate that there has been a sudden change in the way the infection spreads and so has few implications for overall policy. Figures for new cases in Germany show a suddent increase towards the end of June, caused by an outbreak at a meat processing factory in Westphalia. Estimates produced by the RKI at the time showed a big increase in R_t, accompanied by what seems to us to be a rather narrow credible interval. Figure 9 compares the model-based reproduction number estimate, $\widetilde{R}^e_{t,4}$, with the four day and overlapping seven day moving average estimates, $\widehat{R}_{t,4}$ and $\widehat{R}_{t,7,4}$. The $\widehat{R}_{t,4}$ estimates are very erratic and seriously affected by the failure to take account of the daily pattern. Estimates for Sundays and Mondays are typically lower. The peak in $\widehat{R}_{t,4}$ has Wed-Sat in the numerator. Although $\widehat{R}_{t,7,4}$ irons out some of the daily movement, the estimate of R_t is still affected. The model-based $\widetilde{R}^e_{t,4}$ evolves more smoothly.

After June the data gives no indication of a sustained increase in new cases. The jump in estimates of R_t, particularly $\widehat{R}_{t,4}$, can safely be classed as a spike.

The model was fitted using data from March 25 to June 26. The fit was good with very little evidence of residual serial correlation; the $Q(15)$ statistic is 9.58. A Gaussian distribution seems a good approximation as the Bowman-Shenton test statistic, which is asymptotically distributed as χ^2_2 under the null hypothesis, is only 0.77. Figure 10 provides graphical confirmation. The estimate of q was 0.0026.

Covid Economics 51, 7 October 2020: 36-73

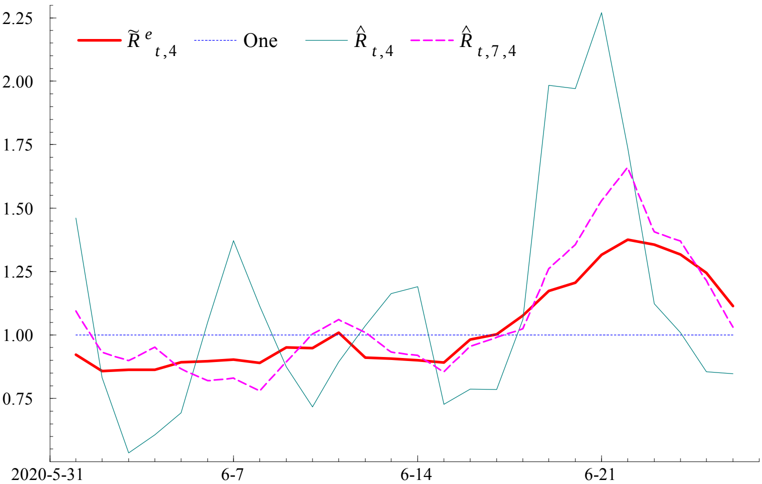


Figure 9: R_t for German new cases

The SD of the posterior distribution of γ_t is 0.0276. If $R_t = 1 + 4\gamma_t$ its SD is 0.110. When $R_t = \exp(4g_{y,t})$ setting $E_t(R^e_{t,4}) = 1$ gives $SD_t(R^e_{t,4}) =$ 0.111 and a 68% credible interval of $[0.895\,, 1.11\,7]$. It makes little difference whether R_t is taken to be normal or lognormal. As regards the contribution of g_t to the variability $g_{y,t}$, the June 26th value of $g_{T|T}$ was only 0.0030 and $SD(g_T)$ was less than one per cent of the SD of γ_t.

Covid Economics 51, 7 October 2020: 36-73

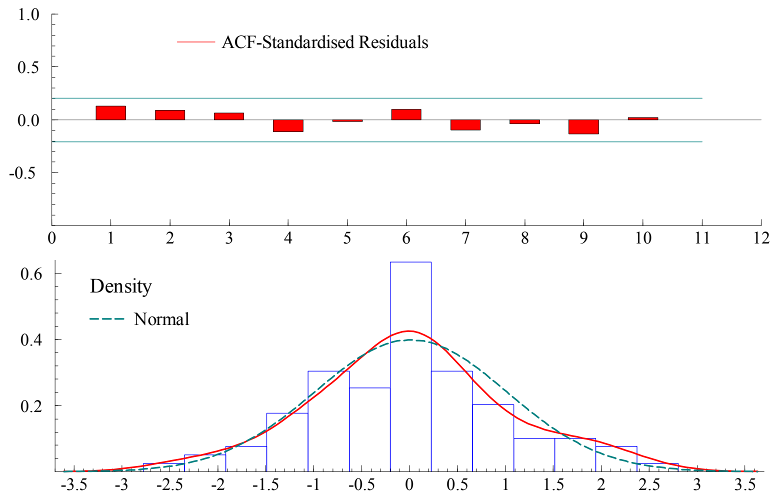


Figure 10: Residual correlogram and histogram of residuals from model for Germany

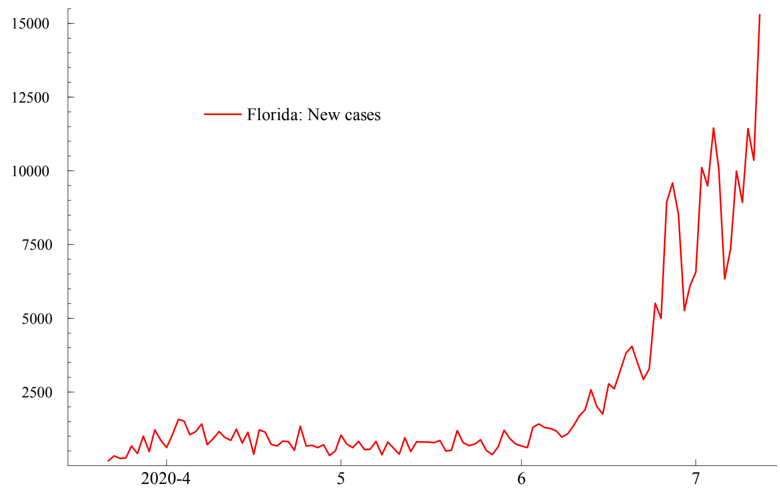


Figure 11: Daily new cases in Florida from March 22nd to July 12th.

6.2 Waves

The US state of Florida, the third most populous in US with a population of around 20 million, provides an example of a second wave. Figure 11 shows daily new cases[6] from early March until July 19th. There is a peak in early April followed by a steady decline. This is similar to the pattern shown in Figure 1 for Germany and reflects the fact that Florida, like Germany, was in lockdown during April. After April restrictions in Florida were eased. There was a leveling out in May, followed by a sharp rise in June.

Figure 12 shows the logarithm of the growth rate of the number of confirmed cases, deaths and fraction of positives, starting March 22nd. (Before March 22nd the data are very erratic.) After May there was an increase in testing. However, the growth rate in tests is roughly constant from the end of May onwards and this shows up in Figure 12 where the logarithm of the growth in the proportion of positives follows a similar path to that of the

[6]Data on Florida are sourced from:
https://covidtracking.com/data

Covid Economics 51, 7 October 2020: 36-73

Covid Economics 51, 7 October 2020: 36-73

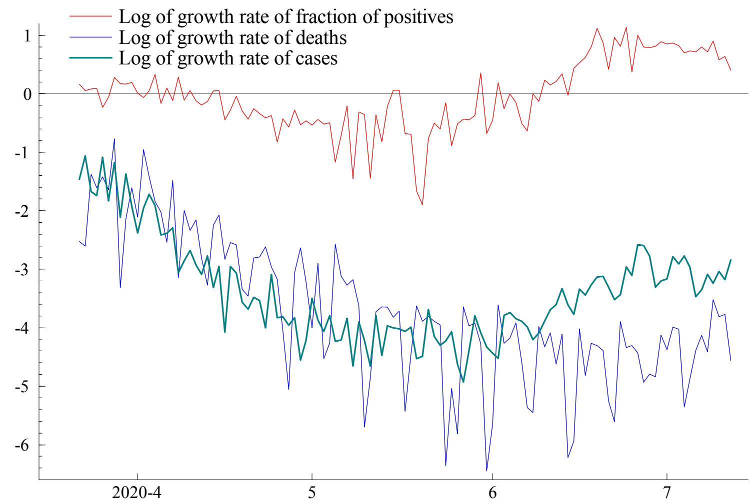


Figure 12: Logarithm of the growth rate of the total number of confirmed cases and deaths in Florida, together with the logarithm of the growth rate of the fraction of positives out of total tested

logarithm of the growth in total cases[7]. This suggests that confirmed cases are still a good indicator of the path of new infections and hence of $R_{t.}$

Figure 12 also shows the logarithm of the growth of deaths. This series seems to follow a similar path to that of confirmed cases but with a lag of two to three weeks. More recent data confirms that this is indeed the case.

Fitting the dynamic Gompertz model, with a daily component, to data on confirmed cases from March 22nd to July 12th gave residuals with very little residual serial correlation as the $Q(16)$ statistic was only 8.42. The Bowman-Shenton test statistic, which is asymptotically distributed as χ^2_2 under the null hypothesis, was only 0.11 so a Gaussian distribution cannot be rejected. The signal-noise estimate, q, was 0.0014. The graph of the filtered estimates of γ_t and g_t is shown in Figure 13 and the resulting nowcasts of R_t, computed as $\widetilde{R}^e_{t,4} = \exp(4g_{y,t|t})$, are shown in Figure 14. At the beginning of June, $\gamma_{t|t}$

[7]Growth rate for proportion of positives = Growth rate for confirmed cases - growth rate for total tested.

Covid Economics 51, 7 October 2020: 36-73

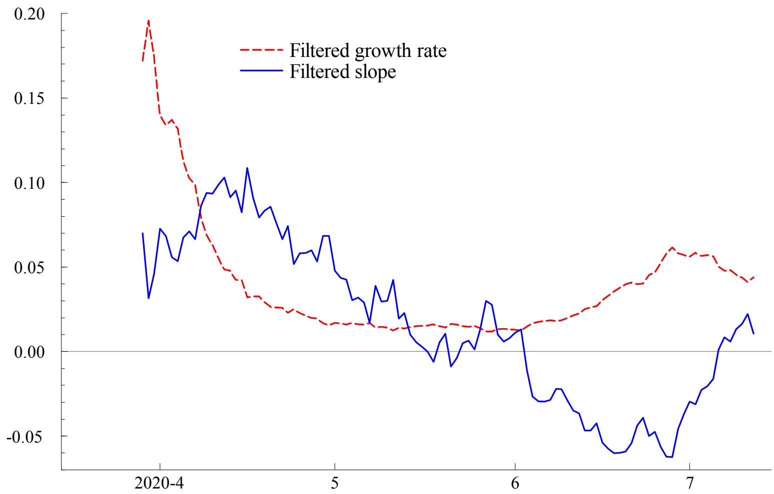


Figure 13: Filtered estimates of the growth rate (g_t) and slope (γ_t) for confirmed cases in Florida.

falls sharply and there is an attendant rise in $g_{t|t}$. The fall in $\gamma_{t|t}$ continues until the end of June when it changes direction and $g_{t|t}$ peaks[8]. Thereafter $\gamma_{t|t}$ and $g_{t|t}$ start to move closer together. These movements are reflected in $\widetilde{R}^e_{t,4}$ which rises to 1.5 at the end of June and then falls so that at the end of the sample $\widetilde{R}^e_{T,4} \simeq 1.1$.

The estimates of γ_t and g_t in Figure 13 are very different from those for Germany in that the estimated $g_{t|t}$ no longer becomes negligible with time. Indeed from the start of June it is greater than $\gamma_{t|t}$, as evidenced by the fact that R_t is bigger than one. For example on 12th July $g_{t|t}$ is 0.044 and $\gamma_{t|t}$ is 0.011. Nevertheless its contribution to the variability of g_{yt} is still negligible. The SD of the posterior distribution of γ_t is 0.0275 while that of δ_t is 0.1296 translating into a SD of 0.0057 for g_t. If the covariance term is ignored, the SD of g_{yt} is 0.0281, only a little above the SD of γ_t.

Re-estimating the model with another week of data has the estimate of

[8]A sequence of negative prediction errors dominate the effect of $\gamma_{t|t}$, which is still negative, and consequently $g_{t|t}$ falls.

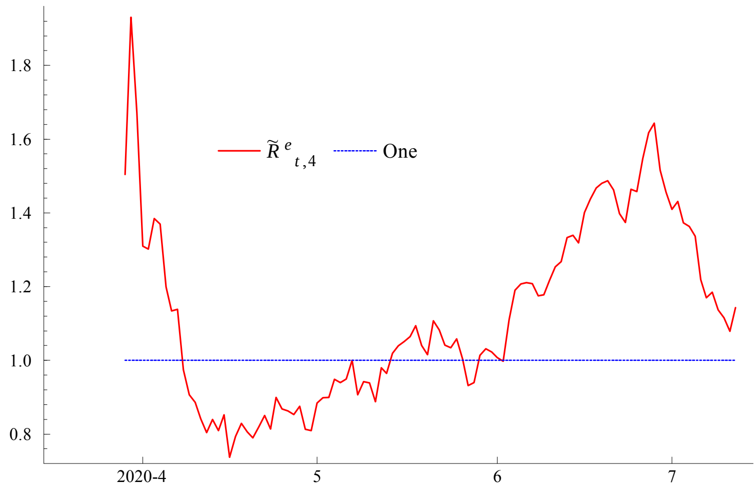


Figure 14: Time series model estimates, $\widetilde{R}_{t,4}$, from new cases in Florida

Covid Economics 51, 7 October 2020: 36-73

the last γ_t, that is γ_T, up to 0.031 but the SD is little changed at 0.027. The estimate of g_T is 0.034 giving an $\widetilde{R}^e_{T,4}$ of 1.01. Finally estimating using data up to August 12th leaves the estimate of γ_T virtually unchanged but, because the estimate of g_T is smaller at 0.011, $\widetilde{R}^e_{T,4}$ is below one, with a value of 0.91.

7 Conclusion

A new class of time series models are able to track the progress of an epidemic by providing nowcasts and forecasts of the daily number of new cases and deaths. Estimates of the instantaneous reproduction number R_t can be computed as a by-product, using a formula that links it to the estimated growth rate of new cases or deaths, based on assumptions made about the serial interval distribution. Credible (confidence) intervals for estimates of the growth rate of new cases and of R_t can be constructed.

Current methods for tracking R_t do not pay due attention to the time series properties of the data, whereas our method is based on time series techniques that have been shown to be effective in a range of disciplines, particularly econometrics. The dynamic response depends on a signal-noise ratio that is estimated from the data rather than being inferred from knowledge about the serial interval of infections. An important element in time series methodology is diagnostic checking and the fit of the model. We show how diagnostic methods can be applied in the context of epidemics and in doing so we raise questions about some of the assumptions, explicit or implicit, that are currently made in the estimation of R_t. The way in which the proposed model performs in tracking spikes and waves for COVID-19 is illustrated with examples using data from Germany and Florida.

We stress again that computing R_t is a by-product of our approach. Estimating R_t plays no part in nowcasting or forecasting daily cases and deaths. If R_t is required we can not only estimate it but can also make forecast it. Information on R_0 could be used at the start of an epidemic, but with a dynamic time series model its impact soon wears off. This leads us to question whether R_t is really needed at all. Although R_t has an appealing interpretation, the accuracy of its estimates can never be checked as its true value is never known. When $\widehat{g}_{yt}$, the estimate of the growth rate of new cases, is

Covid Economics 51, 7 October 2020: 36-73

CEPR PRESS

positive, it can be expressed in a way that is straightforward to interpret simply by presenting the doubling time, $\ln 2/\widehat{g}_{yt}$. So why bother with R_t ?

There is one outstanding issue to which there is no easy answer. In the early days of the coronavirus pandemic, the data on new cases was based for hospital admissions in many countries. This is no longer the case now, with increased testing of the population, including the testing that is part of track and trace. We acknowledge the sensitivity of estimates to changes in the count of cases which is solely due to changes in the rate of testing and detection. It remains a general limitation that the unbiasedness of estimates over any time window depends on the assumption that the rates of testing and detection are constant. When this is not the case, computing coherent estimates of R_t over time can be problematic; see Abbott et al (2020, pp.4-5).

Acknowledgements

We would like to thank Mark Salmon and Qingyuan Zhao for helpful comments and suggestions. Jonas Knecht supplied valuable research assistance.

8 References

Abbott S, Hellewell J, Thompson RN et al. (2020). *Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts* [version 1; peer review: awaiting peer review]. Wellcome Open Research, 5:112 (https://doi.org/10.12688/wellcomeopenres.16006.1)

Bettencourt LM and R.M. Ribeiro (2008). *Real time Bayesian estimation of the epidemic potential of emerging infectious diseases.* PLoS One 3, e2185. doi:10.1371/journal.pone.0002185

Birrell, P., Blake, J., van Leeuwen, E., PHE joint Modelling Cell, Gent, N. and D. De Angelis (2020). *Real-time Nowcasting and Forecasting of COVID-19 Dynamics in England: the first wave.* Mimeo, MRC Cambridge.

Chowell, G., Nishiura, H. and L. M. A. Bettencourt (2007) *Comparative estimation of the reproduction number for pandemic influenza from daily case notification data.* J. Royal Society Interfaces, 4(12), 155-166. https://doi.org/10.1098/rsif.2006.0161

Chowell, G., Sattenspiel, L., Bansal, S. and C. Viboud. (2016) Mathematical models to characterize early epidemic growth: a review. Phys Life Rev, 18, 66-97.

Covid Economics 51, 7 October 2020: 36-73

Covid Economics 51, 7 October 2020: 36-73

Cori, A., Ferguson, N.M., Fraser, C. and S. Cauchemez. (2013). *A new framework and software to estimate time-varying reproduction numbers during epidemics.* Am. J. Epidemiol. 178, 1505–1512.

Durbin, J. and S. J. Koopman (2012). *Time Series Analysis by State Space Methods.* Oxford University Press, Oxford

Gostic, K., McGough, L., Baskerville, E., Abbott, S., Joshi, K., Tedijanto, C., Kahn, R., Niehus, R., Hay, J., De Salazar, P., Hellewell, J., Meakin, S., Munday, J., Bosse, N., Sherratt, K., Thompson, R., White, L., Huisman, J., Scire, J., Bonhoeffer, S., Stadler, T., Wallinga, J., Funk, S., Lipsitch, M. and S. Cobey. (2020). *Practical considerations for measuring the effective reproductive number, R_t.* medRxiv preprint. Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/2020.06.18.20134858

Harvey A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge: Cambridge University Press.

Harvey, A. and Kattuman, P. (2020). *Time Series Models Based on Growth Curves with Applications to Forecasting Coronavirus.* Harvard Data Science Review. Special issue 1 - COVID -19.

https://hdsr.mitpress.mit.edu/pub/ozgjx0yn

Koopman, S.J. , R. Lit, and A.C. Harvey (2020). *STAMP 8.4 Structural Time Series Analyser, Modeller and Predictor .* London: Timberlake Consultants Ltd.

Koopman, S.J. and A.C. Harvey (2003). *Computing observation weights for signal extraction and filtering.* Journal of Economic Dynamics and Control, 27, 1317-33.

Levitt, M., Scaiewicz, A. and F. Zonta (2020) *Predicting the Trajectory of Any COVID-19 Epidemic From the Best Straight Line.* medRxiv preprint doi: https://doi.org/10.1101/2020.06.26.20140814

Lit, R., Koopman, S.J. and A.C. Harvey (2020). *Time Series Lab - Score Edition.* https://timeserieslab.com

Thompson, R.N. , J.E. Stockwind, R.D. van Gaalene, J.A. Polonskyf, Z.N. Kamvarg, P.A. Demarsh, E. Dahlqwist, S. Li, E. Miguel, T. Jombartg, J. Lessler and S. Cauchemez (2019). *Improved inference of time-varying reproduction numbers during infectious disease outbreaks.* Epidemics

Viboud, C., Simonsen, L. and Chowell, G. (2016) *A generalized growth model to characterize the early ascending phase of infectious desease outbreaks.* Epidemics, 15, 27-37.

Wallinga, J. and M. Lipsitch (2006). *How Generation Intervals Shape the Relationship Between Growth Rates and Reproductive Numbers.* Proc.R.

Soc. B. 274, 599-604.

A Growth curves

The generalized logistic model (GL) is

$$\mu(t)=\overline{\mu}/(1+(\gamma_0/\kappa)e^{-\gamma t})^{\kappa},\quad \gamma_0,\gamma,\kappa>0, \tag{25}$$

where γ is a growth rate parameter. When γ_0 is determined by the value of the curve at $t=0$, it is $\gamma_0=\kappa\left[(\overline{\mu}/\mu(0))^{1/\kappa}-1\right]$. Differentiation yields

$$\ln\mu'(t)=\rho\ln\mu(t)+\delta-\gamma t, \tag{26}$$

where $\mu'(t)=\ln d\mu(t)/dt$ is the 'incidence curve', $\delta=\ln(\gamma_0\overline{\mu}^{-1/\kappa}\gamma)$ and $\rho=(\kappa+1)/\kappa$ so $0<\kappa<\infty$ implies $1<\rho<\infty$. Alternatively, because $\mu'(t)=g(t)\mu(t)$,

$$\ln g(t)=(\rho-1)\ln\mu(t)+\delta-\gamma t, \tag{27}$$

where $g(t)$ is the growth rate of $\mu(t)$.

The Gompertz curve, which can be obtained by letting $\kappa\to\infty$ in equation (25) is

$$\mu(t)=\overline{\mu}\exp(-\gamma_0 e^{-\gamma t}),\quad \gamma_0,\gamma>0,\qquad -\infty<t<\infty, \tag{28}$$

with $\gamma_0=\ln(\overline{\mu}/\mu(0))$. In this case,

$$\ln\mu'(t)=\ln\mu(t)+\delta-\gamma t, \tag{29}$$

and

$$\ln g(t)=\delta-\gamma t, \tag{30}$$

where $\delta=\ln\gamma_0\gamma$.

Viboud et al (2016) propose $\mu'(t)=\alpha\mu(t)^{\beta}$, with β being a deceleration parameter in the range $0\leq\beta\leq1$, for the ascending phase of an epidemic; see also Chowell et al (2016). This implies

$$\ln\mu'(t)=\delta+\beta\ln\mu(t) \tag{31}$$

where $\delta=\ln\alpha$. When $\beta=1$ there is exponential growth. Equation (29) gives exponential growth when $\gamma=0$ but the introduction of the time trend gives sub-exponential growth.

Covid Economics 51, 7 October 2020: 36-73

B R for a growth curve

Before proceeding it is worth getting an idea of potential movements in R_t by seeing how it relates to the incidence curve of a deterministic growth curve.

The point of inflexion on the growth curve is the point at which the incidence curve, $\mu'(t)$, peaks. It follows from equations (12) and (27) that the peak for the GL is when $g(t) = \gamma/\rho$ and is at time $\ln \gamma_0/\gamma = (\delta - \ln \gamma)/\gamma$. The incidence curve declines more slowly than it ascends when $\kappa > 1$.

For the Viboud et al (2016) model, $\gamma = 0$, so $(1/\tau) \ln R(t) = \beta g(t)$ giving $R(t) = \exp(\tau\beta g(t)) = \exp(\tau\beta\alpha\mu(t)^{\beta-1})$. When $\beta = 1$, $R(t) = \exp(\tau\alpha) = \exp(\tau \exp \delta)$. The Gompertz curve allows movement away from exponential growth to take place by introducing a time trend into equation (31). The instantaneous R curve then declines as

$$R(t) = \exp(\tau(\exp(\delta - \gamma t) - \gamma)). \tag{32}$$

The parameter δ, rather than γ_0, is treated as the initial condition (if needed, $\gamma_0 = (1/\gamma) \exp \delta$) and can be defined as soon as the epidemic takes off and has a measurable growth rate. At this point $t = 0$ and $R(0) = \exp(\tau(-\gamma + \exp \delta))$. The thin line in Figure 15 shows $R(t)$ when $\gamma = 0.1$ and $\delta = -0.91$; the value of δ was set by assuming that at the start of the epidemic $R(0) = 5$ and $\gamma = 0$. The associated incidence curve, that is $\mu'(t) = \mu(t) \exp(\delta - \gamma t)$ where $\mu(t)$ is as in equation (28) with $\overline{\mu} = 100$ and γ_0 is set to $(1/\gamma) \exp \delta$, can be seen to peak when $R(t) = 1$ just before $t = 14$. By contrast the decline in the upper thick line, where $\gamma = 0.02$, is very slow. The aim of policy is to shift this line downwards by increasing γ.

As $t \to \infty$, $R(t) \to \exp(-\tau\gamma)$ so when $\gamma = 0.1$ and $\tau = 4$, $R(t) \to 0.67$. This feature of $R(t)$ could be misleading because at some point the number of people infected will be so small that the epidemic effectively ends.

A formulation such as in equation (32) is not actually consistent with some standard equations for $R(t)$ given by considerations of sigmoid curves. In particular $R(t) = (1 - F(t))R(0)$ where $F(t)$ is a standardized growth curve with $\overline{\mu} = 1$; see Bettencourt and Ribeiro (2008, p3). This equation implies that $R(t) \to 0$ as $t \to \infty$.

Covid Economics 51, 7 October 2020: 36-73

Covid Economics 51, 7 October 2020: 36-73

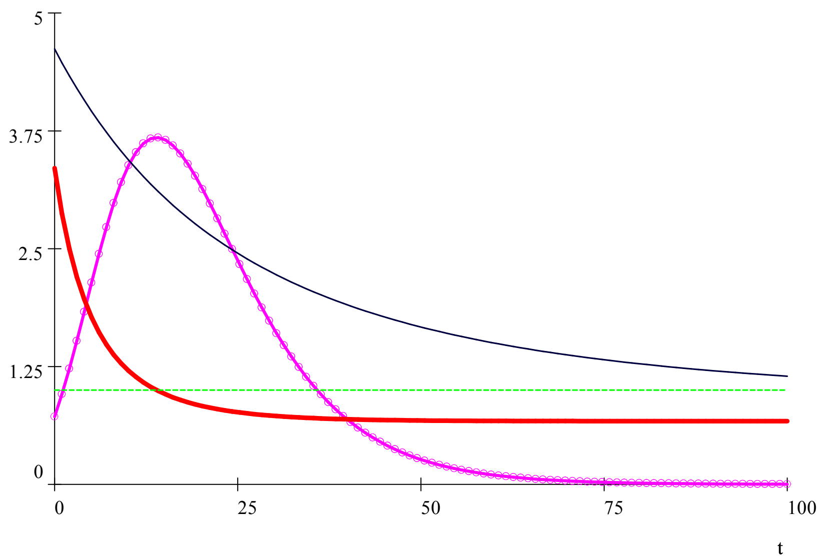


Figure 15: Gompertz incidence curve for $\gamma = 0.1$ (dots, magenta) and $R(t)$ (red). The thin (blue) line shows $R(t)$ when $\gamma = 0.02$.

CEPR PRESS

Covid Economics 51, 7 October 2020: 36-73

C Small numbers: the negative binomial distribution

When y_t is small, it may be better to specify its distribution, conditional on past values, as discrete. The usual choice is the negative binomial which, when parameterized in terms of a time-varying mean, $\xi_{t|t-1}$, and a fixed positive shape parameter, υ, has probability mass function

$$p(y_t) = \frac{\Gamma(\upsilon + y_t)}{y_t!\Gamma(\upsilon)} \xi_{t|t-1}^{y_t} (\upsilon + \xi_{t|t-1})^{-y_t} (1 + \xi_{t|t-1}/\upsilon)^{-\upsilon}, \qquad y_t = 0, 1, 2, ..,$$

with $Var_{t-1}(y_t) = \ \xi_{t|t-1} + \xi^2_{t|t-1}/\upsilon$. The dynamics are constructed as

$$\ln \xi_{t|t-1} = \ln Y_{t-1} + \delta_{t|t-1}, \quad t = 3, ..., T, \tag{33}$$

and

$$\begin{aligned} \delta_{t|t} &= \delta_{t|t-1} + \alpha_1 u_t, \quad 0 \leq \alpha_1 \leq 1 \\ \delta_{t+1|t} &= \delta_{t|t} - \gamma_{t|t-1} \\ \gamma_{t+1|t} &= \gamma_{t|t} = \gamma_{t|t-1} + \alpha_2 u_t, \quad \alpha_2 = \alpha_1^2/(2 - \alpha_1) \end{aligned} \tag{34}$$

with u_t being the standardized score

$$u_t = y_t/\xi_{t|t-1} - 1. \tag{35}$$

Predictions of future observations can be made with in the same way as for the (5).

The Poisson distribution is a special case of the Negbin obtained by letting $\upsilon \to \infty$. The score remains as in equation (35). ML estimation of both Negbin and Poisson models can be carried out using the TSL package of Lit, Koopman and Harvey (2020) and a likelihood ratio test for the Poisson distribution computed. (Because υ is on the boundary of the parameter space, the asymptotic distribution of the LR statistic is a mixture of χ^2_0 and χ^2_1.)

When the numbers of new cases are relatively large, the results of the Negbin and Gaussian models may not be too different. In the Gaussian model, the conditional distribution of y_t is lognormal and the score is $v_t = \ln y_t - \ln \xi_{t|t-1} = \ln(y_t/\xi_{t|t-1})$, where $\ln \xi_{t|t-1}$ is as in Equation (33). When Δy_t is small relative to $\xi_{t|t-1}$, $\ln(y_t/\xi_{t|t-1}) \simeq y_t/\xi_{t|t-1} - 1$ so the two filters will produce similar results for a given value of α_1. The same will be true of the forecasts of future values of y_t. At the same time the conditional distribution of y_t will be such that the standard deviation in a Negbin model with moderate sized υ is roughly proportional to its mean, just as in the lognormal distribution.

Remark 8 *For the Negbin distribution there is no unobserved components model corresponding to the filter for the level and slope. If such a model were to be formulated it would be parameter-driven and could only be estimated by a computationally intensive procedure as in Durbin and Koopman (2012).*

Covid Economics 51, 7 October 2020: 36-73

CEPR PRESS

Covid Economics **Issue 51, 7 October 2020**

Crowding out bank loans: Liquidity-driven bond issuance[1]

Olivier Darmouni[2] and Kerry Y. Siani[3]

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

According to conventional wisdom, banks play a special role in providing liquidity in bad times, while capital markets are used to fund investment in good times. Using micro-data on corporate balance sheets following the COVID-19 shock, we provide evidence that instead, the corporate bond market is central to firms' access to liquidity, crowding out bank loans even when the banking sector is healthy. We first show that, contrary to good times, bond issuance is used to increase holdings of liquid assets rather than for real investment. Second, most issuers, including many riskier "high-yield" firms, prefer issuing bonds to borrowing from their bank. Over 40% of bond issuers leave their credit line untouched in 2020Q1. Moreover, a large share of bond issuance is used to repay existing bank loans. This liquidity-driven bond issuance questions the comparative advantage of banks in liquidity provision, and suggests that the V-shaped recovery of bond markets, propelled by the Federal Reserve, is unlikely to lead to a V-shaped recovery in real activity.

Covid Economics 51, 7 October 2020: 74-133

1 For helpful comments, we thank Wei Jiang, Yiming Ma, Giorgia Piacentino, Patricia Moser, Laura Veldkamp, and Stijn Van Nieuwerburgh.
2 Associate Professor at Columbia Business School.
3 Doctoral student at Columbia Business School.

CEPR PRESS

Covid Economics 51, 7 October 2020: 74-133

Introduction

Liquidity has always been a central topic in corporate finance, and the recent "dash for cash" that followed the COVID outbreak has put this issue at the forefront of the crisis [Acharya et al., 2012, Li et al., 2020]. The textbook view emphasizes the role of the banking sector in providing liquidity in bad times, while capital markets fund investment in good times. Banks have a comparative advantage for providing liquidity both because credit lines are committed in advance and because bank loans are funded by deposits [Holmström and Tirole, 1998, Kashyap et al., 2002, Gatev and Strahan, 2006]. However, corporate bond issuance reached historical heights in the spring of 2020, even though the banking sector was healthy. This surge was partly due to a spectacular change in the Federal Reserve credit policy that supported the corporate bond market directly for the first time. This episode raises the question: What is the role of the bond market in providing liquidity in bad times? How do firms choose between borrowing from banks versus the bond market? What are the implications for monetary policy?

This paper sheds light on these questions through the lens of bond issuance, i.e. the primary market for corporate bonds. While it is clear that the Federal Reserve has revitalized *markets*[1], there are still some open questions regarding the net effects on *firms* and the real sector. Our approach is to understand aggregate issuance dynamics through the lens of micro-data on bond issuers' balance sheet. Importantly, we examine data that include the latter part of the crisis through June 2020, not only the March-April period. We provide evidence that, contrary to the textbook view emphasizing the role of banks, the corporate bond market is central to firms' access to liquidity, in two ways. First, unlike in good times, bond issuance was used to increase holdings of liquid assets rather than real investment. Second, bond issuance crowded out bank loans although the banking sector was healthy, even for many "high-yield" riskier firms. These findings question the comparative advantage of banks in

[1]See for example Haddad et al. [2020], Boyarchenko et al. [2020], Falato et al. [2020], Kargar et al. [2020], O'Hara and Zhou [2020].

Covid Economics 51, 7 October 2020: 74-133

liquidity provision, and have implications for the real effects of bond issuance and the new Federal Reserve credit policy.

We first show that, propped up by the Fed, the bond market lent extensively to firms in this period. While spreads rose before falling, closely following secondary markets (see Figure 1), the dynamics of volume were extraordinary. Both investment-grade (IG) and high-yield (HY) markets reached historical heights in the post-March 2020 period. As of end of May 2020, investment grade (high yield) issuance by U.S. firms reached \$500 billion (\$103 billion), compared to \$200 billion (\$72 billion) over the same period last year. This amounted to a remarkable "V-shaped recovery" in bond markets in a matter of weeks, including for riskier firms that were shut out for no more than a few weeks. Interestingly, HY bond issuance extended beyond issuers eligible for direct Federal Reserve purchases, suggesting an important role for ETF purchases and a broad commitment to "backstop" the market.

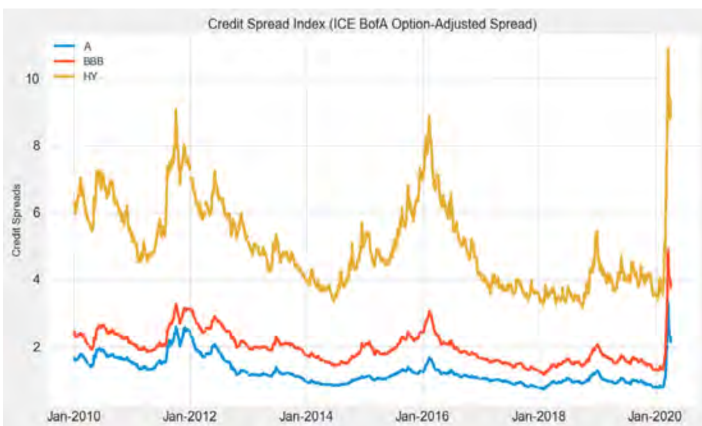


Figure 1 – IG vs. HY corporate bond spreads

Source: ICE BofA US High Yield Index Option-Adjusted Spread: spreads between an OAS index of all bonds in the respective rating category and the corresponding spot Treasury curve. “HY", or high yield, indicates bonds rated BB or below based on an average of Moody’s, S&P, and Fitch. Credit spreads are in percentage points. Retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/BAMLH0A0HYM2, https://fred.stlouisfed.org/series/BAMLC0A3CA, https://fred.stlouisfed.org/series/BAMLC0A4CBBB, July 8, 2020

While this surge in issuance is striking, there are still open questions about how it ultimately affected firms and the real sector. We argue that a necessary first step to trace out the real effects of issuance flows is to understand how firms' balance sheets are affected. For this purpose, we link the issuance data with firm-level financial statements and capital structure information. While these data are imperfect, balance sheet adjustments are nevertheless indicative of underlying economic forces. Broadly speaking, we are interested in understanding (i) what firms do with the funds raised from the bond market, relative to normal times, and (ii) how bond issuance interacted with bank financing.

We start by documenting that during COVID, firms used the bond market differently than in normal times. First, while in normal times, firms follow an issuance pattern and raise bonds when they have lower cash balances and debt coming due, firms issuing during COVID raise bond capital earlier in their bond financing cycle and have less debt coming due. Second, we document that after issuance, COVID-era issuers are more likely to hoard cash rather than invest in real assets. In addition, firms were less likely to payout to equity holders after issuing during COVID. To document these facts, we examine quarterly changes in firms' assets following issuance during the COVID period: March 1 - May 31 2020. We draw comparisons to normal times using data from 2010-2019. A reasonable concern with identifying unique aspects of COVID issuance is that the firms that were able to issue bonds during this period may be different than bond issuers in normal times. To overcome this issue, we control for potential bond market selection bias towards stronger firms by estimating within-firm regressions that further account for macroeconomic fluctuations by absorbing industry-quarter fixed effects. Our findings lend credence to the hypothesis that a large share of issuance was "precautionary" and thus unlikely to be immediately reinvested in the firms. For example, Chevron issued $650 million in bonds on March 24th, but is cutting its 2020 capital spending plan by $4 billion. Moreover, by comparing bond issuer balance sheet adjustments during COVID to those during the 2008 financial crisis, we find that the use of bond capital to increase liquidity is a new phenomenon, suggesting a secular shift in the use of corporate bonds.

We then document that, even though the banking sector was healthy, bond issuance

Covid Economics 51, 7 October 2020: 74-133

crowded out bank loans, in two ways. We first show that many firms left their existing credit lines untouched while issuing bonds instead. For example, Chevron had $5 billion in credit line available at the beginning of 2020, yet it still issued $650 million in bonds. Strikingly, this behavior includes many riskier HY firms: almost 40% of HY issuers received no new net bank funding between January and March. Only 21% had maxed out their credit line by end of March, and the average draw-down rate was below 50%. Many of these riskier firms had available "dry powder" from banks, arranged before the crisis, that they did not use. The pattern is even stronger for IG firms, which represent the bulk of aggregate issuance in this period, with over 60% not drawing on their existing credit lines. In aggregate, the amount of undrawn bank credit available at the beginning of 2020 was larger than the total funds raised from bond issuance. HY issuers in our matched sample issued $90 billion in bonds while having $142 billions of undrawn credit available. The gap is even larger for IG issuers.

Second, we show that a large share of issuers that did borrow from their bank early in the crisis repaid by issuing a bond in the following weeks. While over 60% of HY firms that issued in March-July received bank funds in March, two-thirds of these repaid some amount after their bond issuance. Over 40% actually repaid their credit line in full, and only a few borrowed additional funds from banks in the second quarter. For example, Kraft Heinz, which was downgraded from IG to junk in February 2020, drew $4 billion from its credit line between February and March. In May, it issued $3.5 billion in bonds (up from a planned $1.5 billion, due to strong investor demand) and used these funds to repay its credit line. In six months, the share of Kraft's credit coming from banks went from zero to 12% and then back to zero. We find that Kraft is far from an isolated example: among HY issuers repaying bank loans, the median firm paid back 100% of its Q1 borrowing, representing 60% of their bond issuance. In aggregate, a full quarter of HY firms' bond proceeds went to pay back bank loans. The pattern is similar for IG firms, although a smaller share drew on their credit lines in the first place. We estimate that at least $70 billion was repaid by bond issuers to banks between April and July 2020. Moreover, the majority of the Federal Reserve single-name corporate

Covid Economics 51, 7 October 2020: 74-133

Covid Economics 51, 7 October 2020: 74-133

bond portfolio consists of issuers that had access to bank funds which they did not draw.[2]

The last part of the paper discusses the interpretation of these findings and their implications. We argue that the importance of *liquidity-driven bond issuance* in bad times is key to explaining the events of the first half of 2020. Akin to a "permanent cash-flow hypothesis," firms want to borrow today an amount of liquidity equal to the cash-flow short-fall they expect in the near future. Liquidity-driven debt issuance thus spikes *because* the real recovery is slow, not in spite of it. On the other hand, investment-driven debt issuance is delayed. Moreover, firms borrow from the source for which frictions are smaller, which by revealed preference was the bond market and not bank loans. There are at least three reasons why that was the case in the spring of 2020. First, bond financing is more *committed* for a long period of time: it typically has a longer maturity and no maintenance covenants that banks can use to renegotiate credit [Sufi, 2009]. For smaller and more short-lived shocks in normal times, credit line draw-downs can be more attractive given their lower setup costs.[3] Second, although not fueled by deposits, investor demand for bonds was strong in these times: order books for bonds in the primary market were high and bond funds quickly recovered from the market turbulence and net outflows in March documented in Falato et al. [2020]. Moreover, the cost of funds for bonds might have fallen disproportionately relative to loans, with many issuers borrowing at historically low rates. The Federal Reserve's unprecedented support to "backstop" the bond market played an important role in reducing the cost of bond capital, while spreads on credit lines do not adjust until maturity unless they are renegotiated.

Lastly, we draw some implications for central bank intervention. First, our evidence that the corporate bond market is a key source of liquidity in bad times supports direct intervention in this market in addition to traditional lender of last resort policies geared towards the banking sector. However, it is important to account for the crowding out of bank loans when evaluating the aggregate effects of these new public programs on the real economy. For

[2]Based on Federal Reserve portfolio as of July 31, 2020, as reported on August 10, 2020. https://www.federalreserve.gov/monetarypolicy/smccf.htm

[3]Note however that in these times commercial paper tend to dominate credit lines for firms that have access to that market.

the majority of issuers, propping up bond markets does not alleviate a hard credit constraint, since they have available bank funding. Preventing large credit line draw-downs is nevertheless valuable for at least three reasons: it guarantees a longer-term funding source for firms, it helps weaker issuers to "keep their powder dry," and it reduces balance sheet constraints on banks. However, how to weigh these benefits against potential losses on central bank bond holdings or risks of asset prices distortion is an outstanding question that is an important area for future research.

Related Literature Our first contribution is to provide evidence that the bond market is a key source of liquidity in bad times for large firms. Broadly speaking, the conventional view suggests that banks are the primary source of funds in bad times. Banks hold large amounts of deposits [Kashyap et al., 2002], receive deposit inflow in bad times [Gatev and Strahan, 2006], and can provide liquidity insurance by arranging funding ahead of time via credit lines [Holmström and Tirole, 1998, Acharya et al., 2018]. While it is well established that large firms tend to prefer to borrow from capital markets for investment purposes, our evidence shows they also prefer it for liquidity purposes. Like bank credit, the bond market plays a dual role that varies over the cycle: it funds investment in good time and builds liquidity buffers in bad times.[4]

Many studies have used the recent COVID crisis as a testing ground for how firms access liquidity in bad times. However, although existing works tend to focus on the first part of the crisis (until early April), we also examine the latter part of the period, spanning May to July. This period is crucial to understand the underlying economics: it allows us to study the crowding out of bank loans as well as the behavior of HY firms that almost exclusively issued in the second quarter. Li et al. [2020] document the importance of bank lending for corporate liquidity at the height of the crisis. Looking at the whole period of March - July, we find

[4]Corporate liquidity management is a central topic in finance research, and has received considerable attention. See for instance Almeida et al. [2004], Eisfeldt and Muir [2016], Bolton et al. [2011], Graham and Leary [2018], Acharya et al. [2012], Opler et al. [1999], Bates et al. [2009], Denis and Sibilkov [2010], Riddick and Whited [2009], Foley et al. [2007] or Almeida et al. [2014] for a survey.

Covid Economics 51, 7 October 2020: 74-133

Covid Economics 51, 7 October 2020: 74-133

that (1) not all firms seeking liquidity drew down on bank loans, and (2) many firms later tap bond markets to replace their outstanding bank loans. We relate closely to Acharya and Steffen [2020b] which studies how the prevalence of cash, credit line draw-downs and bond issuance vary in the cross-section of firms, highlighting the significant impact of credit risk on corporate cash holdings up until mid-April. We further map out balance sheet adjustments of these bond issuers in the months following the initial crisis, allowing us to infer how firms deployed bond capital. We also relate to recent work on secondary bond markets [Haddad et al., 2020, Kargar et al., 2020, O'Hara and Zhou, 2020, Falato et al., 2020], bond issuance [Boyarchenko et al., 2020, Halling et al., 2020], credit lines draw-down [Greenwald et al., 2020, Chodorow-Reich et al., 2020] and the value of financial flexibility [Fahlenbrach et al., 2020]. Brunnermeier and Krishnamurthy [2020a], Brunnermeier and Krishnamurthy [2020b] and Crouzet and Tourre [2018] show that the effect of the crisis as well as the appropriate response depends on the underlying frictions in corporate financing.

The focus of this paper is the role of the bond market in bad times. It thus relates to a long line of work on the choice of bank vs. bond financing [Bolton and Scharfstein, 1996, Diamond, 1991, Rajan, 1992]. The conventional view, based on the Great Recession, is that firms substitute towards bonds and away from loans in bad times because banks' balance sheets weaken, driving down loan supply [Becker and Ivashina, 2014, Crouzet, 2017, De Fiore and Uhlig, 2015, Schwert, 2018, Adrian et al., 2013]. The COVID crisis, which did not originate from the banking sector, shows that this is not the only force at play in driving firms' preference for bonds.

Finally, our paper contributes to the ongoing debate about the efficacy of monetary policy, and specifically measures aimed at the corporate bond market.[5] Our evidence supports intervention in the corporate bond market, given that it is a key source of liquidity in bad times, extending the traditional lender of last resort policy beyond the banking sector. However, the

[5]This includes work on quantitative easing [Grosse-Rueschkamp et al., 2019, Ertan et al., 2019, Giambona et al., Todorov, 2020, Arce et al., 2018, Lhuissier and Szczerbowicz, 2018, De Santis and Zaghini, 2019, Siani, 2019] as well as monetary policy more generally [Kashyap et al., 1996, Crouzet, 2019, Darmouni et al., 2019, Ippolito et al., 2018, Holm-Hadulla and Thürwächter, 2020, Bolton and Freixas, 2006, Elliott et al., 2019].

Covid Economics 51, 7 October 2020: 74-133

crowding out of bank loans we document matters when evaluating the aggregate effects of these new public programs on the real economy.

1 Background and Data

The scope of this paper is to understand the role that the corporate bond market plays for firm liquidity in both good and bad times. We use the period surrounding COVID as a testing ground for bad times. Due to mandated COVID-related lockdowns, many firms faced significant reductions in operating income in spring 2020. (De Vito and Gomez [2020], OECD [2020]) As cash-generating operations halted, firms resorted to a variety of measures to alleviate severe cash shortfalls. Many large firms, such as General Electric, Boeing, and Airbnb slashed operating expenses by cutting down their workforce, while others suspended dividends.[6] In late March, media outlets reported that more than 130 companies in the U.S. and Europe drew down over $125 billion in bank debt.[7] Acharya and Steffen [2020b] find that by April 9, close to 70% of originally available credit lines were drawn down. Moreover, Li et al. [2020] document that the weekly growth rate in bank lending hit over 6%, the highest rate in recorded history.

Internal funds and bank lending are natural sources of liquidity in bad times. The textbook view emphasizes the role of the banking sector to fund negative liquidity shocks [Kashyap et al., 2002, Bolton et al., 2016], either because (1) credit lines are arranged in advance (liquidity insurance) [Holmström and Tirole, 1998], or (2) deposit inflows may increase in bad times (flight to quality) [Gatev and Strahan, 2006]. What is novel about the COVID period, however, is how heavily and widely firms relied on bond capital for liquidity needs. Despite significant volatility in credit spreads, firms issued bonds at record volumes in March - May

[6]Sources: "GE to Cut 10% of Aviation Workforce as Coronavirus Grounds Airliners"; *Wall Street Journal*, 03/23/2020; "Boeing Cuts Its Workforce Due To The Coronavirus Crisis", *NPR*, 04/29/2020; "Airbnb Cuts 1,900 Jobs, 25% Of Its Workforce, As Pandemic Freezes Travel" *NPR*, 05/05/2020. Many firms reduced payouts to equity: Ford Motor Co. and Freeport-McMoRan Inc. suspended dividend payments while AT&T halted share repurchases. "Companies Race for Cash in Coronavirus Crisis", *Wall Street Journal*, 03/23/2020

[7]"Dash for cash: companies draw $124bn from credit lines", *Financial Times*, 03/25/2020.

Covid Economics 51, 7 October 2020: 74-133

2020. At the start of the crisis, only firms with higher credit ratings were able to issue bonds at substantially elevated spreads [Acharya and Steffen, 2020b], while riskier firms with sub-investment grade credit ratings were shut out of the market. That led to an unprecedented change in the Federal Reserve credit policy with the introduction of programs to support this market directly for the first time in spring 2020.[8] What followed was a remarkable recovery of both investment grade and high yield corporate bond issuance. In this paper, we track this wave of bond capital raising and compare how firms use bond proceeds during the liquidity crunch of COVID versus during normal times.

To this end, we construct a panel data set covering all U.S. non-financial bond issuers from January 2000 to June 2020. Bond-level issuance data comes from Mergent FISD, which includes detailed issuance-level data on corporate bond offerings, and is combined with bond auction data from Credit Flow Research (CFR). We restrict the sample to U.S. dollar bonds of at least \$100 million face value issued by firms that report in U.S. dollars. In line with much of the empirical literature on corporate bond issuance, we exclude financial, sovereign, and utility issuers. We further exclude convertible bonds, capital impact bonds, community bonds, PIK securities, and bonds issued directly in exchange for an identical bond.[9] We merge the issuance data with quarterly balance sheet data from Compustat and hand-collected debt composition from Capital IQ. For the analysis on firm balance sheet adjustments, we include only those issuers we can match to Compustat. The filters leave us with 1,491 unique issuers issuing 9,699 bonds for which we have quarter-end balance sheet data. For the firms that issued during COVID, 25% have data available for the second quarter following issuance; for the rest of the firms, second quarter data has not yet been reported.[10] Tables 1 and 2 display

[8]Through the Primary Market Corporate Credit Facility (PMCCF) and Secondary Market Corporate Credit Facility (SMCCF), the Federal Reserve pledged a combined \$750 billion to the purchase of corporate bonds and bond ETFs. The facility is backed by \$75 billion of equity capital from the Department of Treasury. While the primary focus of these programs was to purchase investment grade securities, on April 9, 2020, the Fed announced that high yield ETFs and certain recently downgraded firms would also be eligible. See the Fed's statement and Boyarchenko et al. [2020] for an analysis of the program and corresponding announcement effects.

[9]Bonds associated with the T-Mobile / Sprint acquisition in April 2020 are also excluded

[10]We are able to match 85.6% of bonds in our sample to firms in Compustat. 48% of unmatched bonds are foreign issuers. The rest do not have reported financials in Compustat in the quarters of issuance. For balance

summary statistics of our baseline sample.[11] In a typical week in 2019, a median of five IG firms issued $6 billion in bonds while a median of five HY firms issued $3.5 billion. In normal times, the median bond size is $500 million with an eight year tenor and yield of 5.125%.

2 Bond Issuance during COVID

2.1 Record aggregate issuance volumes

A striking fact about the COVID episode is that the dynamics of volume were extraordinary. Propped up by the Federal Reserve announcements, the bond market lent extensively to firms. Figures 2 and 4 show that both investment-grade (IG) and high-yield (HY) markets reached historical heights in March and April 2020. The riskiest U.S. firms issued over $110 billion in high yield bonds in January-May 2020, compared to $89 billion in the same period in 2019, despite a three-week hiatus in March 2020. Similarly, IG bond issuance hit over $500 billion in volume issued by 204 unique firms by May, compared with $200 billion over the same period in 2019.

IG firms began issuing early on during the COVID crisis despite significant spikes in credit spreads (see Figure 4). Figure 5 shows weekly spread dynamics separately for investment-grade (IG) and high-yield (HY) issuers during this period. Average weekly spreads for IG bond issuers increased by over 200 bps from February to March before falling back to levels slightly elevated above pre-COVID levels.[12]

While IG issuance was robust even prior to the Fed's announcement to support corporate bond markets, Figure 4 shows that high yield issuers were buoyed significantly by the Fed.

sheet analyses, we include only the 91% of matched issuing firms that either report financial statements in U.S. dollars or are domiciled in the U.S.

[11]Firms that issue in bond markets are on the larger end of the distribution of all firms. In 2019, the median bond issuer had $10.8 billion in total assets and $1.4 billion in quarterly revenues at year end, compared to the median Compustat firm with $1.7 billion in assets and $216 million in quarterly revenues.

[12]Compared to credit spread spikes around the 2008 financial crisis, these fluctuations are moderate. Between November 2008 and January 2009, average IG bond spreads on new bonds averaged well over 600 basis points in some weeks. See Figure 3 for more historical context.

Covid Economics 51, 7 October 2020: 74-133

Table 1 – Summary statistics: bond issuance, 2019-2020

	Num Offerings	Amount (Bn)	Tenor	Rating	Credit Spread	Yield
IG Issuance: 2019						
10%	2	1.5	9.3	13.9	96	2.94%
50%	5	6.0	13.5	14.7	142	3.78%
90%	10	22.7	19.3	16.5	185	4.47%
IG Issuance: Weeks since March 2020						
2020-03-02	11	7.8	12.7	14.5	141	2.46%
2020-03-09	3	3.9	12.2	14.2	211	2.91%
2020-03-16	10	44.5	16.1	17.1	270	3.95%
2020-03-23	28	64.2	13.2	16.0	273	3.68%
2020-03-30	15	53.4	15.3	15.5	321	4.05%
2020-04-06	12	22.7	10.8	15.3	314	3.82%
2020-04-13	10	26.4	11.7	15.6	231	3.18%
2020-04-20	16	20.7	10.3	14.6	277	3.50%
2020-04-27	22	66.0	14.1	15.8	215	3.12%
2020-05-04	26	53.9	12.3	15.2	255	3.29%
2020-05-11	19	36.7	14.7	14.8	251	3.55%
2020-05-18	11	35.5	16.5	16.1	170	2.83%
2020-05-25	8	10.6	14.8	15.4	162	2.44%
HY Issuance: 2019						
10%	1	1.4	6.8	7.9	315	5.07%
50%	5	3.5	7.8	9.3	381	6.23%
90%	10	8.4	9.2	10.1	539	7.56%
HY Issuance: Weeks since March 2020						
2020-03-02	2	1.8	10.0	10.0	368	4.69%
2020-03-30	4	5.8	4.5	10.2	787	8.23%
2020-04-06	3	1.6	5.0	7.0	814	8.62%
2020-04-13	8	12.8	5.5	10.8	641	7.38%
2020-04-20	15	10.4	5.2	9.4	702	7.36%
2020-04-27	6	3.1	5.0	8.5	554	7.14%
2020-05-04	8	7.7	8.4	10.8	599	7.14%
2020-05-11	11	8.1	6.2	8.2	662	7.23%
2020-05-18	8	5.0	6.4	9.8	517	7.44%
2020-05-25	7	8.6	6.1	9.6	651	7.80%

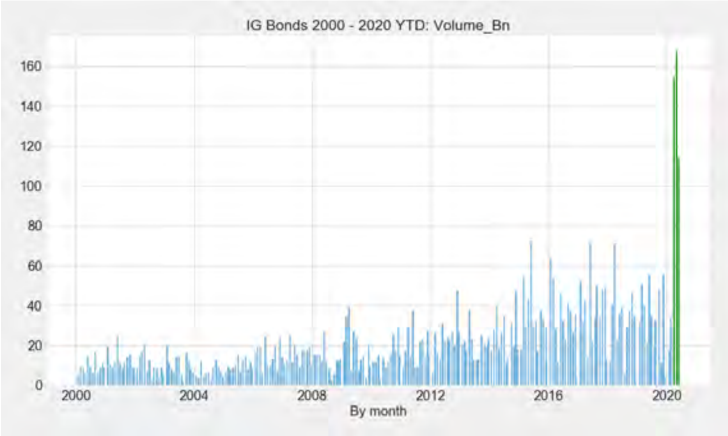
Source: Mergent FISD, http://bv.mergent.com/view/scripts/MyMOL/index.php, retrieved July 30, 2020.
Note: Includes all USD corporate bond issuance of over $100 million in size issued by U.S. domiciled companies or companies that report in U.S. dollars. Excludes sovereign, supra-sovereign, financial, and utility offerings, convertible notes, impact bonds, bonds issued directly in exchange of existing bonds, PIK notes, and reopening issuance of existing bonds. Variables are average across week, except number of offerings and amount issued, which are sum across weeks.

Table 2 – Summary statistics: bond issuers, 2017-2020

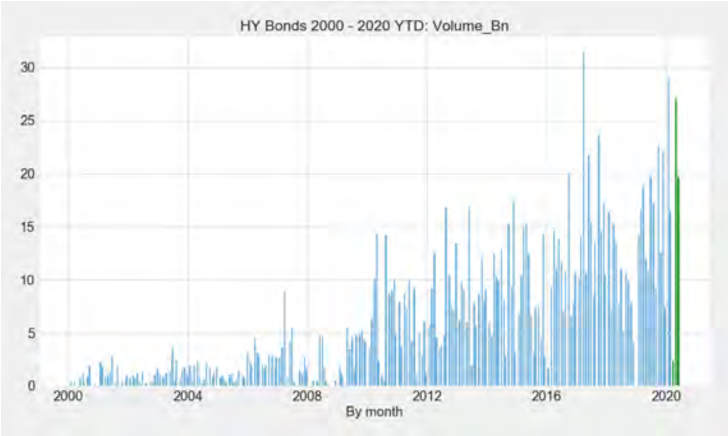
	Normal times			Covid times		
	10%	50%	90%	10%	50%	90%
Balance sheet metrics						
Cash/Assets (prior Q4)	0.6%	5.1%	20.6%	0.9%	4.8%	18.8%
Cash/Assets (Q1)	0.5%	4.8%	20.4%	1.7%	7.9%	21.1%
Debt/Assets (prior Q4)	16.1%	37.5%	64.0%	21.7%	38.8%	65.0%
Debt/Assets (Q1)	17.8%	38.6%	62.2%	24.3%	41.9%	67.2%
Current debt/Debt (prior Q4)	0.0%	2.9%	15.3%	1.2%	5.5%	16.0%
Log assets (prior Q4)	7.3	9.1	11.1	8.5	9.8	11.5
Cash flow metrics						
Sales growth	-17%	-1%	14%	-26%	-6%	8%
Profit growth	-186%	-25%	116%	-301%	-31%	64%
Cash flow growh	-137%	-42%	69%	-146%	-64%	29%
Cash growth	-46%	-1%	96%	-19%	29%	303%
Bond metrics						
Amount per bond (MM)	300.0	500.0	1100.0	400.0	650.0	1330.0
Credit spread (bps)	92.6	235.0	513.4	148.5	292.5	713.9
Yield	3.264%	5.125%	7.762%	2.328%	3.831%	8.225%
Tenor (years)	5.0	8.0	30.0	5.0	10.0	30.0
Coupon	3.115%	5.000%	7.675%	2.250%	3.800%	8.375%
Rating	7.0	12.0	16.0	9.0	14.0	17.0
Days since last issuance	180.0	561.0	2153.0	140.0	407.0	1656.2
Days to next maturity	91.2	1121.0	2964.4	52.8	411.0	1844.0

Source: Mergent FISD, http://bv.mergent.com/view/scripts/MyMOL/index.php, retrieved July 30, 2020 and Compustat. **Note:** Includes all USD corporate bond issuance of over $100 million in size issued by U.S. domiciled companies or companies that report in U.S. dollars. “COVID" refers to bond issuers from March 1 - May 29, 2020. “Normal" refers to bond issuers from March 1 - May 29, 2017-2019. Growth variables are measured from Q4 of prior year to Q1 in year of issuance. Excludes sovereign, supra-sovereign, financial, and utility offerings, convertible notes, impact bonds, bonds issued directly in exchange of existing bonds, PIK notes, and reopening issuance of existing bonds. See Table 16 for mapping of credit ratings to the numerical aggregation shown here.

Covid Economics 51, 7 October 2020: 74-133



(a) IG Bond Issuance Volume since 2000



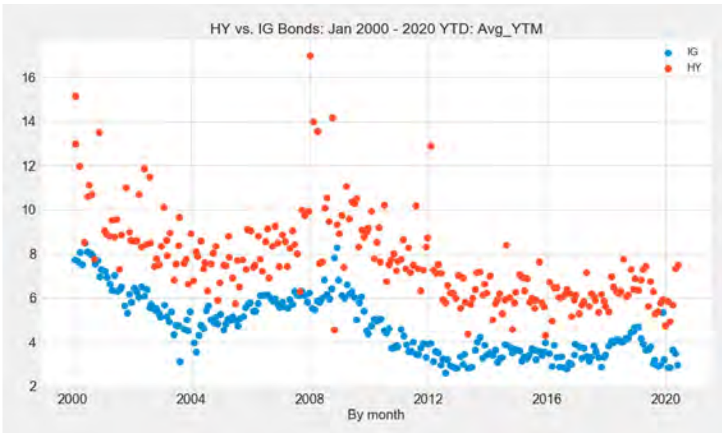
(b) HY Bond Issuance Volume since 2000

Figure 2 – Comparing IG vs. HY spreads and yields at issuance, since 2020

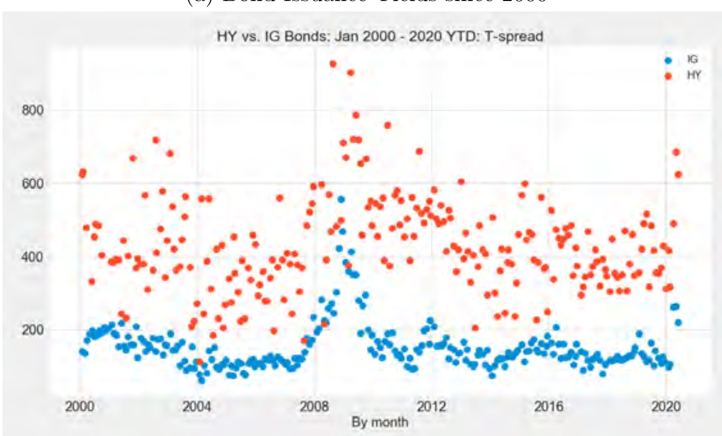
Source: Mergent FISD, http://bv.mergent.com/view/scripts/MyMOL/index.php, retrieved July 30, 2020.

Covid Economics 51, 7 October 2020: 74-133

Covid Economics 51, 7 October 2020: 74-133



(a) Bond Issuance Yields since 2000

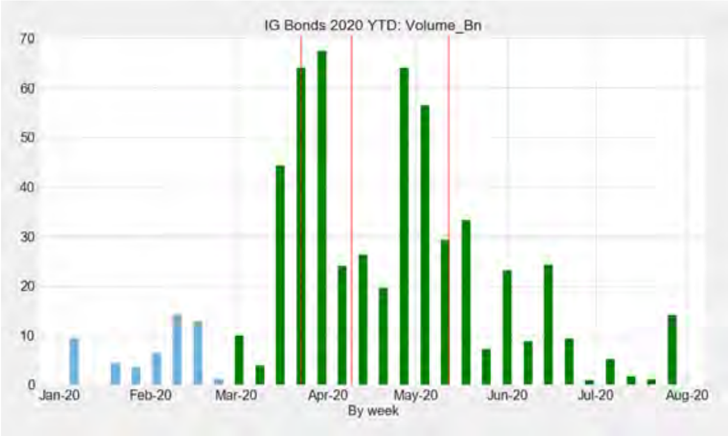


(b) Bond Issuance Credit Spreads since 2000 (bps)

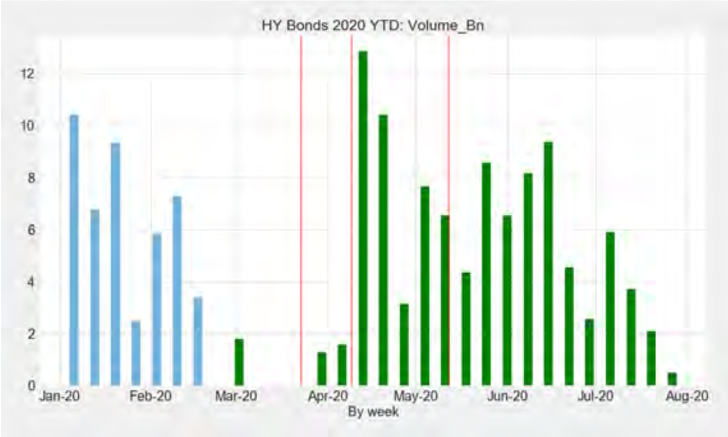
Figure 3 – Comparing IG vs. HY spreads and yields at issuance, since 2000

Source: Mergent FISD, http://bv.mergent.com/view/scripts/MyMOL/index.php, retrieved July 30, 2020

Covid Economics 51, 7 October 2020: 74-133



(a) IG Bond Issuance Volume since 2020

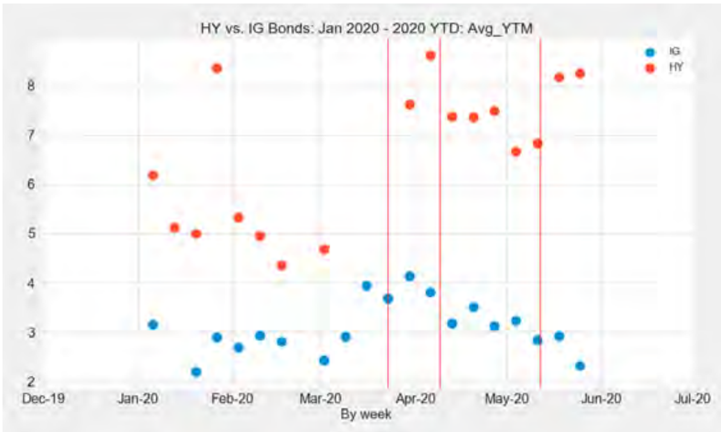


(b) HY Bond Issuance Volume since 2020

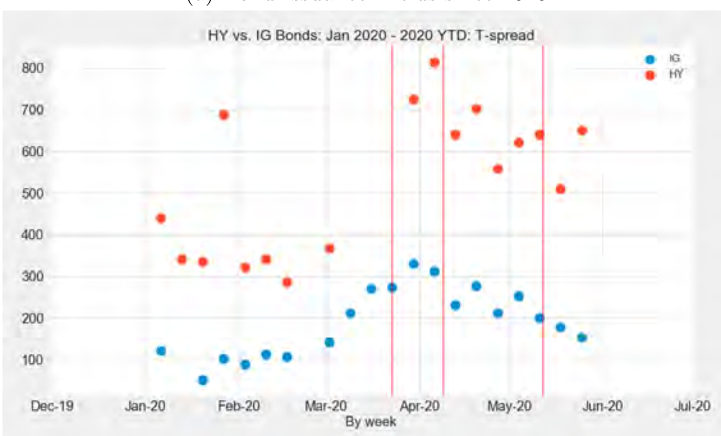
Figure 4 – Comparing IG vs. HY spreads and yields at issuance, since 2020

Source: Mergent FISD, http://bv.mergent.com/view/scripts/MyMOL/index.php, retrieved July 30, 2020. Note red lines correspond to March 23, 2020 (first Fed announcement to buy corporate bonds); April 9, 2020 (first Fed announcement to buy high yield corporate bonds); and May 12, 2020 (start of Fed bond buying program).

Covid Economics 51, 7 October 2020: 74-133



(a) Bond Issuance Yields since 2020



(b) Bond Issuance Credit Spreads since 2020 (bps)

Figure 5 – Comparing IG vs. HY spreads and yields at issuance, since 2020

Source: Mergent FISD, http://bv.mergent.com/view/scripts/MyMOL/index.php, retrieved July 30, 2020. Note red lines correspond to March 23, 2020 (first Fed announcement to buy corporate bonds); April 9, 2020 (first Fed announcement to buy high yield corporate bonds); and May 12, 2020 (start of Fed bond buying program).

However, note that the increased volume of HY issuance extended beyond issuers eligible for direct Federal Reserve purchases. Indeed, direct purchase of HY bonds is restricted to "fallen angels" issuers that were classified as investment grade as of March 22, 2020.[13] As of the end of May, the only firms eligible under the Fed's bond-buying criteria that had actually issued bonds were Ford Motor Co. and Macy's Inc. This gap suggests an important role for ETF purchases and a broad commitment to "backstop" the market, in line with some of the evidence in Boyarchenko et al. [2020].

2.2 Bond issuance to raise liquidity in bad times

In this section, we aim to explore whether the bond market provided liquidity to firms during the COVID crisis. To do this, we link issuance data with firm-level financial statements and capital structure information. While these data are imperfect, balance sheet adjustments are nevertheless indicative of underlying economic forces. Broadly speaking, we are interested in understanding which firms issued during COVID in 2020, why, and whether there are differences from normal times.

2.2.1 Which firms issued bonds in Spring 2020?

Firms that issued bonds in Spring 2020 were different from the typical bond issuer in normal times. A priori, the selection into issuance is unclear: weaker firms might need more funds, but might be excluded from the market. Figure 6 compares characteristics of firms issuing bonds during the height of the COVID crisis to those issuing bonds in normal times. It is apparent that 2020 issuers are larger and started the year with more cash on their balance sheets. Moreover, Figure 7 shows that firms that issued during COVID were better rated issuers that were more likely to have issued recently. These results are in line with Halling et al. [2020], who highlight the prominence of experienced bond issuers during this period,

[13] According to the Fed disclosure on April 9, in order to be eligible, a firm needed a plurality of agencies to rate it IG (BBB- and above) as of March 22, and a plurality of agencies to rate it BB-/Ba3 or above at the time of the Fed purchase. See the Fed announcement for more details.

Covid Economics 51, 7 October 2020: 74-133

Covid Economics 51, 7 October 2020: 74-133

and are consistent with a narrative that bond markets were only willing to lend to firms on the safer end of the spectrum during the period of market turmoil.

Because we want to study why firms choose to issue bonds in normal times vs. in COVID times, we need to separate out potential selection effects. To do this, we run regressions that include firm fixed effects to account for bond market selection bias towards stronger firms. We further include industry-quarter fixed effects to absorb any industry-specific shocks to demand for capital.

$$\begin{aligned} Y_{f,q-1} &= \alpha_f + \alpha_{ind \times q-1} + \beta_0 \mathbf{1}\{issue_{fq}\} \\ &\quad + \beta_1 \mathbf{1}\{issue_{fq}\} \times \mathbf{1}\{\text{COVID issue}_{fq}\} + \gamma' X_{f,q-1} + \epsilon_{f,q-1} \end{aligned} \tag{1}$$

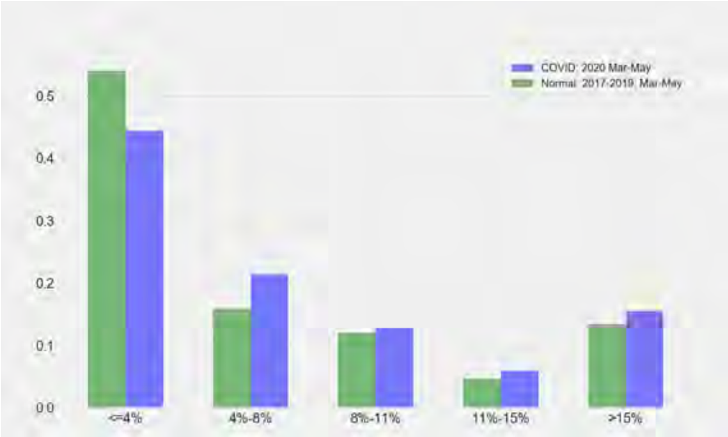
Table 3 present estimates of how different balance sheet characteristics are for firms issuing in normal times vs. during COVID. Estimates of β_0 represent how characteristic $Y_{f,q-1}$ differs between a firm's balance sheet prior to a bond issuance in normal times and the same firm's balance sheet in any other quarter not immediately surrounding a bond issuance. Estimates of β_1 represent the incremental effect of issuing a bond during COVID on the respective balance sheet metric $Y_{f,q-1}$.

First, we find that firms that issued during COVID are even more cash-rich in the previous quarter than they are prior to bond issuance in normal times. The estimated $\hat{\beta}_1$ for the cash-assets regression is positive with a magnitude that is both statistically and economically significant. Moreover, Table 3 shows that firms usually issue bonds when they have not issued bonds in many quarters and when the proportion of current debt is high. This is evidence of a regular cadence in bond issuance. During COVID, however, issuing firms enter the quarter with less current debt coming due. This suggests that rollover risk was less likely to be the primary decision factor for issuance during COVID than initially thought.[14] COVID issuers also have issued more recently than they had prior to bond issuance in normal times (see also Figure 7). These results are consistent with the notion of bond issuance for the purpose of

[14] "Will the coronavirus trigger a corporate debt crisis?", *Financial Times*, 03/12/2020.

Covid Economics 51, 7 October 2020: 74-133

(a) Log Assets, Q4



Cash as % of Total Assets, Q4

Figure 6 – Balance sheet characteristics for bond issuers: COVID vs. Normal times

Source: Compustat and Mergent FISD, http://bv.mergent.com/view/scripts/MyMOL/index.php, retrieved July 30, 2020.

Notes: Blue bars are 2019 Q4 characteristics of firms issuing in March-May 2020. Green bars are 2019 Q4 characteristics of firms issuing in March - May 2017-2019

CEPR PRESS

Table 3 – Firm balance sheets prior to issuance

	(1) Cash/Assets$_{t-1}$	(2) Current debt/Debt$_{t-1}$	(3) Leverage$_{t-1}$	(4) Qtrs since last bond$_{t-1}$
Issue$_t$	-0.00538***	0.0148***	-0.000319	0.0892***
	(0.00121)	(0.00363)	(0.00192)	(0.0143)
COVID×Issue$_t$	0.0166***	-0.0169**	-0.00990	-0.184***
	(0.00342)	(0.00778)	(0.00966)	(0.0309)
ROA$_{t-1}$	0.0606	-0.140*	-0.428***	0.797*
	(0.0773)	(0.0751)	(0.131)	(0.406)
Firm FE	✓	✓	✓	✓
Ind x Yr-Qtr FE	✓	✓	✓	✓
Observations	33943	22510	32665	7454
R-squared	0.782	0.314	0.796	0.814

Notes: Observations are at the firm-quarter level, from 2010Q1-2020Q2, for firms that issue bonds. Regression (4) includes only firm-quarters 2015Q1-2020Q2 due to data limitations. We exclude firm-quarters where the firm has just issued a bond in the prior 3 quarters. "Cash/assets" is che/at. "Current debt/debt" is debt due within one year ($dd1$) divided by total debt. "Leverage" is total debt / total assets. Total debt is $dltt + dlc$, total LT debt plus debt in current liabilities. Ratios are all winsorized to the 1%. "Qtrs since last bond" indicate log of the the number of quarters since the last issuance by the same firm. $Issue_t$ is an indicator variable that equals one if that firm issues a bond in the following quarter. Firm controls include return on assets (operating income before depreciation divided by total assets, or $oibdp/at$). We include firm (gvkey) fixed effects and industry (naic2) x year-quarter fixed effects. Standard errors, in parentheses, are clustered as the industry level.

Covid Economics 51, 7 October 2020: 74-133

Covid Economics 51, 7 October 2020: 74-133

unanticipated liquidity provision.

2.2.2 What did firms do with bond capital in COVID vs. normal times?

Next, we explore how firms deploy bond capital in times of crisis. We document that, during COVID, firms used the bond market differently than in normal times. To do this, we examine quarterly changes in firms' assets before and following issuance in 2020.[15] We draw comparisons to normal times using data from 2010-2019. In normal times, bonds are often used for long-term investment and acquisitions, or even to finance payouts to share holders (Farre-Mensa et al. [2018], Acharya and Plantin [2020]). However, we find that spring 2020 issuers were more likely to hoard cash rather than invest in real assets.

More precisely, we run three empirical tests. First, we run a simple *frequency analysis*, where we count the number of bond-issuing firms with significant decreases and increases in various balance sheet characteristics following issuance in normal times vs. in COVID times. See Table 7 for a summary. Second, we run an *event study analysis* by regressing firm balance sheet characteristics on dummy variables for each of the four quarters leading up to issuance and the two quarters following issuance.

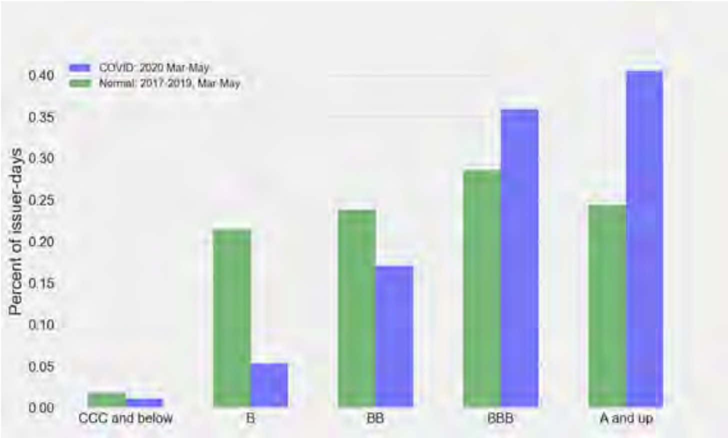
$$Y_{fq} = \sum_{m=-4}^{1} \beta_m Issue_{f,q+m} + \alpha_f + \alpha_{ind \times year} + \gamma' X_{fq} + \epsilon_{fq} \tag{2}$$

We run the regression separately for IG and HY firms, and for issuance during normal times vs. issuance during COVID. Then we plot the time dummy coefficients, β_m, to visualize the pre- and post-trends of balance sheet characteristics in COVID times vs. normal times. Results are in Figures 13 and 8.

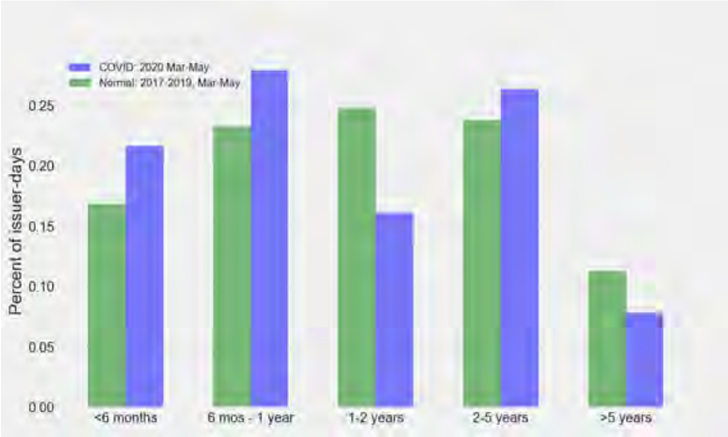
Next, we explore within-firm *post-issuance balance sheet adjustments* in normal times versus COVID times. This allows us to compare the magnitudes of balance sheet adjustments of

[15]Note we are restricted to bond issuers for which we have end-of-quarter balance sheet data in 2020Q1 and 2020Q2

CEPR PRESS



(a) Ratings of bonds issued



(b) Days since last issuance

Figure 7 – Bond issuer characteristics: COVID vs. Normal times

Source: Mergent FISD, http://bv.mergent.com/view/scripts/MyMOL/index.php, retrieved July 30, 2020.
Notes: Blue bars are characteristics of firms issuing in March-May 2020. Green bars are characteristics of firms issuing in March - May 2017-2019

Covid Economics 51, 7 October 2020: 74-133

Covid Economics 51, 7 October 2020: 74-133

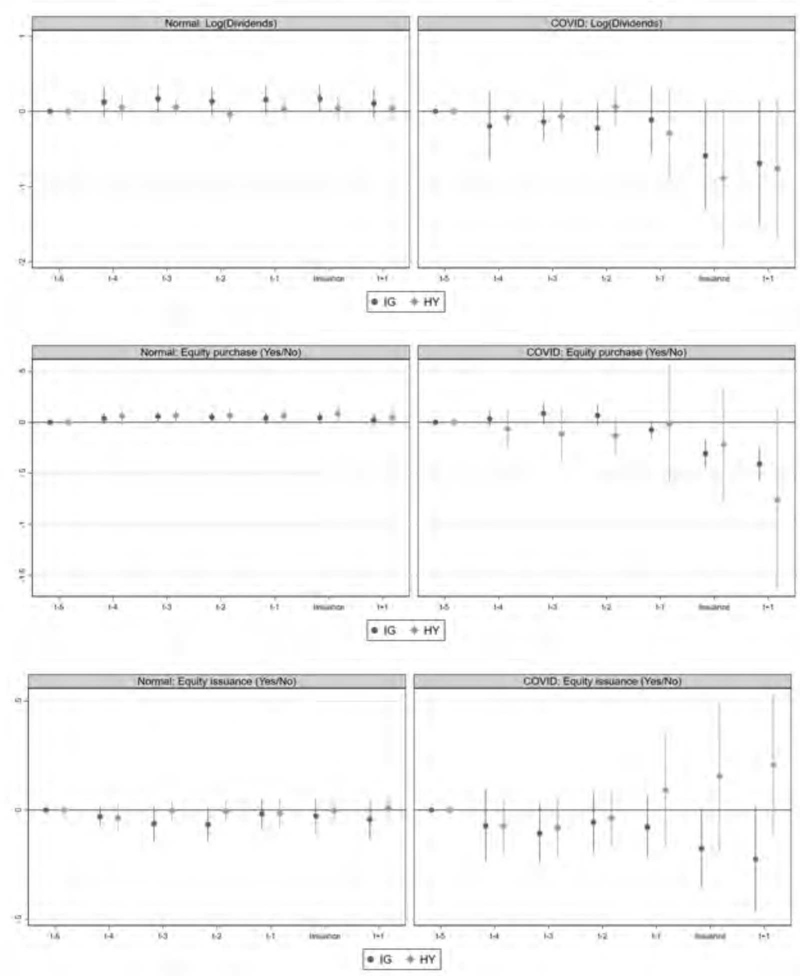


Figure 8 – Coefficient plots: Equity payout items

Notes: Each point is an estimate of β_{t+m} from the regression $Y_{fq} = \sum_{m=-4}^{1} \beta_m Issue_{f,t+m} + \alpha_f + \alpha_{ind\times year} + \gamma' X_{fq} + \epsilon_{fq}$, with 95% confidence intervals. The blue points are investment grade firms (rated BBB- and above), while the red points are high yield firms (rated below BBB-). Observations are firm-quarters up to five quarters prior to a bond issuance and two quarters following a bond issuance. ep_dummy is an indicator variable that equals 1 if there were positive share repurchases ($prstkc > 0$) in that quarter, and 0 otherwise. eq_iss is an indicator variable that equals 1 if there was positive equity issuance ($sstk > 0$) in that quarter, and 0 otherwise. "Normal" times includes bonds issued between 2017-2019, while "Covid" times includes bonds issued between March 1 - May 31, 2020.

issuance in good vs. bad times within firm. To do this, we run the following regression:

$$\begin{aligned} Y_{f,q+m} = \alpha_f + \alpha_{ind \times q+m} + \beta_0 \mathbf{1}\{issue_{fq}\} \\ + \beta_1 \mathbf{1}\{issue_{fq}\} \times \mathbf{1}\{\text{COVID issue}_{fq}\} \\ + \gamma' X_{f,q+m} + \epsilon_{f,q+m} \qquad \forall m = 1, 2 \end{aligned} \tag{3}$$

That is, we look at how a firm's balance sheet changes up to two quarters following a bond issuance, absorbing the firm fixed effect and accounting for industry specific shocks (we further control for return on assets and quarterly stock market returns to account for unobservable firm-specific shocks that could affect Y). We can thus interpret the $\hat{\beta}_1$ estimate as the incremental effect on Y of an issuance during COVID, m quarters after bond issuance. Table 4 and 5 show the results, and below we discuss the main findings.

Greater increase in cash holdings. First, in the frequency analysis, we find that firms issuing during COVID were much more likely to end the quarter with a significant ($> 10\%$) increase in their cash balance (See Figure 9). With the event study analysis in Figure 13, we show that this result persists into the second quarter following bond issuance, and that this persistence of cash hoarding is more pronounced for COVID issuance. In normal times, cash holdings are elevated at the end of the quarter of issuance, but decrease by the second quarter following issuance. In COVID times, however, cash holdings continue to be elevated in the second quarter following issuance. For high yield firms, this result is even more dramatic. Further, we find that HY firms have elevated cash in the quarter prior to bond issuance. At the start of the COVID crisis, there was virtually no HY issuance (see Figure 4), thus the increased cash prior to bond issuance likely reflects that these riskier firms found alternative sources of cash (such as drawing down on a bank loan) before they were able to access bond markets.

From the post-issuance balance sheet adjustments regressions in Tables 4 and 5, we find that the effect of bond issuance during COVID on a firm's cash to assets ratio within firm is nearly doubled in the first quarter relative to normal times. Moreover, in the second quarter

Covid Economics 51, 7 October 2020: 74-133

Covid Economics 51, 7 October 2020: 74-133

Table 4 – Firm balance sheets right after issuance

	(1) Cash/Assets$_{t+1}$	(2) Leverage$_{t+1}$	(3) Non-cash assets$_{t+1}$	(4) Current debt/Debt$_{t+1}$	(5) Equity payout$_{t+1}$
Issue$_t$	0.0118***	0.0249***	0.0686***	-0.00375***	0.169***
	(0.00122)	(0.00236)	(0.00995)	(0.00108)	(0.0346)
COVID×Issue$_t$	0.0194***	-0.0197**	-0.0275	-0.0218***	-0.432**
	(0.00494)	(0.00859)	(0.0478)	(0.00512)	(0.158)
ROA$_{t+1}$	0.0696	-0.414***	-1.637*	-0.172**	6.458***
	(0.0755)	(0.122)	(0.792)	(0.0752)	(1.174)
Qtrly stock return$_{t+1}$	0.0000151	-0.0000776	-0.000291	-0.0000587***	-0.000192
	(0.0000106)	(0.0000767)	(0.000274)	(0.00000690)	(0.000130)
Firm FE	✓	✓	✓	✓	✓
Ind x Yr-Qtr FE	✓	✓	✓	✓	✓
Observations	34385	33063	34385	23488	31568
R-squared	0.783	0.795	0.946	0.311	0.766

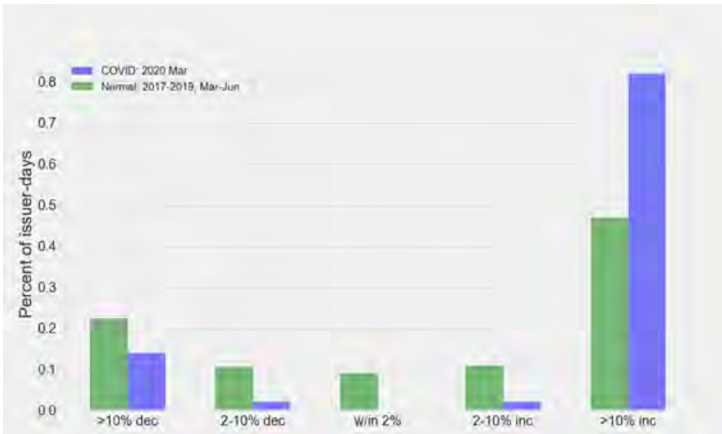
Notes: Observations are at the firm-quarter level, from 2010Q1-2020Q2. We exclude observations where the firm issued in the previous 2 quarters. $Issue_t$ is an indicator variable that equals one if that firm issues a bond in that quarter; $COVID \times Issue_t$ is an indicator variable for the reference bond being issued after March 1, 2020 and before June 1, 2020. "Cash/Assets" is che/at, winsorized at the 1% level. "Leverage" is total debt / total assets. "Non-cash assets" is the log of $(at - che)$. "Current debt / debt" is debt due within one year ($dd1$) divided by total debt. Total debt is $dltt + dlc$, total LT debt plus debt in current liabilities. "Equity payout" is the log of (equity purchases + dividend payments), or $log(prstkc + dv)$. Firm controls include return on assets (operating income before depreciation divided by total assets, or $oibdp/at$) and one quarter stock market return $((prrc_t/prrc_{t-1}) - 1)$. We include firm (gvkey) fixed effects and industry (naic2) x quarter fixed effects. Standard errors, in parentheses, are clustered as the industry level.

Covid Economics 51, 7 October 2020: 74-133

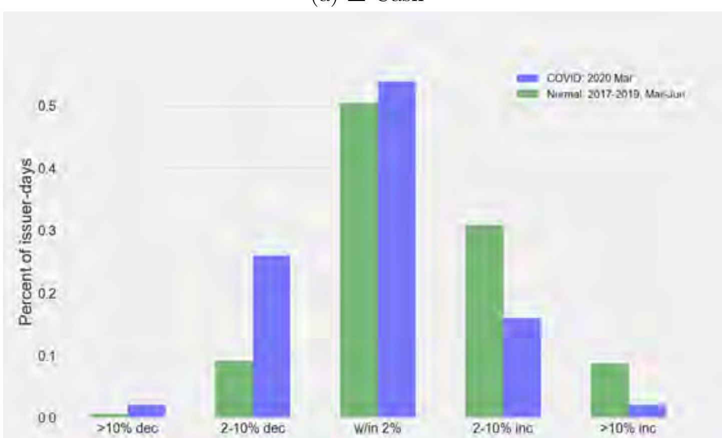
Table 5 – Firm balance sheets two quarters after issuance

	(1) Cash/Assets$_{t+2}$	(2) Leverage$_{t+2}$	(3) Non-cash assets$_{t+2}$	(4) Current debt/Debt$_{t+2}$	(5) Equity payout$_{t+2}$
Issue$_t$	0.00156	0.0248***	0.0843***	-0.0144***	0.124***
	(0.00100)	(0.00245)	(0.0110)	(0.00111)	(0.0235)
COVID×Issue$_t$	0.0183***	-0.0136	-0.122	-0.0113	-0.424**
	(0.00631)	(0.0187)	(0.0873)	(0.0115)	(0.151)
ROA$_{t+2}$	0.0975	-0.405***	-1.656**	-0.162**	6.297***
	(0.0804)	(0.126)	(0.784)	(0.0716)	(1.089)
Qtrly stock return$_{t+2}$	0.0000129	-0.0000772	-0.000296	-0.0000612***	-0.000173
	(0.0000118)	(0.0000742)	(0.000281)	(0.00000689)	(0.000126)
Firm FE	✓	✓	✓	✓	✓
Ind x Yr-Qtr FE	✓	✓	✓	✓	✓
Observations	33482	32200	33482	22872	30766
R-squared	0.790	0.796	0.945	0.313	0.763

Notes: Observations are at the firm-quarter level, from 2010Q1-2020Q2. We exclude observations where the firm issued that quarter or two quarters prior. $Issue_t$ is an indicator variable that equals one if that firm issues a bond in that quarter; $COVID \times Issue_t$ is an indicator variable for the reference bond being issued after March 1, 2020 and before June 1, 2020. "Cash/Assets" is che/at, winsorized at the 1% level. "Leverage" is total debt / total assets. "Non-cash assets" is the log of $(at - che)$. "Current debt / debt" is debt due within one year ($dd1$) divided by total debt. Total debt is $dltt + dlc$, total LT debt plus debt in current liabilities. "Equity payout" is the log of (equity purchases + dividend payments), or $log(prstkc + dv)$. Firm controls include return on assets (operating income before depreciation divided by total assets, or $oibdp/at$) and one quarter stock market return $((prrc_t/prrc_{t-1}) - 1)$. We include firm (gvkey) fixed effects and industry (naic2) x quarter fixed effects. Standard errors, in parentheses, are clustered as the industry level.



(a) Δ Cash



(b) Δ Non-Cash Assets

Figure 9 – Balance sheet adjustments for bond issuers, COVID vs. normal times

Source: Compustat and Mergent FISD, http://bv.mergent.com/view/scripts/MyMOL/index.php, retrieved July 30, 2020.

Notes: Blue bars are characteristics of firms that issued in March 2020. Green bars are characteristics of firms issuing in March - June, 2017-2019. Balance sheet adjustment compares quarter end prior to bond issuance to quarter end immediately following bond issuance.

Covid Economics 51, 7 October 2020: 74-133

after issuance, there is a statistically and economically significant increase in cash hoarding activity following bond issuance during COVID relative to normal times.

Did the increase in cash reserves result from better operating performance? To test this, we compare balance sheet adjustments across corporate issuers based on their exposure to COVID. If the relative increase in cash reserves reflects superior operating results, we would expect to see a greater increase in cash by firms less affected by COVID shutdowns. Table 6 displays the results, classifying firms into most and least affected sectors based on sector-level employment changes from the BLS. In the most affected sectors, 81% of March 2020 issuance resulted in a significant increase in cash, relative to only 61% in normal times, a large increase of 20pp. Least affected sectors exhibit a more limited difference, from 64% to 72%. That is, the cash increases are likely to be the result of increased savings rather than improved operational performance during this period. These findings lend credence to the hypothesis that a large share of issuance was "precautionary", particularly for firms that experienced worse operational performance due to COVID [Acharya et al., 2012].[16]

Less increase in real investment. Are bond proceeds used towards reinvestment in real activity? To explore this question, we use increases in non-cash assets as a proxy for investment in operating activity. In the frequency analysis reported in Table 7, we find that in normal times, 58% of IG issuers increase non-cash assets by the second quarter following issuance; however, in COVID times, only 18% of issuers did. The results from the post-issuance balance sheet adjustments regressions in Tables 4 and 5 allow us to compare the differences within firm. We find that firms show an increase in non-cash assets of 7% and 8% in the first and second quarter end, respectively, following a bond issuance in normal times. COVID-era issuance, however, is not followed by an increase in reinvestment. Instead, there is a negative (but statistically insignificant) effect on non-cash assets in the two quarters following bond issuance during COVID. Evidently, bond capital raised during COVID is unlikely to

[16]In addition, in the least affected sectors, there was more increase in leverage than in normal times, which can potentially worsen debt overhang going forward. [Brunnermeier and Krishnamurthy, 2020a, Crouzet and Tourre, 2018]

Covid Economics 51, 7 October 2020: 74-133

Table 6 – Comparing exposed vs. less exposed issuers, normal vs. COVID times

	Percent of firms issuing in given period with:					
	Normal issuers			**COVID issuers**		
Least affected firms						
	Increase	**Same**	**Decrease**	**Increase**	**Same**	**Decrease**
Cash	64%	6%	31%	72%	0%	28%
Non-cash assets	39%	51%	11%	20%	51%	29%
Leverage	53%	24%	23%	66%	16%	18%
Net working capital	55%	3%	42%	53%	0%	47%
Most affected firms						
Cash	61%	8%	31%	81%	0%	19%
Non-cash assets	41%	51%	8%	4%	69%	28%
Leverage	50%	24%	26%	50%	26%	24%
Net working capital	50%	6%	44%	26%	0%	74%

Notes: Exposure to COVID is measured by the firm's industry. Most affected firms belong to a NAIC3 industry that had a higher than median loss in employment from January - April 2020. Least affected firms belong to a NAIC3 industry that had a lower than median loss in employment from January - April 2020. COVID issuers are firms issuing in March 2020. Normal issuers are firms issuing in March 1 - June 30, 2017-2019. Normal firms are also categorized into "least affected" and "most affected" based on how impacted their NAIC3 category was during COVID. An "increase" ("decrease") is when the balance sheet characteristic has a quarter-on-quarter change of greater (less) than 2%; "Same" refers to a smaller than 2% change in either direction.

be immediately reinvested in the firms. An illuminating example is Chevron, which raised \$650 million in bond capital on March 24th, and explicitly said that it would not use these funds for investment. Instead, it plans to reduce its 2020 capital spending plan by \$4 billion (or 20%) in response to the crisis. Chevron CEO said: "We are taking actions expected to preserve cash, support our balance sheet strength, lower short-term production, and preserve long-term value." This suggests that the fast rebound in bond issuance will not lead to a correspondingly quick rebound in investment, output or employment.

Equity payouts and issuance: Next, we explore whether firms use bond proceeds to pay out shareholders. The last column in the post-issuance balance sheet adjustments regressions of Table 5 show that in normal times, firms follow up bond issuance with an increase in total equity payouts.[17] However, during COVID, bond issuers were more likely to reduce equity payouts in the quarters following a bond issuance. We further employ the event study analysis from specification (2) and map out the evolution of a firm's dividend payouts and share repurchases relative to the bond issuance timeline. Figure 8 show that the reduction in equity payouts includes both decreases in dividends and equity repurchases following bond issuance, consistent with news of firms slashing dividends to manage cash flow during the crisis.[18] These reductions in equity payouts that coincide with bond issuance are consistent with the hypothesis that these issuers were liquidity constrained during COVID.

As further evidence that issuers during COVID were facing liquidity shortfalls in greater magnitude than issuers in normal times, we find that some bond issuers during COVID also tapped equity markets. Indeed, in our sample of bond issuers from March to May 2020, 52% of bond issuers also issued equity. From our event study analysis in Figure 8, we find that while in normal times, equity issuance does not necessarily coincide with bond issuance, the riskiest bond issuers during COVID were more likely to issue equity right before or concurrently with

[17]Total equity payouts is defined as equity purchases plus dividend payments minus equity issuance proceeds, scaled by total assets. See Table 15 for more details on variable definitions

[18]Many firms reduced payouts to equity: Ford Motor Co. and Freeport-McMoRan Inc. suspended dividend payments while AT&T halted share repurchases. "Companies Race for Cash in Coronavirus Crisis", *Wall Street Journal*, 03/23/2020

Covid Economics 51, 7 October 2020: 74-133

Covid Economics 51, 7 October 2020: 74-133

Table 7 – Comparing IG vs. HY exposed issuers, normal vs. COVID and 2008 Crisis times

	Percent of firms issuing in given period with:					
	IG issuers			**HY issuers**		
Normal times						
	Increase	Same	Decrease	Increase	Same	Decrease
Cash change (Q1)	62%	11%	38%	61%	6%	37%
Cash change (Q2)	62%	9%	39%			
Non-cash assets (Q1)	47%	52%	9%	46%	46%	10%
Non-cash assets (Q2)	58%	36%	12%			
Leverage (Q1)	63%	27%	16%	55%	33%	14%
Leverage (Q2)	57%	24%	26%			
COVID times						
Cash change (Q1)	76%	4%	24%	74%	5%	21%
Cash change (Q2)	88%	4%	8%			
Non-cash assets (Q1)	16%	60%	27%	8%	48%	44%
Non-cash assets (Q2)	18%	47%	35%			
Leverage (Q1)	65%	23%	15%	50%	26%	24%
Leverage (Q2)	61%	22%	17%			
Crisis times						
Cash change (Q1)	67%	3%	33%	86%	0%	14%
Cash change (Q2)	77%	2%	22%			
Non-cash assets (Q1)	39%	44%	20%	14%	45%	41%
Non-cash assets (Q2)	60%	23%	18%			
Leverage (Q1)	57%	15%	31%	25%	35%	40%
Leverage (Q2)	52%	8%	42%			

Notes: IG firms are rated BBB- and above. COVID issuers are firms issuing in March 1 -May 31 2020. Normal issuers are firms issuing in March 1 -May 31, 2017-2019. Crisis issuers are firms issuing in March 1 -May 31, 2009. An "increase" ("decrease") is when the balance sheet characteristic has a quarter-on-quarter change of greater (less) than 2%; "Same" refers to a smaller than 2% change in either direction.

a bond issuance. This suggests that these firms went to great lengths to raise cash during COVID.

2.2.3 How did use of bond capital during COVID compare to the 2008 crisis?

Has the way firms use bond capital changed since the financial crisis? To explore this question, we extend the dataset back to 2000, and look at post-issuance balance sheet adjustments during the 2008 financial crisis. We find some evidence that the use of bond capital for liquidity purposes is unique to the COVID crisis, suggesting a secular shift in the way firms use bond markets. Table 8 and 9 show the changes in balance sheet characteristics for firms issuing between September 1, 2008 and June 30, 2009, using the baseline specification. While firms did increase cash in the two quarters following bond issuance in normal times from 2000-2019, firms issuing during the 2008 financial crisis actually shed more cash in the quarters following the crisis. This suggests that cash hoarding exhibited by COVID-era issuers did not occur during the 2008 crisis. If anything, there was a small (though statistically insignificant) increase in non-cash assets following bond issuance, suggesting that crisis-era firms did use bond capital to reinvest in real activity.

Overall, we find that, unlike in normal times, issuers used bond capital for liquidity purposes during COVID. In COVID times, firms issued bonds earlier than usual, and paid out less to shareholders. Importantly, COVID-era bonds were used to build up cash reserves rather than to reinvest in the firm. These results show that, at a broad level, credit can play a dual role: it can fund investment or it can build liquidity buffers, a pattern that has been well documented for bank credit. We argue that the bond market also plays both roles, and that the relative importance of these roles changes with the state of the economy.

Covid Economics 51, 7 October 2020: 74-133

Covid Economics 51, 7 October 2020: 74-133

Table 8 – Firm balance sheets right after issuance: 2008 Crisis

	(1) Cash/Assets$_{t+1}$	(2) Leverage$_{t+1}$	(3) Non-cash assets$_{t+1}$	(4) Equity payout$_{t+1}$
Issue$_t$	0.00892***	0.0217***	0.0615***	0.164***
	(0.000748)	(0.00283)	(0.00542)	(0.0278)
Crisis×Issue$_t$	-0.0104***	-0.000260	0.0351	-0.00376
	(0.00281)	(0.0106)	(0.0328)	(0.178)
ROA$_{t+1}$	0.0278	-0.471**	-2.258**	9.162***
	(0.0716)	(0.177)	(0.986)	(2.133)
Qtrly stock return$_{t+1}$	0.00000746	-0.0000507	-0.000132	-0.000219***
	(0.00000614)	(0.0000392)	(0.000127)	(0.0000651)
Firm FE	✓	✓	✓	✓
Ind x Yr-Qtr FE	✓	✓	✓	✓
Observations	28119	27025	28119	25551
R-squared	0.776	0.794	0.946	0.765

Notes: Observations are at the firm-quarter level, from 2000Q1-2019Q4. We exclude observations where the firm issued in the previous 2 quarters. $Issue_t$ is an indicator variable that equals one if that firm issues a bond in that quarter; $Crisis \times Issue_t$ is an indicator variable for the reference bond being issued after September 1, 2008 and before June 30, 2009. "Cash/Assets" is che/at, winsorized at the 1% level. "Leverage" is total debt / total assets. "Non-cash assets" is the log of $(at - che)$. "Current debt / debt" is debt due within one year ($dd1$) divided by total debt. Total debt is $dltt + dlc$, total LT debt plus debt in current liabilities. "Equity payout" is the log of (equity purchases + dividend payments), or $log(prstkc + dv)$. Firm controls include return on assets (operating income before depreciation divided by total assets, or $oibdp/at$) and one quarter stock market return $((prrc_t/prrc_{t-1}) - 1)$. We include firm (gvkey) fixed effects and industry (naic2) x quarter fixed effects. Standard errors, in parentheses, are clustered as the industry level.

Covid Economics 51, 7 October 2020: 74-133

Table 9 – Firm balance sheets two quarters after issuance: 2008 Crisis

	(1) Cash/Assets$_{t+2}$	(2) Leverage$_{t+2}$	(3) Non-cash assets$_{t+2}$	(4) Equity payout$_{t+2}$
Issue$_t$	0.00892***	0.0217***	0.0615***	0.164***
	(0.000748)	(0.00283)	(0.00542)	(0.0278)
Crisis×Issue$_t$	-0.0104***	-0.000260	0.0351	-0.00376
	(0.00281)	(0.0106)	(0.0328)	(0.178)
ROA$_{t+2}$	0.0278	-0.471**	-2.258**	9.162***
	(0.0716)	(0.177)	(0.986)	(2.133)
Qtrly stock return$_{t+2}$	0.00000746	-0.0000507	-0.000132	-0.000219***
	(0.00000614)	(0.0000392)	(0.000127)	(0.0000651)
Firm FE	✓	✓	✓	✓
Ind x Yr-Qtr FE	✓	✓	✓	✓
Observations	28119	27025	28119	25551
R-squared	0.776	0.794	0.946	0.765

Notes: Observations are at the firm-quarter level, from 2000Q1-2019Q4. We exclude observations where the firm issued that quarter or two quarters prior. $Issue_t$ is an indicator variable that equals one if that firm issues a bond in that quarter; $Crisis \times Issue_t$ is an indicator variable for the reference bond being issued after September 1, 2008 and before June 30, 2009. "Cash/Assets" is che/at, winsorized at the 1% level. "Leverage" is total debt / total assets. "Non-cash assets" is the log of $(at - che)$. "Current debt / debt" is debt due within one year ($dd1$) divided by total debt. Total debt is $dltt + dlc$, total LT debt plus debt in current liabilities. "Equity payout" is the log of (equity purchases + dividend payments), or $log(prstkc + dv)$. Firm controls include return on assets (operating income before depreciation divided by total assets, or $oibdp/at$) and one quarter stock market return $((prrc_t/prrc_{t-1}) - 1)$. We include firm (gvkey) fixed effects and industry (naic2) x quarter fixed effects. Standard errors, in parentheses, are clustered as the industry level.

CEPR PRESS

Covid Economics 51, 7 October 2020: 74-133

3 The Crowding Out of Bank Loans

Another important aspect to consider is that the bond market should not be analyzed in a vacuum. Indeed, bond issuers have access to both bond and loan markets. The conventional view tends to argue that banks are the main source of funds in bad times, over capital markets [Gatev and Strahan, 2006, Kashyap et al., 2002, Bolton et al., 2016].

How did bond issuers use loan markets during this episode? To investigate this question, we match our issuance data with information on each issuer's debt composition from Capital IQ. These data contain information on amount outstanding of different debt instruments, including revolving credit, term loans, leases and commercial paper. It also includes information on undrawn credit lines, important sources of liquidity for firms [Sufi, 2009, Li et al., 2020, Acharya and Steffen, 2020b], that were available as the COVID crisis unfolded. Note that the debt composition data is reported only at quarter end, so we approximate flows by computing differences between quarters. We break down the analysis into two steps: (i) the first quarter of 2020 (early part of the crisis) and (ii) the second quarter of 2020 (later part of the crisis). We collect data for all firms that issued bonds in March-June 2020 that report their financial statements in U.S. dollars in our analysis. Our merged sample includes 304 firms for which Capital IQ reports data for the first quarter following issuance. We have data for the second quarter for 90% of these firms; for the rest of the firms, this data has not yet been released.

3.1 Bank Borrowing in 2020 Q1

We first show that many issuers left their existing credit lines untouched in the first quarter of 2020, even though the shock did not originate in the banking sector. We find that in aggregate, most (if not all) funds raised in the bond market could have potentially been raised by drawing on outstanding credit lines with banks. For more information on credit line usage during the first part of the crisis, see Acharya and Steffen [2020b] for a study of all public firms and Li et al. [2020] for a study of banks. Greenwald et al. [2020] and Chodorow-Reich

CEPR PRESS

Covid Economics 51, 7 October 2020: 74-133

et al. [2020] study a large panel of firms including SMEs.

As an example, Chevron had $5 billion of its credit line available at the beginning of 2020, yet it still issued $650 million in bonds. We show that Chevron was far from an isolated case, and strikingly, this behavior includes many riskier HY firms. Table 10 tracks the change in debt composition during the first quarter of 2020. The first three rows show the share of firms that, respectively, (i) maxed out of their credit lines (i.e., have revolving credit outstanding larger than 90% of their available credit as of end of 2019), (ii) drew on their credit lines without maxing out, and (iii) did not draw on their credit line. Note that because the data consists of stocks of debt outstanding reported quarterly, these numbers are not completely free of measurement error.[19] The fourth row reports the share of firms that did not receive bank funding, in net, in the first quarter, aggregating revolving credit, term loans and leases. The fifth row reports average draw-down rates, defined as the ratio of additional revolving credit over available credit at the end of 2019.

Looking at the sample of all HY firms that issued between March and June, 38% received no new net bank funding between January and March. Only 21% had maxed out their credit line by end of March, and the average draw-down rate was 41%. Looking beyond credit lines and including new term loans and leases does not change the picture: 38% did not receive new net bank funding in the first quarter that covers the height of the crisis. This implies that many of these riskier firms had available "dry powder" from banks, arranged before the crisis, that they decided not to use early in the first part of the crisis, even though they did not issue any bonds until later in the crisis. The pattern is even more striking when looking at IG firms that issued in March or April, although there is still a risk gradient within this group. Among firms rated BBB (the riskiest IG issuers), 55% left their credit line untouched and 43% did not get additional bank funds, in net, in the first quarter of 2020. Their average

[19]First, our definition of "maxing out" can occasionally incorrectly include firms that signed new credit lines during the COVID crisis. In our exploration, this measurement problem seems to be more pronounced for IG firms. For instance, MacDonald's signed a new credit line of $10B, of which it drew $1B. Second, we can only observe quarter-end balance. If a firm drew on its credit line on March 1st and repaid it by March 31st, our data would not capture this behavior.

Table 10 – Bank borrowing in 2020Q1 for bond issuers

	HY Share	IG, BBB Share	IG, A or above Share
Maxed out CL	0.21	0.081	0.034
Drew some CL	0.42	0.37	0.15
Did not draw CL	0.37	0.55	0.81
No net bank funds	0.38	0.43	0.64
Av. drawdown rate	0.41	0.18	0.081

Notes: This table classifies bond issuers based on changes in outstanding debt for different credit instruments during 2020Q1, based on Capital IQ Capital Structure Summary table. Row 1 defines issuers that maxed out credit lines if the increase in Revolving Credit is at least 90% of Undrawn Revolving Credit at the end of 2019. Row 2 defines issuers that drew some of their credit lines if this ratio is between 90% and 0%. Row 3 defines issuers that did not draw if this ratio is 0% or less. Row 4 defines issuers with no net bank funding if there was no increase in the sum of Revolving Credit, Term Loans and Capital Leases. Row 5 defines the draw-down rate as the ratio as the increase in Revolving Credit over Undrawn Revolving Credit at the end of 2019. HY issuers include all U.S. HY firms that issued a bond between March and June that we could merge with Capital IQ information. IG issuers include all U.S. IG firms that issued a bond between March and April that we could merge with Capital IQ information.

Covid Economics 51, 7 October 2020: 74-133

draw-down rate is only 18%. For the safest firms, rated A or above, 81% left their credit line untouched and the draw-down rate was only 8% on average. This difference is consistent with Acharya and Steffen [2020b].

In aggregate, the amount of undrawn bank credit available at the beginning of 2020 was larger than the total funds raised from bond issuance. Table 11 shows the aggregate flows by different types of debt instruments. Our sample of IG firms raised a total of $466 billion in bonds and $125 billion in loans in this period, despite having $685 billion of credit line available at the start of the year. IG firms also borrowed in the form of term loans and commercial paper during this time, although to a much smaller extent than bond issuance. While the gap is smaller for HY firms, it appears that a large majority of funds raised in the bond market could nevertheless have come instead from drawing on existing credit lines. Indeed, HY issuers in our matched sample issued $90 billion while having $142 billion of undrawn credit. Their aggregate draw-downs reach only $53 billion. Figure 10 illustrates this unused aggregate dry powder visually.

Covid Economics 51, 7 October 2020: 74-133

Table 11 – Debt Composition: Aggregate Flows over 2020Q1

	HY Billions of USD	IG, BBB Billions of USD	IG, A or above Billions of USD
Bond issuance	89.8	217.2	248.7
Credit line	53.3	72.7	7.32
Term loan	-1.66	25.2	22.1
Commercial paper	-1.84	4.41	21.1
Undrawn credit EOY 2019	142.3	441.0	243.6

Notes: This table classifies aggregate debt flows based on FISD bond issuance data (Row 1) as well as changes in outstanding debt for other credit instruments during 2020Q1 based and Capital IQ Capital Structure Summary table (Rows 2,3 and 4). Undrawn credit EOY 2019 is the outstanding available Undrawn Revolving Credit at the end of 2019. HY issuers include all U.S. HY firms that issued a bond between March and June that we could merge with Capital IQ information. IG issuers include all U.S. IG firms that issued a bond between March and April that we could merge with Capital IQ information.

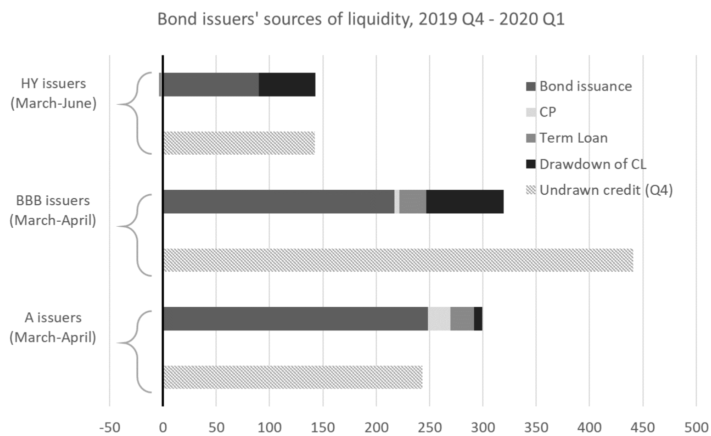


Figure 10 – Visualizing dry powder: Debt Composition Aggregate Flow

Notes: This figure classifies aggregate debt flows based on FISD bond issuance data as well as changes in outstanding debt for other credit instruments during 2020Q1 based on Capital IQ Capital Structure Summary table. Undrawn credit EOY 2019 is the outstanding available Undrawn Revolving Credit at the end of 2019. See Table 11 for underlying numbers.

Covid Economics 51, 7 October 2020: 74-133

3.2 Repaying Bank Loans in 2020 Q2

Next, we examine whether firms use proceeds from bond issuance to repay bank loans. To this end, we investigate changes in firms' debt composition during the second quarter of 2020. The latter part of the COVID period, from April to June, has received less attention, but is particularly revealing.

We find that a large share of issuers that did borrow from their banks early in the crisis repaid by issuing a bond in the following quarter. Panel A for 12 shows the share of bond issuers that repaid in Q2 at least some amount of their Q1 credit line draw-down. Among all HY issuers, two-thirds of these repaid some amount after their bond issuance. In fact, over 40% actually repaid their credit line *in full*, and only a few borrowed additional funds from banks in the second quarter. For example, Kraft Heinz, which was downgraded from IG to junk in February 2020, drew $4 billion from its credit line between February and March. In May, it issued $3.5 billion of bonds (up from a planned $1.5B, due to strong investor demand) and used these funds to repay its credit line in its entirety. In six months, the share of Kraft Heinz's credit coming from banks went from zero to 12% and then back to zero. Kraft Heinz is far from an isolated example: Panel B of Table 12 shows the distribution of credit line repayment as a fraction of either Q1 draw-down or bond issuance, conditional on repaying. Among HY issuers repaying bank loans, the median firm paid back 100% of its Q1 borrowing, representing 60% of their bond issuance. These patterns are similar for IG firms, although a smaller share drew on their credit lines in the first place. 84% of firms that drew-down repaid their bank, with the median also repaying 100%. There is little difference between issuers rated BBB and A or above.

Figure 11 illustrates the cross-section of repayment behavior by plotting credit line draw-down in Q1 against draw-down in Q2 for each firm in our sample. A negative value indicates that the firm paid down a portion of the outstanding credit line. Strikingly, many firms are exactly on the negative forty-five degree line, denoting full repayment within three months. A noticeable number of firms repaid even more, using bonds to pay down bank debt that

Covid Economics 51, 7 October 2020: 74-133

Table 12 – Crowding out bank loans

Panel A: Share of bond issuers repaying credit lines in Q2

	Mean
HY	
Share Repaid some credit line in Q2, condtional on Q1 draw-down	0.61
Share Repaid all credit line in Q2, condtional on Q1 draw-down	0.42
IG, BBB	
Share Repaid some credit line in Q2, condtional on Q1 draw-down	0.84
Share Repaid all credit line in Q2, condtional on Q1 draw-down	0.55
IG, A or above	
Share Repaid some credit line in Q2, condtional on Q1 draw-down	0.85
Share Repaid all credit line in Q2, condtional on Q1 draw-down	0.69

Panel B: Fraction of credit line repayment conditional on repaying

	Mean	25%	50%	75%
HY				
Q2 CL repayment/Q1 CL drawdown (%)	160.1	80	100	120
Q2 CL repayment/Bond issuance (%)	87.3	30	60	108.3
IG, BBB				
Q2 CL repayment/Q1 CL drawdown (%)	154.6	72.0	100	104
Q2 CL repayment/Bond issuance (%)	76.5	30.2	65.8	100
IG, A or above				
Q2 CL repayment/Q1 CL drawdown (%)	101.9	100	100	100
Q2 CL repayment/Bond issuance (%)	81.6	17.6	50	125.0

Notes: Panel A displays the share of bond issuers that repaid some of their credit line balance 2020Q2, based on Capital IQ Capital Structure Summary table, separately by high-yield and investment grade issuers. Panel B displays the fraction of credit line repayment in 2020Q2 relative to 2020Q1 credit line draw-downs (Row 1) or bond issuance in 2020 since March (Row 2), conditional on repaying some positive amount in 2020Q2. The sample includes all U.S. firms that issued a bond between March and June that we could merge with Capital IQ information for Q1 and Q2.

preceded the COVID crisis. Many firms repaid partially, with only a few borrowing more in the second quarter.

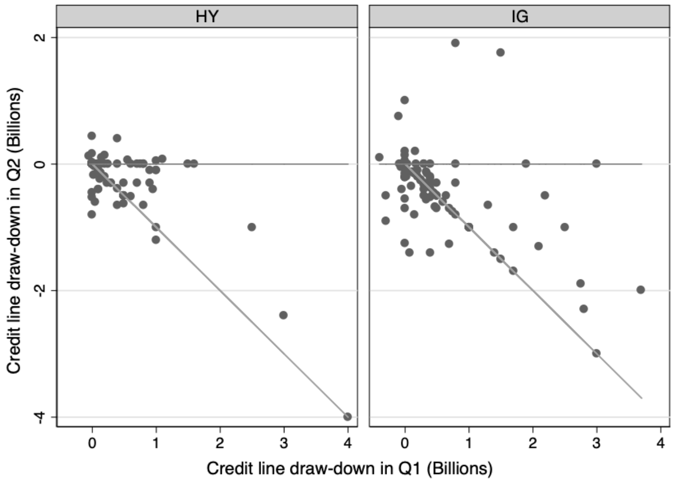


Figure 11 – Visualizing crowding out: Credit line draw-downs in 2020Q2 vs. 2020Q1

Note: This figures plots credit line repayment in 2020Q2 against 2020Q1 credit line draw-downs, based on Capital IQ Capital Structure Summary table, separately by high-yield and investment grade issuers. For ease of interpretion, the figure also displays the negative 45 degree line (exact repayment in Q2) and horizontal line (no change in credit line in Q2). Excludes large outliers Volkswagen, Ford, and GM.

Table 13 displays the aggregate amount of bond proceeds used to repay bank loans between the first and second quarter of 2020. We find that in aggregate, a full quarter of HY firms' bond proceeds went to pay back bank loans, repaying just below half of what was borrowed in the first quarter. The pattern is similar for IG firms, although weaker since a smaller share drew on their credit lines in the first place. BBB firms repaid about 47% of their bank borrowing from Q1, while firms rated A or above repaid over 80%. We estimate that at least \$70 billion was repaid by bond issuers to banks between April and June 2020. This flow of repayment is consistent with aggregate bank lending shown in Figure 12: total bank commercial and industrial lending has fallen by \$180 billion between mid-May and end of June.

This represents a remarkable pattern of debt substitution, whereby firms borrow from bond

Covid Economics 51, 7 October 2020: 74-133

Table 13 – Crowding out of bank loans: Aggregate Flows over 2020Q1 vs. 202Q2

	HY Billions of USD	IG, BBB Billions of USD	IG, A or above Billions of USD
Bond issuance since March 2020	80.4	228.1	241.5
Credit line Q1	49.6	73.3	11.1
Credit line Q2	-19.1	-33.9	-9.61
Term loan Q1	-2.16	24.2	22.6
Term loan Q2	-2.25	-10.2	1.95

Notes: This table classifies aggregate debt flows based on based on FISD bond issuance data (Row 1) as well as changes in outstanding debt for credit lines and term loans based and Capital IQ Capital Structure Summary table. Rows 2 and 4 displays the change between 2019Q4 and 202Q1. Rows 2 and 4 displays the change between 202Q1 and 202Q2. The sample includes all U.S. firms that issued a bond between March and June that we could merge with Capital IQ information for Q1 and Q2.

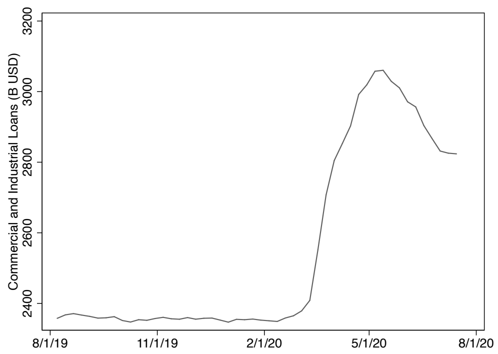


Figure 12 – Aggregate Commercial Lending

Source: Board of Governors of the Federal Reserve System (US), Commercial and Industrial Loans, All Commercial Banks [TOTCI], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/TOTCI, July 29, 2020.

investors and pay back their banks. Results from our regression analysis (3) with firm leverage ratios as the outcome variable suggest that this debt substitution phenomenon is likely unique to bad times. Firms issuing bonds in normal times experience an uptick in leverage in the quarters following raising bond capital (see Table 4 and Table 5). However, firms issuing during COVID times notably have an economically significant incremental decline in leverage in the quarter of bond issuance, nearly wiping out any increase in leverage from the bond

Covid Economics 51, 7 October 2020: 74-133

issuance itself.[20] The event study in Figure 13 shows a similar story, where COVID issuance, unlike issuance in normal times, is followed by statistically insignificant changes in total debt. Altogether, these results suggest that during COVID, firms issue bonds partially to pay down bank loans, rather than to lever up for investment purposes. These firms thus display a revealed preference for outstanding bond capital as a source of liquidity over bank credit.

4 Discussion and Implications

4.1 Interpreting the Findings

How do our findings square with existing views of liquidity provision and corporate borrowing? Our evidence raise two questions: What explains the spike in debt markets activity while the real activity is far from having recovered? And why did bond issuance crowd out bank loans?

We argue that the importance of liquidity-driven bond issuance in bad times is key to explaining the events of the first half of 2020. Note first that credit plays a dual role: firms can borrow for *investment* reasons to fund long-term projects, or they can borrow for *liquidity* reasons to withstand temporary cash-flow shocks. In the context of the COVID crisis, the cash-flow shortfall is expected to last for possibly a few years and is large in absolute value. However, for most industries the net present value of firms' profits was much less affected, consistent with the quick stock market recovery. The optimal firm's response would thus be to borrow today an amount of liquidity equal to the short-fall, consistent with a "permanent cash-flow hypothesis." Liquidity-driven debt issuance thus spikes *because* the real recovery is slow, not in spite of it. In fact, the larger the shock, the larger the issuance volume. On the other hand, investment-driven debt issuance will be delayed.

Moreover, it is well established that for investment purposes, large firms tend to prefer to borrow from capital markets rather than from banks, because they are safer and more

[20] Recall that HY issuers are not included in Table 5 since they issued in 20201Q2 and the Q3 data is not widely available yet.

Covid Economics 51, 7 October 2020: 74-133

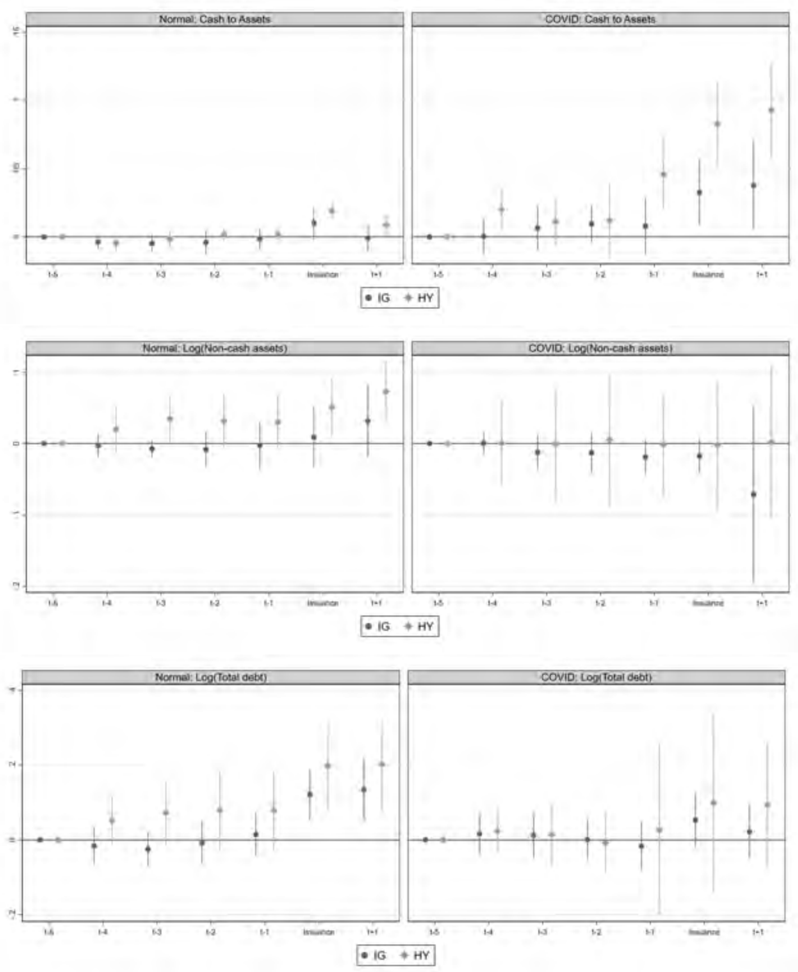


Figure 13 – Coefficient plots: Balance sheet items

Notes: Each point is an estimate of β_{t+m} from the regression
$Y_{fq} = \sum_{m=-4}^{1} \beta_m Issue_{f,t+m} + \alpha_f + \alpha_{ind \times year} + \gamma' X_{fq} + \epsilon_{fq}$, with 95% confidence intervals. The blue points are investment grade firms (rated BBB- and above), while the red points are high yield firms (rated below BBB-). Observations are firm-quarters up to five quarters prior to a bond issuance and two quarters following a bond issuance. *cash_ assets_ w* is *cheq/atq*, winsorized at the 1% level. *log_non_cash_ta* is the log of $(atq - cheq)$. Total debt is $dlttq + dlcq$, total LT debt plus debt in current liabilities. "Normal" times includes bonds issued between 2017-2019, while "Covid" times includes bonds issued between March 1 - May 31, 2020.

Covid Economics 51, 7 October 2020: 74-133

transparent [Petersen and Rajan, 1994, Holmstrom and Tirole, 1997]. However, banks are viewed to have a comparative advantage over capital markets for providing liquidity in bad times. A first argument is that there are synergies with their deposit franchise, either because a flight to safety leads to deposit inflows in bad times [Gatev and Strahan, 2006] or because banks are required to hold large balances of liquid assets [Kashyap et al., 2002]. A second argument is that funding can be committed in advance in the form of credit lines: liquidity insurance allows for more funding relative to a "wait-and-see" policy [Holmström and Tirole, 1998]. In fact, while large U.S. public firms borrow heavily from the bond market for investment, it is well known that virtually all of these firms have standing credit line agreements with banks [Sufi, 2009]. However, our evidence of borrowing patterns during bad times shows that bond markets seem to often dominate bank loans for liquidity purposes as well, despite the fact that bond issuance is not funded by deposits and is not arranged in advance. We thus contribute to the debate on what makes banks "special."[21]

Importantly, we show that bonds were revealed preferred to bank loans in bad times, even though the banking sector was healthy and lending extensively [Li et al., 2020]. We thus also contribute to the debate on what drives the cyclicality of bank versus bond credit. The conventional view, based on the Great Recession, is that firms substitute towards bonds and away from loans in bad times because banks' balance sheets weaken, driving down loan supply [Becker and Ivashina, 2014, Schwert, 2018, Adrian et al., 2013]. While variation in bank loan supply is indisputably important in general, the recent episode suggests it is not the only force at play. More generally, existing works tend to stress the role of credit line draw-downs, leaving little place for liquidity-driven bond issuance [Ivashina and Scharfstein, 2010, Greenwald et al., 2020]. Liquidity provision being different in the current crisis relative to the Great Recession is consistent with the finding in Acharya and Steffen [2020b] that well-rated firms drew-down less in 2020 relative to 2008-09. While we cannot completely rule

[21]Note that this should not be interpreted as saying that bond investors should provide revolving credit lines instead of banks. Simply that the comparative advantage of banks for liquidity provision is smaller than previously thought.

Covid Economics 51, 7 October 2020: 74-133

out that some large firms wanted to draw but were denied funds by their banks, existing evidence points to this supply restriction being limited in 2020. Li et al. [2020] show that banks experienced massive increases in deposits and cash that reached twice as much as the already extraordinary increase in aggregate lending, and that liquidity and capital posed no constraint on banks, in stark contrast to what happened during the 2008 crisis. Moreover, Greenwald et al. [2020] and Chodorow-Reich et al. [2020] show that large firms were able to draw unlike smaller firms.

Why then did firms rely so much on bonds for liquidity in the spring of 2020? We provide a few possible non-mutually exclusive explanations. First, recessions typically imply cash-flow shocks that last for as long as a few years, and firms thus prefer sources of funds that are *committed* for a long period of time. Loans are less attractive in that respect: they have (1) shorter maturities and (2) more restrictive covenants. While the typical loan maturity is four years (see Schwert [2018]), the median IG (HY) bond issued in 2019 is 13 (8) years (see Table 1). Moreover, it is well known that loans have covenants that give lenders discretion to reduce credit [Sufi, 2009, Chodorow-Reich and Falato, 2017, Lian and Ma, 2018, Greenwald et al., 2019, Acharya et al., 2014]. On the other hand, bonds include different covenants that are less intrusive and much more rarely violated passively by borrowers [Green, 2018].[22] For smaller cash-flow shocks in normal times, credit lines could be more attractive as they have smaller set up costs.[23]

[22]In practice, this difference is often measured through the relative prevalence of "maintenance" relative to "incurrence" covenants. Roberts and Schwert [2020] provide the following example: "Consider a leverage covenant restricting the debt-to-EBITDA ratio to remain below four. With a maintenance covenant, should the borrower's debt-to-EBITDA ratio rise above four for any reason, the borrower would be considered in violation of the covenant and in technical default, absent a waiver from the lender. With an incurrence covenant, the borrower must take an action (e.g., issue debt) that generates a debt-to-EBITDA ratio greater than four in order to be in violation. For instance, if the borrower's debt-to-EBITDA ratio rises above four because of an earnings shock, the borrower would not be in violation of the incurrence covenant." Cov-light loans tend to have less maintenance covenants, but they also tend to be term loans. Table 17 in the Appendix shows the distribution of covenants in our sample, comparing common IG and HY bond covenants in normal times vs. COVID times. Both HY and IG bonds have few maintenance covenants. Incurrence covenants are more common in HY bonds, while some became more common across rating categories after the start of COVID.

[23]Note however that commercial paper often tends to dominate credit lines for the safest firms that have access to that market.

Covid Economics 51, 7 October 2020: 74-133

Second, even if bond issuance is not funded by counter-cyclical deposits, investor demand for bonds remained strong during the COVID episode. Using granular data on order books, we find that the demand for IG bonds was high even before the first Fed announcement on March 23rd. Figure 14 shows that the (weekly average) ratio of total order book to amount issued for each bond remained elevated throughout March. Moreover, while Falato et al. [2020] document unprecedented outflows from corporate bond funds in March and early April, the phenomenon was short-lived. Following the Federal Reserve's announced intent to support corporate bond markets on April 9, there were significant net inflows to both HY and IG bond funds that remained very large through August (see Figure 15). A potential rationale is that corporate bonds represent an ideally positioned asset class: they balance (1) investor demand for safe assets and (2) reach for yield. First, corporate bonds, particularly issued by highly rated firms, have limited downside.[24] The Fed's stated support helped buoy the safety of corporate bonds further. At the same time, these bonds pay a high spread over Treasuries, providing an attractive alternative for safe investments in a time when interest rates are at historical lows. For instance, Table 1 shows that median IG bonds pay 142 basis points above U.S. Treasury yields, while HY bonds exceed respective risk-free benchmarks by 381 basis points.

Finally, the macroeconomic environment, and especially the actions of the Federal Reserve, might have reduced the cost of bond financing relatively more than the cost of bank loans. Spreads on credit lines are set ahead of time and thus do not adjust until the agreement expires, unless there is a renegotiation. On the other hand, bond issuance is priced in real time, and many issuers have benefited from historically low rates after the Federal Reserve's unprecedented support to "backstop" the market [Boyarchenko et al., 2020].[25]

[24]The annual default rate for all rating categories BB and above has been well below 1% since 2003. Source: S&P Global "Default, Transition, and Recovery: 2019 Annual Global Corporate Default And Rating Transition Study", April 29, 2020

[25]Note that we are not claiming that bonds are necessarily cheaper than loans, simply that the bond-loan spread could have shifted during this time. In fact, bonds tend to have a higher rate than loans, in large part because they are junior to bank loans. Schwert [2020] uses firm-level variation to estimate the level of the spread in a sample of U.S. firms. Estimating the change in this spread at a high-frequency is however challenging.

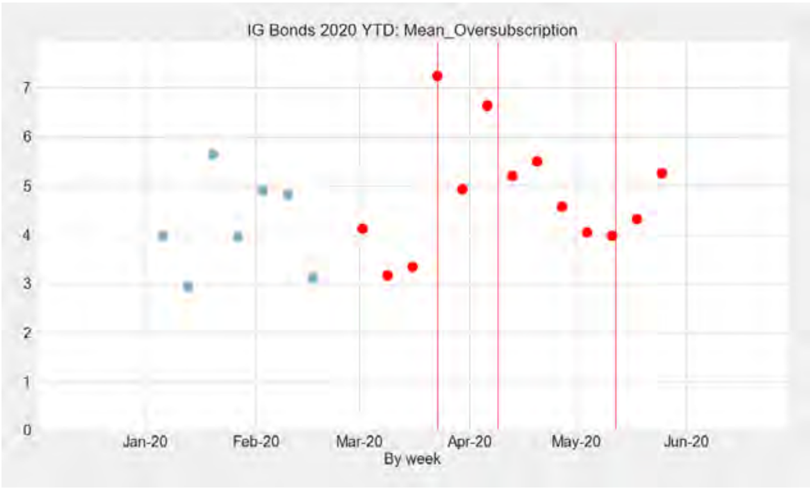


Figure 14 – Oversubscription of orderbooks

Source: Credit Flow Research.
Note: Oversubscription is computed as the ratio of the size of the order book divided by the amount issued, as reported by underwriters of the bond. Reported are weekly averages for each bond's oversubscription ratio

Interestingly, these explanations are consistent with the timing of bank and bond financing over the first half of 2020. Many firms, especially riskier ones, first drew down on their credit lines (when bond markets were closed to high yield firms) and then issued bonds in the second quarter. These dynamics likely arose because (i) in the second quarter, it became clear that the recession would last more than a few months; and (ii) the central bank started to support the market explicitly in April.

4.2 Implications for Monetary Policy

Finally, our findings have important implications for central bank intervention. First, our evidence that the corporate bond market is a key source of liquidity in bad times supports intervention in this market directly. The need to extend traditional lender of last resort policies geared toward the banking sector has been recognized by central banks around the world, and has led to a drastic innovation of credit policy by the Federal Reserve in particular. The

Covid Economics 51, 7 October 2020: 74-133

Covid Economics 51, 7 October 2020: 74-133

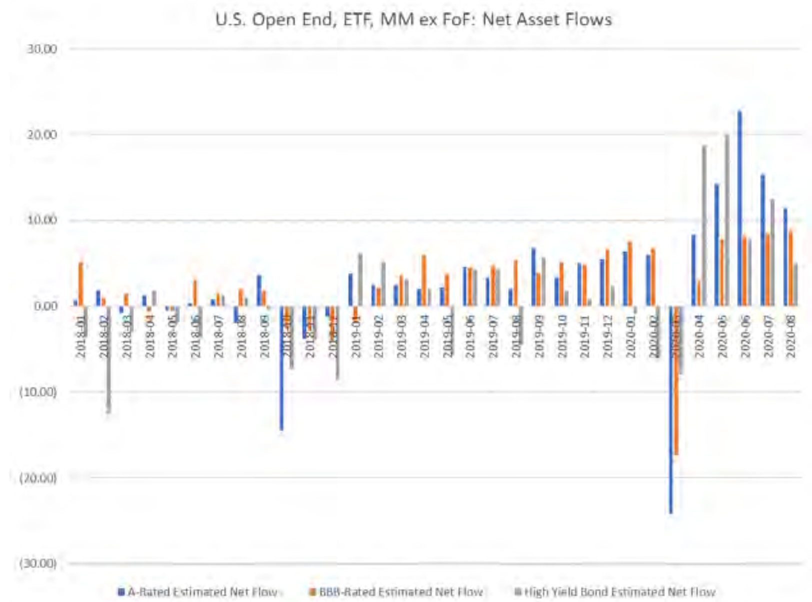


Figure 15 – Monthly net fund flows

Source: Morningstar Direct.

Note: Monthly net fund flows for U.S. open end funds, ETFs, money markets. Excludes funds of funds. Recorded in billions of U.S. dollars. Accessed September 18, 2020.

Covid Economics 51, 7 October 2020: 74-133

European Central Bank (ECB) has also extended the set of eligible corporate bonds that can be pledged as collateral for central bank funding.

However, it is important to account for the crowding out of bank loans when evaluating the aggregate effects of these new public programs on the real economy. For the majority of issuers, propping up bond markets does not alleviate a hard credit constraint, since they have available bank funding. One dollar of bond issued does not equal one dollar of new net borrowing by the firm. Bond markets should not be considered in isolation: bond issuers are active in both bond and loan markets, and they strategically substitute between the two.

Preventing large credit line draw-downs is nevertheless valuable for at least three reasons. First, bond capital guarantees longer-term funding sources for firms, while bank loans have a higher risk of being withdrawn by banks after, for instance, a violation of a maintenance covenant even if the borrower has not "misbehaved." Second, they help weaker issuers "keep their powder dry". Many HY issuers drew down on their credit lines during the weeks in which the bond market was in distressed. Having the option to access liquidity quickly is valuable for these firms. Third, it reduces balance sheet constraints on banks [Grosse-Rueschkamp et al., 2019, Acharya and Steffen, 2020a]. If banks are close to their constraint, that could help credit flow towards smaller firms which are dependent on bank credit.

However, how to weigh these benefits against potential losses on central bank bond holdings or risk of asset price distortions is difficult and an important area for future research. Intuitively, supporting the riskiest segments of the bond markets seem to have the largest benefit, but also the largest potential costs. One observation is that the Federal Reserve's purchase of single-name corporate bonds has skewed towards the largest and safest firms. There are some concerns about whether these firms are the most "constrained." For instance, Table 14 shows that the majority of the Federal Reserve single-name corporate bond portfolio consists of issuers that had access to bank funds that they did not draw on during the crisis. Brunnermeier and Krishnamurthy [2020a] argues that, in the absence of clear frictions, an "IG-corporate QE" program would have limited real effects on these firms.

Covid Economics 51, 7 October 2020: 74-133

Table 14 – Bank borrowing for bond issuers: by share of Fed's single name bond portfolio

	Top 30 (Share)	Others (Share)	Not Purchased by Fed (Share)
Maxed out CL	0.091	0.065	0.16
Did not draw CL	0.76	0.55	0.46
No net bank funds	0.58	0.46	0.42
Av. drawdown rate	0.086	0.18	0.32

Notes: This table classifies bond issuers based on changes in outstanding debt for different credit instruments during 2020Q1, based on Capital IQ Capital Structure Summary table. Column 1 includes the thirty largest bond issuers in our sample in terms of share of the Federal Reserve's single-name bond portfolio holdings as of July 31, 2020. Column 2 includes other bond issuers that are part of the Federal Reserve's single-name bond portfolio holdings as of July 2020. Column 2 includes the remaining bond issuers in our sample.Row 1 defines issuers that maxed out credit lines if the increase in Revolving Credit is at least 90% of Undrawn Revolving Credit at the end of 2019. Row 2 defines issuers that drew some of their credit lines if this ratio is between 90% and 0%. Row 3 defines issuers that did not draw if this ratio is 0% or less. Row 4 defines issuers with no net bank funding is there was no increase in the sum of Revolving Credit, Term Loans and Capital Leases. Row 5 defines the draw-down rate as the ratio as the increase in Revolving Credit over Undrawn Revolving Credit at the end of 2019. The sample includes all U.S. firms that issued a bond between March and July that we could merge with Capital IQ information for Q1 and Q2. Fed purchases are collected from https://www.federalreserve.gov/monetarypolicy/smccf.htm as of August 10, 2020

CEPR PRESS

Covid Economics 51, 7 October 2020: 74-133

5 Conclusion

While the textbook view emphasizes the role of the banking sector in providing funds in bad times, the corporate bond market was at the center of the recent COVID crisis. This paper sheds light on the role of the bond market in providing liquidity in bad times through the lens of bond issuance. Using micro-data on firm balance sheets, we track the usage of bond capital during COVID times and compare to normal times. We show that, propped up by the Fed, the bond market lent extensively to firms in this period, with both investment-grade and high-yield markets reaching historical heights. However, we argue that this V-shaped recovery of bond markets is unlikely to lead to a V-shaped recovery in real activity by documenting two facts on balance sheet adjustments.

First, firms used the bond market differently than in normal times: COVID issuers are more likely to hoard cash rather than invest in real assets. By comparing to the use of bond capital during COVID to the 2008 crisis, we find that this is a new phenomenon, suggesting a shift in the use of corporate bonds. Second, the majority of issuers prefer to issue bonds rather than receive bank loans during these times, exhibiting a revealed preference for bond capital. We find two ways in which bank loans are crowded out by bond issuance. One, firms chose to issue bonds even while their existing credit lines were untouched. Banks had significant capital committed to firms that chose to issue bonds rather than tap the committed capital. Two, firms that did draw down on bank loans then issued bonds in order to pay down the bank loans. In bad times, bond issuance can be liquidity-driven rather than investment-driven, and crowd out bank loans.

These results have important implications for unconventional monetary policy. Our findings that bond capital provides a key source of liquidity to firms in bad times support central bank intervention in corporate bond markets in times of corporate liquidity crises. However, for the majority of issuers, supporting bond markets does not alleviate a hard credit constraint, since they have access to committed bank capital. Instead, the bond market support may help prevent a large aggregate draw-down on bank capital, thus reducing balance sheet

CEPR PRESS

constraints for banks. Relaxing liquidity constraints for banks can thus benefit firms that do not have access to bond markets. We leave the question of how to weigh the benefits of leaving more dry powder for the banks against potential losses from asset price distortion for further research. The rich interactions between corporate debt and the macro-economy is a promising agenda going forward [Brunnermeier and Krishnamurthy, 2020b].

Covid Economics 51, 7 October 2020: 74-133

References

Viral Acharya and Sascha Steffen. Stress tests" for banks as liquidity insurers in a time of covid. *VoxEU. org, March*, 22, 2020a.

Viral Acharya, Sergei A Davydenko, and Ilya A Strebulaev. Cash holdings and credit risk. *The Review of Financial Studies*, 25(12):3572–3609, 2012.

Viral V Acharya and Guillaume Plantin. Monetary easing, leveraged payouts and lack of investment. 2020.

Viral V Acharya and Sascha Steffen. The risk of being a fallen angel and the corporate dash for cash in the midst of covid. *CEPR COVID Economics*, 10, 2020b.

Viral V Acharya, Heitor Almeida, Filippo Ippolito, and Ander Perez-Orive. Bank lines of credit as contingent liquidity: A study of covenant violations and their implications. 2014.

Viral V. Acharya, Heitor Almeida, Filippo Ippolito, and Ander Perez-Orive. Credit lines and the liquidity insurance channel. Working Paper, 2018.

Tobias Adrian, Paolo Colla, and Hyun Song Shin. Which financial frictions? parsing the evidence from the financial crisis of 2007 to 2009. *NBER Macroeconomics Annual*, 27(1):159–214, 2013.

Heitor Almeida, Murillo Campello, and Michael S Weisbach. The cash flow sensitivity of cash. *The Journal of Finance*, 59(4):1777–1804, 2004.

Heitor Almeida, Murillo Campello, Igor Cunha, and Michael S Weisbach. Corporate liquidity management: A conceptual framework and survey. *Annu. Rev. Financ. Econ.*, 6(1):135–162, 2014.

Óscar Arce, Ricardo Gimeno, and Sergio Mayordomo. The effects of the eurosystem's corporate sector purchase programme on spanish companies. Working Paper, 2018.

Thomas W Bates, Kathleen M Kahle, and René M Stulz. Why do us firms hold so much more cash than they used to? *The journal of finance*, 64(5):1985–2021, 2009.

Bo Becker and Victoria Ivashina. Cyclicality of credit supply: Firm level evidence. *Journal of Monetary Economics*, 62:76–93, 2014.

Patrick Bolton and Xavier Freixas. Corporate finance and the monetary transmission mechanism. *Review of Financial Studies*, 19(3):829–870, 2006.

Patrick Bolton and David S. Scharfstein. Optimal debt structure and the number of creditors. *Journal of Political Economy*, 104(1):1–25, 1996.

Patrick Bolton, Hui Chen, and Neng Wang. A unified theory of tobin's q, corporate investment, financing, and risk management. *The journal of Finance*, 66(5):1545–1578, 2011.

Patrick Bolton, Xavier Freixas, Leonardo Gambacorta, and Paolo Emilio Mistrulli. Relationship and transaction lending in a crisis. *Review of Financial Studies*, 29(10):2643–2676, 2016.

Nina Boyarchenko, Anna T Kovner, and Or Shachar. It's what you say and what you buy: A holistic evaluation of the corporate credit facilities. 2020.

Markus Brunnermeier and Arvind Krishnamurthy. Corporate debt overhang and credit policy. In *Conference draft, Brookings Papers on Economic Activity*, 2020a.

Markus Brunnermeier and Arvind Krishnamurthy. The macroeconomics of corporate debt. *The Review of Corporate Finance Studies*, 2020b.

Gabriel Chodorow-Reich and Antonio Falato. The loan covenant channel: How bank health transmits to the real economy. Technical report, National Bureau of Economic Research, 2017.

Gabriel Chodorow-Reich, Olivier Darmouni, Stephan Luck, and Matthew Plosser. Bank liquidity provision across the firm size distribution. Working Paper, 2020.

Nicolas Crouzet. Aggregate implications of corporate debt choices. *Review of Economic Studies*, 85 (3):1635–1682, 2017.

Nicolas Crouzet. Credit disintermediation and monetary policy. *Working Paper*, 2019.

Nicolas Crouzet and Fabrice Tourre. Can the cure kill the patient? corporate credit interventions and debt overhang. Working Paper, 2018.

Olivier Darmouni, Oliver Giesecke, and Alexander Rodnyansky. The bond lending channel of monetary policy. *Available at SSRN 3419235*, 2019.

Fiorella De Fiore and Harald Uhlig. Corporate debt structure and the financial crisis. *Journal of Money, Credit and Banking*, 47(8):1571–1598, 2015.

Roberto A. De Santis and Andrea Zaghini. Unconventional monetary policy and corporate bond issuance. 2019.

Antonio De Vito and Juan-Pedro Gomez. Estimating the covid-19 cash crunch: Global evidence and policy. *Journal of Accounting and Public Policy*, page 106741, 2020.

David J Denis and Valeriy Sibilkov. Financial constraints, investment, and the value of cash holdings. *The Review of Financial Studies*, 23(1):247–269, 2010.

Douglas W. Diamond. Monitoring and reputation: The choice between bank loans and directly placed debt. *Journal of Political Economy*, 99(4):689–721, 1991.

Andrea L Eisfeldt and Tyler Muir. Aggregate external financing and savings waves. *Journal of Monetary Economics*, 84:116–133, 2016.

David Elliott, Ralf R. Meisenzahl, José-Luis Peydró, and Bryce C. Turner. Nonbanks, banks, and monetary policy: Us loan-level evidence since the 1990s. *Working Paper*, 2019.

Aytekin Ertan, Anya Kleymenova, and Marcel Tuijn. Financial intermediation through financial disintermediation: Evidence from the ecb corporate sector purchase programme. *Fama-Miller Working Paper*, pages 18–06, 2019.

Rüdiger Fahlenbrach, Kevin Rageth, and René M Stulz. How valuable is financial flexibility when revenue stops? evidence from the covid-19 crisis. Technical report, National Bureau of Economic Research, 2020.

Covid Economics 51, 7 October 2020: 74-133

Antonio Falato, Itay Goldstein, and Ali Hortaçsu. Financial fragility in the covid-19 crisis: The case of investment funds in corporate bond markets. Technical report, National Bureau of Economic Research, 2020.

Joan Farre-Mensa, Roni Michaely, and Martin C Schmalz. Financing payouts. *Ross School of Business Paper*, (1263), 2018.

C Fritz Foley, Jay C Hartzell, Sheridan Titman, and Garry Twite. Why do firms hold so much cash? a tax-based explanation. *Journal of financial economics*, 86(3):579–607, 2007.

Evan Gatev and Philip E Strahan. Banks' advantage in hedging liquidity risk: Theory and evidence from the commercial paper market. *The Journal of Finance*, 61(2):867–892, 2006.

Erasmo Giambona, Rafael Matta, Jose-Luis Peydro, and Ye Wang. Quantitative easing, investment, and safe assets: The corporate-bond lending channel.

John R Graham and Mark T Leary. The evolution of corporate cash. *The Review of Financial Studies*, 31(11):4288–4344, 2018.

Daniel Green. Corporate refinancing, covenants, and the agency cost of debt. *Covenants, and the Agency Cost of Debt (December 18, 2018)*, 2018.

Daniel Greenwald et al. Firm debt covenants and the macroeconomy: The interest coverage channel. *Manuscript, July*, 2019.

Daniel L Greenwald, John Krainer, and Pascal Paul. The credit line channel. Federal Reserve Bank of San Francisco, 2020.

Benjamin Grosse-Rueschkamp, Sascha Steffen, and Daniel Streitz. A capital structure channel of monetary policy. *Journal of Financial Economics*, 133(2):357–378, 2019.

Valentin Haddad, Alan Moreira, and Tyler Muir. When selling becomes viral: Disruptions in debt markets in the covid-19 crisis and the fed's response. Technical report, National Bureau of Economic Research, 2020.

Michael Halling, Jin Yu, and Josef Zechner. How did covid-19 affect firms' access to public capital markets? *The Review of Corporate Finance Studies*, 2020.

Federic Holm-Hadulla and Claire Thürwächter. Heterogeneity in corporate debt structures and the transmission of monetary policy. Working Paper, 2020.

Bengt Holmstrom and Jean Tirole. Financial intermediation, loanable funds, and the real sector. *the Quarterly Journal of economics*, 112(3):663–691, 1997.

Bengt Holmström and Jean Tirole. Private and public supply of liquidity. *Journal of Political Economy*, 106(1):1–40, 1998.

Filippo Ippolito, Ali K. Ozdagli, and Ander Perez-Orive. The transmission of monetary policy through bank lending: The floating rate channel. *Journal of Monetary Economics*, 95:49–71, 2018.

Victoria Ivashina and David Scharfstein. Bank lending during the financial crisis of 2008. *Journal of Financial economics*, 97(3):319–338, 2010.

Covid Economics 51, 7 October 2020: 74-133

Mahyar Kargar, Benjamin T Lester, David Lindsay, Shuo Liu, and Pierre-Oliver Weill. Corporate bond liquidity during the covid-19 crisis. *Covid Economics*, 27:31–47, 2020.

Anil K. Kashyap, Jeremy C. Stein, and David W. Wilcox. Monetary policy and credit conditions: Evidence from the composition of external finance: Reply. *American Economic Review*, 86(1): 310–314, 1996.

Anil K Kashyap, Raghuram Rajan, and Jeremy C Stein. Banks as liquidity providers: An explanation for the coexistence of lending and deposit-taking. *The Journal of finance*, 57(1):33–73, 2002.

Stéphane Lhuissier and Urszula Szczerbowicz. Monetary policy and corporate debt structure. Working Paper, 2018.

Lei Li, Philip E Strahan, and Song Zhang. Banks as lenders of first resort: Evidence from the covid-19 crisis. *The Review of Corporate Finance Studies*, 2020.

Chen Lian and Yueran Ma. Anatomy of corporate borrowing constraints. *Unpublished working paper*, 2018.

OECD. The world economy at risk. *OECD Economic Outlook, Interim Report March*, 2020.

Maureen O'Hara and Xing Alex Zhou. Anatomy of a liquidity crisis: Corporate bonds in the covid-19 crisis. *Available at SSRN 3615155*, 2020.

Tim Opler, Lee Pinkowitz, René Stulz, and Rohan Williamson. The determinants and implications of corporate cash holdings. *Journal of financial economics*, 52(1):3–46, 1999.

Mitchell A Petersen and Raghuram G Rajan. The benefits of lending relationships: Evidence from small business data. *The journal of finance*, 49(1):3–37, 1994.

Raghuram G. Rajan. Insiders and outsiders: The choice between informed and arm's-length debt. *Journal of Finance*, 47(4):1367–1400, 1992.

Leigh A Riddick and Toni M Whited. The corporate propensity to save. *The Journal of Finance*, 64 (4):1729–1766, 2009.

Michael R Roberts and Michael Schwert. Interest rates and the design of financial contracts. Technical report, National Bureau of Economic Research, 2020.

Michael Schwert. Bank capital and lending relationships. *The Journal of Finance*, 73(2):787–830, 2018.

Michael Schwert. Does borrowing from banks cost more than borrowing from the market? *The Journal of Finance*, 75(2):905–947, 2020.

Kerry Siani. Global demand spillovers in corporate bond issuance: The effect of underwriter networks. *Available at SSRN 3508937*, 2019.

Amir Sufi. Bank lines of credit in corporate finance: An empirical analysis. *The Review of Financial Studies*, 22(3):1057–1088, 2009.

Karamfil Todorov. Quantify the quantitative easing: Impact on bonds and corporate debt issuance. *Journal of Financial Economics*, 135(2):340–358, 2020.

Covid Economics 51, 7 October 2020: 74-133

Internet Appendix

Table 15 – Variable Definitions

Variable	Data source	Compustat code	Definition
Cash	Compustat	che	Cash and short term investments
Cash_assets	Compustat	$\frac{che}{at}$	Cash and short term investments, scaled by total assets. (winsorized at 1%)
Non-cash assets	Compustat	$at - che$	Total assets - cash and short term investments
Leverage	Compustat	$(dltt + dlc)/at$	Total debt / Total assets (winsorized at 1%)
Cpltd_debt	Compustat	$dd1/(dltt + dlc)$	Current portion of LT debt / Total debt (winsorized at 1%)
Net equity payout	Compustat	$(prstkcq - sstkq)/atq$	Net equity purchases, scaled by total assets (winsorized at 1%)
Total payout	Compustat	$\frac{prstkcq - sstkq + dvq}{atq}$	(Net equity purchases + dividends)/total assets (winsorized at 1%)
Log_dvy_q	Compustat	$log(dvq)$	Log of (Quarterly dividend payments (winsorized at 1%))
Log_gross_payout	Compustat	$log(prstkcq + dvq)$	Log of (Equity purchases + dividend payments)
ROA	Compustat	$\frac{oibdpq}{at}$	Operating income / total assets (winsorized at 1%)
Cash flow	Compustat	oancf	Quarterly operating cash flow
Profit	Compustat	ni	Quarterly net income
Credit spread	Mergent FISD	treasury_spread	Credit spread relative to benchmark US Treasury (basis points)
Yield to maturity	Mergent FISD	offering_yield	Yield to maturity on bond at issuance
Rating	Mergent FISD		Credit rating at issuance by Moody's, S&P, and Fitch: median if 3 ratings, minimum if 2 ratings; see Table 16

Notes: Quarterly ratios are winsorized at 1%.

Covid Economics 51, 7 October 2020: 74-133

CEPR PRESS

Covid Economics 51, 7 October 2020: 74-133

Table 16 – Credit Rating Legend

Moody's	S&P	Fitch	Numerical
Aaa	AAA	AAA	22
Aa1	AA+	AA+	21
Aa2	AA	AA	20
Aa3	AA-	AA-	19
A1	A+	A+	18
A2	A	A	17
A3	A-	A-	16
Baa1	BBB+	BBB+	15
Baa2	BBB	BBB	14
Baa3	BBB-	BBB-	13
Ba1	BB+	BB+	12
Ba2	BB	BB	11
Ba3	BB-	BB-	10
B1	B+	B+	9
B2	B	B	8
B3	B-	B-	7
Caa1	CCC+	CCC+	6
Caa2	CCC	CCC	5
Caa3	CCC-	CCC-	4
Ca	CC	CC	3
C	C	C	2
C	D	D	1

Table 17 – Covenants

	IG: normal	HY: normal	IG: covid	HY: covid
Maintenance covenants:				
Bondholder_Protective_RatingDeclineTrigger	0.0%	0.0%	0.0%	0.0%
Issuer_Restrictive_MaintenanceNetWorth	0.0%	0.0%	0.0%	0.0%
Issuer_Restrictive_NetEarningsTest	0.0%	0.0%	0.0%	0.0%
Incurrence covenants:				
Bondholder_Protective_AssetSaleClause	0.0%	25.8%	0.7%	49.3%
Bondholder_Protective_ChangeofControlPut	62.9%	93.0%	52.9%	92.8%
Issuer_Restrictive_ConsolidationMerger	72.8%	78.0%	91.7%	95.7%
Issuer_Restrictive_DividendsRelatedPayments	0.0%	37.1%	0.5%	36.2%
Issuer_Restrictive_SaleofAssets	70.2%	74.7%	91.5%	95.7%

Notes: Computes percentage of bonds that report covenants that have each covenant. Source: Mergent FISD, http://bv.mergent.com/view/scripts/MyMOL/index.php, retrieved September 20, 2020.

Covid Economics Issue 51, 7 October 2020

Job loss and behavioral change: The unprecedented effects of the India lockdown in Delhi[1]

Kenneth Lee,[2] Harshil Sahai,[3] Patrick Baylis[4] and Michael Greenstone[5]

Date submitted: 28 September 2020; Date accepted: 30 September 2020

Covid Economics 51, 7 October 2020: 134-158

On March 24, 2020, India's Prime Minister announced the world's largest COVID-19 lockdown, bringing to a near-halt the economic and social lives of more than one billion Indian residents. This paper quantifies the economic impacts and behavioral changes induced by this unprecedented policy using two unique data sources: Facebook mobility data and a representative sample of previously surveyed low income Delhi households. Compliance with the lockdown was widespread: intra-city movement declined by 80% following the announcement. The economic consequences have been accordingly severe, with income and days worked falling by 86 and 72% respectively. Nevertheless, observance of public health directives was high: mask usage rose by 73 percentage points and handwashing became nearly universal, while time spent outdoors and smoking both declined. We also show how government-provided social assistance may have averted more dire predictions of widespread famine, resource scarcity, access to medical care, and security. But the declines in mental health and the near-exhaustion of personal savings, amidst a rising infection rate, indicate an important and evolving role for policy-makers as the crisis continues.

1 We are grateful to Chethana A., Rongmon Deka, Shipra Karan, Ian Pitman, Meghna Singh, Rathan Sudheer, and Ramya Teeparthi for excellent research assistance. All errors are our own. This work was supported by the Tata Center for Development at the University of Chicago.

2 Energy Policy Institute at the University of Chicago.
3 University of Chicago.
4 University of British Columbia.
5 University of Chicago.

1 Introduction

In response to the rapid spread of COVID-19, Prime Minister Narendra Modi announced the world's largest coronavirus lockdown on March 24, 2020, forcing nearly 1.3 billion people to halt virtually all social and economic activities. The policy immediately altered patterns of daily life across India, as factories, non-essential shops, religious buildings, parks, and transportation hubs were abruptly closed. In this paper, we describe the effects of the lockdown on a representative sample of mostly poor and non-migrant workers in Delhi. We use data collected before and during the COVID-19 crisis to document how the lockdown affected economic and behavioral outcomes.

We contribute to a rapidly growing literature describing the economic impacts of COVID-19 and the lockdowns and social distancing policies that have been introduced around the world to combat its spread. This literature, along with the daily reports in the news media, has placed a spotlight on the enormous effects of the pandemic on employment, consumption, economic output, and the environment, particularly in developed countries.[1] However, there is comparatively less information thus far on the microeconomic impacts of the COVID-19 lockdowns in developing countries like India, and even less on compensatory changes in health-related behaviors.[2] Personal habits and behaviors, like frequent handwashing, social distancing, and mask-wearing in public, have been identified as key actions that individuals could take to reduce the likelihood of infection.

It is widely believed that the impacts of COVID-19 will be the most severe for poor individuals living in the densely populated urban settlements and slums of the developing world. On Day 1 of the India lockdown, for example, some predicted that the resulting economic collapse would be especially devastating for the bottom 50 percent of workers in the informal sector (e.g. Singh et al. 2020). Within two weeks into the lockdown, stories of casual workers in Delhi being forced to survive on food distributed by the government had become common (e.g. Yadav 2020).

There are several reasons why the urban poor would be especially vulnerable to the health and economic impacts of COVID-19 and the policies designed to slow its spread. For instance, the urban

[1]For example, in the United States, Baker et al. (2020) show how households radically altered their spending in response to the pandemic; Bartik et al. (2020) describe the impact of COVID-19 on employment and business closures using surveys of small businesses; and Coibion et al. (2020) describe the impacts on employment using surveys of households. In the area of energy and the environment, Burlig, Cicala, and Sudarshan (2020) measure the impact of COVID-19 on electricity consumption and particulate pollution in various countries around the world.

[2]In contemporaneous work in India, Afridi, Dhillon, and Roy (2020) use cross-sectional survey data to reveal large, negative labor impacts for a sample of primarily self-employed and wage laborers in Delhi, a population group that is similar to the one studied in this paper; while Bertrand, Krishnan, and Schofield (2020) use data from a nationally representative survey to show that 84 percent of Indian households experienced a decrease in income and have limited resources to continue coping with the economic situation. Deshpande (2020) uses nationally representative household panel data to show that Indian women were more likely to be unemployed than Indian men after the lockdown.

Covid Economics 51, 7 October 2020: 134-158

Covid Economics 51, 7 October 2020: 134-158

poor are likely to face higher rates of transmission risk, not only because they live in overcrowded neighborhoods, but also because the quality of water and sanitation in these neighborhoods tends to be poor (Marx, Stoker, and Suri 2013). In addition, they may face greater mortality risk due to higher rates of pre-existing health conditions (Zheng et al. 2020). Moreover, in Delhi, most people have endured lifelong exposures to extreme levels of air pollution, and higher COVID-19 mortality rates have been associated with higher levels of air pollution (Wu et al. 2020). Finally, the urban poor may be more prone to the economic shocks associated with extreme social distancing mandates, due to their likely employment in wage-paying occupations which tend to require physical work, and low levels of savings which are necessary for smoothing out short-term fluctuations in income.

We report on the effects of the India lockdown in Delhi. Using Facebook mobility data, we first show that general intra-city movement dropped to less than 20 percent of normal following the lockdown announcement. We then shift our focus to our representative sample of 1,392 mostly poor and non-migrant workers in Delhi, consisting of people living in low-income, informal settlements and municipal public bus commuters. Using survey data collected both before and during the lockdown, we show that weekly income and days worked fell by 86.2 and 72.2 percent, respectively, by mid-May. While the impacts were equally large across all income quartiles in our sample, they were larger for those employed in daily wage-paying occupations (as opposed to occupations that pay monthly salaries). Weekly income and days worked increased as lockdown restrictions were relaxed, but they remained well below baseline levels in mid-June.

At the same time, we observe widespread adoption of recommended public health directives. For instance, mask usage increased by 72.8 percentage points (pp); time spent indoors increased by 50.5 pp; smoking decreased by 12.8 pp; and regular hand washing increased by 10.0 pp. The magnitudes of these effects are large, and in certain cases, rival or exceed the impacts of past interventions (to increase handwashing or encourage mask usage, for example) documented in the development economics literature (see, e.g., Kremer and Zwane 2007; Clasen et al. 2014; Hussam et al. 2019; Baylis et al. 2020). Focusing on mask usage in particular, we use a combination of survey data and Twitter data to highlight the potential role of extreme fear and unparalleled media coverage in driving these large shifts in behavior.

In our data, we do not observe substantive changes over the course of the lockdown in rates of hunger, product scarcity, ability to access medical care, or security. We also find suggestive evidence that the negative income shocks were mitigated by the Delhi government's widespread food assistance programs, which were accessed by nearly half of our sample. Our data suggest that although the India lockdown had a large effect on movement and finances, it did not prevent people from accessing bare essentials. That said, there are warning signs in the data. Relatively high rates of depression, challenges in food supply chains, and dwindling levels of savings pose serious concerns about the cumulative effects of subsequent lockdowns. In addition, it remains to be seen whether the large behavioral changes we observe will persist as the novelty, fear, and media coverage of COVID-19 subside, or if access to basic necessities is substantially compromised.

In the next section, we describe the timeline of India's lockdown and describe its impact on general intra-city movement in Delhi. We then describe our data collection and discuss the representativeness of our sample. Next, we summarize the leading patterns that emerge from our data. The final section concludes and highlights some of the questions that could be answered through further surveys and research.

2 Context

Timing of the India lockdown The first case of COVID-19 in India was reported on January 30, 2020. By mid-March, there were just over 100 confirmed cases in the country. As the fear of contagion grew, India acted decisively, and at an early stage relative to its case count. On March 19, Prime Minister Modi announced a 14-hour voluntary public curfew (the Janata curfew) that would take place on March 22. Although the Janata curfew was voluntary, it was believed to be widely followed (as reported by HT Correspondents 2020). Then, on March 24, the Prime Minister ordered all of India to be locked down, severely restricting the movement of and halting nearly all social and economic activities for 1.3 billion people. The lockdown, which was initially planned for 21 days, came into effect just hours following the announcement. On April 14, the Prime Minister extended the lockdown for another 19 days, until May 3; and then on May 1, he extended it again for another 14 days, to May 17. On May 3, the Delhi government announced that it would begin to allow certain relaxations permitted by the Ministry of Home Affairs, including for example, allowing non-essential workers and private offices to run at limited capacity, and greater movement of private vehicles. On May 30, the central government announced plans for the gradual easing of the lockdown, which would begin on June 8. In summary, the majority of Delhi residents lived under lockdown from March 22 to June 8 with gradual easing of restrictions starting in May, even as strict measures continue to be

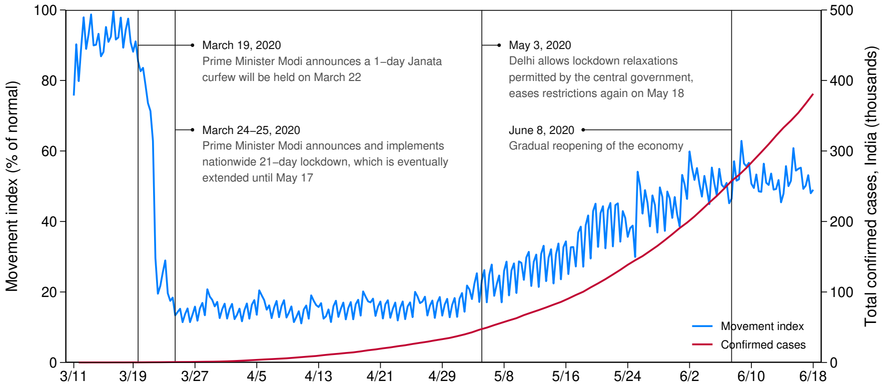
Covid Economics 51, 7 October 2020: 134-158

imposed in certain areas[3].

The sharp reduction in intra-city movement The India lockdown had a profound impact on social distancing. This can be demonstrated by the effect it had on reducing intra-city movement in Delhi, a plausible proxy for social distancing. To show this, we employ data capturing the movement of over two million Facebook users in Delhi who have enabled location services on their mobile applications. The data, which is made available to researchers and non-governmental organizations through an agreement with Facebook Data for Good, allow us to track intra-city movement at a fine spatial granularity and high temporal frequency (Maas et al. 2019).

Specifically, we aggregate Facebook data from the user-level to 166 administrative regions in the Delhi-National Capital Territory. The data capture the number of users that move from one region to another, during three equally spaced windows each day, representing on average 750 region-to-region movement vectors (ranging from 0.1 to 30 kilometers) every eight hours.

Figure 1 — Impact of the lockdown on general intra-city movement in Delhi



Source: Movement data is from Facebook Data for Good (Maas et al. 2019); Data on COVID-19 cases is from covid19india.org.
Notes: We plot the movement index, which summarizes the intra-city movements of over two million users of the Facebook mobile app (which tracks locational data), against the total number of confirmed COVID-19 cases in India. Specifically, the movement index is constructed by calculating the average number of people who moved from one part of the city to another as a fraction of a baseline value for the period. Baseline values are defined as the average movement for the same time-of-day and day-of-week, in the 45 days preceding February 24, 2020, the start of available Facebook data.

In Figure 1, we demonstrate the reduction in movement over the course of the India lockdown,

[3]Containment zones, specific areas demarcated by local authorities according to health guidelines, remain under lockdown restrictions. At the time of this writing, Delhi has over 900 active containment zones.

Covid Economics 51, 7 October 2020: 134-158

CEPR PRESS

Covid Economics 51, 7 October 2020: 134-158

relative to a baseline that is defined as average movement for the same time-of-day and day-of-week, during the 45 days preceding February 24, 2020. We restrict our focus to a subset of 78 regions for which there is a balanced panel.[4] The sharp reduction in movement first appeared after March 19, following the announcement of the Janata Curfew. By March 22, movement was down to around 30 percent of normal. By March 24, it was down to between 10 and 20 percent of normal, where it remained until May 3, when some of the lockdown restrictions first began to be lifted. By late-May, movement reached around 50 percent of normal, where it remained until mid-June.

These data correspond to people who use smartphones, the Facebook mobile application, and have access to reliable internet. It is likely that the data is representative of a higher income group than the population we study in this paper, which we describe in detail in the following section. That said, in our results, we show how respondents in our sample reported similar reductions in social interactions, consistent with the pattern shown in Figure 1.

3 Data

Sample selection and data collection The sampling frame for this study consists of mostly poor and non-migrant workers living in Delhi, and some of the surrounding urban areas. The sampling frame combines two subsamples established in recent years. The first subsample, which we refer to as "low-income neighborhoods" (n = 3,018), was created in 2018 and captures individuals residing in poor, informal settlements across Delhi.

To create this sample, we first consulted the list of Jhuggie Jhopri ("J.J.") Squatter Settlements/Clusters provided by the Delhi Urban Shelter Improvement Board. At the time, this was the only publicly available list of slum clusters or squatter settlements in Delhi. Using this list, we randomly selected a number of sampling points (i.e., locations where enumerators could begin administering in-person surveys) located around the center of each J.J. cluster, in proportion to the cluster's population size. We excluded sampling points that were deemed to no longer be slums or squatter settlements (due to urban development), using a combination of satellite images and in-person checks. This left us with roughly 1,200 sampling points, around which our team of enumerators enrolled individuals into our study and administered a brief social and economic questionnaire. These surveys were carried out between October and December 2018.

The second subsample in our sampling frame, which we refer to as "public bus commuters" (n

[4]An unbalanced panel produces qualitatively similar results.

= 2,110), was created in 2019 and captures individuals who use the public bus system in Delhi, and its neighboring, satellite cities, Gurgaon and Noida. To create this sample, we randomly selected 120 bus stops operated by the Delhi Transport Corporation, 18 bus stops from routes operated by the Noida Metro Rail Corporation, and 79 bus stops from routes operated by Gurgaon Metropolitan City Bus Limited. Our team of enumerators then visited each of these bus stops, repeatedly and at randomized points in the day, to administer a brief social and economic questionnaire. These surveys were carried out between October and December 2019. Importantly, we asked several identical questions on demographic characteristics and health behaviors in both surveys [5].

In total, there are 5,128 individuals in our sampling frame, including 3,018 individuals from low-income neighborhoods, and 2,110 public bus commuters. We pool together survey data from these samples to represent the pre-COVID-19, or baseline, levels of the economic and behavioral variables.

In the days following the announcement of the India lockdown, we launched additional rounds of telephone surveys, carried out between March 27 and April 19 ("Round 1"), April 25 and May 13 ("Round 2"), and May 26 and June 18 ("Round 3"), respectively. In each round, our team of enumerators contacted individuals from the sampling frame in a randomized order and administered a brief social and economic questionnaire, tailored to capture COVID-19-related outcomes.

In our final dataset, we include 1,392 individuals who are considered to be active participants in the labor force, meaning they were employed in at least one survey round over the course of the study. Note that we surveyed 1,744 individuals in total, including both active participants in the labor force and non-workers (e.g., homemakers, students, etc.), and when we include the full sample in our analyses, the findings are the same. That said, for the purposes of this paper, we focus on the workers for whom there is both employment and behavior data. We address selection concerns due to attrition from the sampling frame and attrition across survey rounds by presenting below alternative methods to calculate our main results and comparisons of baseline characteristics between individuals who did and did not consent[6].

Sample characteristics and representativeness Broadly, the population group we study can be characterized as representative of mostly poor, non-migrant, male workers (of working age) living in

[5]Both subsamples were created for Baylis et al. (2020), which estimates the demand for clean air in Delhi.

[6]We were able to successfully contact 68 percent of the sampling frame, the remainder of which could not be enrolled for a number of reasons, including unanswered calls, and non-functioning phone numbers. Of those we contacted, 66 percent consented to be surveyed. We further restrict this pool to individuals whose names matched and who were living in Delhi, Gurgaon, or Noida as of March 2020. In the dataset of 1,392 active participants in the labor force, all were surveyed at baseline, 1,052 were surveyed in Round 1; 752 were surveyed in Round 2; and 640 were surveyed in Round 3.

Covid Economics 51, 7 October 2020: 134-158

CEPR PRESS

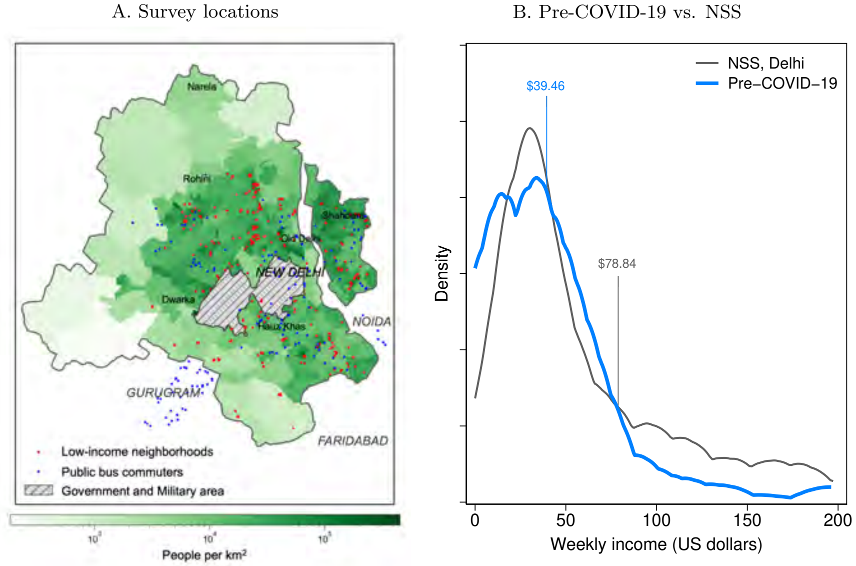
Delhi. In Appendix Table A1, column 1, we summarize the key characteristics at baseline for these individuals. The majority of individuals are long-term residents of Delhi (87.3 percent) and can thus be considered non-migrants. Only 20.3 percent of our sample reported that they had either migrated or were planning to migrate at some point during the lockdown. Compared to developed country settings, they have relatively low levels of educational attainment (less than half reported completing secondary school) and earn low levels of income (the average weekly income is \$39.46, which is roughly \$2,052 per annum). Although 32.4 percent held salaried jobs at baseline (i.e., occupations that pay income on a monthly basis, as opposed to on a daily basis), 29.1 percent held jobs commonly associated with the lower rungs of the income ladder, including auto rickshaw drivers, street vendors, skilled laborers, construction workers, and domestic workers. Baseline levels of health also appear to be poor, with 42.0 percent of respondents in low-income neighborhoods reporting cardiorespiratory health symptoms, possibly the result of living in an environment with extreme levels of air pollution.

In columns 2 and 3, we compare the characteristics of individuals from the low-income neighborhoods to those of the public bus commuters. Respondents in low-income neighborhoods are, on average, more likely to be women (18.6 versus 4.9 percent), less likely to have completed secondary school (35.0 percent versus 65.2 percent), and earn less (\$28.57 versus \$53.17 per week). Despite these differences, the general employment profile is similar, as a substantial share of respondents in both subsamples held occupations that involved low-skill labor and/or required a physical presence.

Covid Economics 51, 7 October 2020: 134-158

Covid Economics 51, 7 October 2020: 134-158

Figure 2 — Sample representativeness



Source: Panel survey data collected between 2018 and 2020 from active participants in the labor force; 2011-12 National Sample Survey (NSS) 68th Round (Delhi sample).
Notes: In panel A, we display the approximate locations in which we enrolled respondents into our study. Red points show the locations of respondents in the low-income neighborhoods. Blue points show the locations of the public bus stops where we enrolled public bus commuters. In panel B, we plot and compare the baseline (i.e., pre-COVID-19) epanechnikov kernel (bandwidth 10) of income over the past week (converted to U.S. dollars) to that of the National Sample Survey (NSS) Delhi sample, after inflating the NSS data using CPI figures from the World Bank. Sample means are shown in the figure.

In Figure 2, we summarize the representativeness of our sample. In Panel A, we plot our sampling locations over a map of the population density in Delhi. The low-income neighborhoods we survey, identified by the red circles, cover some of the most densely populated areas of the city. In panel B, we plot and compare the baseline (i.e., pre-COVID-19) Epanechnikov kernel (bandwidth 10) of weekly income (converted to U.S. dollars) to that of the National Sample Survey (NSS) Delhi sample, inflated to 2019. At baseline, mean weekly income in our sample is $39.46, compared to $78.84 in the NSS Delhi sample, which falls in between the 25th and 50th percentile of income, suggesting that we are capturing a slightly poorer segment of the bulk of the overall population.

Covid Economics 51, 7 October 2020: 134-158

4 Results

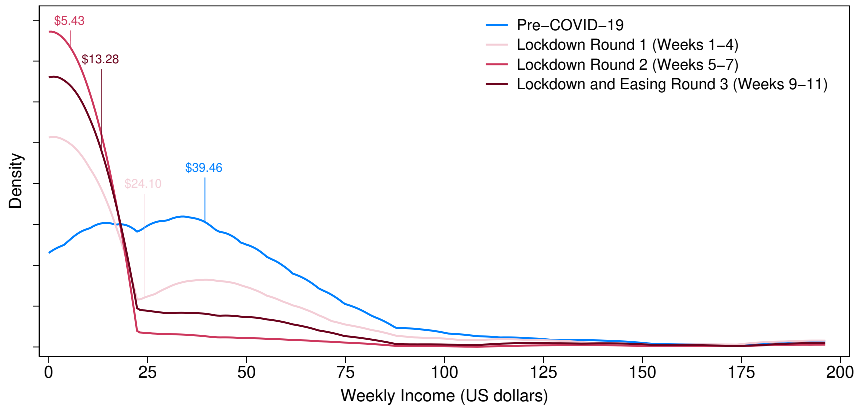
We estimate the effects of the India lockdown on our sample of mostly poor and non-migrant workers in Delhi through simple pre-post comparisons of key variables. This approach leaves open the possibility of confounding factors. However, we believe there is a limit to these concerns, considering how rapidly the India lockdown was introduced, and the sheer magnitude of the changes we observe. In our view, the only reasonable explanation for the large-scale changes described below is the India lockdown, and the general atmosphere of extreme fear and intense media coverage of COVID-19.

Over the eleven weeks following the announcement of the India lockdown, three leading patterns emerge from our data: there are large reductions in employment and income; widespread compliance with public health instructions; and limited effects on hunger, access to health care, and security, thus far.

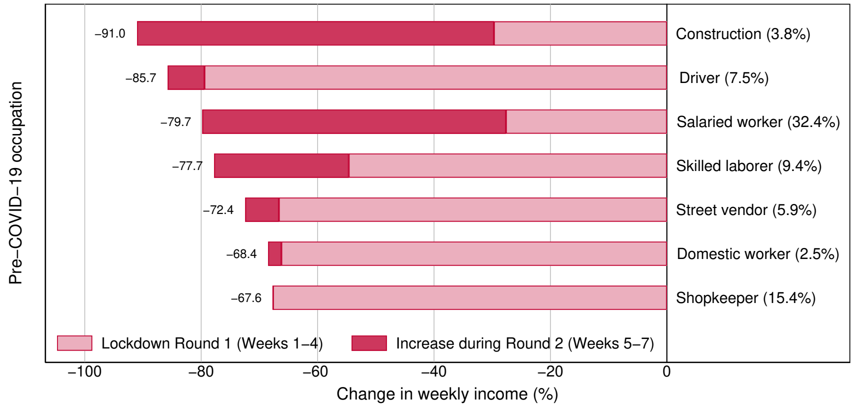
Covid Economics 51, 7 October 2020: 134-158

Figure 3 — Impact of the lockdown on weekly income

A. Weekly income



B. Change in weekly income, by occupation



Source: Panel survey data collected between 2018 and 2020 from active participants in the labor force. *Notes*: In Panel A, we plot epanechnikov kernels (bandwidth 10) of income over the past week (converted to U.S. dollars) at baseline ("Pre-COVID-19"), during weeks 1 to 4 of the COVID-19 lockdown ("Round 1"), during weeks 5 to 7 ("Round 2"), and during weeks 9 to 11 ("Round 3"). Sample means are shown in the figure. In Panel B, we show the change in weekly income based on occupation categories recorded at baseline. For each occupation, we show the sample share in parentheses.

An unprecedented decline in employment and earnings In Figure 3, we illustrate the impact of the India lockdown on weekly income for the workers in our sample. In Panel A, we plot the

Covid Economics 51, 7 October 2020: 134-158

distributions of income at baseline, recorded between 2018 and 2019, and income during Rounds 1, 2, and 3 of the lockdown. Compared to baseline, mean weekly income fell by 38.9 percent in Round 1 (i.e., weeks one to four), and by 77.5 percent in Round 2 (i.e., weeks five to seven), as more and more workers are forced into unpaid leave. By Round 2, approximately 90 percent of survey respondents reported that their weekly income had fallen to zero. Relatedly, average days worked over the past week fell from 4.3 to 1.2 days, a decline of 72.2 percent, by Round 2. Although income and days worked increased during Round 3, they remained 66.3 percent and 42.8 percent below baseline levels, respectively[7]. Broadly, these results are consistent with those found in contemporaneous work by Afridi, Dhillon, and Roy (2020) and Bertrand, Krishnan, and Schofield (2020).

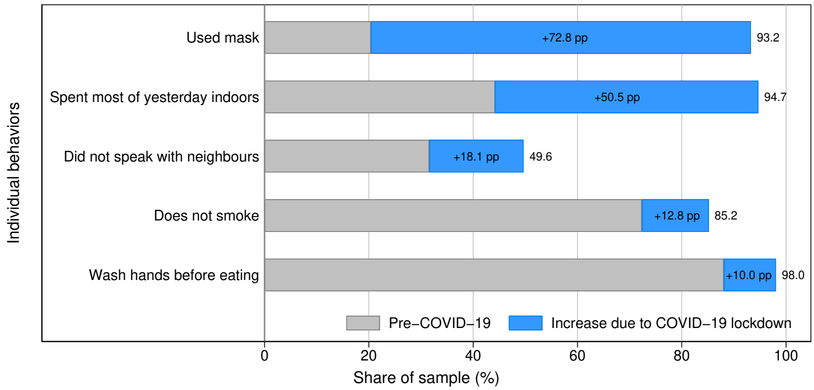
In Panel B, we report the change in weekly income for different groups of people sorted by their reported occupation at baseline. In the first four weeks of the lockdown, nearly all non-salaried occupations, including skilled laborers, drivers, street vendors, and domestic workers, experienced losses of 30 percent or more, while salaried workers suffered much less. However, by weeks five to seven of the lockdown, workers across all categories reported income losses of more than 60 percent[8].

Widespread compliance with public health directives The India lockdown has been accompanied by a number of public health messaging campaigns to raise awareness of COVID-19 and encourage people to adopt certain public health behaviors, including regular handwashing, mask-wearing in public, social distancing, and abstinence from smoking and spitting, among other recommendations. In Delhi, these messages have been delivered in various ways. For example, voice recordings have been played at the beginning of all private telephone calls; large billboards have been displayed in the streets; and in some areas, auto rickshaws have been repurposed to patrol the streets while playing health messages through loudspeakers.

[7]In Round 3, 45% of respondents said it was unlikely that their income levels would return to pre-lockdown levels in the following two months.

[8]All occupation groups experienced an increase in income during weeks nine to eleven of the lockdown, with the greatest increase for drivers, construction workers, and skilled laborers.

Figure 4 — Impact of the lockdown on compliance with public health directives



Source: Panel survey data collected between 2018 and 2020 from active participants in the labor force.
Notes: We present the effects of the lockdown on compliance with public health directives, comparing baseline mean to the average of Round 1 and Round 2 means. The outcome, "Did not speak with neighbors" is based on retrospective information collected in Rounds 2 and 3, where we asked respondents the following: "During a typical month prior to March 1, 2020, before the lockdown, how often did you talk with any of your neighbors?" For all other outcomes, we utilize data collected both at baseline and during the lockdown.

In Figure 4, we summarize the impact of the India lockdown on several health-related behaviors that we documented both before and after the onset of COVID-19. In some cases, we observe dramatic changes relative to baseline. Recent mask usage, for example, increased by 72.8 percentage points. The share of respondents stating that they "spent most of yesterday indoors" increased by 50.5 percentage points. In Appendix Figure A1, we compare the time spent indoors outcome to the Facebook movement index for each round. We see that both outcomes show similar declines in movement. Regular handwashing increased by 10.0 percentage points, becoming nearly universal. Even the share of respondents reporting that they do not smoke improved by 12.8 percentage points[9].

Some of the behavioral gains can be attributed to the short-term bans introduced while the lockdown was underway. On April 15, for instance, the Delhi government temporarily banned the sale of tobacco and announced that everyone would be required to wear a mask in public.[10] In our data, however, rates of non-smoking and mask usage had already shown dramatic changes in the weeks leading up to April 15, and are similar before and after the introduction of these temporary orders.

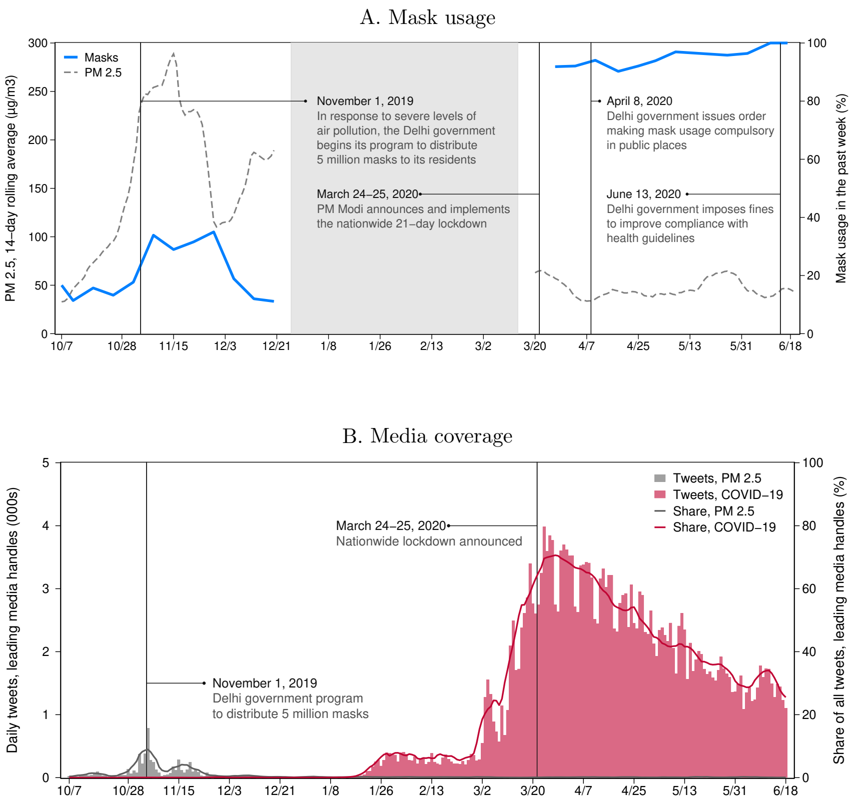
[9]We compare the baseline sample mean to the average of Round 1 and Round 2 sample means for each outcome. Results are qualitatively similar when we include Round 3 data, as reported in Appendix Table A2. The only outcome which differs significantly is time spent outdoors, which increased in Round 3 as movement restrictions weres relaxed.
[10]Anecdotally, the ban on tobacco sales was not well-enforced.

These behavioral effects are notable considering that many defensive health behaviors have low rates of takeup in the developing world (Dupas and Miguel 2017). In certain cases, the effect sizes rival or exceed the impacts of past interventions (to increase handwashing or encourage mask usage, for example) documented in the development economics literature (see, e.g., Kremer and Zwane 2007; Clasen et al. 2014; Hussam et al. 2019; Baylis et al. 2020).

What causes such uniform compliance? The potential role of fear and media in mask usage In highly polluted cities across the world, face masks had increasingly been seen as a viable defensive measure against the harmful effects of PM 2.5 (see, e.g., Langrish et al. 2012; Cherrie et al. 2018). During the week of November 1, 2019, the peak air pollution period in India, the Delhi government ambitiously distributed five million masks to its residents to help people defend themselves against PM 2.5 concentrations that were surpassing 300 $\mu g/m^3$ (i.e., 30 times higher than the WHO standard).

Covid Economics 51, 7 October 2020: 134-158

Figure 5 — Media coverage as a potential factor driving the increase in mask usage, relative to during the peak air pollution period



Source: Panel survey data collected between 2018 and 2020 from active participants in the labor force; U.S. Embassy, New Delhi, air pollution monitor; Twitter.
Notes: In Panel A, we plot the share of our sample reporting mask usage over the past week. The shaded region indicates the period in which there is no data. In Panel B, we plot Twitter data scraped from the top 50 media handles in Delhi. We focus on tweets containing the following keywords/phrases: pollution, delhi air, delhis air, delhi s air, airquality, air quality, airpurifier, air purifier, airemergency, air emergency, toxicair, toxic air, smog, stubbleburning, stubble burning, cropburning, crop burning, oddeven, odd even, pm2.5, pm25, pm10; and covid, corona, virus, pandemic, quarantine. At their peak, on March 27, 2020, COVID-19 tweets accounted for 70.6 percent of all tweets. In comparison, on November 3, 2019, air pollution tweets accounted for 8.8 percent of all tweets.

In Figure 5, Panel A, we plot mask usage in our sample during the most recent peak period of air pollution (between October and December 2019), and during the India lockdown. Although mask usage increased in the weeks following the government distribution of masks, it did not reach 35 percent of the sample. In contrast, mask usage became nearly universal during the COVID-19 crisis, despite relatively low levels of air pollution.

It is likely that the behavioral response to the current crisis, relative to the air pollution crisis, has been driven by a combination of factors, chief among them the extreme level of concern and fear about the coronavirus, and the unprecedented level of media coverage of the pandemic. For example, 79.7 percent of respondents in our sample reported feeling "extremely concerned" about COVID-19. In comparison, less than 40 percent of respondents felt this way about air pollution during the peak pollution period between October and December 2018. High levels of fear have been well-documented in India. For instance, in some areas, extreme levels of concern led people to not only comply with the recommended health behaviors during the lockdown, but to also form volunteer squads to ensure that others complied (as reported in Gettleman and Raj 2020).

In addition, media coverage of COVID-19 has vastly overshadowed coverage of all other crises in recent memory. In Figure 5, Panel B, we present the daily number of tweets related to COVID-19 and air pollution, produced by 50 leading Indian media and newspaper handles, most of which are based in Delhi (e.g., Times of India, Hindustan Times, etc.). During Delhi's most recent peak air pollution period, there is a sharp increase in pollution-related coverage. At the time, air pollution was one of the most discussed topics in Delhi. Yet the cumulative media coverage of COVID-19 has been many times greater. At their peak, on November 3, 2019, air pollution tweets accounted for 8.8 percent of all tweets, according to our measure. In contrast, on March 27, 2020, just a couple days into the lockdown, COVID-19 tweets accounted for 70.6 percent of all tweets. Based on this measure, from March 25 to May 13, COVID-19 accounted for 55.9 percent of all media coverage, although its share of total coverage has declined. Taken together, these patterns raise the question of whether compliance with public health directives can remain high, even as the novelty, fear, and media coverage of COVID-19 subside.

Levels of hunger, health care access, and insecurity are not alarming, but concerns remain

When the lockdown was first announced, many observers were concerned about the emergency issues (e.g., hunger) that could arise if economic activity were to shut down completely, particularly in a setting with such high population density and poverty rates. Migrant laborers, for example, were heavily and almost immediately impacted. Just days into the lockdown, tens of thousands of migrant laborers and daily wage workers attempted to leave Delhi and return to their villages of origin, either by bus or on foot, in a migration that resulted in numerous deaths (e.g., as reported in Pal and Siddiqui

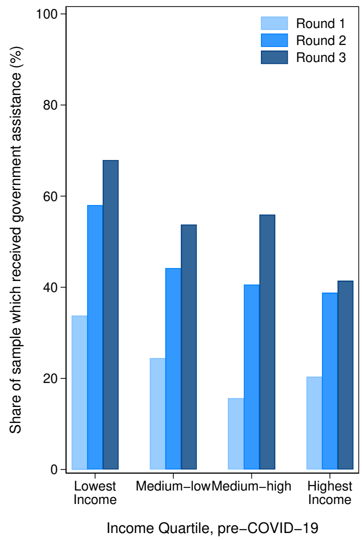
Covid Economics 51, 7 October 2020: 134-158

CEPR PRESS

2020).

In order to mitigate anticipated hunger needs, the Delhi government set up over 500 assistance centers across the city to provide access to food (typically rice and lentils) for those who needed it. These centers appear to have been well-placed, as the average distance between the individuals in our sample and the nearest government assistance center is only 640 meters. In our data, which covers mostly poor and non-migrant workers in Delhi, it appears that negative impacts on consumption may have been mitigated by this effort. Roughly 34 percent of our sample reported accessing these centers.

Figure 6 — Negative income shocks and access to government benefits



Source: Panel survey data collected between 2018 and 2020 from active participants in the labor force.
Notes: We present the share of individuals that reported receiving government benefits during the lockdown. Individuals are grouped into quartiles by baseline income, and the sample consists of individuals who experienced a negative income shock. To calculate negative income shocks, we take the difference between baseline income and post-lockdown income, averaging over rounds as necessary. 95 percent of government benefits consisted of food assistance.

In Figure 6, we report the share of individuals that reported receiving government assistance during the lockdown (the vast majority of which was in the form of food). We focus on the subsample of individuals who experienced a negative income shock. The share of recipients is larger for lower income groups (e.g., 64 percent of individuals in the lowest quartile of income reported receiving assistance), suggesting that the Delhi government relief efforts were effective in reaching those in greatest need of assistance.

Covid Economics 51, 7 October 2020: 134-158

We also do not observe high, or even moderately high, levels of hunger, scarcity, inability to access medical care, or insecurity, as reported in Appendix Figure A2. During the first seven weeks of the lockdown, few respondents experienced hunger (3.0 percent reported skipping a meal recently); there were only moderate levels of product scarcity (17.2 percent indicated they were unable to purchase an item); and there were only a handful of reports of crime (4.3 percent stated that they had experienced a recent crime incident). Among individuals who attempted to access medical care (or knew of someone personally who attempted to access care), only 4.7 percent reported any challenges.

Nevertheless, whether these patterns will continue even in the immediate future is uncertain. Nearly half of the respondents in our sample reported some difficulties in food supply chains, ranging from mobility restrictions to inadequate quantities, business closures, and high prices. In Round 2, only 25.2 percent of respondents stated that they had savings remaining, and out of these individuals, 44.1 percent estimated that they would run out of savings soon, a finding that is consistent with the work of Bertrand, Krishnan, and Schofield (2020)[11]. We also note concerning levels of mental and emotional well-being during the lockdown. 43.6 percent of respondents reported feeling either depressed, bothered, that they lacked focus, or that everything was an effort, for the majority of the week leading up to the survey.

Selection and attrition To address potential concerns of selection due to differences between the study sample and the sampling frame, and attrition across survey rounds, we present simple means, inverse-probability weighted means, and simple means for a balanced panel for the main economic and social costs, and behavioral changes in Appendix Table A2. For each outcome and each round, the first column reports simple means for the main sample, the second column reports weighted means for the main sample, and the third column reports simple means for the balanced panel. The results are qualitatively similar across each method.

In Appendix Table A3, we report differences in baseline characteristics between individuals who did and did not consent to survey, across all rounds. While there are significant differences across several characteristics, including gender, education, and types of occupation, weekly income and health behaviors are similar in both groups. Further, as reported in Appendix Table A2, we show that these differences do not appear to impact qualitative conclusions.

[11]The corresponding figures for Round 3 are 32.0 percent and 48.8 percent.

Covid Economics 51, 7 October 2020: 134-158

CEPR PRESS
Covid Economics 51, 7 October 2020: 134-158

5 Discussion

Taken together, our data suggest that for our representative sample of mostly poor and non-migrant workers in Delhi, the India lockdown imposed harsh financial consequences. At the same time, it induced substantial compliance with public health directives important in limiting the spread of COVID-19, and these changes were potentially driven by high levels of fear and media exposure. Thus far, a worst case scenario (involving alarming rates of hunger and lack of access to health care, for example) has largely been avoided, possibly due to the success of the Delhi government's assistance programs. But with diminishing levels of savings, there remain serious concerns about the near-term future, particularly as infection rates rise. In addition, it remains to be seen whether the positive behavioral changes we observe will persist, as fear and media coverage subside once the lockdown is lifted and elements of daily life return to normal. These results suggest that the government should maintain a position to rapidly expand cash, medical, and food assistance, particularly in future lockdowns.

Further research is urgently required to understand what emergency relief programs could have the greatest impact for this population group; what factors would allow these behavioral changes, and in particular those relating to social distancing, to persist; the degree to which the risk of infection will impact the trust that people have in others as well as in authority figures, and the social, economic, and political consequences of those changes; and how the cumulative economic costs to this population group can be reversed over time, among others.

Despite the debilitating economic costs of the India lockdown, the potential value of the lives saved from the social distancing policies, and other behavioral changes that accompanied the lockdown, may also be large. This is especially the case in settings with underlying vulnerabilities, such as high levels of pre-existing health conditions, poor water and sanitation and other infrastructure, and high population densities. Ultimately, future decisions to implement extreme social distancing mandates will involve complex trade-offs between the costs of higher infection rates in a world that remains open, and the economic and non-economic costs in one that is periodically closed. A key factor influencing this trade-off is the extent to which people follow the public health directives.

Covid Economics 51, 7 October 2020: 134-158

References

Afridi, Farzana, Amrita Dhillon, and Sanchari Roy. 2020. "How has Covid-19 crisis affected urban poor? Findings from a phone survey." *Ideas for India.*

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." *The Review of Asset Pricing Studies.*

Bartik, Alexander, Marianne Bertrand, Zoë B. Cullen, Edward L. Glaeser, Michael Luca, and Christopher T. Stanton. 2020. " The impact of COVID-19 on small business outcomes and expectations." *Proceedings of the National Academy of Sciences.*

Baylis, Patrick, Michael Greenstone, Kenneth Lee, and Harshil Sahai. 2020. "Pollution Masks and the Demand for Clean Air: Experimental Evidence from Delhi." Unpublished manuscript.

Bertrand, Marianne, Kaushik Krishnan, and Heather Schofield. 2020. "How are Indian households coping under the COVID-19 lockdown? 8 key findings." *Energy Policy Institute of Chicago.*

Burlig, Fiona, Steve Cicala, and Anant Sudarshan. 2020. "Impact of COVID-19 on Electricity Consumption and Particulate Pollution."

Cherrie J.W., Andrew Apsley, Hilary Cowie, Susanne Steinle, William Mueller, Chun Lin, Claire J. Horwell, Anne Sleeuwenhoek, and Miranda Loh. 2018. "Effectiveness of Face Masks Used to Protect Beijing Residents Against Particulate Air Pollution." *Occupational and Environmental Medicine.*

Clasen, Thomas, Sophie Boisson, Parimita Routray, Belen Torondel, Melissa Bell, Oliver Cumming, Jeroen Ensink, Matthew Freeman, Marion Jenkins, Mitsunori Odagiri, Subhajyoti Ray, Antara Sinha, Mrutyunjay Suar, Wolf-Peter Schmidt. 2014. "Effectiveness of a Rural Sanitation Programme on Diarrhoea, Soil-transmitted helminth Infection, and Child Malnutrition in Odisha, India: a Cluster-Randomised Trial." *The Lancet Global Health.*

Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber. 2020. "Labor Markets During the COVID-19 Crisis: A Preliminary View." *NBER Working Paper Series.*

Deshpande, Ashwini. 2020. "The Covid-19 Pandemic and Lockdown: First Effects on Gender Gaps in Employment and Domestic Work in India." *Centre for Economic Data and Analysis Discussion Paper Series.*

Dupas, Pascaline, and Edward Miguel. 2017. "Impacts and Determinants of Health Levels in Low-Income Countries." *Handbook of Field Experiments.*

Gettleman, Jeffrey, and Suhasini Raj. 2020. "Powered by Fear, Indians Embrace Coronavirus Lockdown." *New York Times.*

HT Correspondents. "Delhi observes Janta Curfew, city of nearly 20 million comes to a standstill." *Hindustan Times.*

Hussam, Reshmaan, Atonu Rabbani, Giovanni Reggiani, and Natalia Rigol. 2019. "Rational Habit Formation: Experimental Evidence from Handwashing in India." *Harvard Business School Working Papers.*

Kremer, Michael, and Alix Peterson Zwane. 2007. "What Works in Fighting Diarrheal Diseases in Developing Countries? A Critical Review." *The World Bank Research Observer.*

Langrish, Jeremy P., Xi Li, Shengfeng Wang, Matthew M. Y. Lee, Gareth D. Barnes, Mark R. Miller, Flemming R. Cassee, Nicholas A. Boon, Ken Donaldson, Jing Li, Liming Li, Nicholas L. Mills, David E. Newby, and Lixin Jiang. 2012. "Reducing Personal Exposure to Participate Air Pollution Improves Cardiovascular Health in Patients with Coronary Heart Disease." *Environmental Health Perspectives.*

Maas, Paige, Andreas Gros, Laura McGorman, P. Alex Dow, Shankar Iyer, Wonhee Park, Chaya Nayak. 2019. "Facebook Disaster Maps: Aggregate Insights for Crisis Response & Recovery" *Proceedings of the 16th ISCRAM Conference.*

Marx, Benjamin, Thomas Stoker, and Tavneet Suri. 2013. "The Economics of Slums in the Developing World." *Journal of Economic Perspectives.*

Pal, Alasdair, and Danish Siddiqui. 2020. "Special Report: India's migrant workers fall through cracks in coronavirus lockdown." *Reuters.*

Singh, Karan Deep, Vindu Goel, Hari Kumar, and Jeffrey Gettleman. 2020. "India, Day 1: World's Largest Coronavirus Lockdown Begins." *New York Times.*

Wu, Xiao, Rachel C. Nethery, M. Benjamin Sabath, Danielle Braun, and Francesca Dominici. 2020. "Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study." *medRxiv.*

Yadav, Anumeha. 2020. "India: Hunger and uncertainty under Delhi's coronavirus lockdown." *Al Jazeera.*

Zheng, Ying-Ying, Yi-Tong Ma, Jin-Ying Zhang, and Xiang Xie. 2020. "COVID-19 and the cardiovascular system." *Nature Reviews Cardiology.*

Covid Economics 51, 7 October 2020: 134-158

CEPR PRESS

Covid Economics 51, 7 October 2020: 134-158

Appendix

Table A1 — Characteristics of active participants in the labor force

	Full sample	Low-income neighbor-hoods	Public bus commuters	p-value of diff.
	(1)	(2)	(3)	(4)
Panel A: Demographic characteristics				
Female (%)	12.7	18.6	4.9	< 0.01
Age (years)	35.4	35.8	34.8	0.11
Completed secondary school (%)	47.9	35.0	65.2	< 0.01
Household size	5.1	5.4	4.8	< 0.01
Long-term Delhi resident (%)	87.3	94.7	77.2	< 0.01
Panel B: Common (primary) occupations				
Salaried job (%)	32.4	18.8	50.6	< 0.01
Business (%)	15.4	18.2	11.6	< 0.01
Skilled labour (%)	9.4	10.5	7.9	0.10
Driver (%)	7.5	8.2	6.7	0.32
Street vendor (%)	5.9	8.3	2.7	< 0.01
Homemaker (%)	5.4	8.4	1.3	< 0.01
Construction worker (%)	3.8	4.9	2.4	0.01
Student (%)	3.4	2.5	4.5	0.04
Domestic worker (%)	2.5	4.1	0.3	< 0.01
Panel C: Economic indicators				
BPL card holder (%)	19.7	22.6	15.4	< 0.01
Weekly individual income (USD)	39.46	28.57	53.17	< 0.01
Weekly household income (USD)	58.44	45.13	73.38	< 0.01
Government water source	21.55	23.6	18.8	0.03
Distance to gov't assistance (km)	646.4	623.0	690.5	0.19
Sample size	1,392	797	595	

Notes: Active participants in the labor force are defined as those individuals who were employed in at least one survey round during the course of the study. Columns 2 and 3 report sample means for the slum residents and public bus commuters samples, respectively. Column 4 reports p-value of the difference between the means. In panel C, we include sample means for BPL ("Below Poverty Line") card holder and Government water source (which captures whether the individual's primary drinking water source is a Delhi Jal Board tanker) as additional potential indicators of low-income status.

Covid Economics 51, 7 October 2020: 134-158

Table A2 — Comparison between simple and weighted means

	Round 1			Round 2			Round 3		
	Simple	Weighted	Balanced	Simple	Weighted	Balanced	Simple	Weighted	Balanced
Panel A: Economic and Social Costs									
Weekly individual income (USD)	24.1	25.3	23.5	5.4	5.6	6.0	13.2	14.1	11.9
Days worked in the past week	1.0	1.1	0.9	1.1	1.2	1.2	2.5	2.5	2.2
Struggled with mental well-being (%)	43.3	43.0	42.7	42.3	42.1	47.0	45.7	45.5	47.3
Challenges accessing food markets (%)	–	–	–	48.6	48.0	50.4	51.1	50.4	50.7
Unable to purchase an item (%)	21.8	21.7	18.1	15.8	16.0	15.4	11.2	11.3	12.8
Panel B: Behavioral Changes									
Wore mask in the past week (%)	92.6	91.7	92.8	93.9	93.4	94.5	96.6	96.8	97.3
Spent yesterday outdoors (%)	4.2	4.3	3.7	6.2	5.9	8.6	18.2	18.6	19.3
Did not speak with neighbours (%)	–	–	–	49.6	50.5	47.2	51.8	52.6	52.4
Smoker (%)	15.4	15.1	17.6	14.1	13.5	15.3	16.5	16.6	14.9
Wash hands before eating (%)	98.7	98.7	99.3	97.4	97.3	98.0	97.5	97.1	98.0
Sample size	1,052	1,048	226	752	751	226	640	637	226

Notes: This table presents sample means for each round by three methods. For each round, the first column reports simple means for the full sample, the second column reports weighted means for the full sample, and the third column reports simple means for the balanced sample (those who consented to survey in each round). Weights are the inverse of the predicted probability of consenting to survey. Probability of consenting to survey is predicted by running a logistic regression on baseline characteristics. Panel A reports economic and social costs and Panel B reports behavioral changes.

CEPR PRESS

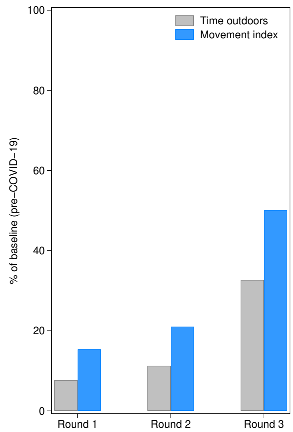
Covid Economics 51, 7 October 2020: 134-158

Table A3 — Characteristics of individuals who did or did not consent, across all rounds

	Did not consent (1)	Consented (2)	p-value of diff. (3)
Female (%)	37.7	23.9	0.00
Age (years)	35.5	35.1	0.33
Completed secondary school (%)	51.5	46.7	0.02
Household size	5.2	5.2	0.55
Long-term Delhi resident (%)	84.0	87.6	0.01
Homemaker (%)	22.6	16.6	0.00
Salaried job (%)	29.3	25.9	0.06
Business (%)	11.3	12.3	0.46
Skilled labour (%)	4.3	7.5	0.00
Street vendor (%)	3.6	4.7	0.18
Driver (%)	3.7	6.0	0.01
Student (%)	7.0	6.5	0.63
Construction worker (%)	2.1	3.0	0.16
Domestic worker (%)	1.3	2.0	0.16
Unemployed (%)	12.2	11.5	0.58
Weekly individual income (USD)	39.2	37.3	0.32
Weekly household income (USD)	63.7	58.9	0.14
Days worked in the past week	3.3	3.9	0.00
BPL card holder (%)	15.8	20.3	0.01
Owns a mask (%)	0.4	0.3	0.12
Wore mask in the past week (%)	18.2	21.4	0.19
Spent yesterday outdoors (%)	53.6	55.3	0.47
Smoker (%)	26.7	26.4	0.89
Wash hands before eating (%)	87.3	88.1	0.56
Sample size	947	1,744	

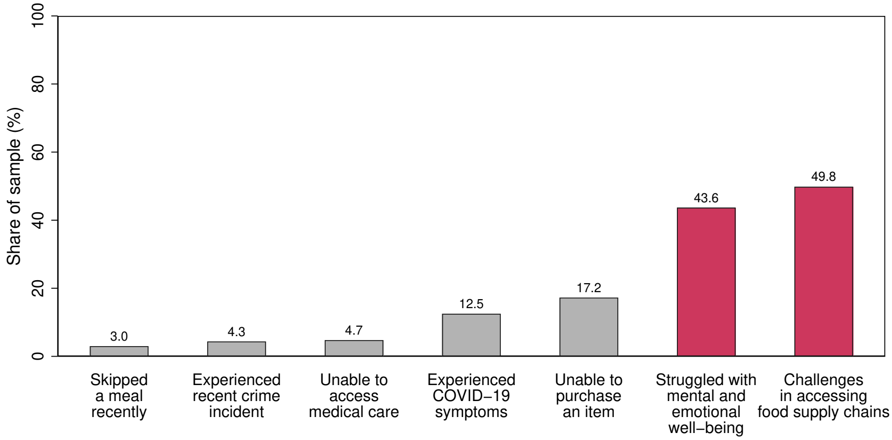
Notes: This table presents differences in baseline characteristics between individuals who consented to survey at least once across Rounds 1, 2, and 3, and those individuals who did not consent.

Figure A1 — Time indoors versus Facebook movement index



Source: Panel survey data collected between 2018 and 2020 from active participants in the labor force. Movement data is from Facebook Data for Good (Maas et al. 2019).
Notes: We plot self-reported time spent indoors and the Facebook movement index for each survey round.

Figure A2 — Non-economic costs experienced during the lockdown



Source: Panel survey data collected between 2018 and 2020 from active participants in the labor force.
Notes: Data collection began on March 27, 2020, two days after the start of the lockdown. For each indicator, we plot the average result over seven weeks of data. The mental and emotional well-being outcome is defined as the share of individuals who reported feeling either depressed, bothered, that they lacked focus, or that everything was an effort, for the majority of the week leading up to the survey. Challenges accessing food supply chains include: mobility restrictions, inadequate quantities available, closed markets, and price shocks.

Covid Economics 51, 7 October 2020: 134-158

CEPR PRESS

Covid Economics **Issue 51, 7 October 2020**

Coping during COVID-19: Family businesses and social assistance in Nigeria[1]

Elvis Korku Avenyo[2] and Gideon Ndubuisi[3]

Date submitted: 30 September 2020; Date accepted: 1 October 2020

The COVID-19 pandemic and the consequent Government-imposed restrictions have altered the way of life around the globe. Using the new "Nigeria Baseline COVID-19 National Longitudinal Phone Survey 2020", this paper contributes to the nascent literature on the Economics of COVID-19 by examining the impact of changes in income and social assistance due to the pandemic on the coping strategies of family business owners. We find that family business owners who experienced a reduction in income and those that received social assistance due to the pandemic are likely to increase their coping level. We discuss the policy implications of these findings.

Covid Economics 51, 7 October 2020: 159-184

1 We would like to thank the Editor, Charles Wyplosz, and an anonymous reviewer for constructive comments. All errors remain solely ours.

2 Oxford Department of International Development, University of Oxford and and South African Research Chair in Industrial Development, University of Johannesburg.

3 UNU-MERIT/Maastricht University.

Covid Economics 51, 7 October 2020: 159-184

1. Introduction

The COVID-19 pandemic is changing the world around us. While the first laboratory index case was identified in Wuhan, China, in December 2019, the virus has spread to about 216 countries. The current global laboratory-confirmed cases are about 14.3 million, while the death-related cases are about 600,000 as of July 22, 2020 (WHO, 2020a), making the virus the worst pandemic of the modern age since the Spanish Flu of 1918. Given the health challenges the virus poses and the lack of a clinically approved drug/vaccine for its cure or prevention, governments across the globe have enforced border shutdowns, travel restrictions and quarantine, social distancing, and in some cases, lockdown measures in a bid to "flatten the curve". While the socio-economic impact of the virus and Government-imposed regulations are still less understood, they are unequivocally wreaking socio-economic havoc worldwide.[1] It is projected that this could lead to the worst worldwide recession since the 1929 Great Depression (Barker *et al.,* 2020; Bartik *et al.,* 2020; IMF, 2020; Khan *et al.,* 2020). Hence, assessing the impacts of the COVID-19 pandemic on economies and vulnerable groups is fundamental to informing and tailoring governments' and the international community's responses to the pandemic.

In this paper, we examine the impact of the COVID-19 pandemic on individuals that operate family businesses. In particular, using a real time survey on COVID-19, our study examines the impact of changes in income and social assistance on the coping strategies of individuals who operate family-owned businesses. The ubiquity of family-owned businesses across the globe is well recognized in the literature (Xi *et al.,* 2015; Kraus *et al.,* 2020). In developing economies, where institutions are underdeveloped, firms that are predominantly operating as family-owned businesses are a norm rather than the exception. As a result, family-owned businesses are the highest employer of labor, wealth creation, and innovation in the private sector of these economies (Kraus *et al.,* 2020). Despite these, there is little understanding of how owners of family businesses are coping with and are adjusting to COVID-19 induced shocks, and the possible role of social assistance in mitigating the adverse effects of the pandemic. Instead, the extant literature on the economic effects of COVID-19 has primarily focused on the labor market-related effects (e.g., Adams-prassl *et al.,* 2020; Beine *et al.,* 2020; Chiou & Tucker 2020; Dingel & Neiman 2020; Kalenkoski & Pabilonia, 2020). Few studies focus on business establishment, examining how they are affected by the COVID-19 induced shocks (Avenyo *et al.,* 2020; Bartik *et al.,* 2020; Fairlie, 2020; Kraus *et al.,* 2020).

Insights from the literature on "crisis management" suggest that family-owned businesses may be better at handling crises due to their socio-emotional endowments (Gómez-Mejía *et al.,* 2007, 2011; Faghfouri *et al.,* 2015). However, their limited financial capital and resources

[1] For an extensive review of the economic impact of COVID-19 across sectors, see Nicola *et al.* (2020).

Covid Economics 51, 7 October 2020: 159-184

make them more vulnerable to external threats and crises (Cater & Schwab, 2008; Herbane, 2013; Kraus *et al.,* 2013; Roux-Dufort, 2007). Also, family-owned businesses have the unique characteristic that external crisis such as the COVID-19 pandemic hits their owners twice: once as a private citizen and second as business owners (Runyan, 2006; Kraus *et al.,* 2020). Hence, it is safe to argue that the COVID-19 pandemic threatens to affect family-owned businesses and household members that operate these businesses. Along this line, adopting the requisite strategies in the face of crisis becomes paramount for their survival and the individuals that own and operate them. Given their enormous contributions to the economy, this has important implications for spurring growth and development in the post-COVID-19 pandemic.

Therefore, to address our research question on the effect of changes in income due to the pandemic, and COVID-19 induced social assistance on the coping mechanisms of owners of family businesses, we utilize the "Nigeria Baseline COVID-19 National Longitudinal Phone Survey 2020" that was recently published by the World Bank. To capture the COVID-19 shock on family-business owners, we explore ordinal variations in income change, while we measured social assistance using a survey question that asks whether households have received any social assistance (direct food, cash transfers, or other in-kind transfers) from external bodies like government, cooperatives or non-government organizations (NGO) since mid-March. Our results indicate that owners of family businesses who experienced a decrease in income due to the COVID-19 shock increased their coping strategies. We then examine the role of social assistance, distinguishing between government and non-governmental organizations' assistance. Our results show that assistance from various sources enabled owners of family businesses to adopt higher coping strategies, although the effect tends to be higher for social assistance from the Government. Our results are suggestive of the essential role of social assistance in times of crisis, while underscoring the differential response of family business owners in times of crisis. Hence, to reduce income inequality and spur a more-inclusive growth in the post-COVID-19 pandemic era, policies such as social assistance during crisis must make concerted efforts to target the most affected.

The rest of the paper is organized as follows. The next section presents a synopsis on COVID-19 in Nigeria. Section 3 describes the data and estimation strategy. Section 4 presents and discusses the results while we conclude in section 5.

2. Synopsis of COVID-19 in Nigeria

The COVID-19 virus is a highly communicable and infectious disease that is caused by severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2) that is now deemed a pandemic by the World Health Organization (Huang *et al.,* 2020; Harapan *et al.,* 2020; Xiong *et al.,* 2020). Its symptoms include fever, cough, shortness of breath, sore throat, runny nose, and sneezing, among others (Huang *et al.*, 2020; Harapan *et al.*, 2020; Ohia *et al.*, 2020). The COVID-19 virus spreads primarily through close contacts with droplets of saliva or discharge from the nose when an infected person coughs or sneezes COVID-19 (Huang *et al.*, 2019; Harapan *et al.*, 2020; WHO, 2020c). As at the time of writing, there are still no clinically approved drugs/vaccines for its cure or prevention. Hence, governments across the globe have resulted in measures such as lockdown, social distancing, self-isolation or self-quarantine, and observation of simple hygiene habits such as regular washing of hands and wearing face masks.

While the first global reported case on COVID-19 was in December 2019, the first recorded laboratory index case in Nigeria was in Lagos State on February 27, 2020 (NCDC, 2020a). It also happens to be the first reported case in Africa. The country's index case is reported to be imported by an immigrant who albeit resides in Nigeria. The index case led to the activation of the Level 3 Emergency Operations Centre, the country's highest emergency level. Before the reported index case, however, the Nigerian government through the Nigeria Centre for Disease Control (NCDC)[2] had put in place measures aimed at controlling and containing the outbreak in the country. Such measures include, but are not limited to, issuing technical guidelines and response plans, training health workers, and supporting laboratories for COVID-19 testing and campaign to educate Nigerians on how to protect themselves from the disease setting-up isolation centers (NCDC, 2020b).

Despite the supposed Nigerian government preparedness, the number of laboratory-confirmed cases and COVID-19 death-related cases in the country have been on the rise, as shown in Figures 1A and 1B. As of July 22, 2020, the number of laboratory-confirmed cases in the country is 37,801, with 805 deaths (NCDC, 2020c). The 36 states in the country, including the capital territory, have recorded at least one confirmed case. As shown in Figure 2, Lagos state, the country's economic capital city, has the highest number of laboratory-confirmed and death-related cases, with the respective figures being 6,065 and 72. On the other hand, Kogi State has the lowest number of confirmed cases (5) with two recorded death-related cases, while Taraba, although, has 54 reported cases, has recorded zero deaths at the time of writing (NCDC, 2020c).

Like many other countries across the globe, the Nigerian Government in a bid to prevent the further spread of the virus had introduced different counter-measures such as the closing of schools, cancellation of public events, restrictions on public gatherings including burials and faith gatherings, stay at home recommendations, social distancing, restrictions on internal movements, and border closures. While some of these restrictions have been eased, school

[2] NCDC is the government agency responsible for leading "the prevention, detection, investigation, monitoring, and control of communicable diseases".

Covid Economics 51, 7 October 2020: 159-184

closure and intra-state-border movement (for non-essential workers) are still in vogue. The government has also initiated different measures to cushion the adverse effects induced by the virus. Some of such measures include a fiscal stimulus package by the country's apex bank in the sum of $138.89 million credit facility to households and SMEs most affected by the pandemic, $277.78 million loans to the health sector and $2.78 billion to the manufacturing sector. Also, the interest rates on all the CBN interventions have been revised downwards from 9 to 5 percent, and a one-year moratorium on CBN intervention facilities has been introduced (PWC, 2020; Dixit *et al.*, 2020; KPMG, 2020; Adegite & Abu, 2020).

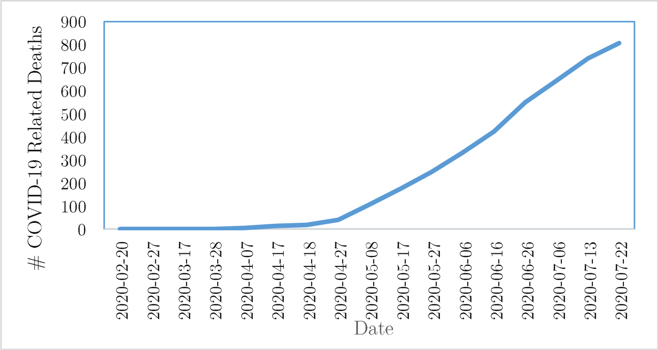


Figure 1A: Total COVID-19 Laboratory Confirmed Cases in Nigeria
Data Source: Our World in Data[3]

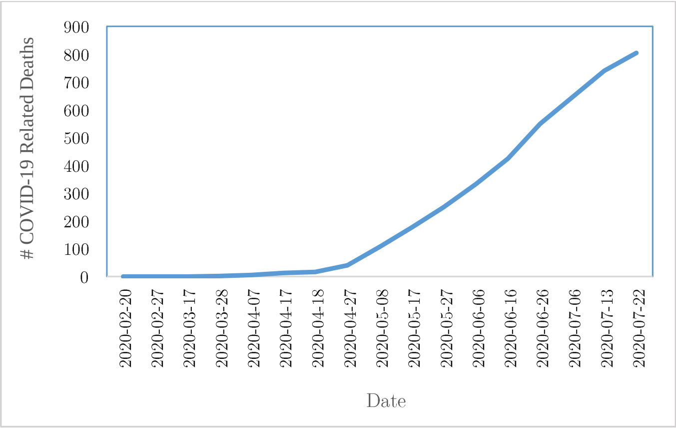


Figure 1B: Total Number of COVID-19 Death-Related in Nigeria
Data Source: Our World in Data[4]

[3] https://github.com/owid/covid-19-data/tree/master/public/data
[4] https://github.com/owid/covid-19-data/tree/master/public/data

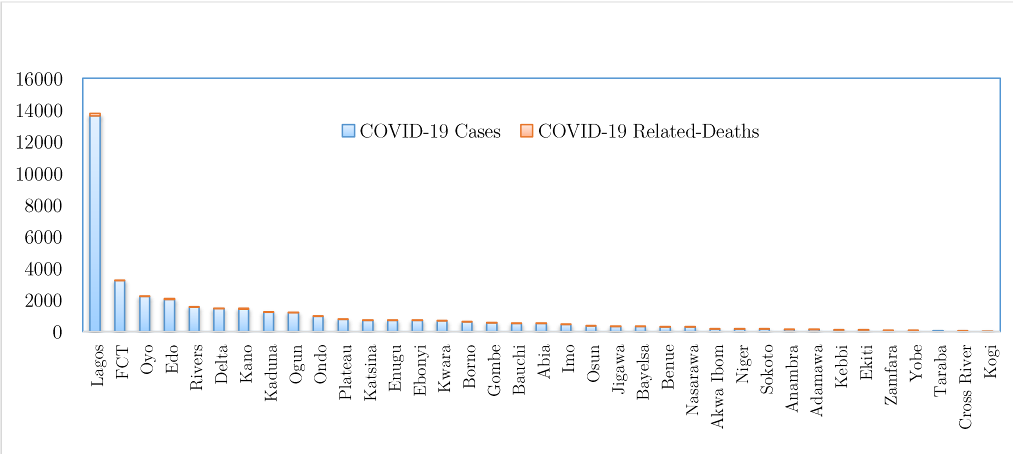


Figure 2: Spatial Distribution of COVID-19 Laboratory Confirmed and Death-Related Cases across Nigerian States
Data Source: NCDC (2020c)

3. Methodology

3.1. Data

This paper is based on data from the national representative sample of households in the "Nigeria Baseline COVID-19 National Longitudinal Phone Survey 2020" (COVID-19 NLPS) conducted by the World Bank. The COVID-19 NLPS is phone-based, that is, all information was collected via phone calls. Given the significant consequences of COVID-19 on the Nigerian economy, the largest economy in Africa, the main purpose of the survey is to provide real-time evidence on the socio-economic effects of the pandemic on households for the better and effective decision-making process to mitigate the crisis.[5] The survey covers a key range of issues, including questions on the spread of COVID-19, employment, and income, among others. The data has the added advantage that we can observe individuals within a household, including their activities, such as whether they operate a family business or otherwise. The survey also includes specific questions on several coping/shocks of households and the available social safety nets to mitigate the effect of COVID-19 (see National Bureau of Statistics, 2020). We use the version of the data published on 2020-06-04.

[5] For further details on the description and documentation of the survey, see https://microdata.worldbank.org/index.php/catalog/3712/study-description.

Covid Economics 51, 7 October 2020: 159-184

Table 1 presents selected questions and the distribution of selected key household, owner, and family business attributes in the COVID-19 NLPS data set. In this paper, we consider households and members of households that have operated family business since the beginning of March. In our sample, 9,503 household members who responded to have operated family business since the beginning of March. The COVID-19 NLPS data set asked whether if the respondent's household has been affected by COVID-19 shock since mid-March, and how households cope with the shock. Our sample shows that 3,018 respondents, representing about 32%, indicated that their households had been affected by COVID-19 shock. Of this, about 31% of individuals relied on reducing food consumption to cope with the shock of COVID-19. Other main coping strategies of individuals in our sample include reliance on savings (26%), reduction in non-food consumption (18%), and assistance from family and friends (11%). Our sample also shows that the majority of individuals experienced a reduction in income, with about 77% of respondents indicating a falling income since mid-March. This evidence is also observed in the growing literature on COVID-19, suggesting that COVID-19 is leading to job losses and reductions in incomes of workers (e.g., see Balde *et al.*, 2020; Campello *et al.*, 2020). Of key importance to this paper is the question of whether members of households have received any assistance, and the main sources of assistance received since mid-March. Our sample shows that about 5% of members of households have received a form of assistance in either direct food, cash transfers, or other in-kind transfers (excluding food), with the main source of assistance coming from the Government at the federal, state, and local levels.

Based on the question on how household members cope with COVID-19, we constructed a variable called 'Level of coping' that takes the value of 0 if the member has not been affected and has not introduced any coping strategy (None); 1 if the member has been affected and introduced only 1 coping strategy (Low); 2 if the member has been affected and introduced 2 coping strategies (Medium); and 3 if the member has been affected and introduced more than 3 coping strategies (High). This categorization helps to better capture vulnerable household members, particularly those who indicate to have been affected adversely by COVID-19. Also, we group the main sources of assistance as 'Government' (federal, state and local), "Community/NGOs" (Community organization/ cooperative, and NGOs), and "Religious bodies".

Figures 3A-3D show the graphical representation of a family business owner's coping levels and key individual and household attributes as presented in Table 1. Figure 3A shows that the majority of owners, about 77%, did not have any coping strategy, while about 19%, 4%, and 0.3 % had low, medium, and high levels of coping, respectively. As noted earlier, only a few owners received any assistance (about 5%), while the majority of them who received assistance got it mainly from the Government (about 50 %) (See figure 3C). This is not surprising because, at the unset of the outbreak in the country (which is what our dataset captures), different individuals and organizations that are into philanthropy and charity in the country donated to the government, making the government almost the sole channel of assistance. Figure 3B also shows that assistance was not exclusive to owners with positive coping levels, as about 3% with no coping levels also received assistance. Figures 3B, 3D, and 3E show the distribution of our key attributes across coping levels of family business owners. Figure 3B shows that the proportion of owners in high coping levels were the main recipients of assistance (about 33%) followed by those in medium coping levels. This is indicative of a possible positive correlation between the two variables. About 90% of owners in high coping levels received assistance mainly from the Government, while those in medium coping levels received assistance mainly from the Community/NGOs. Lastly, the largest proportion of owners who experienced a reduction in their income have high coping levels, followed by medium, low, and none coping levels in that sequence. Owners that have experienced an increase and the same income levels are found mainly with "None" coping level (see Figure 3D). We present the basic descriptive statistics of these variables in the section below.

Covid Economics 51, 7 October 2020: 159-184

Covid Economics 51, 7 October 2020: 159-184

Table 1: Household, individual and family business attributes

Attributes	% of household
Family business: Since the beginning of 2020, did you or any HH member operate a family business?	N=24,124
1. Yes	9503
Coping: How did your HH cope with the [SHOCK]? (Yes)	N=3,018
1. Received assistance from friends & family	10.77
2. Borrowed from friends & family	9.68
3. Credited purchases	4.54
4. Delayed payment obligations	1.76
5. Sold harvest in advance	4.71
6. Credited purchases	4.54
7. Delayed payment obligations	1.76
8. Sold harvest in advance	4.71
9. Sale of assets (agric and no-agric)	4.87
10. Engaged in additional income-generating avenues	7.65
11. Relied on savings	25.71
12. Reduced food consumption	30.45
13. Reduced non-food consumption	18.03
14. Took a loan from a financial institution	0.20
Income: Since mid-March, has income from [SOURCE]?	N= 3,078
1. Increased	5.39
2. Same	17.58
3. Reduced	77.03
Assistance: Since mid-March, has HH received any assistance in the form of [ASSISTANCE]?	N=9,503
1. Yes	4.65
2. No	95.35
Source of assistance: What was the main source of this [ASSISTANCE]? (Yes)	N= 438
1. Federal government	10.96
2. State government	23.74
3. Local government	6.39
4. Community organization/ cooperative	14.84
5. NGO	7.31
6. Religious bodies	19.86

Covid Economics 51, 7 October 2020: 159-184

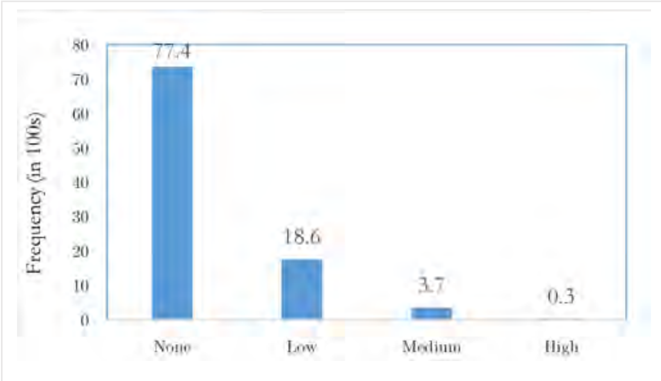


Figure 3A: Coping levels of individuals who have operated family businesses

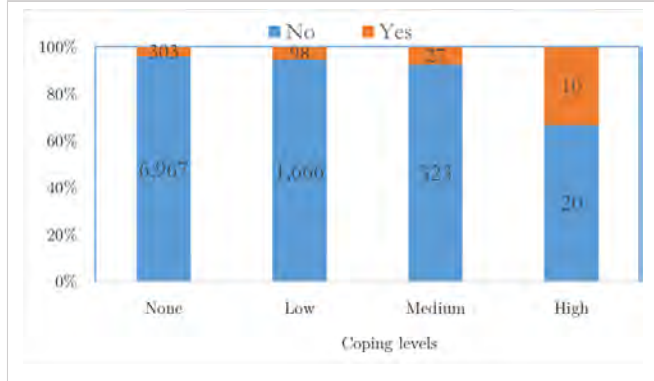


Figure 3B: Coping levels and assistance by individuals who have operated family businesses

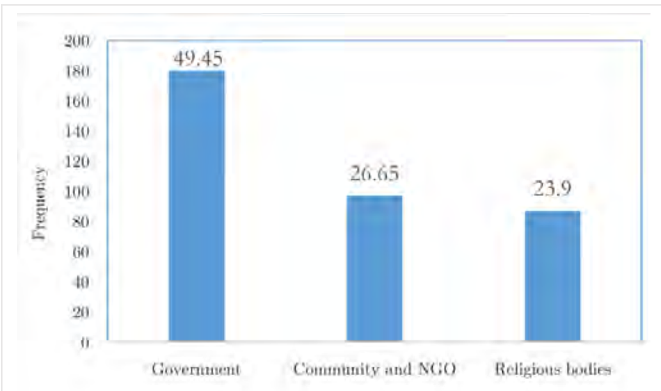


Figure 3C: Sources of assistance received by individuals who have family operated family businesses

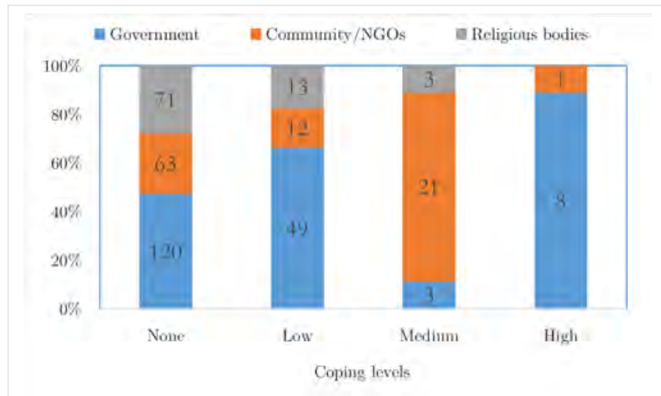


Figure 3D: Sources of assistance and coping levels of individuals who have operated family businesses.

CEPR PRESS

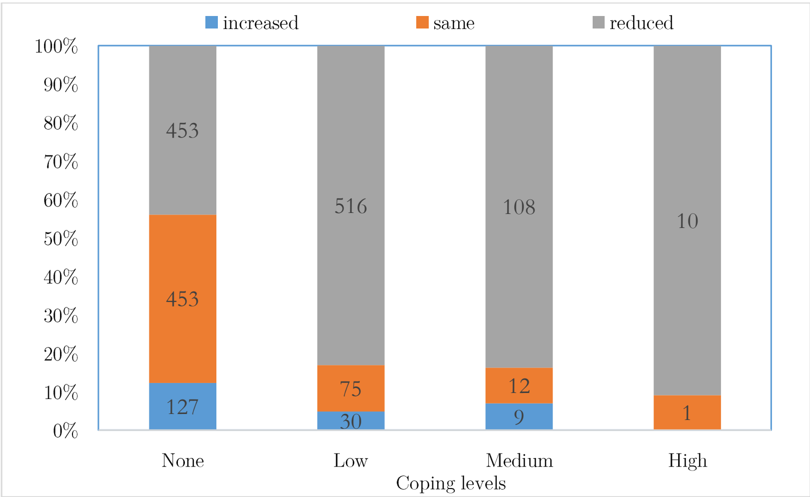


Figure 3E: Changes in income and coping levels of individuals who have operated family businesses

3.2. Variables and descriptive statistics

As noted earlier, this paper examines how changes in income and social assistance given due to COVID-19 affect the coping mechanisms and strategies of individuals who operate family business since the beginning of March in Nigeria. To this end, we first restrict our sample to individuals who operate a family business (owners) and then define the outcome variable as the number of coping strategies adopted by this restricted sample. In particular, our dependent variable measures how individuals who operate a family business cope with the shock of COVID-19 since mid-March. It is as an ordered categorical variable between 0 – 3, assuming 0 if the household of the individual has no coping strategy (None), 1 if the household of the individual has introduced only 1 coping strategy (Low), 2 if the household of the individual has introduced 2 coping strategies (Medium), and 3 if the household of the individual has introduced more than 3 coping strategies (High).

To measure changes in income of the individual since mid-March, we use a variable in the COVID-19 NLPS data set to generate 3 levels of changes income: 1 if the individual experienced an increase; 2 if the individual experienced the same level; and 3 if the individual experienced a reduction in income. The empirical literature shows the worsening levels of income of workers resulting from job losses and business closures due to lockdown and restrictions on the movement of people (Balde *et al.*, 2020; Beatric *et al.*, 2020; Chetty *et al.*, 2020; Islam *et al.*, 2020). Several empirical findings suggest that social safety nets are

Covid Economics 51, 7 October 2020: 159-184

instrumental in helping individuals and households mitigate the negative consequences of the COVID-19 pandemic (Balde *et al.*, 2020; Beaunoyer *et al.*, 2020). We employ a measure of assistance that indicates whether the household has received any assistance in food, cash transfers, or other in-kind transfers (excluding food) from the Government, Community organization or cooperative, NGO, or Religious bodies since mid-March. To further analyze the possible heterogeneity in the effect of the main source of assistance on how individuals in households cope with the shock of COVID-19, we constructed and defined 'Source of assistance' as 0 if the individual's household has received no assistance, 1 if the household has received assistance mainly from Government, 2 if the household has received assistance mainly from a Community organization or cooperative or NGO, and 3 if the household has received assistance mainly from Religious bodies.

We control for other household, individual, and family business attributes in the model. The household's location is expected to affect the coping strategies of its members, given that COVID-19 is more prevalent in urban areas. As a result, we expect individuals operating family business and are located in urban areas to have higher coping levels than those situated in rural areas. Females tend to have higher care responsibilities at the household level than males (Balde, 2020; Balde *et al.*, 2020; Avenyo *et al.*, 2020). With the increase in the care needs of households due to the global health crisis by COVID-19, females affected by the shock are expected to have higher coping levels. Individuals who worked in mid-March are expected to have lower coping levels since they may be less likely to use their savings, borrowed from friends and family, and reduced consumption of food and non-food items during the period. Households with a larger number of working members may have less dependency and hence lower levels of coping due to COVID-19. Having younger children in a household suggest higher levels of vulnerability and hence higher coping level. Empirical evidence suggests that some sectors are more vulnerable to COVID-19 than others (Balde *et al.*, 2020). We also control for 3 categories of the sector of the family business: Agriculture, fishing and hunting; Manufacturing/Mining; and Services. Finally, we control for the 36 states and the capital territory that constitute the Federal Republic of Nigeria. Table 2 presents the summary statistics of all variables introduced in our model.[6]

[6] See Table A1 in the appendix for the definition of all variables in our model.

Covid Economics 51, 7 October 2020: 159-184

Table 2: Definition of variables and sample descriptive statistics

Variable name	Mean	SD	Min	Max
Coping	0.296	0.560	0	3
Income	2.716	0 .558	1	3
Assistance	0.047	0.211	0	1
Source of assistance	1.744	0.819	0	3
Residence	0.393	0.489	0	1
Gender	0.499	0.500	1	0
Worked in mid-March	0.823	0.382	1	0
Sector	2.969	0.222	1	3
HH size	12.713	2.707	12	34
Children in primary school	0.944	0.230	0	1
State			1	37

3.3. Model and estimation strategy

In line with the discussions above, and given that our dependent variable (coping) is grouped into four ordered categories ranging from 0-3, (0= no coping strategy, 1= one coping strategy, 2= two coping strategies, 3= three or more coping strategies), we follow Gebreyeesus & Mohnen (2013) to formulate ordered probit econometric model, where we define Coping* as:

$$Coping_i^* = x_i\beta + \mu_i \qquad (1)$$

$$Prob\ Coping_i = 0|x_i \ = \Phi(\varepsilon_1 - x_i\beta) \qquad (2i)$$

$$Prob\ Coping_i = 1|x_i \ = \Phi\ \varepsilon_2 - x_i\beta \ - \Phi(\varepsilon_1 - x_i\beta) \qquad (2ii)$$

$$Prob\ Coping_i = 2|x_i \ = \Phi\ \varepsilon_3 - x_i\beta \ - \Phi\ \varepsilon_2 - x_i\beta \qquad (2iii)$$

$$Prob\ Coping_i = 3|x_i \ = 1 - \ \Phi(\varepsilon_3 - x_i\beta) \qquad (2iv)$$

where $Coping_i^*$ is an unobserved latent variable. x_i is a vector of all changes in income, assistance, and all other household, individual, family business, and fixed effects such as State of the household. β is a set of unknown coefficients of x_i to be estimated, while μ_i is assumed to be an error term with classical characteristics. Φ is a cumulative normal distribution function (c.d.f) of ε_i. ε_1, ε_2 and ε_3 are the threshold parameters.

One of our main variables of interest, assistance, is likely to be explained by other factors that may not directly explain coping levels. To test this, we formulate a structural model in conditional mixed process (cmp) (Roodman, 2011) whereby we jointly estimate the coping (oprobit) and assistance (probit) equations using maximum likelihood as:

Covid Economics 51, 7 October 2020: 159-184

$$Coping_i^* = Q_i\beta_1 + Assistance_i\beta_2 + \varepsilon_1 \quad (3i)$$
$$Assistance_i = Z_i\delta + \varepsilon_2 \quad (3ii)$$

where Q_i is the set of all explanatory variables (a subset of x in equation 1 that excludes assistance) that may explain coping levels across households, while Z is a vector of all factors that explain the probability of households receiving assistance. We introduce "shock" as our exclusion restriction since households exposed to the COVID-19 shock may be expected to be vulnerable and, therefore, more likely to receive assistance.

4. Results and Discussion

Table 3 shows the estimation results from the simultaneous equation (i.e., equation 3) using a recursive model. All estimations are clustered at the household level and are weighted using sampled household weights[7] provided in the data. To test the parallel regression assumption and to compare our results, we first estimate and present results using ordered probit and then the estimations for the recursive model (ordered probit simultaneous equation FIML). The ordered probit results are reported in columns 1-2, while our preferred results from the ordered probit simultaneous equation model are reported in columns 3-6. We also report results for the reduced form regressions for assistance below columns 3-6. Columns 3-4 report results when we use assistance in our estimation, while columns 5-6 consider the main sources of assistance. Comparing the estimation results from the ordered probit (column 2) and ordered probit simultaneous equation (column 4) shows that the magnitude of both the assistance and changes in income (reduced) coefficients are larger and statistically significant at 5%. These suggest that ordered probit estimation biases our estimates downwards. All other results remain similar qualitatively. We, therefore, proceed to interpret only the results from the recursive model.

Our preferred estimation results showing the effect of assistance and change in income on the level of coping strategies are reported in columns 3-4. The extended results reported in column 4 show that household members operating a family business who experienced a reduction in their income tend to adopt higher levels of coping strategies than those with increased income levels. As noted earlier, family-owned businesses have the unique characteristic that external crisis such as the COVID-19 pandemic hits their owners twice: once as a private citizen and second as business owners (Runyan, 2006; Kraus *et al.*, 2020). Hence, the results reported in columns 3-4 are intuitive since a fall in family business owners' income due to COVID-19 poses a risk to the survival of the business and their households.

[7] For details on how the weights are generated in the survey, see https://microdata.worldbank.org/index.php/catalog/3712/study-description#metadata-disclaimer_copyright

Covid Economics 51, 7 October 2020: 159-184

Covid Economics 51, 7 October 2020: 159-184

Hence, the immediate response to this existential threat is to adopt different coping strategies. Our results also show that assistance is positive and statistically significant, suggesting that individuals that receive any form of assistance tend to have higher coping levels as compared with those without assistance. This result is in line with our conjecture as assistance may serve as a buffer to those affected by shocks, in this case, COVID-19. This form of social safety net may help affected people to better cope and mitigate the negative consequences of the COVID-19 shock.

Given that the reported coefficients from oprobit models are not "easily interpretable" (Gebreyeesus & Mohnen, 2013), we report the computed average marginal effects of all our explanators for each level of coping in Table 4. Columns 1-4 report the average marginal effects of assistance and all other explanatory variables on each specific level of coping. Based on the average marginal effects, family business owners that experienced a reduction in income tend to have a lower probability (about 11%, of being in the "none" coping level) and have a higher probability of being in the low and medium levels of coping by 6.9% and 3.3% compared with their counterparts that have experienced increases in income levels. We do not find any significant statistical difference between family business owners in high coping levels that have experienced a reduced income and those that experienced an increase in income. Also, compared with those that experienced increases in income, we do not find statistically significant differences between having the same income and all the levels of coping. Family business owners in households that received assistance have a lower probability of being in the "none" coping level by about 53% and raises the probability of being in the low, medium, and high coping levels by 33%, 16%, and 3.5% respectively. These suggest that assistance helps to cushion family business owners, their households, and their businesses in extension, and enables assistance receiving household members who operate family businesses to better cope with the vagaries of COVID-19.

Columns 5-8 report the average marginal effects when we consider the main source of assistance on each coping level. Compared with those who do not receive any form of assistance, the results show that assistance from government reduces the probability of being in none coping level by 55%, and increases the probability of being in the low, medium and high coping levels by 35%, 16%, and 3.4%, respectively. Similar results are also obtained for those who received assistance from NGOs and religious bodies. The results thus reiterate the importance of social assistance in cushioning the negative COVID-19 induced shocks.

In terms of other explanatory variables, the results show that having children in primary school has a positive and significant effect on coping levels. Comparing with those who have no children in primary school, the average marginal effects show that family business owners with children in primary school tend to have a lower probability of being in the 'none' coping level by 16%, while it increases the probability of being in the lower, medium and high coping levels by 10%, 5%, and 1% respectively. This supports our motivation that children may require additional care and needs leading to higher coping levels. Also, with the closure of schools in the country, parents are forced to stay home to take care of their children, especially those in primary school and below. This may have a further dampening effect on households' wealth leading to the increasing need for coping strategies. Compared with the agricultural sector, family businesses in manufacturing tend to have lower coping levels, with average marginal effects showing that manufacturing family businesses tend to have a lower probability of being none, low and medium levels of coping by 77%, and 81% respectively. This may be due to the lower human interaction compared with service and agricultural sector based family businesses where interaction may be more intensive. We do not find a significant statistical difference between service sector-based family businesses and that of those based in the agricultural sector.

In an unreported result, we estimate several regressions interacting assistance with the different levels of income. While the coefficients of assistance and changes in income remain statistically similar to those presented and discussed above, the interacting terms are mostly insignificant. These results are not reported in the paper.

Table 3: Effect of assistance and changes in income on coping strategies of family businesses

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation method	Ordered probit		Ordered probit simultaneous equation FIML			
			Coping			
Assistance	0.371**	0.210	0.443**	1.872***		
	(2.11)	(1.23)	(2.05)	(13.75)		
Type of assistance[a]						
Government					0.292*	1.928***
					(1.67)	(9.23)
NGO					0.862**	1.950***
					(2.42)	(8.47)
Religious body					-0.120	1.484***
					(-0.47)	(5.58)
Change in income[b]						
Same		0.102		0.071		0.052
		(0.54)		(0.38)		(0.28)
Reduced		0.243		0.390**		0.365**
		(1.50)		(2.24)		(2.14)
HH size (log)		-0.332		-0.377		-0.387
		(-0.78)		(-0.94)		(-0.96)
Children in primary sch.		0.596***		0.557**		0.561**

Covid Economics 51, 7 October 2020: 159-184

Covid Economics 51, 7 October 2020: 159-184

		(2.59)		(2.42)		(2.44)
Gender		-0.002		0.006		0.010
		(-0.02)		(0.08)		(0.12)
Urban		0.158		0.188		0.174
		(1.45)		(1.49)		(1.40)
Worked in March		0.082		0.039		0.042
		(0.73)		(0.34)		(0.37)
Sector[c]						
Manufacturing/Mining		-0.765*		-0.806*		-0.805*
		(-1.78)		(-1.90)		(-1.89)
Services		-0.081		-0.102		-0.103
		(-0.28)		(-0.35)		(-0.36)
Assistance						
Change in income						
Same			-0.940***	-0.488**	-0.952***	-0.570**
			(-3.85)	(-2.24)	(-3.84)	(-2.38)
Reduced			-0.839*	-0.366	-0.856*	-0.474
			(-1.77)	(-1.15)	(-1.81)	(-1.31)
HH size (log)			1.937	2.399**	1.925	2.373**
			(1.45)	(2.49)	(1.44)	(2.46)
Children in ps			0.823*	0.715*	0.827*	0.730*
			(1.86)	(1.69)	(1.85)	(1.71)
Shock of COVID			0.325	1.112***	0.254	1.078***
			(1.00)	(4.99)	(0.69)	(4.54)
Gender			-0.157	-0.135	-0.160	-0.127
			(-0.86)	(-0.95)	(-0.87)	(-0.79)
Urban			-0.185	-0.233	-0.185	-0.255
			(-0.48)	(-0.88)	(-0.48)	(-0.88)
Worked in March			0.195	0.171	0.190	0.167
			(0.64)	(0.48)	(0.62)	(0.46)
Rho_1_2_cons			-0.284	-1.663***	-0.228	-1.601***
			(-1.25)	(-4.77)	(-0.93)	(-4.45)
cut_1_1_cons	0.757***	1.732	0.760***	1.777	0.757***	1.719
	(18.40)	(1.42)	(18.30)	(1.50)	(18.34)	(1.46)
cut_1_2_cons	1.790***	2.860**	1.794***	2.863**	1.796***	2.810**
	(25.13)	(2.26)	(25.39)	(2.34)	(25.77)	(2.30)
cut_1_3_cons	3.064***	4.114***	3.064***	3.936***	3.085***	3.901***
	(15.85)	(3.38)	(15.98)	(3.32)	(15.30)	(3.30)
N	9414	2796	9414	2789	9414	2789

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. [a] No assistance as base outcome; [b] Increased income as base outcome; [c] Agriculture sector as base outcome;. All extended regressions include state fixed effects.

Title 4: Average marginal effect by coping strategy category

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Based on the estimation on Table 3 column 4				Based on the estimation on Table 3 column 6			
Coping level	None	Low	Medium	High	None	Low	Medium	High
Assistance	-0.528***	0.333***	0.160***	0.035**				
	(-14.51)	(10.19)	(6.71)	(2.34)				
Type of assistance								
Government					-0.546***	0.348***	0.164***	0.034**
					(-9.35)	(7.41)	(5.91)	(2.38)
NGO					-0.552***	0.352***	0.166***	0.034**
					(-8.54)	(7.19)	(5.68)	(2.30)
Religious body					-0.420***	0.268***	0.126***	0.026**
					(-5.71)	(5.74)	(4.20)	(2.10)
Change in income								
Same	-0.020	0.013	0.006	0.001	-0.015	0.009	0.004	0.001
	(-0.38)	(0.38)	(0.38)	(0.37)	(-0.28)	(0.28)	(0.28)	(0.27)
Reduced	-0.110**	0.069**	0.033**	0.007	-0.103**	0.066**	0.031**	0.006
	(-2.26)	(2.23)	(2.18)	(1.61)	(-2.16)	(2.15)	(2.10)	(1.53)
HH size (log)	0.106	-0.067	-0.032	-0.007	0.110	-0.070	-0.033	-0.007
	(0.94)	(-0.96)	(-0.89)	(-0.93)	(0.97)	(-0.98)	(-0.91)	(-0.95)
Children in ps	-0.157**	0.099**	0.048**	0.011*	-0.159**	0.101**	0.048**	0.010*
	(-2.43)	(2.37)	(2.32)	(1.77)	(-2.45)	(2.39)	(2.34)	(1.76)
Gender	-0.002	0.001	0.001	0.000	-0.003	0.002	0.001	0.000
	(-0.08)	(0.08)	(0.08)	(0.08)	(-0.12)	(0.12)	(0.12)	(0.12)
Urban	-0.053	0.033	0.016	0.004	-0.049	0.031	0.015	0.003
	(-1.48)	(1.47)	(1.44)	(1.31)	(-1.39)	(1.39)	(1.35)	(1.24)
Worked in March	-0.011	0.007	0.003	0.001	-0.012	0.008	0.004	0.001
	(-0.34)	(0.34)	(0.34)	(0.34)	(-0.37)	(0.37)	(0.37)	(0.37)
Sector								
Manufacturing/Mining	0.227*	-0.143*	-0.069*	-0.015	0.228*	-0.145*	-0.069*	-0.014
	(1.90)	(-1.91)	(-1.79)	(-1.50)	(1.90)	(-1.91)	(-1.78)	(-1.49)
Services	0.029	-0.018	-0.009	-0.002	0.029	-0.019	-0.009	-0.002
	(0.35)	(-0.35)	(-0.35)	(-0.35)	(0.36)	(-0.36)	(-0.35)	(-0.35)

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To further check our results' robustness, we estimate and report unweighted regressions for all our preferred results in Table 3. We report the results for this exercise in Table 5. As the results indicate, the results reported in the Table, are qualitatively similar to our preferred results in Table 3 (column 3-6) above, thereby reinforcing the robustness of our results and findings that assistance and changes in income of family business owners are key factors influencing the coping strategies of Nigerians households in the era of the COVID-19 pandemic.

Covid Economics 51, 7 October 2020: 159-184

Table 5: Effect of assistance and changes in income on coping strategies of family businesses (unweighted)

	(1)	(2)	(3)	(4)
Estimation method	Ordered probit simultaneous equation FIML			
	Coping			
Assistance	0.443**	1.872***		
	(2.05)	(13.75)		
Type of assistance[a]				
Government			0.361*	2.281***
			(1.81)	(4.99)
NGO			0.568**	2.112***
			(2.24)	(8.11)
Religious body			-0.139	1.551***
			(-0.64)	(5.38)
Change in income[b]				
Same		0.071		0.165
		(0.38)		(0.87)
Reduced		0.390**		0.406**
		(2.24)		(2.40)
HH size (log)		-0.377		-0.432
		(-0.94)		(-1.18)
Children in ps		0.557**		0.434**
		(2.42)		(2.30)
Gender		0.006		0.020
		(0.08)		(0.30)
Urban		0.188		0.237**
		(1.49)		(2.46)
Worked in March		0.039		0.027
		(0.34)		(0.23)
Sector[c]				
Manufacturing/Mining		-0.806*		-0.711*
		(-1.90)		(-1.96)
Services		-0.102		-0.084
		(-0.35)		(-0.37)
Assistance				
Change in income				
Same	-0.940***	-0.488**	-0.673***	-0.491**
	(-3.85)	(-2.24)	(-2.80)	(-2.38)
Reduced	-0.839*	-0.366	-0.609**	-0.406*
	(-1.77)	(-1.15)	(-2.08)	(-1.78)
HH size (log)	1.937	2.399**	2.707**	2.503***
	(1.45)	(2.49)	(2.46)	(2.97)
Children in ps	0.823*	0.715*	0.291	0.204
	(1.86)	(1.69)	(0.71)	(0.47)
Shock of COVID	0.325	1.112***	0.052	1.087***

Covid Economics 51, 7 October 2020: 159-184

	(1.00)	(4.99)	(0.17)	(4.35)
Gender	-0.157	-0.135	-0.166	-0.005
	(-0.86)	(-0.95)	(-1.10)	(-0.02)
Urban	-0.185	-0.233	-0.071	-0.075
	(-0.48)	(-0.88)	(-0.24)	(-0.31)
Worked in March	0.195	0.171	-0.120	-0.077
	(0.64)	(0.48)	(-0.43)	(-0.27)
Rho_1_2_cons	-0.284	-1.663***	0.033	-1.363***
	(-1.25)	(-4.77)	(0.14)	(-3.14)
cut_1_1_cons	0.760***	1.777	0.766***	1.586
	(18.30)	(1.50)	(27.42)	(1.51)
cut_1_2_cons	1.794***	2.863**	1.768***	2.604**
	(25.39)	(2.34)	(36.17)	(2.48)
cut_1_3_cons	3.064***	3.936***	2.763***	3.558***
	(15.98)	(3.32)	(16.73)	(3.45)
N	9414	2796	9414	2796

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *t* statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. [a] No assistance as base outcome; [b] Increased income as base outcome; [c] Agriculture sector as base outcome;. All extended regressions include state fixed effects.

5. Conclusion

The impact of COVID-19 on human lives and businesses have necessitated several coping strategies by national governments, international organizations, and individuals to curb the pandemic's negative effects. Family businesses provide a critical economic contribution to employment and income and buffer households from shock and uncertainty. Nonetheless, the literature also highlights the vulnerability of family businesses to shocks. The COVID-19 pandemic and the resultant fall in business activities threaten to upend the livelihoods of many family-business-owned households in developing countries. Despite, the growing literature on the micro-economic impacts of COVID-19 largely lacks evidence on how family businesses and their owners are coping with the crisis, and how available social safety nets help these owners to ameliorate their wellbeing.

In this paper, we attempt to understand how social assistance and changes in income due to COVID-19 are influencing the coping strategies of individuals operating family businesses. We use new data on the "Nigeria Baseline COVID-19 National Longitudinal Phone Survey 2020" by the World Bank, and estimate opbrobit simultaneous equation model in our empirical analysis. We find two main things. First, the paper shows that family business owners' income changes influence how they cope with the COVID-19 pandemic, with family business owners experiencing a reduction in income having the highest level of coping. Second, family business owners whose households receive assistance tend to better mitigate the negative consequences of COVID-19 by increasing their level of coping. Our

Covid Economics 51, 7 October 2020: 159-184

results also indicate that family business owners with children in primary school have higher levels of coping than their counterparts without children in primary school. Manufacturing family businesses also tend to have lower coping levels than those in the agricultural sector. These results hold when we also consider the main sources of assistance received by households.

From a policy perspective, our results identify some of the most vulnerable groups to the COVID-19 pandemic in Nigeria, which include households with children in primary schools, households with family business owners experiencing a reduction in income, and/or are in high-risk sectors. It also underscores the importance of social safety nets such as assistance to households in mitigating and coping with the negative consequences of COVID-19. To reduce income inequality and spur economic growth and development in the post-COVID-19 pandemic, the need for targeted social assistance to this vulnerable group(s) cannot be overemphasized. Hence, it may be of utmost policy relevance for the government to support family-owned businesses and their owners to better mitigate the pandemic's negative consequences on vulnerable households and their members.

The analysis in this paper is based only on the "Nigeria Baseline COVID-19 National Longitudinal Phone Survey 2020". A natural extension of the paper could be to use additional data from Round 2, and Round 3 surveys. Despite, the paper provides some insights on how changes in income and social assistance due to COVID-19 affect the coping strategies of individuals who operate family-owned businesses in Nigeria.

Covid Economics 51, 7 October 2020: 159-184

Covid Economics 51, 7 October 2020: 159-184

References

Adams-Prassl, A., Boneva, T., Golin, M., Rauh, C. (2020). Inequality in the impact of the coronavirus shock: Evidence from real-time surveys. IZA DP No. 13183.

Adegite, V., & Abu, N. (2020). Covid-19 pandemic—Nigeria's fiscal and economic measures. BloombergTax. Retrieved from https://news.bloombergtax.com/daily-tax-report-international/insight-covid-19-pandemic-nigerias-fiscal-and-economic-measures

Avenyo, E., Francois, J., & Zinyemba, T. (2020). *COVID-19, lockdowns, and Africa's informal sector: Lessons from Ghana*. UNU-MERIT Working Paper No. 2020-028

Balde, R. (2020). Unpaid care work inequality and women's employment outcomes in Senegal. in: Konte, Maty Nyasha Tirivayi (eds.), Women and Sustainable Human Development: Empowering Women in Africa, Palgrave Macmillan, Geneva.

Balde, R., Boly, M., & Avenyo, E. (2020). *Labor market effects of COVID-19 in sub-Saharan Africa: An informality lens from Burkina Faso, Mali, and Senegal*. UNU-MERIT Working Paper No. 2020-022.

Bartik, A., Bertrand, M., Cullen, Z., Glaeser, E., Luca, M., & Stanton, C. (2020). *How are small businesses adjusting to COVID-19? Early evidence from a survey*. NBER Working Paper No. 26989.

Baker, S., Bloom, N., Davis, S., Kost, K., Sammon, M., & Viratyosin, T. (2020). *The unprecedented stock market impact of COVID-19*. National Bureau of Economic Research No. 26945.

Beaunoyer, E., Dupere, S., & Guitton, M. (2020). COVID-19 and digital inequalities: Reciprocal impacts and mitigation strategies. *Computers in Human Behavior*, 111, 106424

Beine, M., Bertoli, S., Chen, S., Ambrosio, C., Islam, N., Koulovatianos, C., Leduc, K., Lorenz, N., Peluso, E., Peroni, C., Picard, P., Pieretti, P., Tenikue, M., Theloudis, A., & Kerm, P. (2020). *Economic effects of Covid-19 in Luxembourg: First RECOVid working note with preliminary estimates*. Working Paper, Research Luxembourg.

Campello, M., Kankanhalli, G., & Muthukrishnan, P. (2020). *Corporate hiring under COVID-19: Labor market concentration, downskilling, and income inequality*. NBER Working Paper No. 27208

Cater, J., & Schwab, A. (2008). Turnaround strategies in established small family firms. *Family Business Review*, 21, 31-50.

Chetty, R., Friedman, J., Hendren, N., & Stepner, M. (2020). *How did COVID-19 and stabilization policies affect spending and employment? A new real-time economic tracker based on private-sector data*. NBER Working Paper No. 27431

Chiou, L., & Tucker, C. (2020). *Social distancing, internet access, and inequality*. NBER Working Paper No. 26982.

Cho, S., Lee, J., & Winters, J. (2020). *Employment impacts of the Covid-19 pandemic across metropolitan status and size*. IZA Discussion Paper No. 13468

Dingel, J., & Neiman, B. (2020). *How many jobs can be done at home?* Becker Friedman Institute White Paper, June 2020.

Dixit, S., Ogundeji, Y., Onwujekwe, O. (2020). How well has Nigeria responded to COVID-19. BROOKINGS. Retrieved from https://www.brookings.edu/blog/future-development/2020/07/02/how-well-has-nigeria-responded-to-covid-19/.

Faghfouri, P., Kraiczy, N., Hack, A., & Kellermanns, F. (2015). Ready for a crisis? How supervisory boards affect the formalized crisis procedures of small and medium-sized family firms in Germany. *Review of Managerial Science*, 9, 317-338.

Fairlie, R. (2020). *The impact of Covid-19 on small business owners: Evidence of early-stage losses from the April 2020 current population survey*. NBER Working Paper No. 27309.

IMF. (2020). The Great Lockdown. Retrieved from
https://www.imf.org/en/Publications/WEO/Issues/2020/04/14/World-Economic-Outlook-April-2020

Gebreeyesus, M., & Mohnen, P. (2013). Innovation performance and embeddedness in networks: evidence from the Ethiopian footwear cluster. *World Development*, *41*, 302-316.

Gomez-Mejia, L., Cruz, C., Berrone, P., & De Castro, J. (2011). The bind that ties: Socioemotional wealth preservation in family firms. *Academy of Management Annals*, 5, 653-707.

Gomez-Mejıa, L., Haynes, K., Nunez-Nickel, M., Jacobson, K., & Moyano-Fuentes, J. (2007). Socioemotional wealth and business risks in family-controlled firms: evidence from Spanish olive oil mills. *Administrative Science Quarterly*, 52, 106-137.

Harapan, H., Itoh, N., Yufika, A., Winardi, W., Keam, S., Te, H., & Mudatsir, M. (2020). Coronavirus disease 2019 (COVID-19): A literature review. *Journal of Infection and Public Health, 13*(5), 667-673.

Herbane, B. (2013). Exploring crisis management in UK small- and medium-sized enterprises. *Journal of Contingencies and Crisis Management*, 21, 82-95.

Huang, C., Wang, Y., … & Cao, B. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet,* 395, 497–506.

Kahn, L., Lange, F., & Wiczer, D. (2020). *Labor demand in the time of COVID-9: Evidence from vacancy postings and UI claims.* National Bureau of Economic Research No. 27061.

Kalenkoski, C., & Pabilonia, S. (2020). *Initial impact of the COVID-19 pandemic on the employment and hours of self-employed coupled and single workers by gender and parental status*. IZA DP No. 13443

Covid Economics 51, 7 October 2020: 159-184

KPMG. (2020). Nigeria: Government and institution measures in response to COVID-19. Retrieved from https://home.kpmg/xx/en/home/insights/2020/04/nigeria-government-and-institution-measures-in-response-to-covid.html

Kraus, S., Clauss, T., Breier, M., Gast, J., Zardini, A, Tiberius, V. (2020). The economics of COVID-19: Initial empirical evidence on how family firms in five European countries cope with the corona crisis. *International Journal of Entrepreneurial Behavior & Research*, https://doi.org/10.1108/IJEBR-04-2020-0214

Kraus, S., Moog, P., Schlepphorst, S., & Raich, M. (2013). Crisis and turnaround management in SMEs: a qualitative-empirical investigation of 30 companies. *International Journal of Entrepreneurial Venturing*, 5, 406-430.

National Bureau of Statistics (2020). Nigeria COVID-19 National Longitudinal Phone Survey (COVID-19 NLPS) 2020. Dataset downloaded from https://microdata.worldbank.org/index.php/catalog/3712/study-descriptionon [04/06/2020].

NCDC. (2020a). COVID-19 outbreak in Nigeria- Situation Report S/N: 001, February 29, 2020. Retrieved from https://ncdc.gov.ng/diseases/sitreps/?cat=14&name=An%20update%20of%20COVID-19%20outbreak%20in%20Nigeria

NCDC. (2020b). 100 Days of Nigeria COVID-19 Response. Retrieved from https://ncdc.gov.ng/news/253/100-days-of-nigeria-covid-19-response

NCDC. (2020c). COVID-19 Nigeria - Wednesday 3:23 pm July 22 2020. Retrieved from: https://covid19.ncdc.gov.ng/#!

Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-Jabir, A., Iosifidis, C., Agha, M., & Agha, R. (2020). The socioeconomic implications of the coronavirus pandemic (COVID-19): A review. *International Journal of Surgery*, 78, 185–193.

Ohia, C., Bakarey, A. S., & Ahmad, T. (2020). COVID-19 and Nigeria: Putting the realities in context. *International Journal of Infectious Diseases, 95*, 279–281

PWC (2020). Government's COVID-19 response measures. Retrieved from https://www.pwc.com/ng/en/covid-19/government-covid-19-response-measures.html.

Runyan, R. (2006). Small business in the face of crisis: identifying barriers to recovery from a natural disaster. *Journal of Contingencies and Crisis Management*, 14, 12-26.

Roodman, D. (2011). Fitting fully observed recursive mixed-process models with cmp. *The Stata Journal, 11*(2):159–206.

Roux-Dufort, C. (2007). Is crisis management (only) a management of exceptions? *Journal of Contingencies & Crisis Management*, 15, 105-114.

Covid Economics 51, 7 October 2020: 159-184

Xi, J., Kraus, S., Filser, M., & Kellermanns, F. (2015). Mapping the field of family business research: past trends and future directions. *International Entrepreneurship and Management Journal*, 11, 113-132.

Xiong, T-Y., Redwood, S., Prendergast, B., & Chen, M. (2020). Coronaviruses and the cardiovascular system: Acute and long-term implications. European Heart Journal, 41, 1798–800.

WHO. (2020a). Coronavirus disease (COVID-19) - Situation Report – 182. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200720-covid-19-sitrep-182.pdf?sfvrsn=60aabc5c_2

WHO. (2020b). Coronavirus. Retrieved from https://www.who.int/health-topics/coronavirus#tab=tab_1

Appendix

Table A1: Definition of variables and descriptive statistics

Variable name	Definition
Coping	A categorical variable that takes the value between 0 - 3 to measure how individuals who operate a family business cope with the shock of COVID-19 since mid-March: 0 if the household of the individual has no coping strategy (None), 1 if the household of the individual has introduced only 1 coping strategy (Low), 2 if the household of the individual has introduced 2 coping strategies (Medium), and 3 if the household of the individual has introduced more than 3 coping strategies (High).
Income	A categorical variable that measures changes in the income of individuals since mid-March.: 1, 2, or 3 if the household experienced an increase, the same, or a reduction in income.
Assistance	A dummy variable that assumes the value of 1 if the household has received any assistance in direct food, cash transfers, or other in-kind transfers (excluding food) from the Government, Community organization or cooperative, NGO, or Religious bodies since mid-March.
Source of assistance	A categorical variable that shows the main source of assistance received by the household since mid-March: 0 if no assistance, 1 if assistance is received from Government, 2 if assistance is received from a community organization or cooperative or NGO, and 3 if assistance is received from Religious bodies.
Residence	A dummy showing the location of the household, with 1 if the household resides in an urban area and 0 if rural.
Gender	A dummy showing the sex of the respondent, with 1 if male and 0 if female.
Working mid-March	A dummy variable that assumes the value of 1 if the respondent was working was working in mid-March and 0 if otherwise.
Sector	A categorical variable that best describes the sector of the family business with Agriculture, fishing and hunting (0.97) assuming the value of 1; Manufacturing/Mining (1.10) assuming the value of 2, and Services (97.92) assuming the value of 3.
HH size	A continuous variable indicating the number of members in the household.
Children in primary school	A dummy variable that assumes the value of 1 if the respondent had any children attending primary/secondary school before closed and 0 if otherwise.
Shock	A dummy variable that assumes the value of 1 if the household been affected by a shock since mid-March and 0 if otherwise.
State	A categorical variable that shows the 37 States in Nigeria.

Covid Economics 51, 7 October 2020: 159-184