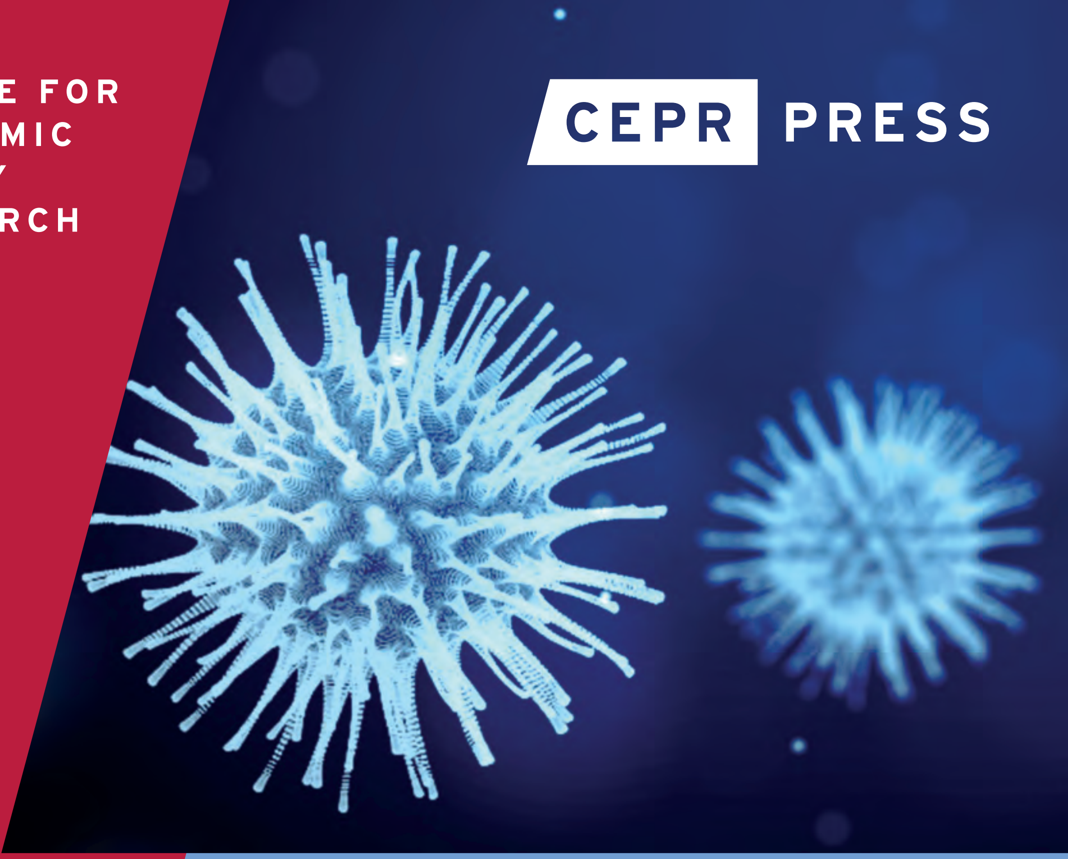
CENTRE FOR ECONOMIC POLICY RESEARCH

CEPR PRESS



COVID ECONOMICS

VETTED AND REAL-TIME PAPERS

ISSUE 6
17 APRIL 2020

Covid Economics
Vetted and Real-Time Papers

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

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

Contact: Submissions should be made at https://portal.cepr.org/call-papers-covid-economics-real-time-journal-cej. 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.

Ethics

Covid Economics will publish 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 journal's articles.

Editorial Board

Covid Economics
Vetted and Real-Time Papers

Issue 6, 17 April 2020

Contents

CEPR PRESS

Covid Economics **Issue 6, 17 April 2020**

US economic activity during the early weeks of the SARS-Cov-2 outbreak[1]

Daniel Lewis,[2] Karel Mertens[3] and James Stock[4]

Date submitted: 10 April 2020; **Date accepted:** 11 April 2020

This paper describes a weekly economic index (WEI) developed to track the rapid economic developments associated with the response to the novel Coronavirus in the United States. The WEI shows a strong and sudden decline in economic activity starting in the week ending March 21, 2020. In the most recent week ending April 4, the WEI indicates economic activity has fallen further to -8.89% scaled to 4 quarter growth in GDP.

1 Bi-weekly updates of the Weekly Economic Index are available at https://www.jimstock.org/. The views expressed are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York, the Federal Reserve Bank of Dallas or the Federal Reserve System. We are grateful to Mihir Trivedi and Eric Qian for research assistance, to Tyler Atkinson for useful suggestions, and to Mark Booth for sharing the tax collections data.

2 Economist, Federal Reserve Bank of New York.
3 Senior Economic Policy Advisor, Federal Reserve Bank of Dallas.
4 Professor of Political Economy, Harvard University.

Covid Economics 6, 17 April 2020: 1-21

Economists are well-practiced at assessing real economic activity based on a range of familiar aggregate time series, such as the unemployment rate, industrial production, or GDP growth. However, these series represent monthly or quarterly averages of economic conditions, and are only available at a considerable lag, after the month or quarter ends. When the economy hits sudden headwinds such as the COVID-19 pandemic, conditions can evolve rapidly. How can we monitor the high-frequency evolution of the economy in "real time"?

To address this challenge, this paper develops a Weekly Economic Index (WEI) that measures real economic activity at a weekly frequency and that can be updated relatively quickly.[2] Few of the government agency data releases macroeconomists often work with are available at weekly or higher frequency. Our weekly series instead come mostly from private sources such as industry groups that collect data for the use of their members, or from commercial polling companies. Financial data, such as stock market prices and interest rates, are also available at high frequency. We do not use financial data in the construction of the WEI, as our objective is to obtain a direct measure of real activity, and not of financial conditions.

To compute our index, we extract the first principal component from 10 weekly time series, using the sample from January 2008 to present. We scale our baseline index to four-quarter GDP growth, such that a reading of 2 percent in a given week means that if the week's conditions persisted for an entire quarter, we would expect, on average, 2 percent growth relative to a year previous.[3] The top panel in **Figure 1** plots the WEI based on data through April 9, 2020. The trough of the Great Recession is clearly visible, as well as the subsequent recovery. The WEI index also shows a modest decline during the 2015-2016 mini-recession, during which the energy and agricultural sectors as well as certain segments of the manufacturing economy experienced substantial slowdowns in growth.

The bottom panel in **Figure 1** shows the evolution of the WEI from January 2019 to its most recent value. As is clear from the figure, developments

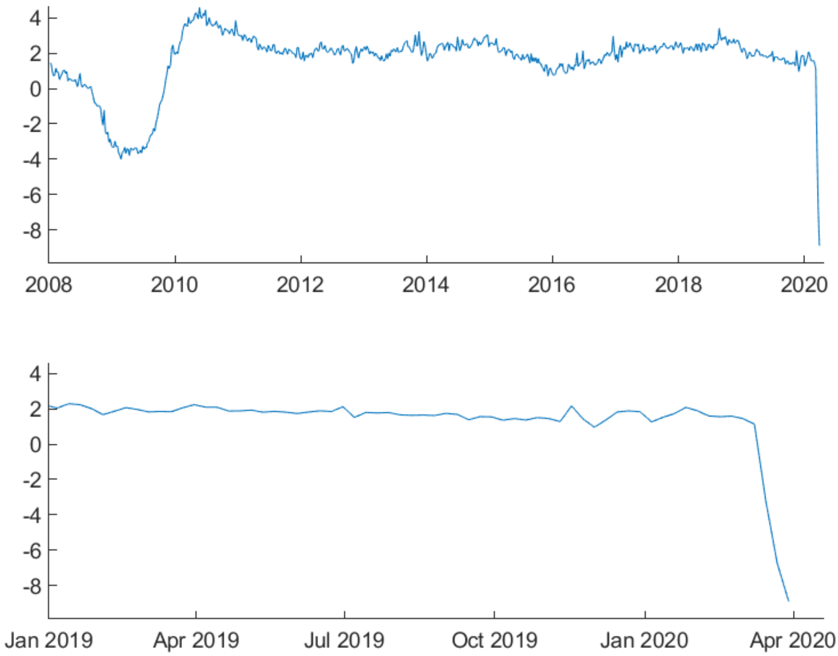
---

[2] The WEI builds on earlier work by Stock (2013). A preliminary version of WEI was presented in a NY Fed blog post, see Lewis, Mertens and Stock (2020).
[3] Specifically, the mean and standard deviation of the Weekly Economic Index have been adjusted so that they match the mean and standard deviation of the four-quarter growth of GDP from 2008 through the fourth quarter of 2019.

related to the Coronavirus pandemic have led the index to fall to levels below those of 2008 in recent weeks. Specifically, the WEI registers a strong and sudden decline in economic activity starting in the week ending March 21, 2020, falling to -3.23%. For reference, the WEI stood at 1.58 for the week ending February 29. The week ending March 21 saw an unprecedented 3.28 million initial UI claims (seasonally adjusted), a sharp decline in consumer confidence and fuel sales, and a more modest decline in steel production. There was also a countervailing surge in retail sales, as consumers took to stores to stock up on consumer staples. In the week ending March 28, the WEI fell to -6.75%. This further decline was driven by another sharp increase in unemployment insurance initial claims, which came in at 6.65 million (seasonally adjusted), far surpassing the prior week's record-setting release. The drop was reinforced by a major decline in fuel sales in response to stay-at-home orders and other restrictions, a fall in steel production, and a surge in continuing unemployment insurance claims (7.46 million seasonally adjusted), as well as modest decreases in electricity output, rail traffic, temporary and contract employment, and consumer confidence. In the most recent week ending April 4, the WEI fell to -8.89%. This week's decrease was again driven by initial unemployment insurance claims (6.61 million seasonally adjusted) and sharp decreases in fuel sales and steel production, and reinforced by falls in rail traffic, electricity output, and tax withholdings, while retail sales stalled.

To track the rapidly evolving economic fallout of the Coronavirus pandemic, the WEI is updated weekly every Tuesday and Thursday. The weekly updates contain preliminary estimates for the prior week based on the available data. The latter are based on estimated historical relationships between the WEI and the series available at the time of the update. The final values of the WEI are available after two weeks.

Figure 1: Weekly Economic Indicator (WEI)



Notes: Based on data available through April 9, 2020. The units are scaled to 4-quarter GDP growth.

The rest of this paper describes the underlying weekly data series as well as the details behind the construction of WEI. We also document the close relationship between the WEI and widely used lower frequency indicators of aggregate economic activity in the US, such as real GDP growth and industrial production.

I. The Weekly Data Series

Table 1 below lists the series we use to construct our baseline WEI. These include a measure of same-store retail sales, an index of consumer sentiment, initial and continued claims for unemployment insurance, an index of temporary and contract employment, tax collections from paycheck withholdings, a measure of steel production, a measure of fuel sales, a measure of railroad traffic, and a measure of electricity consumption. Unless

Covid Economics 6, 17 April 2020: 1-21

the source already provides year-on-year growth rates, we transform all series to represent 52-week percentage changes, which also eliminates most seasonality in the data. **Figure 2** plots all the transformed series that serve as inputs to the index.

Table 1: Weekly Variables

Series	Native Units	Notes
Redbook Research: Same Store, Retail Sales Average, Y/Y % Chg.	NSA, Y/Y % Chg.	**The index is sales-weighted, year-over-year same-store sales growth for a sample of large US general merchandise retailers representing about 9,000 stores. By dollar value, the Index represents over 80% of the "official" retail sales series collected by the Department of Commerce. http://www.redbookresearch.com/**
Rasmussen Consumer Index	Index	**Daily survey of 1500 American adults Sun-Thurs. Index is a 3-day moving average based on five questions about the current state of both the economy and personal finances, whether the economy and personal finances are getting better or worse, and whether the economy is in a recession. https://www.rasmussenreports.com/**
Unemployment Insurance: Initial Claims	NSA, Thous.	**Number of claims filed by unemployed individuals after separation from an employer. Data collected from local unemployment offices. https://oui.doleta.gov/unemploy/**
Insured Unemployment (Continued Claims)	NSA, Thous.	**Number of continued claims filed by unemployed individuals to receive benefits. Data collected from local unemployment offices. https://oui.doleta.gov/unemploy/**
American Staffing Association Staffing Index	NSA, Jun-12-06=100	**The ASA Staffing Index tracks temporary and contract employment trends. Participants include a stratified panel of small, medium, and large staffing companies that together provide services in virtually all sectors of the industry. They account for about one-third of industry sales offices. https://americanstaffing.net/**
Federal Withholding Tax Collections	Y/Y % Chg.	**Treasury receipts of income and payroll taxes withheld from paychecks. The series is filtered for daily volatility patterns and adjusted for tax law changes. https://taxtracking.com/**

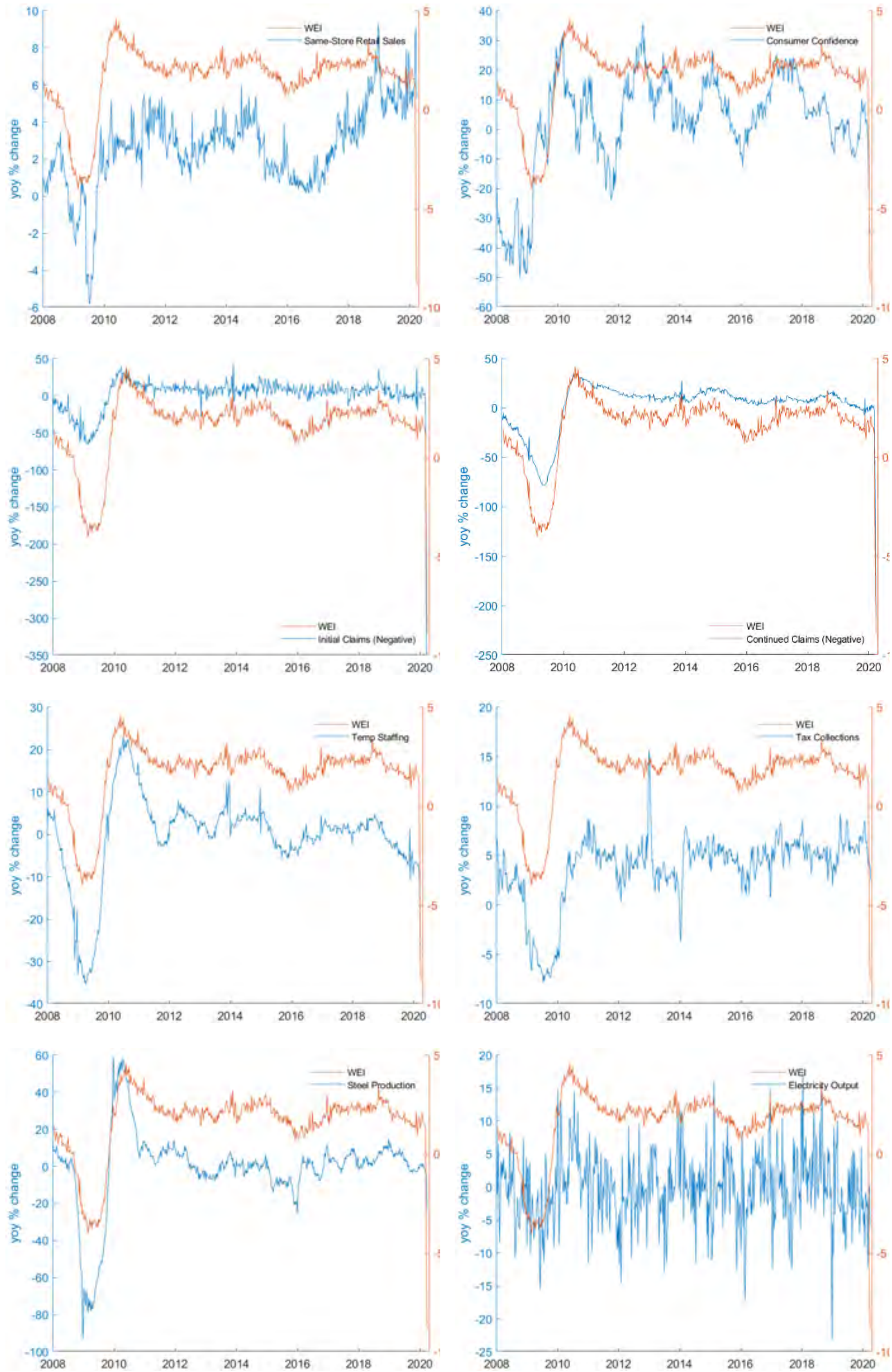
Covid Economics 6, 17 April 2020: 1-21

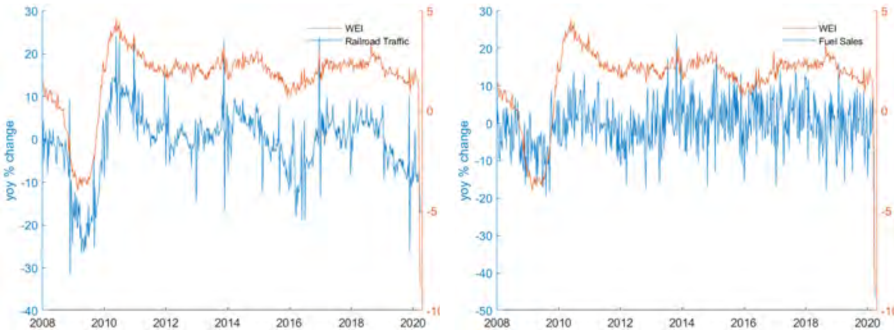
Raw Steel Production	NSA, Thous. Net Tons	**Weekly production tonnage provided from 50% of the domestic producers combined with monthly production data for the remainder. https://www.steel.org/industry-data**
US Fuel Sales to End Users	NSA, EOP, Thous. barrels/day	**Weekly product supplied minus change in stock of finished gasoline and distillate fuels. This estimates wholesale gasoline + diesel sales to retailers and large end users (e.g., UPS). Published by the U.S. Energy Information Administration in the Weekly Petroleum Status Report. https://www.eia.gov/petroleum/supply/weekly/**
U.S Railroad Traffic	NSA, units	**Total carloads and intermodal units reported by railroad companies to the Association of American Railroads https://www.aar.org/data-center/**
Electric Utility Output	**NSA, Gigawatt Hours**	**Total output for U.S. (excluding Alaska and Hawaii) investor-owned electric companies. https://www.eei.org/**

As can be seen in **Figure 2**, some of the weekly series exhibit considerable noise from week to week, such that gleaning broader trends from any one series can be difficult. The series, however, also display a clear cyclical pattern, which suggests that they might usefully be combined into a single index. The WEI is computed from these ten series using the method of principal component analysis. The first principal component of these ten series provides an estimate of a signal about the economy which is common to all variables. By construction, the Weekly Economic Index is a weighted average of the ten series. The mathematics of principal components analysis is summarized next.

Covid Economics 6, 17 April 2020: 1-21

Figure 2 Weekly Variables and WEI





Notes: Based on data available through April 9, 2020. For sources, see **Table 1**.

Covid Economics 6, 17 April 2020: 1-21

II. Construction of the Weekly Economic Index[4]

A leading framework for the construction of an economic index from multiple time series is the so-called dynamic factor model, developed by Geweke (1977) and Sargent and Sims (1977). The dynamic factor model posits the existence of a small number of unobserved or latent series, called factors, which drive the co-movements of the observed economic time series. Application of dynamic factor models to estimating economic indexes range from the construction of state-level indexes of economic activity (Crone and Clayton-Matthews, 2005) to large-scale indexes of economic activity (for example, the Chicago Fed National Activity Index, or CFNAI). See Stock and Watson (2016) for a review.

The premise of a dynamic factor model is that a small number – in our application, a single – latent factor, f_t, drives the co-movements of a vector of N time-series variables, X_t. The dynamic factor model posits that the observed series is the sum of the dynamic effect of the common factor and an idiosyncratic disturbance, e_t, which arise from measurement error and from special features that are specific to an individual series:

$$X_t = \lambda(L)f_t + e_t \qquad (1)$$

[4] Parts of this section are adapted from the appendix in Stock (2013).

where L is the lag operator. The elements of the $N\times 1$ vector of lag polynomials λ (L) are the dynamic factor loadings, and $\lambda_i(L)f_t$ is called the common component of the i^{th} series. The dynamic factor can be rewritten in static form by stacking f_t and its lags into single vector F_t, which has dimension up to the number of lags in λ(L):

$$X_t = \Lambda F_t + e_t \tag{2}$$

where Λ is a matrix with rows being the coefficients in the lag polynomial λ(L).

The two primary methods for estimating the unobserved factor f_t are by principal components and using state space methods, where the factor is estimated by the Kalman filter. Broadly speaking, early low-dimensional applications used parametric state-space methods and more recent high-dimensional applications tend to use nonparametric principal components or variants. We used both methods in developing the WEI, but found the results using the parametric DFM to be sensitive to specification details (lags, sample length, etc.), so principal components estimation is used in this paper.

Covid Economics 6, 17 April 2020: 1-21

Covid Economics 6, 17 April 2020: 1-21

Table 2: PCA Results

Series	Weights Baseline	Weights Trimmed (ALS)
Same-Store Retail Sales	0.29	0.29
Consumer Confidence	0.23	0.21
Tax Collections	0.30	0.31
Initial Claims	-0.38	-0.38
Continued Claims	-0.41	-0.41
Temp Staffing	0.40	0.40
Steel Production	0.37	0.36
Fuel Sales	0.17	0.18
Electricity Output	0.12	0.13
Railroad Traffic	0.34	0.34
Total variance explained	54.4%	52.7%

Notes: Estimation sample is first week of 2008 through last week of February 2020. The first column uses all observations. The second column is based on a trimmed sample in which outliers were removed. In this case, the weights are estimated using alternating least squares, see for instance Stock and Watson (2002b).

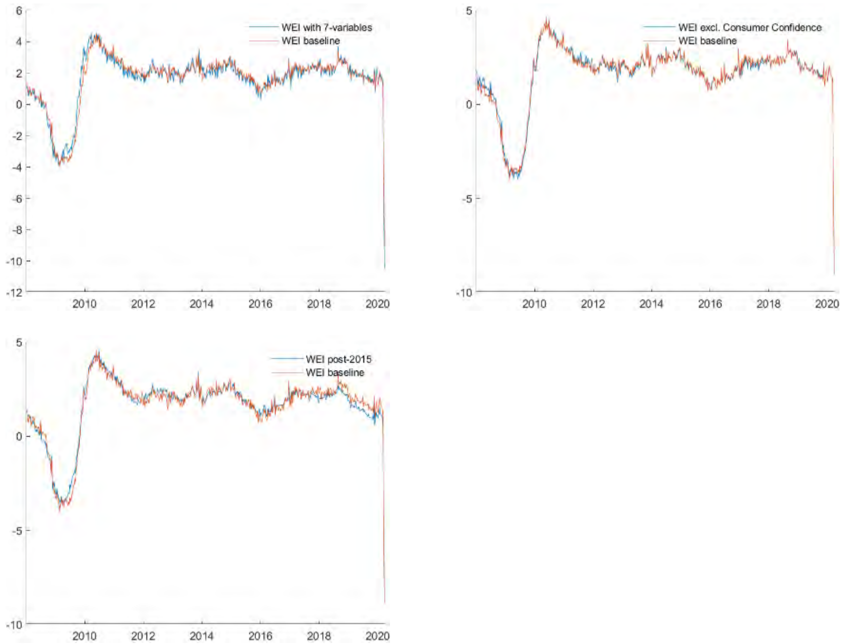
An alternative approach to using high-frequency data for real-time monitoring ("nowcasting") is to focus on forecasting a specific economic release, such as the monthly change in employment, and to construct a model that updates those forecasts as new data comes in. The dynamic factor model and its state space implementation is useful for this purpose because a single model automatically adapts to new data becoming available to estimate the variable of interest. For applications of dynamic factor models to nowcasting, see Giannone, Reichlin and Small (2008) and Aruoba, Diebold and Scotti (2009).

Table 2 provides the weights associated with the first principal component, as well as the total variance explained based on the 10 weekly series described above. The first column provides the weights using the full sample between the first week of January 2008 and the last week of February, 2020. The second column shows the weights over the same sample period, but after treating outliers in the weekly series as missing observations. Removing outliers overall has little effect on the weights, and for WEI we therefore use the full-data weights. We find that WEI explains 54% of the overall variance of the underlying series.

Robustness The WEI is robust to changes in the details of its construction. Subtracting or adding individual series has little effect on the overall path; the same is true for estimating the weights on each series using only more recent data. The left panel of **Figure 3** compares our baseline index to one with a subset of 7 variables (omitting railroad traffic, tax withholdings and continuing claims). The middle panel plots a version in which we omit consumer sentiment. Both figures illustrate that the common signal is not driven by the precise choice of series. The right panel of **Figure 3** plots the baseline WEI against a series computed with weights estimated using only data from 2015 onward, showing that the relationship between these series has been fairly constant during and after the Great Recession.

Covid Economics 6, 17 April 2020: 1-21

Figure 3: Robustness Checks



Notes: Based on data available through April 9, 2020.

Covid Economics 6, 17 April 2020: 1-21

Real Time Updating The WEI is published every Tuesday and Thursday following the releases of the underlying weekly data. Each WEI release reports a preliminary estimate for the prior week based on the data available at that point. The Tuesday update reflects incoming data on retail sales, consumer confidence and steel production for the prior week, as well as data on temp staffing for the week before the prior week. The Thursday update reflects incoming data on fuel sales, tax collections, initial claims, electricity use, and railroad traffic for the prior week, as well as data on continued claims for the week before the prior week. Final values for the WEI are available with a two-week lag.

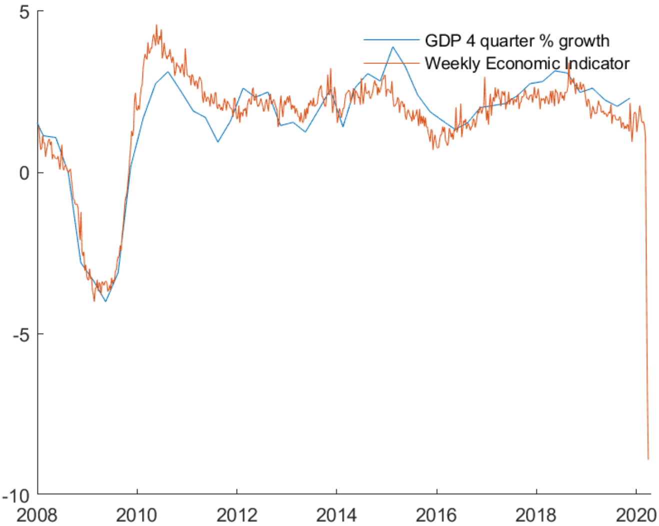
Our procedure to produce the preliminary values is based on forecasting regressions for the WEI on two lags of WEI as well as the non-missing data. For example, in the Thursday update, where we miss the latest data points for two series, we use the estimated value from the forecasting regression of the WEI on its lags and the current values of the eight available series. The forecasting regressions use the same estimation sample (Jan 2008 to Feb 2020) as the one used to estimate weights in **Table 2**. The Tuesday update misses observations of two consecutive weeks of continuing claims. In that case, we first produce a preliminary WEI for two weeks prior based on this procedure, and subsequently repeat the forecasting step for the latest week using the preliminary WEI as input for the forecast.

III. Relationship Between the Weekly Index and Lower Frequency Measures

Figure 4 plots the WEI together with the four-quarter growth rate of real GDP. The latter series is used to scale the baseline index. A reading of 2 percent in a given week indicates that if the week's conditions persisted for an entire quarter, real GDP would be 2 percent greater than the same quarter a year ago. The panels in the first row of **Figure 5** plots the index against the monthly change in nonfarm payrolls and the twelve-month percentage change in industrial production (IP). The figure shows that the index tracks both payroll changes and IP growth closely. The bottom panels of **Figure 5** plot the index against two additional monthly activity indicators: capacity utilization and the ISM manufacturing index. The index also tracks these series relatively well. The close relationship with the lower frequency measures indicates that, despite the noise inherent in the raw high-frequency data, our methodology to combine these data into a weekly index produces an informative and timely signal of real economic activity.

Covid Economics 6, 17 April 2020: 1-21

Figure 4: WEI and GDP growth



Notes: Based on data available through April 9, 2020.

Figures 4 and 5 help to illustrate two important differences between our index and a nowcast, like those for GDP growth produced by the Federal Banks of New York, Atlanta or St. Louis. First, a nowcast focuses on a single important target series, and uses the information contained in intermediate data to predict that series. In contrast, while we report the WEI in GDP growth units, this is simply an *ex post* normalization; the WEI does not focus on a single outcome by targeting either a consumption variable or a production variable—both are important to get a sense of real activity. Second, most nowcasts (including those of the New York, Atlanta and St Louis Feds) focus on lower-frequency targets like GDP growth, which are very informative about the economy. But, since GDP is a quarterly variable, such models are not equipped to highlight variation from one week to the next (see also McCracken, 2020). The goal of these nowcasts is only to predict average variation in the target series over thirteen weeks, which they generally do well.

Covid Economics 6, 17 April 2020: 1-21

Figure 5: Relationship Between WEI and Other Monthly Activity Measures

Notes: Based on data available through April 9, 2020.

This said, it is useful to examine the predictive content of the WEI for relevant lower frequency activity indicators. A complication in comparing weekly and monthly data is the non-alignment of the calendar. To address this non-alignment, we introduce the concept of "pseudo-weeks", which divide the month into four weeks, the first starting on the first day of the month, the first three having seven days (and thus 5 weekdays and 2 weekend days), and the final pseudo-week running from 22nd through final day of the month (so including between 7 and 10 days). Each day of the month naturally falls into a calendar week of the original WEI, so we compute the pseudo-week WEI as an average of the WEI of the constituent days. With these pseudo-weeks, we have an approximate measure of the signal provided by the index after the first, second, third, and fourth weeks of the month. We also calculate a monthly WEI by computing the average WEI for all constituent days.

Covid Economics 6, 17 April 2020: 1-21

Payroll Employment We first explore predictive power for changes in payroll employment. Specifically, we begin by computing a monthly regression,

$$\Delta Y_t = c + \beta WEI_t^{monthly} + \sum_{s=1}^{2} \gamma_s \Delta Y_{t-s} + u_t, \qquad (4)$$

where Y_t is monthly private payroll employment. We compute heteroskedasticity and autocorrelation robust standard errors using the EWC estimator recommended by Lazarus et al (2018). Column (I) of **Table 3** reports the results; the WEI is a highly significant predictor of employment changes, with an R^2 of 0.83 (regressing on WEI alone gives 0.66).

We run an additional regression, reported in column (II), adding the change in employment from the ADP release as a control. The ADP release, which precedes the official payroll numbers by two days, is known to be highly informative for the eventual BLS release. We find that despite this strong relationship, the WEI provides additional information, above and beyond the ADP release.

Next, we turn to intra-month regressions. Week by week, we run "nowcasting" regressions based on the information flow from the WEI. These take the form

$$\Delta Y_t = c + \sum_{i=1}^{\overline{w}} \beta_i WEI_t^{w_i} + \sum_{s=1}^{2} \gamma_s \Delta Y_{t-s} + u_t, \ \overline{w} = 1,2,3,4\,; \qquad (5)$$

where $WEI_t^{w_i}$ is the average WEI for the i^{th} pseudo-week of month t. For employment, since the payroll survey is conducted during the second week of the month, we consider the last two pseudo-weeks of the prior month and the first two pseudo-weeks of the current month. The results are reported in columns (III) to (VI). In regression (VI), we find that the second pseudo-week of the month, that on which the payroll survey is focused, is a significant predictor of employment changes. Moreover, from the last week of the prior month onwards, the weekly information provided by the WEI is jointly significant (from the F-test that all weekly coefficients are zero).

Industrial Production The WEI also helps to nowcast industrial production (IP). While **Figure 5** shows a clear relationship between 12-month percentage changes in IP and the WEI, we now consider the more conventional monthly percentage change. We regress this first on the monthly WEI and lagged IP

Covid Economics 6, 17 April 2020: 1-21

growth according to (4), where Y_t is monthly log IP. Column (I) of **Table 4** shows that the monthly average WEI (and lags) explains 17% of variation in IP growth, about two weeks before the official release (still 16% dropping lags of IP growth). We then proceed with the weekly nowcasting regressions, following (5). We find that, from the second week of the month onwards, the flow of information from the WEI is a significant predictor of monthly IP growth; the explained variation rises from 15% to 28%. The most recent week is a significant positive predictor of IP growth, while the first week is a negative predictor, since it is closely related to production in the prior month.

Table 3: Employment Regression Results

Regressors	(I)	(II)	(III)	(IV)	(V)	(VI)
$WEI_t^{monthly}$	21.90***	10.54*				
	(6.78)	(6.05)				
ΔADP_t		0.84***				
		(0.09)				
WEI week 2, current month						111.51*
						(60.87)
WEI week 1, current month					51.13	-46.11
					(39.06)	(62.10)
WEI week 4, past month				54.02	-2.32	8.64
				(42.37)	(67.83)	(68.51)
WEI week 3, past month			10.77	-42.09	-33.66	-56.13
			(10.30)	(49.64)	(50.62)	(54.78)
F-test: weekly coefficients = 0			1.09	6.41	6.55	3.61
			(0.32)	(0.01)	(0.01)	(0.05)
F-test: weekly coefficients equal				0.50	0.79	0.76
				(0.62)	(0.53)	(0.58)
SER	96.42	73.20	97.84	97.01	96.40	94.39
Adjusted R^2	0.83	0.90	0.83	0.83	0.83	0.84

*Notes: HAR standard errors computed using the EWC estimator of Lazarus et al (2018). Results starred at the 1%, 5%, and 10% levels, ***,**,*.*

GDP growth Finally, the WEI also aids in nowcasting GDP growth. To show this, we first regress GDP growth on the quarterly WEI, following (4), where ΔY_t is 4-quarter GDP growth (percent) and we replace the monthly average

Covid Economics 6, 17 April 2020: 1-21

WEI with the quarterly average WEI. The results in Column (I) of **Table 5** show that the quarterly WEI is a significant predictor of GDP growth, with 89% of variation explained (85% without lagged GDP growth), nearly a month before the advance release. We then regress the 4-quarter growth rate on the flow of information from the WEI, starting with the WEI for just the first month of the quarter, and so on, following

$$\Delta GDP_t = c + \sum_{i=1}^{\overline{m}} \beta_i WEI_t^{m_i} + \sum_{s=1}^{2} \gamma_s \, \Delta GDP_{t-s} + u_t, \ \overline{m} = 1,2,3 \ ; \qquad (6)$$

where ΔGDP_t is 4-quarter GDP growth (percent). Columns (II) to (IV) report the results. For the first two months of the quarter, the most recent month's WEI is a significant (positive) predictor of growth, with the adjusted R^2 rising from 0.86 to 0.92. Data on the final month does not appear to add much additional information, although the coefficients on monthly WEI are jointly significant for all specifications. We conclude that a strong signal of GDP growth is available from the WEI from the second month of the quarter, nearly two months before the advance release.

IV. Conclusion

In normal times, familiar macroeconomic aggregates provide accurate descriptions of economic conditions with a modest delay. When conditions evolve rapidly from day to day and week to week, as is the case in the current environment, less familiar sources of data can provide an informative and timely signal of the state of the economy. The WEI provides a parsimonious summary of that signal.

Covid Economics 6, 17 April 2020: 1-21

Table 4: Industrial Production Regression Results

Regressors	(I)	(II)	(III)	(IV)	(V)
$WEI_t^{monthly}$	0.17 (0.11)				
WEI week 4, current month					0.99** (0.36)
WEI week 3, current month				0.57* (0.31)	-0.11 (0.26)
WEI week 2, current month			0.76** (0.31)	0.18 (0.33)	-0.05 (0.27)
WEI week 1, current month		0.14 (0.10)	-0.60** (0.28)	-0.57** (0.26)	-0.63** (0.25)
F-test: weekly coefficients = 0		1.94 (0.19)	2.99 (0.09)	1.92 (0.19)	2.78 (0.09)
F-test: weekly coefficients equal		0.00	2.50 (0.13)	1.84 (0.20)	1.80 (0.21)
SER	0.67	0.68	0.66	0.65	0.62
Adjusted R^2	0.17	0.15	0.18	0.20	0.28

*Notes: HAR standard errors computed using the EWC estimator of Lazarus et al (2018). Results starred at the 1%, 5%, and 10% levels, ***,**,*.*

Table 5: GDP Regression Results

Regressors	(I)	(II)	(III)	(IV)
$WEI_t^{quarterly}$	0.70** (0.22)			
WEI month 3				0.25 (0.57)
WEI month 2			1.66*** (0.45)	1.25 (0.86)
WEI month 1		0.64* (0.28)	-1.12** (0.42)	-0.97** (0.37)
F-test: weekly coefficients = 0		5.42 (0.06)	12.81 (0.01)	8.14 (0.04)
F-test: weekly coefficients equal			4.28 (0.08)	2.87 (0.17)
SER	0.55	0.63	0.48	0.48
Adjusted R^2	0.89	0.86	0.92	0.91

*Notes: HAR standard errors computed using the EWC estimator of Lazarus et al (2018). Results starred at the 1%, 5%, and 10% levels, ***,**,*.*

Covid Economics 6, 17 April 2020: 1-21

References

Aruoba, S.B., F.X. Diebold, and C. Scotti, (2009), "Real-Time Measurement of Business Conditions," *Journal of Business & Economic Statistics* 27, 417-427.

Bai, J., (2003), "Inferential Theory for Factor Models of Large Dimensions," *Econometrica*, 71, 135-172.

Bai, J., and S. Ng, (2006), "Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions," *Econometrica*, 74,1133-1150.

Chamberlain, G., and M. Rothschild, (1983), "Arbitrage Factor Structure, and Mean-Variance Analysis of Large Asset Markets," *Econometrica*, 51,1281-1304.

Crone, T.S. and A Clayton-Matthews, (2005), "Consistent Economic Estimates for the 50 States," *The Review of Economics and Statistics*, 87, 593-603.

Engle, R.F., and M.W. Watson, (1981), "A One-Factor Multivariate Time Series Model of Metropolitan Wage Rates," *Journal of the American Statistical Association*, 76, 774-781.

Geweke, J., (1977), "The Dynamic Factor Analysis of Economic Time Series," in *Latent Variables in Socio-Economic Models*, ed. by D.J. Aigner and A.S. Goldberger, Amsterdam: North-Holland.

Giannone, D., L. Reichlin, and D. Small, (2008), "Nowcasting: The Real-Time Informational Content of Macroeconomic Data," *Journal of Monetary Economics*, 55, 665-676.

Lazarus, E., D.J. Lewis, J.H. Stock, and M.W. Watson, (2018), "HAR Inference: Recommendations for Practice," *Journal of Business and Economic Statistics,* 36, 541-559.

Lewis, D.J., K. Mertens, and J.H. Stock, (2020), "Monitoring Economic Activity in Real Time," *Liberty Street Economics,* March 28, 2020.

https://libertystreeteconomics.newyorkfed.org/2020/03/monitoring-real-activity-in-real-time- the-weekly-economic-index.html

McCracken, M., (2020), COVID-19: Forecasting with Slow and Fast Data, On the Economy Blog, Federal Reserve Bank of St Louis, https://www.stlouisfed.org/on-the-economy/2020/april/covid-19-forecasting-slow-fast-data

Sargent, T.J., (1989), "Two Models of Measurements and the Investment Accelerator," *Journal of Political Economy* 97:251–287.

Sargent, T.J., Sims, C.A., (1977). Business cycle modeling without pretending to have too much a-priori economic theory. In: Sims, C. et al., (Ed.), *New Methods in Business Cycle Research*. Federal Reserve Bank of Minneapolis.

Stock, J.H., (2013) "Economic Activity During the Government Shutdown and Debt Limit Brinkmanship," *Council of Economic Advisors*, Report October 2013.

Stock, J.H., and M.W. Watson, (1989), "New Indexes of Coincident and Leading Economic Indicators," *NBER Macroeconomics Annual 1989*, 351-393.

Stock, J.H., and M.W. Watson, (1999), "Forecasting Inflation," *Journal of Monetary Economics*, 44, 293-335.

Stock, J.H., and M.W. Watson, (2002a), "Forecasting Using Principal Components from a Large Number of Predictors," *Journal of the American Statistical Association*, 97, 1167-1179.

Stock, J.H., and M.W. Watson, (2002b), "Macroeconomic Forecasting Using Diffusion Indices," *Journal of Business and Economic Statistics*, 20, 147-162.

Stock, J.H., and M.W. Watson, (2016), "Factor Models and Structural Vector Autoregressions in Macroeconomics," *Handbook of Macroeconomics: Volume 2* (2016), 415-525.

Covid Economics 6, 17 April 2020: 1-21

Covid Economics Issue 6, 17 April 2020

CEPR PRESS

Health versus wealth: On the distributional effects of controlling a pandemic[1]

Andrew Glover,[2] Jonathan Heathcote,[3] Dirk Krueger[4] and José-Víctor Ríos-Rull[5]

Covid Economics 6, 17 April 2020: 22-64

Date submitted: 10 April 2020; **Date accepted:** 13 April 2020

Many countries are shutting non-essential sectors of the economy to slow the spread of Covid-19. Older individuals have most to gain from slowing virus diffusion. Younger workers in sectors that are shuttered have the most to lose. In this paper we extend a standard epidemiological model of disease progression to include heterogeneity by age, and multiple sources of disease transmission. We then incorporate the epidemiological block into a multi-sector economic model in which workers differ by sector (basic and luxury) as well as by health status. We study optimal mitigation policies of a utilitarian government that can redistribute resources across individuals, but where such redistribution is costly. We show that optimal redistribution and mitigation policies interact, and reflect a compromise between the strongly diverging preferred policy paths across the subgroups of the population.

1 We thank Josep Pijoan-Mas, Marc Suchard and Andrew Atkeson for helpful comments. We thank attendees at virtual talks at the Minneapolis Fed, the VMACS seminar series, and the Bank of England. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Banks of Kansas City and Minneapolis or the Federal Reserve System.

2 Senior Economist, Federal Reserve Bank of Kansas City.
3 Monetary Advisor, Federal Reserve Bank of Minneapolis and CEPR Research Fellow.
4 Walter H. and Leonore C. Annenberg Professor in the Social Sciences and Professor of Economics, University of Pennsylvania and CEPR Research Fellow.
5 Lawrence R. Klein Professor of Economics, University of Pennsylvania and CEPR Research Fellow.

Covid Economics 6, 17 April 2020: 22-64

1 Introduction

The central debate about the appropriate economic policy response to the global COVID-19 pandemic is about how aggressively to restrict economic activity in order to slow down the spread of the virus, and how quickly to lift these restrictions as the pandemic shows signs of subsiding. In this paper, we argue that one reason people disagree about the appropriate policy is that "lock-down" policies have very large distributional implications. These distributional effects mean that different groups prefer very different policies. Standard epidemiological models assume a representative agent structure, in which households face a common trade-off between restrictions on social interaction that slow the virus transmission but which also depress economic activity. In practice, however, the benefits of slower viral transmission are not shared uniformly, but accrue disproportionately to older households, who face a much higher risk of serious illness or death conditional on becoming infected. At the same time, the costs of reduced economic activity are disproportionately born by younger households, who bear the brunt of lower employment. A second very important dimension of heterogeneity is between younger workers employed in different sectors of the economy. Sensible lock-down policies designed to reduce viral spread will naturally focus on reducing activity in sectors in which there is a social aspect to consumption and sectors that produce goods or services perceived to be luxuries. For example, restaurants and bars are likely to be the first to be closed. Because workers cannot easily reallocate across sectors, this implies that lock-down policies will imply extensive redistribution between young households specialized in different sectors. Thus, different groups in the economy (old versus young, workers in different sectors, healthy versus sick) will likely have very different views about the optimal mitigation strategy. Furthermore, lock-down policies create a need for potentially large redistributive public policies. To the extent that these are costly to implement[1], the optimal mitigation policy will in turn depend on the scope for redistributive policies at the micro level.

In this paper we seek to build and then quantitatively implement a framework to model this interaction between macro mitigation- and micro redistribution policies. This requires a structure with i) a household sector with heterogeneous individuals, ii) an epidemiological block that determines their health transitions through a potential epidemic, and iii) a government with tools for mitigation and redistribution, as well as a desire for social insurance.

[1]For example, the revenue for transfer programs needs to raised through distortionary taxation.

Covid Economics 6, 17 April 2020: 22-64

On the household side, we distinguish between three types of people: young workers in a basic sector, young workers in a luxury sector, and old retired people. Output of workers in the two sectors is combined to produce a single final consumption good. Workers are immobile across sectors. The output of the basic sector is assumed to be so essential that it will not make sense to reduce employment and output in that sector in order to reduce the spread of the disease. In contrast, the policy maker has a potential incentive to shut down part of the economic activity in the luxury sector in order to reduce the rate at which infection spreads.

The epidemiological model builds on a standard SIR diffusion framework. We label our variant a *SAFER* model, reflecting the progression of individuals through a sequence of possible health states. Model individuals start out as susceptible S (i.e. healthy, but vulnerable to infection), can then become infected but asymptomatic A, infected with flu-like symptoms F, infected and needing emergency hospital care E, recovered R (healthy and immune), or dead. The three different types of model agents face differential infection risk (workers face more exposure than non-workers) and differential health outcomes conditional on infection (the old are more likely to end up in emergency care). The reason we partition the infected into three groups is because these groups spread the virus in different ways: the asymptomatic are unlikely to realize they are contagious and will continue to work and to consume; those with flu symptoms will stay at home and only infect family members, while those in hospital care may pass the virus to health care workers.

The government has a utilitarian social welfare function and two policy levers at its disposal to maximize social welfare. First, at each date the planner can choose what fraction of activity in the luxury sector to shut down. We call this policy the extent of mitigation. Mitigation slows the spread of the virus (by reducing the rate at which susceptible workers become asymptomatically infected), but it reduces the market income of some workers in the luxury sector to zero. Second, the planner chooses how much income to redistribute from those working toward those that are not, either because they are old, because they are unwell, or because their workplaces have been closed due to mitigation. Redistribution is desirable due to the utilitarian social welfare function, but crucially, we also assume that this redistribution is costly, so that perfect insurance is not optimal. Conditional on a given path for mitigation, the optimal redistribution problem is equivalent to a static social planner problem, with lower aggregate consumption and more consumption inequality across workers the more costly is redistribution. This in turn feeds back adversely on the dynamic incentives for mitigation, implying that a government facing more costly redistribution needs will choose less mitigation dynamically.

In the context of the model with these trade-offs we then compute optimal paths for mitigation, where (currently) the path for mitigation is restricted to be a simple parametric function of time. We find that the optimal path for mitigation is highly sensitive to the relative welfare weights the government attaches to the three types of people in the model. A planner that prioritizes the old chooses extensive and prolonged mitigation, as the old are highly vulnerable to contracting the disease and dying from them. A planner that prioritizes workers in the luxury sector subject to a potential lock down chooses much weaker mitigation as the economic costs of foregone income and thus consumption dominate for this group.

We also consider how the optimal policy for a utilitarian equal-weights planner varies with the cost of redistribution across worker types. We find that the larger is this cost, the more moderate is optimal mitigation. Thus our economy which features key dimensions of inequality implies a more modest shutdown than a representative agent analogue, at the cost of higher mortality during the epidemic.

There is an extraordinary set of papers currently being written about the pandemic. To cite the ones that we are aware of: Atkeson (2020) was perhaps the first to introduce economists to the epidemiology SIR (Susceptible-Infectious-Recovered) class of models. He emphasizes the negative outcomes that arise if and when the fraction of active infections in the population exceeds 1 percent (at which point the health system is predicted to be severely challenged) and 10% (which may result in severe staffing shortages in key financial and economic infrastructure sectors) as well as the cumulative burden of the disease over an 18 month horizon. Greenstone and Nigam (2020) use the state of the art Imperial College epidemiological model (Flaxman et al. (2020)) to compare the paths under moderate social distancing versus no policy action, and use the statistical value of life approach to assess the social cost of no action. They calculate 1.7 million lives saved between March 1 and October 1 from social distancing, 37% of them due to less overcrowding in hospitals.

Eichenbaum et al. (2020) extend the canonical SIR epidemiology model to study the interaction between economic decisions and pandemics. They emphasize how equilibria without interventions lead to sub-optimally severe pandemics, because infected people do not fully internalize the effects of their economic decisions on the spread of the virus. Krueger et al. (2020) argue that the severity of the economic crisis in Eichenbaum et al. (2020) is much smaller if individuals can endogenously adjust the sectors in which they consume. Toxvaerd (2020) characterizes the simultaneous determination of infection and social distancing. Moll

Covid Economics 6, 17 April 2020: 22-64

et al. (2020) develop a version of a HANK model in which agents differ by occupation, and occupations have two key characteristics: how social their consumption is, and how easily work in the occupation can be done at home. They tie demand for social goods and willingness to work in the workplace to fear of contracting the virus, with endogenous feedback to relative earnings by occupation. Bayer and Kuhn (2020) explore how differences in living arrangements of generations within families contribute to the cross country differences in terms of case-fatality rates. They document a strong positive correlation between this variable and the share of working-age families living with their parents. Berger et al. (2020) extend the baseline Susceptible-Exposed-Infectious-Recovered (SEIR) infectious disease model to explore the role of testing and to thereby get a better idea of how to implement selective social separation policies. Fang et al. (2020) quantify the causal impact of human mobility restrictions using the Chinese experience and find that the lock-down was very effective, providing estimates of diffusion under different scenarios. Hall et al. (2020) provide a simple calculation to assess how much people would be willing to pay to have never had the virus (their answer is about a quarter of one year's worth of consumption).

We start by describing how we model the joint evolution of the economy and the population in Section 2. We then turn to describe how we model mitigation and redistribution policies, and how we go about solving for optimal policies in Section 3. The calibration strategy is described in Section 4. The findings are in Section 5.

2 The Model

We first describe the individual state space, describing the nature of heterogeneity by age and health status. We then describe the multi-sector production technology in Section 2.2, describing how *mitigation* shapes the pattern of production. Section 2.3 describes the details of our *SAFER* extension of the standard *SIR* epidemiological model, and the channels of disease transmission.

2.1 Household Heterogeneity

Agents can be young or old, which we denote y and o respectively. We think of the young as aged below 65 and they will comprise $\mu^y = 85$ percent of the population. For simplicity, and given the short time horizon of interest, we abstract from population growth and ignore aging.

Covid Economics 6, 17 April 2020: 22-64

Within each age group, agents are differentiated by health status that can take six different values: susceptible s, asymptomatic a, miserable with flu-symptoms f, requiring emergency care e, recovered r, or dead d. Individuals in the first group have no immunity and are susceptible to infection. The a, f, and e groups all carry the virus – they are subsets of the infected group in the standard *SIR* model – and can pass it onto others. However, they differ in terms of their symptoms. The asymptomatic have no symptoms or very mild ones, and thus spread the virus unknowingly. We model this state explicitly (in contrast to the prototypical *SIR* model) because a significant percentage of individuals infected with COVID-19 experience no symptoms.[2] Those with flu-like symptoms are sufficiently sick to know they are likely contagious and they stay at home and avoid the workplace and market consumption. Those requiring emergency care are hospitalized. The recovered are again healthy, no longer contagious, and are immune from future infection. A worst case virus progression is from susceptible to asymptomatic to flu to emergency care to dead.[3] However recovery is possible from the asymptomatic, flu and emergency-care states.

2.2 Activity: Technology and Mitigation

Young agents in the model are further differentiated by the sector in which they can work. A fraction μ^b of the young work in the basic sector, denoted b, while the rest, $1-\mu^b$, work in a luxury sector, denoted ℓ. We assume that output of the basic sector is so vital that it is never optimal to send home even a subset of b sector workers from working. In contrast, it may be optimal to require some or all of the workers in the ℓ sector to stay at home in order to reduce the transmission of the virus in the workplace. We will call such a policy a (macroeconomic) mitigation policy, m. More precisely, m_t will denote the fraction of luxury workers that are instructed to not go to work at time t. We assume that workers cannot change sectors (at least not during the short time horizon studied in this paper); thus the sector of work is a fixed characteristic of a young individual.

Time starts at $t = t_0$ and evolves continuously. All economic variables, represented by roman

[2]DeCODE, a subsidiary of Amgen, randomly tested 9,000 individuals in Iceland. Of the tests that came back positive (1 percent), half reported experiencing no symptoms.

[3]Note that in the standard SEIR model, agents in the exposed state E have been exposed to the virus and may fall ill, but until they enter the infected state I they cannot pass the virus on. Our asymptomatic state is a hybrid between the E and the I states in the SEIR model: asymptomatic agents have no symptoms (as in the SEIR E state) but can pass the virus on (as in the SEIR I state). Berger et al. (2020) make a similar modeling choice.

Covid Economics 6, 17 April 2020: 22-64

letters, are understood to be functions of time, but we suppress that dependence whenever there is no scope for confusion. Technology parameters are denoted with Greek letters. Generically, we use the letter x to denote population measures, with superscripts specifying subsets of the population. These super-indices index age, sector, and health status, in that order. For example, x^{ybs} is the measure of young individuals working in the basic sector who are susceptible.

We assume a production technology that is linear in labor and thus output in the basic sector is given by the number of young workers employed there;

$$y^b = x^{ybs} + x^{yba} + x^{ybr}. \tag{1}$$

Note that this specification assumes that those individuals carrying the virus but being asymptomatic continue to work.[4] In contrast, those with flu are assumed to stay at home. Output in the luxury sector, in contrast, does depend on the mitigation policy and is given by

$$y^\ell = (1 - m_t) \left(x^{y\ell s} + x^{y\ell a} + x^{y\ell r} \right). \tag{2}$$

We assume that both sectors produce the same good and are perfect substitutes.[5] Under this assumption, total output of the single consumption good is determined by:

$$y = y^b + y^\ell. \tag{3}$$

We assume that a fixed amount of output $\eta\Theta$ is spent on emergency hospital care, where Θ is the capacity of emergency room beds and η is the cost of providing and maintaining one bed.

In practice, different sectors of the economy are heterogeneous with respect to the extent to which production and consumption generate risky social interaction. For example, some types of work and market consumption can easily be done at home, while for others, avoiding interaction is much harder. A sensible shutdown policy will first shutter those sub-sectors of the luxury sector that generate the most interaction. Absent detailed micro data on social interaction by sector, we model this in the following simple way.[6] Assume workers are assigned

[4] One could instead imagine a policy of tracing contacts of infected people, which would allow the planner to keep some portion of exposed workers at home.

[5] We make this assumption primarily for the sake of tractability. If outputs of the two sectors were complementary, there would be changes in relative prices and wages when output of the luxury sector was suppressed.

[6] See Xu et al. (2020) for more detailed evidence on infection patterns in the workplace.

Covid Economics 6, 17 April 2020: 22-64

to a unit interval of sub-sectors $i \in [0,1]$ where sub-sectors are ranked from those generating the least to the most social interaction.

Assume the sector-specific infection-generating rates are $\beta_w^i = 2\alpha_w i$ and $\beta_c^i = 2\alpha_c i$, where (α_w, α_c) are parameters, to be calibrated below, governing the intensity by which meetings among individuals generate infections. When the government asks fraction m_t of luxury workers to stay at home, assume it targets the sub-sectors generating the most interactions, i.e., $i \in [1-m_t, 1]$. The average interaction rates of the sectors that remain are then $\alpha_w(1-m_t)$ and $\alpha_c(1-m_t)$, respectively[7] Because the government cannot shut down any basic sub-sectors of the economy, the economy-wide work-related infection-generating probability is then given by

$$\beta_w(m_t) = \frac{y^b}{y(m_t)}\alpha_w + \frac{y^l(m_t)}{y(m_t)}\alpha_w(1-m_t)$$

with an analogous expression for $\beta_c(m_t)$. The key property of this expression is that as mitigation is increased, the average social interaction-generating rate will fall.

2.3 Health Transitions: The SAFER Model

We now describe the dynamics of individuals across health states. At t_0, the total mass of individuals is one, $x^{yb} + x^{y\ell} + x^o = 1$, where $x^{yb} = \sum_{i\in\{s,a,f,e,r\}} x^{ybi}$, $x^{y\ell} = \sum_{i\in\{s,a,f,e,r\}} x^{y\ell i}$ and $x^o = \sum_{i\in\{s,a,f,e,r\}} x^{oi}$. In the interests of more compact notation, we will also let $x^i = x^{ybi} + x^{y\ell i} + x^{oi}$ for $i \in \{s,a,f,e,r\}$ denote the total number of individuals in health state i. Finally, at any point in time let $x = \sum_{i\in\{s,a,f,e,r\}} x^i = x^{yb} + x^{y\ell} + x^o$ denote the entire living population.

The crucial health transitions, and the ones that can, in our model, be affected by mitigation policies are the movements from the susceptible to the asymptomatic state. These are characterized by equations (4)-(9) below. Equations (4-6) captures the flow of basic sector workers, luxury sector workers, and older individuals out of the susceptible state and into the asymptomatic state. The number of such workers who catch the virus is their original mass (x^{ybs} for young basic sector workers, for example) times the number of virus-transmitting interactions they have (the term in square brackets). We model four sources of possible virus contagion: people can catch the virus from colleagues at work, from market consumption activities, from

[7] $E[\alpha_w i | i \leq (1-m_t)] = \frac{2\alpha_w}{1-m_t}\int_0^{1-m_t} i di = \frac{2\alpha_w}{1-m_t}\frac{(1-m_t)^2}{2} = \alpha_w(1-m_t)$.

family or friends outside work, and from taking care of the sick in hospitals. The four terms in the bracket capture these four sources of infection, which we index w, c, h and e respectively. The flow of new infections for a given type of individual from each of these activities is the product of the number of contagious people they can expect to meet, which we denote $\mu_j(m_t)$ for $j \in \{w, c, h, e\}$, and the likelihood that such meetings result in infection, which is the infection-generating rate described above, $\beta_j(m_t)$. For work and consumption activities, both the number of contagious people in a given setting and the rate at which they transmit the virus potentially depend on the level of economic mitigation m_t.

$$\dot{x}^{ybs} = -\left[\beta_w(m_t)\mu_w(m_t) + \beta_c(m_t)\mu_c(m_t) + \beta_h\mu_h + \beta_e\mu_e\right]\ x^{ybs} \tag{4}$$

$$\dot{x}^{y\ell s} = -\left[\beta_w(m_t)\mu_w(m_t)(1-m_t) + \beta_c(m_t)\mu_c(m_t) + \beta_h\mu_h\right]\ x^{y\ell s} \tag{5}$$

$$\dot{x}^{os} = -\left[\beta_c(m_t)\mu_c(m_t) + \beta_h\mu_h\right]\ x^{os} \tag{6}$$

where the relevant population shares μ in the above expressions are given by:

$$\mu_w(m_t) = x^{yba} + (1-m_t)x^{y\ell a} \tag{7}$$

$$\mu_c(m_t) = x^a y(m_t) \tag{8}$$

$$\mu_h = x^a + x^f \tag{9}$$

$$\mu_e = x^e \tag{10}$$

Consider the first outflow rate in equation (4). The flow of young basic-sector workers getting infected at work, $\beta_w(m_t)\mu_w(m_t)$, is the probability of a virus-spreading interaction per contagious worker, $\beta_w(m_t)$, times the number of contagious workers, which is defined in equation (7). Note that we are assuming that people with symptoms always stay at home (a minimal precaution), and that basic and luxury workers mingle together at work.

The flow of young basic-sector workers getting infected from market consumption, $\beta_c(m_t)\mu_c(m_t)$ is constructed similarly. We assume that the number of consumption-related infections is proportional to the number of asymptomatic individuals in the population and to the level of economic activity, which is identical to the number of workers (see equation 8).[8] Note that we

[8]Note that we have assumed that the number of shopping-related infections for a given type is proportional to economy-wide output, rather than to the type-specific level of consumption. One interpretation of this assumption is that each consumer visits each store in the economy, and faces a similar infection risk irrespective

Covid Economics 6, 17 April 2020: 22-64

are assuming that people with symptoms stay at home and do not go shopping.

The rate at which a young basic worker contracts the virus at home, $\beta_h \mu_h$, depends on the number of contagious workers in the household, μ_h defined in equation (9). Note that both asymptomatic and flu-suffering workers reside at home. Finally, we assume that caring for those requiring emergency care is a task that falls entirely on basic workers. The risk of contracting the virus from this activity is proportional to the number of hospitalized people, $\mu_e = x^e$, with infection-generating rate β_e, which reflects the strength of precautions taken in hospitals.

In parallel to Equation (4), Equation (5) describes infections for the susceptible population working in the luxury sector. The risks of infection from market consumption and at home for this group are identical to those for basic sector workers. However, individuals in this sector work reduced hours when $m_t > 0$ and thus have fewer work interactions in which they could get infected. Furthermore, workers in the luxury sector do not take care of sick patients in hospitals, and thus the last term in Equation (4) is absent in Equation (5). Similar to Equation (4) and Equation (5), Equation (6) displays infections among the old. They only get infected from market consumption and from interactions at home.

The remainder of the epidemiological block is fairly mechanical and simply describes the transition of individuals though the health states (asymptomatic, flu-suffering, hospitalized, and recovered) once they have been infected. The parameters of these dynamic laws in Equation (11) to Equation (22) are allowed to vary by age. Equations (11) to (13) describes the change in the measure of asymptomatic individuals. There is entry into that state from the newly-infected flowing in from the susceptible state (as described above). Exit from this state to developing flu-like symptoms occurs at rate σ^{yaf} (σ^{oaf}) for the young (old), and exit to the recovered state occurs at rate σ^{yar} (σ^{yar}) for the young (old). Note that someone who recovers at this stage will never know that she contracted the virus.

For individuals suffering from the flu, Equations (14) to (16) show that there is entry from the asymptomatic state and exit to the hospitalized state at rate σ^{yfe} for the young, and to the recovered state at rate σ^{yfr}, with analogous expressions for the old. Equations (17) to (19) describes the movements of those in emergency care, showing entry from those with flu-like symptoms, and exits to death and recovery. The death rate is $\sigma^{yed} + \varphi$, while the recovery rate is $\sigma^{yer} - \varphi$, where φ, described below, is a term related to hospital overuse. Equations (20)

of how much they spend. The common infection risk is proportional to the equilibrium number of stores, which in turn is proportional to the aggregate employment level.

Covid Economics 6, 17 April 2020: 22-64

to (22) displays the evolution of the measure of the recovered population, which features only entry and is an absorbing state. So is death, with the evolution of the deceased population being determined by $\dot{x}^{ybd} = (\sigma^{yed} + \varphi)x^{ybe}$, $\dot{x}^{y\ell d} = (\sigma^{yed} + \varphi)x^{y\ell e}$ and $\dot{x}^{od} = (\sigma^{oed} + \varphi)x^{oe}$. We record them separately from the recovered (who work) since the play no further role in the model.

Finally, Equation (23) describes the extent of overuse of the hospital system that has capacity Θ, which we treat as fixed in the time horizon analyzed in this paper. The probability of death conditional on being sick depends on the extent of hospital overuse. In particular, the parameter λ_o controls how much the death rate of the hospitalized rises (and the recovery rate falls) once hospital capacity Θ is exceeded.

$$\dot{x}^{yba} = -\dot{x}^{ybs} - \left(\sigma^{yaf} + \sigma^{yar}\right) x^{yba} \tag{11}$$
$$\dot{x}^{y\ell a} = -\dot{x}^{y\ell s} - \left(\sigma^{yaf} + \sigma^{yar}\right) x^{y\ell a} \tag{12}$$
$$\dot{x}^{oa} = -\dot{x}^{os} - \left(\sigma^{oaf} + \sigma^{oar}\right) x^{oa} \tag{13}$$
$$\dot{x}^{ybf} = \sigma^{yaf} x^{yba} - \left(\sigma^{yfe} + \sigma^{yfr}\right) x^{ybf} \tag{14}$$
$$\dot{x}^{y\ell f} = \sigma^{yaf} x^{y\ell a} - \left(\sigma^{yfe} + \sigma^{yfr}\right) x^{y\ell f} \tag{15}$$
$$\dot{x}^{of} = \sigma^{oaf} x^{oa} - \left(\sigma^{ofe} + \sigma^{ofr}\right) x^{of} \tag{16}$$
$$\dot{x}^{ybe} = \sigma^{yfe} x^{ybf} - \left(\sigma^{yed} + \sigma^{yer}\right) x^{ybe} \tag{17}$$
$$\dot{x}^{y\ell e} = \sigma^{yfe} x^{y\ell f} - \left(\sigma^{yed} + \sigma^{yer}\right) x^{y\ell e} \tag{18}$$
$$\dot{x}^{oe} = \sigma^{ofe} x^{of} - \left(\sigma^{oed} + \sigma^{oer}\right) x^{oe} \tag{19}$$
$$\dot{x}^{ybr} = \sigma^{yar} x^{yba} + \sigma^{yfr} x^{ybf} + (\sigma^{yer} - \varphi)x^{ybe} \tag{20}$$
$$\dot{x}^{y\ell r} = \sigma^{yar} x^{y\ell a} + \sigma^{yfr} x^{y\ell f} + (\sigma^{yer} - \varphi)x^{y\ell e} \tag{21}$$
$$\dot{x}^{or} = \sigma^{oar} x^{oa} + \sigma^{ofr} x^{of} + (\sigma^{oer} - \varphi)x^{oe} \tag{22}$$
$$\varphi = \lambda_o \max\{x^e - \Theta, 0\}. \tag{23}$$

2.4 Preferences

Preferences incorporate utility both from being alive and from being in a specific health state. Lifetime utility for the old is given by

$$E\left\{\int e^{-\rho_o t}\left[u(c_t^o)+\bar{u}+\widehat{u}_t^j\right]\,dt\right\} \tag{24}$$

where expectations are taken with respect to the random timing of death, and where $\bar{u}$ measures the flow utility from being alive (the utility of being dead is implicitly zero). Similarly, $\widehat{u}_t^j$ is the intrinsic utility of being in state health j. We will assume that $\widehat{u}_t^s = \widehat{u}_t^a = \widehat{u}_t^r = 0$, while $\widehat{u}_t^e < \widehat{u}_t^f < 0$. Thus, having flu-like symptoms is bad, and having to be treated in the hospital is very bad. The old value their consumption c_t^o according to the period utility function $u(c_t^o)$ and discount the future at rate ρ_o.

Symmetrically, the young also care about their consumption c_t^y, as well as about their health and about being alive, according to the lifetime utility function:

$$E\left\{\int e^{-\rho_y t}\left[u\left(c_t^y\right)+\bar{u}+\widehat{u}_t^j\right]\,dt\right\}, \tag{25}$$

In our calibration we will impose $\rho_o > \rho_y$, as a simple way to capture higher life expectancy for the young. As a result, while young and old enjoy the same flow value from being alive, the present value of this value will be lower for the old.

Note that workers who experience flu-like symptoms or are in the hospital do not work. Neither does a fraction m of workers in the luxury sector whose workplaces have been shut down by mitigation policy Therefore in equilibrium young workers will experience different consumption depending on whether they work or not. Thus, the expected utility of a worker will depend on the sector in which she works for two reasons. First, sectors differ in the share of economic activity being shut down (and thus, for the individual worker, in the probability of being able to work when healthy). Second, a worker's sector will affect her distribution of health outcomes.[9]

[9]Note that we have not modeled mortality from natural causes. Over the expected length of the COVID-19 pandemic, mortality from natural causes will be small for both age groups.

Covid Economics 6, 17 April 2020: 22-64

3 The Public Sector

In this section we first describe the government policy tools in Section 3.1, and then in Section 3.2 we analyze how public transfers are determined statically to yield a utilitarian period social welfare function. We conclude by posing the dynamic Ramsey optimal policy problem which maximizes the time integral of discounted instantaneous social welfare by choice of the optimal time path of mitigation m_t.

3.1 Transfers

The public sector is responsible for two choices: mitigation (shutdowns) m_t and redistribution to individuals that currently do not or cannot work. We assume that the infection-generating rates within the workplace and outside the workplace (the β's) are determined exogenously outside the model. What the government chooses is the extent to which it imposes a shutdown of economic activity, via m_t, and how much to transfer to those hurt by shutdowns, those that have fallen sick, and those that have retired. In each instant individuals either work (the young who are healthy enough and not subject to mitigation) or do not work (the rest). The second policy choice beyond mitigation is redistribution between working and non-working individuals. All workers share a common consumption level c^w and all individuals not working share a common consumption level c^n.[10] The second policy choice is how much to transfer, in each instant t, from the working to the non-working population. Crucially, we assume that these transfers are costly, denoting by $T(c^n)$ the per-capita cost of transferring consumption c^n to those out of work and without current income. We assume that $T(.)$ is increasing and differentiable.

To simplify notation, denote by $(\mu^n(m,x), \mu^w(m,x))$ the mass of non-working and working people, respectively, as a function of the health population distribution x and current mitigation $m = m_t$.[11] These are defined as

$$\mu^n(m,x) = x^{y\ell f} + x^{y\ell e} + x^{ybf} + x^{ybe} + m\left(x^{y\ell s} + x^{y\ell a} + x^{y\ell r}\right) + x^o \tag{26}$$

$$\mu^w(m,x) = x^{ybs} + x^{yba} + x^{ybr} + [1-m]\left(x^{y\ell s} + x^{y\ell a} + x^{y\ell r}\right) \tag{27}$$

[10]This is the allocation chosen by a government that values all individuals equally (equal Pareto weights). It is also the only allocation that is feasible if the government can observe an individual's income, but not her sector, age, or health status.

[11]We will suppress the dependence on (x, m) when there is no room for confusion.

Covid Economics 6, 17 April 2020: 22-64

$$\nu^w(m,x) \quad = \quad \frac{\mu^w(m,x)}{\mu^w(m,x)+\mu^n(m,x)} \tag{28}$$

where $\nu^w(m,x)$ is the share of working individuals in the population. The aggregate resource constraint can then be written as

$$\mu^w c^w + \mu^n c^n + \mu^n T(c^n) = y - \eta\Theta = \mu^w - \eta\Theta \tag{29}$$

where $y = \mu^w$ since each working individual produces one unit of output.

Notice that there are no dynamic consequences of the transfer choice c^n. In particular, this choice has no impact on any health transitions. We can therefore solve a static optimal transfer problem at each date t (given the current level of mitigation $m = m_t$) that delivers a maximum level of instantaneous social welfare which we denote $W(m,x)$. We turn to derive this expression now.

3.2 The Instantaneous Social Welfare Function

We now derive the instantaneous social welfare function $W(x,m)$, a necessary ingredient into the optimal mitigation problem of the government. The function $W(x,m)$, assuming that all individuals have log-utility and receive the same social welfare weights, is given by

$$W(x,m) = \max_{c^n,c^w} \left[\mu^w \log(c^w) + \mu^n \log(c^n)\right] + (\mu^w + \mu^n)\bar{u} + \sum_{i,j} x^{ij}\widehat{u}^j \tag{30}$$

where the maximization is subject to the aggregate resource constraint (29). Combining the first order conditions with respect to (c^n, c^w) yields

$$\frac{c^w}{c^n} = 1 + T'(c^n). \tag{31}$$

We can use this relation in the resource constraint to obtain

$$\mu^w \left(1 + T'(c^n)\right) c^n + \mu^n c^n + \mu^n T(c^n) = \mu^w - \eta\Theta \tag{32}$$

Covid Economics 6, 17 April 2020: 22-64

Defining net per-capita income $\tilde{y}$ and average transfer costs $t(c^n)$ as

$$\tilde{y} = \nu - \frac{\eta\Theta}{\mu^w + \mu^n} \tag{33}$$

$$t(c^n) = \frac{T(c^n)}{c^n} \tag{34}$$

we can rewrite the resource constraint in per-capita terms by dividing by $\mu^w + \mu^n$

$$c^n \left[1 + \nu T'(c^n) + (1-\nu)t(c^n)\right] = \tilde{y} \tag{35}$$

Thus the optimal solution to the government transfer problem is given by the solution to the following system:

$$c^n \left[1 + \nu T'(c^n) + (1-\nu)t(c^n)\right] = \tilde{y} \tag{36}$$

$$c^w = c^n(1 + T'(c^n))) \tag{37}$$

for an arbitrary differentiable per-capita transfer cost function $T(.)$. We can also express period welfare in per-capita terms, using

$$W(x,m) = [\mu^w + \mu^w]\, w(x,m) \tag{38}$$

$$w(x,m) = \log(c^n) + \nu \log(1 + T'(c^n)) + \bar{u} + \sum_{i,j} \frac{x^{ij}}{\mu^w + \mu^w} \widehat{u}^j, \tag{39}$$

where the only endogenous input in the period welfare function c^n solves equation (36). In particular, note that $\mu^w + \mu^w$ is independent of mitigation and thus we can discuss the impact of mitigation on current welfare in terms of the per-capita welfare function $w(x,m)$.

The per-capita welfare function shows the basic costs from mitigation m. First, it lowers per-capita income, and through it, the level of consumption. This is the $\log(c^n)$ term in $w(x,)$ which is strictly increasing in net income $\tilde{y}$. In the absence of the cost of transfers, this is the only direct effect of current mitigation. Second, the transfer cost to non-working households distorts risk sharing; this is the second term $\nu \log(1 + T'(c^n))$, which is zero if the marginal transfer cost is zero. Note that an increase in mitigation reduces ν and thus the negative impact of mitigation on current welfare is the more severe, the larger is the marginal cost of transfers. This, ceteris paribus, will reduce the incentives of the government to engage in economically

Covid Economics 6, 17 April 2020: 22-64

costly mitigation.

To see the intuition for our results most clearly, assume that the transfer cost is linear such that $T(c^n) = \tau c^n$. In this case the optimal allocation is given by:

$$
\begin{aligned}
c^w &= \tilde{y} \\
c^n &= \frac{\tilde{y}}{1+\tau} \\
w(x,m) &= \log(\tilde{y}) - (1-\nu)\log(1+\tau) + \bar{u} + \sum_{i,j} \frac{x^{ij}}{\mu^w + \mu^w}\widehat{u}^j
\end{aligned}
$$

Thus the negative economic impact of mitigation is given, in this case, by

$$
\frac{\partial w(x,m)}{\partial m} = \frac{\partial \tilde{y}}{\partial m} + (1+\tau)\frac{\partial \nu}{\partial m} < 0, \tag{40}
$$

since both $\frac{\partial \tilde{y}}{\partial m}$ and $\frac{\partial \nu}{\partial m}$ are negative. In addition, we observe that the larger is the marginal cost of transfers τ the more negative is $(1+\tau)\frac{\partial \nu}{\partial m}$. This is how mitigation and redistribution costs interact: the larger is the marginal cost of redistribution, the larger is the economic cost of mitigation $\frac{\partial w(x,m)}{\partial m}$.

In our quantitative exercises we will assume that the transfer cost function per non-worker is given by the quadratic form $T(c^n) = \frac{\tau}{2}\frac{\mu^n}{\mu^w}(c^n)^2 = \frac{\tau}{2}\left(\frac{1-\nu}{\nu}\right)(c^n)^2$ so that total transfer costs are given by $\mu^n T(c^n) = \mu^w \frac{\tau}{2}\left(\frac{\mu^n c^n}{\mu^w}\right)^2$. This functional form is motivated by the idea that each working household has to transfer $\left(\frac{\mu^n c^n}{\mu^w}\right)$ units of consumption to non-working households. Assuming a quadratic cost of extracting resources from workers, the per-worker cost is thus given by $\frac{\tau}{2}\left(\frac{\mu^n c^n}{\mu^w}\right)^2$.[12] Multiplying this by the total number of workers μ^w gives the total transfer cost. For this specification we obtain as optimal allocations to be inserted in the period welfare function above:

$$
c^n = \frac{\sqrt{1 + 2\tau\frac{1-\nu^2}{\nu}\tilde{y}} - 1}{\tau\frac{1-\nu^2}{\nu}} \tag{41}
$$

$$
c^w = c^n(1 + T'(c^n)) = c^n\left(1 + \tau\frac{1-\nu}{\nu}c^n\right) \tag{42}
$$

[12]The quadratic form is chosen for analytical convenience, but is not central for our qualitative arguments.

Covid Economics 6, 17 April 2020: 22-64

Note that $\left(1+\tau\frac{1-\nu}{\nu}c^n\right)$ is the effective price the planner has to pay, on the margin to take one more unit of output from workers to give to non-workers. As transfers and thus non-worker consumption c^n rise, this price effectively rises, reflecting a higher marginal cost to additional redistribution. In addition, since higher mitigation m reduces the share of workers ν and increases the share of non-workers $1-\nu$, the effective price of transfers at the margin increases with mitigation, and the price rises more the higher is τ.

For future reference, we can also construct expected flow utility for each type

$$\begin{aligned} W^{\ell}(x,m) &= \frac{(x^{y\ell n}+x^{y\ell e}+x^{y\ell r})}{x^{\ell}}\left[(1-m)u(c^w)+mu(c^n)+\bar{u}\right] \\ &\quad +\frac{(x^{y\ell f}+x^{y\ell e})}{x^{\ell}}\left[u(c^n)+\bar{u}-\hat{u}\right] \end{aligned}$$

$$W^b(x,m)=\frac{(x^{ybn}+x^{ybe}+x^{ybr})}{x^b}\left[u(c^w)+\bar{u}\right]+\frac{(x^{ybf}+x^{ybe})}{x^b}\left[u(c^n)+\bar{u}-\hat{u}\right]$$

$$W^o(x,m)=u(c^n)+\bar{u}-\frac{(x^{yof}+x^{yoe})}{x^o}\hat{u}$$

3.3 Optimal Policy

We now assume there is a government/planner (we use these names synonymously as there is no time consistency problem) that chooses optimal policy over time by choosing the path of mitigation $m(t)$; the optimal choice of redistribution $T(t)$ is already embodied in the period social welfare function $W(x)$. The policy problem the planner solves is then given by

$$\max_{m(t)}\int_0^{\infty}e^{-\rho t}W(x)\;dt. \tag{43}$$

subject to the laws of motion of the population Equation (4) to Equation (23).

In a first step we will approximate the optimal time path of mitigation by functions that are part of the following parametric class of generalized logistic functions of time:

$$m(t)=\frac{\gamma_0}{1+\exp(-\gamma_1(t-\gamma_2))} \tag{44}$$

Here the parameter γ_0 controls the level of mitigation at $t=0$. The parameter γ_2 governs when

Covid Economics 6, 17 April 2020: 22-64

mitigation is reduced, and the parameter γ_1 commands how swiftly mitigation is reduced. Note that as $t \to \infty$, $m(t) \to 0$.

More generally, the complete characterization of the optimal policy path is the solution to an optimal control problem. We state that problem formally in Appendix A. It shows that the key trade-off with mitigation efforts m is that a marginal increase in m entails static economic costs of $W_m(x, m)$ stemming from the loss of output and thus consumption of all individuals in the economy, as encoded in $y = y(m)$. The dynamic benefit is a favorable change in the population health distribution: an increase in m reduces the outflow of individuals from the susceptible to the asymptomatic state. We plan to solve the complete unconstrained optimal control problem in future versions of this paper.

4 Calibration

There is a long list of parameters to specify. We set the population share of the young, μ^y, to be 85%, which is the current fraction of the US population aged below 65.

Preferences We assume logarithmic utility from consumption:

$$u(c) = \log c.$$

When evaluating welfare, we discount utility for both young and old at a three percent annual rate for the first 500 days. In all the cases we have explored, 500 days is sufficient for the pandemic to have run its course. After that we apply different discount factors to the two groups to compute remaining lifetime utility in the final steady state, as a simple way to take into account shorter remaining life expectancy for the old. In particular, we think of the typical young person being 32.5 years old with 47.5 expected years to live, and the typical old person being 72.5 with 14 years to live, where these life expectancies are taken from actuarial life tables. Given these values, and a pure discount rate of three percent, adjusted discount rates that incorporate differential expected longevity are 4 percent for the young, and 10 percent for the old.

To set the value of life $\bar{u}$ we follow the value of a statistical life (VSL) approach. The Environmental Protection Agency and the Department of Transportation assume a VSL of \$11.5 million (see Greenstone and Nigam 2020). This is a relatively high value, relative to

Covid Economics 6, 17 April 2020: 22-64

values used in other contexts. Assuming an average of 37 residual life years discounted at a 3 percent rate, this translates to an annual flow value of \$515,000, which is 11.4 times yearly per-capita consumption in the United States.

To translate this into a value for $\bar{u}$ we use the standard value of a statistic life calculation,

$$VSL = \frac{dc}{dr}|_{E[u]=k} = \frac{\ln(\bar{c}) + \bar{u}}{\frac{1-r}{\bar{c}}}$$

where $\bar{c}$ is average per-capita model consumption, and r is the risk of death. Setting $VSL = 11.4\bar{c}$ and $r = 0$ gives $\bar{u} = 11.4 - \ln \bar{c}$. Note that this implies an easily interpretable trade-off between mortality risk and consumption. For example, we can ask what reduction in consumption leads an individual indifferent to facing a 1 percent risk of death. The answer is the solution m to

$$\ln(\bar{c}\,(1-m)) + 11.4 - \ln \bar{c} = 0.99\,(\ln(\bar{c}) + 11.4 - \ln \bar{c})$$

which is $m = 1 - \exp(-0.01 \times 11.4) = 10.8\%$.

As another way to get a feeling for what our choice for the value of a statistical life implies, suppose we were to contemplate a shut down that would reduce consumption for six months by 25 percent. By how much would this shut down have to reduce mortality risk for an agent with 10 expected years of life to prefer the shutdown to no shutdown? The answer is the solution x to

$$\frac{1}{20}\ln(1-0.25) + \frac{19}{20}\ln(1) + 11.4 = (1-x)11.4$$

which is 0.13 percent.

For the disutility of having flu, we define $\hat{u}^f$ as

$$\hat{u}^f = -0.3\,(\ln(\bar{c}) + \bar{u})$$

following Hong et al. (2018). We set $\hat{u}^e = -\,(\ln(\bar{c}) + \bar{u})$, so that the flow value of being in hospital is equal to the flow value of being dead (zero).

Sectors To calibrate the employment and output share of the basic sector of the economy, μ^b, we use BLS employment shares by industry. We categorize the following industries

Covid Economics 6, 17 April 2020: 22-64

as basic: agriculture, health care, financial activities, utilities, and federal government. Mining, construction, manufacturing, education, leisure and hospitality are allocated to the luxury sector. The remaining industries are assumed to be a representative mix of basic and luxury. This partition implies that, pre-COVID, the basic sector accounts $\mu^b = 45.4$ percent of the economy.

Redistribution We adopt the quadratic formulation of transfer costs described above. We pick a value for τ using estimates for the excess burden of taxation, which suggest that raising an extra dollar in revenue at the margin (which can be used to increase consumption for non-workers) has a cost on taxpayers of around \$1.38 (Saez et al. (2012)). This suggests $\tau \frac{1-v}{v} c^n = 0.38$. Given the first order condition above, this means that an optimal redistribution scheme would imply $c_n/c_w = 1/1.38 = 0.72$ in pre-COVID times. Moreover, given $\eta\Theta = 0.021$, $\tau \frac{1-v}{v} c^n = 0.38$, and $v = \mu^y = 0.85$, section 3.2 implies $\tau = 3.51$.

Hospital Capacity Tsai et al. (2020) estimate that 58,000 ICU beds are potentially available nationwide to treat COVID-19 patients. However, only 21.5 percent of COVID-19 hospital admissions require intensive care, suggesting that total hospital capacity is around $58,000/0.215 = 270,000$. Tsai et al. (2020) emphasize that this capacity is very unevenly allocated geographically, and in addition, there is significant geographic variation in virus spread. Thus, capacity constraints are likely to bind in more and more locations as the virus spreads. We therefore set $\Theta = 100,000$, so that hospital mortality starts to rise when 0.042 percent of the population is hospitalized. Because the cost of a day in intensive care is around \$7,500, we set $\eta = 50$, so emergency care consumes about 2.1 percent of pre-COVID output.[13] We set the parameter λ_o such that, absent economic mitigation, the death rate in emergency care for the old at the peak of the epidemic is 20 percent above its value when capacity is not exceeded.

Disease Progression There are twelve σ parameters to calibrate, describing transition rates for disease progression, six for each age. These describe the chance of moving to the next worse health status and the chance of recovery at the three infectious stages: asymptomatic, flu-suffering, and hospitalized. We assume that young and old exit each stage at the same rate, but potentially differ in terms of the share of these exits that are into recovery. In particular, the old will be much more likely to require hospital care conditional on developing flu-like symptoms, and more likely to die conditional on being hospitalized.

[13] Total healthcare spending in the United States is 18 percent of GDP. Of this, around 1/3 is spending on hospitals.

Covid Economics 6, 17 April 2020: 22-64

Covid Economics 6, 17 April 2020: 22-64

Putting aside these differences by age for a moment, the six values for σ are identified from the following six target moments: the average duration of time individuals spend in the asymptomatic (contagious but without symptoms for the disease.), flu-suffering (relatively mild symptoms), and emergency-care states, and the relative chance of recovery (relative to disease progression) in each of the three states. Following the literature on COVID-19 models we set the three durations to 5.2, 10, and 8 days, respectively, and with these durations common across age groups. The exit rate to recovery from the asymptomatic state defines the number of asymptomatic cases of COVID-19 and is an important but highly uncertain parameter. We assume that asymptomatic recovery and progression to the flu-suffering state are equally likely.[14]

We let the relative recovery rates from the flu-suffering and emergency care states vary with age, to reflect the fact that infections in older individuals are more much likely to require hospitalization, and hospitalizations are also somewhat more likely to lead to death. We set the recovery rate from flu-suffering to 96% for the young, and to 75% for the old, based on evidence from Table 1 of the Imperial College study (Ferguson et al. (2020)). Similarly, given evidence on differential mortality rates, we set the recovery rates from the emergency care state to 95% for the young and to 80% for the old (assuming no hospital overuse). Given these choices the probability that a newly-infected young individual will ultimately die from COVID-19 is $0.5 \times 0.04 \times 0.05 = 0.1\%$, while the conditional probability, conditional on developing flu symptoms, is 0.2%. The corresponding numbers for an older individual are 2.5% and 5.0%.

Sources of Infection Given the σ parameters, the β parameters determine the rate at which contagion grows over time. We set $\beta_e = 0.36$, implying that at the peak of infections approximately 5% of infections are to health care workers.[15] The values of α_w, α_c, and β_h determine the overall basic reproduction number R_0 value for COVID-19, and the share of disease transmission that occurs at work, market consumption, and non-market settings.

Mossong et al. (2008) find that 35 percent of potentially-infecting inter-person contact happens in workplaces and schools, that 19 percent occurs in travel and leisure activities, and that the remainder is in home and other settings.[16] These shares should be interpreted

[14]Given that the asymptomatic state has roughly half the duration of the flu state, this implies that roughly half of infected agents in the model will be asymptomatic. Recall that in a random sample in Iceland, half of the positive subjects reported no symptoms.

[15]On March 24th, 14 percent of Spain's confirmed cases were health care workers. However, infection rates of health care workers appear lower in other countries. https://www.nytimes.com/2020/03/24/world/europe/coronavirus-europe-covid-19.html

[16]Xu et al. (2020) discuss in detail heterogeneity in contact rates across different types of business (closed

as reflecting behaviors in a normal period of time, rather than in the midst of a pandemic. We associate workplace and school transmission with transmission at work, travel and leisure with consumption-related transmission, and the residual categories with transmission at home. These targets are used to pin down choices for α_w and α_c, both relative to β_h, as follows.

The basic reproduction number R_0 is the number of people infected by a single asymptomatic person. For a single young person, assuming everyone else in the economy is susceptible and zero mitigation ($m = 0$), R_0^y is given by

$$R_0^y = \frac{\alpha_w x^y + \alpha_c \mu^y + \beta_h}{\sigma^{yar} + \sigma^{yaf}} + \frac{\sigma^{yaf}}{\sigma^{yaf} + \sigma^{yar}} \frac{\beta_h}{\sigma^{yfr} + \sigma^{yfe}} + \frac{\sigma^{yaf}}{\sigma^{yaf} + \sigma^{yar}} \frac{\sigma^{yfs}}{\sigma^{yfe} + \sigma^{yfr}} \frac{\beta_e x^{yb}}{\sigma^{yer} + \sigma^{yed}}$$

where this expression exploits the fact that when $m = 0$, $\beta_w(0) = \alpha_w$ and $\beta_c(0) = \alpha_c$.

The logic is that this individual will spread the virus while asymptomatic, flu-suffering, and in hospital – the three terms in the expression. They expect to be asymptomatic for $(\sigma^{yar} + \sigma^{yaf})^{-1}$ days, flu-suffering (conditional on reaching that state) for $(\sigma^{yfr} + \sigma^{yfe})^{-1}$ days, and hospitalized (conditional on reaching that state) for $(\sigma^{yer} + \sigma^{yed})^{-1}$ days. The chance they reach the flu-suffering state is $\frac{\sigma^{yaf}}{\sigma^{yaf}+\sigma^{yar}}$ and the chance they reach the emergency room is the product $\frac{\sigma^{yaf}}{\sigma^{yaf}+\sigma^{yar}} \frac{\sigma^{yfs}}{\sigma^{yfe}+\sigma^{yfr}}$. While asymptomatic, they spread the virus at both work and at home, and pass the virus on to $\alpha_w x^y + \alpha_c \mu^y + \beta_h$ susceptible individuals per day.[17] While flu-suffering, they stay at home and pass the virus to β_h individuals per day. While sick they pass it to $\beta_e x^{yb}$ basic workers per day in hospital.

The reproduction number for an old asymptomatic person is

$$R_0^o = \frac{\alpha_c \mu^y + \beta_h}{\sigma^{oar} + \sigma^{oaf}} + \frac{\sigma^{oaf}}{\sigma^{oaf} + \sigma^{oar}} \frac{\beta_h}{\sigma^{ofr} + \sigma^{ofe}} + \frac{\sigma^{oaf}}{\sigma^{oaf} + \sigma^{oar}} \frac{\sigma^{ofs}}{\sigma^{ofe} + \sigma^{ofr}} \frac{\beta_e x^{yb}}{\sigma^{oer} + \sigma^{oed}}$$

where this formula is similar to the one for the young, except that it recognizes the old pass the virus on less because they do not work. At the same time, however, because the old are less likely to recover once infected, they potentially carry the virus for a longer time, inducing more transmission in hospitals.

office, open office, manufacturing and retail), and a range of interventions that can reduce those rates.

[17] Recall that x^y is the pre-COVID number of workers, and α_w is the probability that transmission occurs when an infected worker meets a susceptible one. Recall that we assume consumption contagion is proportional to output, and pre-COVID output is $\mu^y = x^y$.

Covid Economics 6, 17 April 2020: 22-64

CEPR PRESS

Covid Economics 6, 17 April 2020: 22-64

For the population as a whole, the overall R_0 is a weighted average of these two group-specific values

$$R_0 = \mu_y R_0^y + (1 - \mu_y) R_0^o$$

where μ^y is the fraction of the population that is young.

The share of total transmission that occurs in the workplace from a randomly drawn newly asymptomatic individual is then given by

$$\frac{\text{workplace transmission}}{\text{all transmission}} = \frac{1}{R_0} \mu^y \left(\frac{\alpha_w}{\sigma^{yar} + \sigma^{yaf}} \right),$$

while the share of transmission due to market consumption is

$$\frac{\text{consumption transmission}}{\text{all transmission}} = \frac{1}{R_0} \left[\mu^y \left(\frac{\mu^y \alpha_c}{\sigma^{yar} + \sigma^{yaf}} \right) + (1 - \mu^y) \left(\frac{\mu^y \alpha_c}{\sigma^{oar} + \sigma^{oaf}} \right) \right].$$

Given these three equations, we set the relative values α_w / β_h, α_c / β_h to replicate shares of workplace and consumption transmission equal to 35 and 19 percent. Note that this evidence does not pin down the *levels* of α_w, α_c and β_h, to which we now turn.

History, R_0, and Initial Conditions We will think of a policy-maker choosing a path for mitigation m_t starting from April 12, 2020. The dynamics of the disease going forward, and thus the optimal path for m_t, will be highly sensitive to the distribution of the population by health status at this date: how many people of each type are susceptible, infected, and recovered, and how the infected group is partitioned by stage into asymptomatics, those with flu, and those in hospital. It is not easy to get an accurate cross-sectional picture of the health of the population, given the fact that only a very small share of the population has recently been tested.

In addition, the dynamics of the disease going forward will depend on the basic reproduction number R_0, which in our model is determined at a structural level by the levels of the infection-generating parameters α_w, α_c and β_h. Existing estimates for R_0 for COVID-19, absent additional social distancing measures or economic shutdowns, are in the range of 2 to 4 (e.g., Flaxman et al. (2020)). But given all the precautions that Americans are currently choosing to take, or being required to take, the current effective R_0 is likely much lower. In addition, the fact that a large share of the U.S. economy has been shuttered has likely lowered R_0 still further.

To pin down the April 12 health status distribution and the April 12 level for the infection-generating parameters we take the following approach. First, we will assume that America changed on March 21st. Before that date, people behaved as normal, and none of the economy was shuttered, corresponding to $m = 0$. On March 21 we assume infection-generating rates fell discretely and proportionately to new lower levels $\zeta\alpha_w$, $\zeta\alpha_c$ and $\zeta\beta_h$ with $\zeta < 1$. Of course, in reality changes in social distancing practices and shutdowns happened more gradually, but March 21 seems a reasonable focal date: California announced closure of non-essential businesses on March 19, New York and Illinois on March 20. In addition, and at the same date, we assume that states introduced measures that effectively shut down a fraction $m = 0.5$ of the luxury sector, immediately therefore idling $0.5(1 - \mu^b) = 27.5$ percent of the workforce. It is difficult to assess how much of economy has been affected directly or indirectly by shutdown measures, but our value for m is consistent with the Faria-e-Castro (2020) forecast that US unemployment will rise above 30 percent in the second quarter (Bick and Blandin (2020) estimate that it is already 20 percent).

Of the data on health outcomes we have, perhaps the most reliable are for the number of deaths attributable to COVID-19. We will therefore target three specific moments involving deaths: (1) the cumulative number of deaths up to March 21 (301), the cumulative number as of April 12 ($21,000$) and the moving-average number of deaths per day by April 12 ($2,000$). To hit these target moments, we treat as free parameters (1) β_h – the pre-March 21 infection-generating rate at home, (2) ζ, the proportional amount by which infection-generating rates fall on March 21, and (3) the initial number of infections at the date we start our model simulation, which is February 12.

To understand how this identification scheme works, consider that the death toll rose from 300 to $21,000$ deaths in only three weeks, but the number of daily deaths was not especially high (nor growing especially fast) at the end of this period. This suggests that there were already many infections in the pipeline on March 21, but that those infections did not grow rapidly from March 21 onward, which indicates a low value for ζ. At the same time, a high level of March 21 infections is informative about the level of initial infections on Feb 12. Finally, a large number of infections on March 21, but a low death toll up to March 21, points to high R_0 (and a high β_h) prior to March 21: rapid spread can deliver lots of new infections without (yet) many deaths.

Covid Economics 6, 17 April 2020: 22-64

Table 1: Millions of People in Each Health State

	S	A	F	E	R
March 21	321.84	5.57	1.04	0.01	1.54
April 12	305.39	4.16	3.68	0.15	16.59

This calibration strategy yields an initial effective R_0 prior to March 21 of 3.0, which falls to 0.72 after March 21, reflecting a value for ζ of 0.33. Part of this decline reflects the start of economic mitigation. Absent mitigation (with $m = 0$) the effective R_0 after March 21 would be 1.0. This calibration implies the following distribution of the population by health status:

Thus, the calibration implies that 2.0 percent of the US population was actively infected on March 21, with that number rising to 2.4 percent by April 12, with an additional 5.1 percent having recovered.[18]

For the time path of mitigation, our baseline simulation, designed to approximate current US policy, will assume $m = 0.5$ for 100 days from March 21 onward, followed by $m = 0$ thereafter. This path is implemented in the context of the mitigation function (eq. 44) by setting $\gamma_0 = 0.5$, $\gamma_1 = -0.3$, and $\gamma_2 = 100$.

[18]These numbers are within the range of expert estimates from the COVID-19 survey compiled by McAndrew (2020) at the University of Massachusetts.

CEPR PRESS

Table 2: Epidemiological Parameter Values

	Behavior-Contagion (Before March 31)		
α_w	infection at work	35% of infections	0.24
α_c	infection through consumption	19% of infections	0.11
β_h	infection at home	$R_0 = 3.0$	0.16
β_e	infection in hospitals	5% of infections at peak	1.11
ζ	Scale of social distancing	$R_0 = 1.0$ w/ $m = 0$	0.33
	Disease Evolution		
σ^{yaf}	rate for young asymptomatic into flu	50% flu, 5.2 days	$\frac{0.5}{5.2}$
σ^{yar}	rate for young asymptomatic into recovered		$\frac{0.5}{5.2}$
σ^{oaf}	rate for old asymptomatic into flu	50% flu, 5.2 days	$\frac{0.5}{5.2}$
σ^{oar}	rate for old asymptomatic into recovered		$\frac{0.5}{5.2}$
σ^{yfe}	rate for young flu into emergency	4% hospitalization, 10 days	$\frac{0.04}{10}$
σ^{yfr}	rate for young flu into recovered		$\frac{0.96}{10}$
σ^{ofe}	rate for old flu into emergency	25% hospitalization, 10 days	$\frac{0.25}{10}$
σ^{ofr}	rate for old flu into recovered		$\frac{0.75}{10}$
σ^{yed}	rate for young emergency into dead	0.2% mortality, 8 days	$\frac{0.05}{8}$
σ^{yer}	rate for young emergency into recovered		$\frac{0.95}{8}$
σ^{oed}	rate for old emergency into dead	5.0% mortality, 8 days	$\frac{0.20}{8}$
σ^{oer}	rate for old emergency into recovered		$\frac{0.80}{8}$

Covid Economics 6, 17 April 2020: 22-64

Table 3: Economic Parameters

ρ	pure discount rate	3.0% per year	$\frac{0.03}{365}$
ρ_+^y	effective discount rate of young	4.0% per year	$\frac{0.01}{365}$
ρ_+^o	effective discount rate of old	10% per year	$\frac{0.07}{365}$
$\bar{u}$	value of life	11.4× consumption p.c.	11.24
$\widehat{u}^f$	disutility of flu	lose 30% of baseline utility	-3.37
$\widehat{u}^e$	disutility of emergency care	lose 100% of baseline utility	-11.24
μ^b	size of basic sector	45.4%	0.454
μ^y	share of young	85%	0.85
τ	transfer cost	$0.38 burden of excess taxation	3.51
α_0	initial share mitigated	50%	0.5
α_1	speed of mitigation		−0.3
α_2	time mitigation begins	100 days	100
Θ	hospital capacity	100, 000 beds	0.00042
η	bed cost	$7,500	50
λ_o	impact of overuse on mortality	20% higher mortality at peak	5.30

Covid Economics 6, 17 April 2020: 22-64

CEPR PRESS

5 Findings

We start by describing model outcomes under what we think of as the policies currently in place in the United States. We then turn to optimal mitigation in the next section.

5.1 Benchmark Results

In Figures 1 to 5 we display the population health dynamics from March 21 to the end of 2020. The red dashed lines represent our baseline scenario with 50 percent economic mitigation ($m_t = 0.5$) for 100 days, and social distancing as described in Section 4. The blue solid line is an alternative that shares the same time path of parameters and policies prior to April 12 – including 50 percent mitigation between March 21 and April 12 – but in which mitigation is set to zero from April 12 onwards.

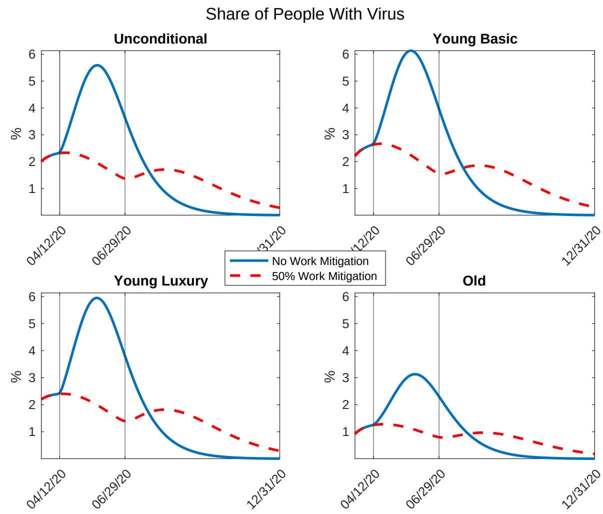


Figure 1: Share of each group infected (asymptomatic + flu + hospitalized).

We start with the currently infected (asymptomatic plus those with flu symptoms and those

Covid Economics 6, 17 April 2020: 22-64

in hospital) in Figure 1. Under our baseline policy, the red dashed line indicates that at April 12 we are already close to the peak of active infections. In constrast, if economic mitigation were to cease enforced starting on April 12, the share of the population actively infected would nearly triple reaching around 5.5% of the population at the end of May. With economic mitigation continuing into June instead, that share never exceeds 2.5% of the population. The figure also shows that the timing of the relaxation of economic mitigation matters greatly: in the benchmark the partial shut-down is lifted at the end of June, and although infections increase mildly the next two months, the peak infection rates are never nearly as large as under the scenario when economic mitigation ends now (April 12).

Covid Economics 6, 17 April 2020: 22-64

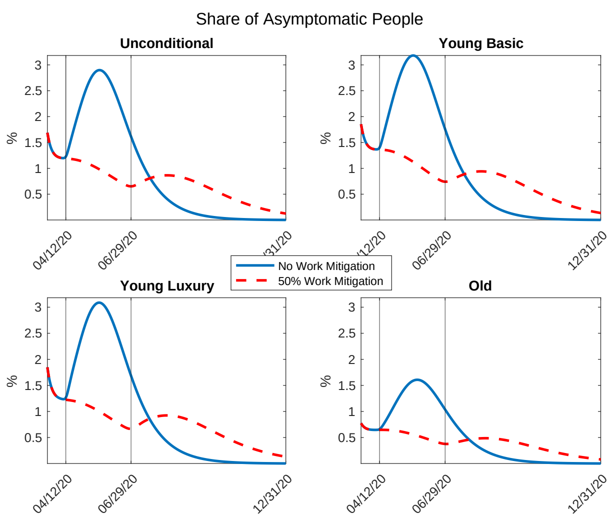


Figure 2: Share of each group infected but asymptomatic

Turning now to the heterogeneity across the population, note that, absent economic mitigation, basic and luxury sector workers are infected at nearly identical rates, while the old – who do not face exposure at work – experience a much lower rate of infection. Economic mitigation

CEPR PRESS

reduces infection rates for all three types. Comparing the two types of workers, the effect of mitigation (i.e., the gap between the blue solid and the red dashed line) is slightly larger for luxury workers – since a share of them stays at home under economic mitigation. But all three groups benefit from economic mitigation to a surprisingly similar extent, reflecting the fact that lower virus spread at work means fewer infected people outside of work, and thus fewer new infections at home and in stores.

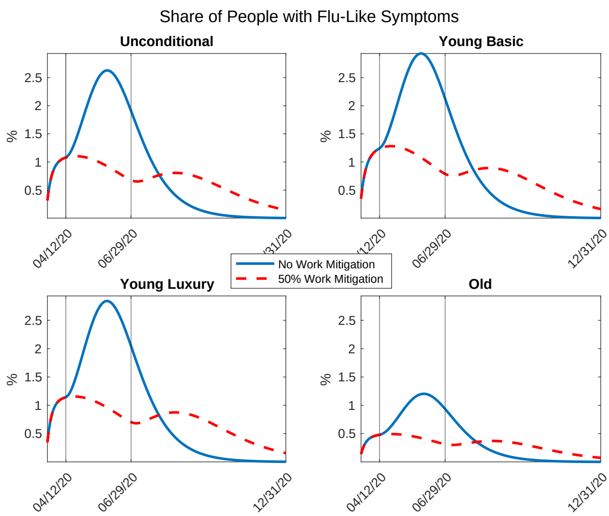


Figure 3: Share of each group with flu symptoms

The next three decompose active infections into the asymptomatic (Figure 2), those suffering from flu-like symptoms (Figure 3), and those in hospital (4). The key observation to note here is that while a smaller share of the old develop mild symptoms, reflecting a lower infection rate (see Figure 3), a much larger share of the old population ends up being severely sick and hospitalized, as the lower right panel of Figure 4 shows. This is true under both mitigation scenarios, but the effect is especially pronounced if economic mitigation is abolished

Covid Economics 6, 17 April 2020: 22-64

early: infections first in the work-place and then at home and during shopping trips sky-rocket, translating into more infections among the old. Although the old are only half as likely to become infected as the young, conditional on becoming infected they are over six times as likely to eventually require hospitalization.

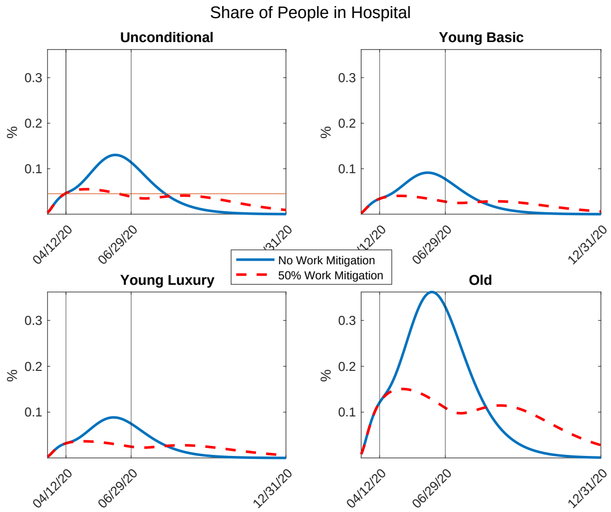


Figure 4: Share of each group hospitalized.

The red horizontal line in the upper left panel of Figure 4 plots hospital capacity, Θ, which we assume to be fixed in the short run. This plot shows another dramatic difference between the two mitigation scenarios. Under the benchmark scenario with 50% economic mitigation until the end of June, the demand for hospital care almost never exceeds capacity, with a modest exception in the weeks of April. However, even then excess demand is never severe, and thus the excess death rate that comes with overstretched hospitals is mild and short-lived. Contrast this to the (counterfactual) scenario when economic mitigation policies are suspended on April 12. The blue line in the upper left panel of Figure 4 indicates that capacity is drastically exceeded for several months in this scenario.

Figure 5 shows daily deaths from COVID-19. Under the baseline policy, with 50 percent

Covid Economics 6, 17 April 2020: 22-64

CEPR PRESS

Covid Economics 6, 17 April 2020: 22-64

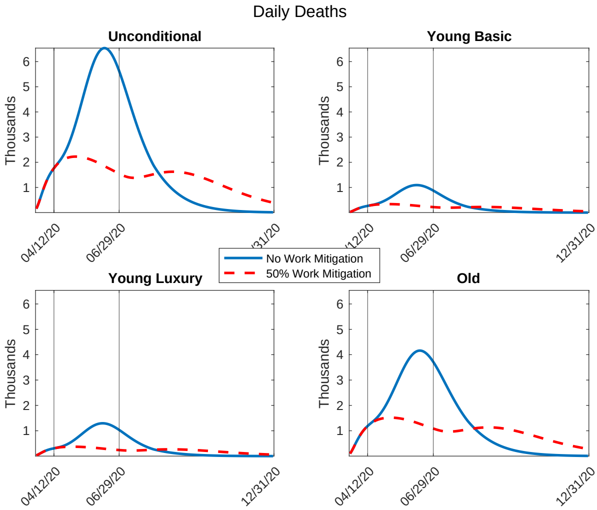


Figure 5: Daily deaths from COVID-19

Covid Economics 6, 17 April 2020: 22-64

mitigation, deaths rise a little but never exceed 2,500 per day. If economic mitigation is ended at Easter, the daily death toll rises dramatically, exceeding 6,000 at the peak. The breakdown across population groups indicates that the virus is predicted to kill more older individual than younger ones, even though the old account for only 15 percent of the population.

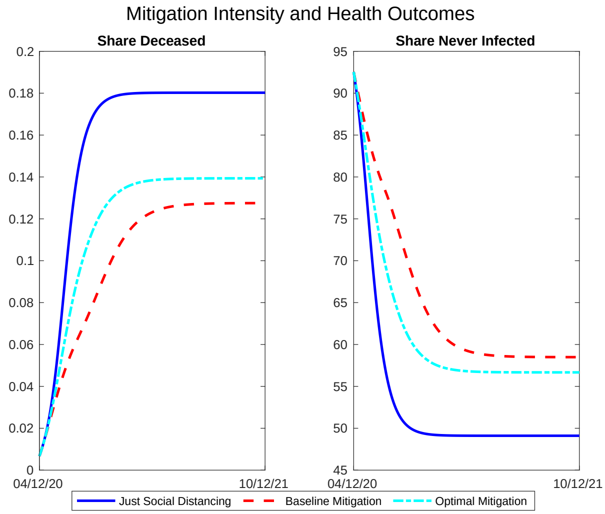


Figure 6: Left panel: share of the initial population deceased. Right panel: share of population never infected (susceptible).

In Figure 6 we display the population health dynamics over the next 18 months, starting on April 12, under three different scenarios. The red dashed line represents the baseline scenario with 50 percent economic mitigation for 100 days. The solid blue *Just Social Distancing* scenario is again the case with no work mitigation. The new cyan dashed line describes the population dynamics associated with the *Optimal Mitigation* path discussed in detail in Section 5.2.

The left panel displays the cumulative share of the population that has died from the virus, and the right panel plots, against time, the share of the population that has not yet been infected (i.e., the susceptible group).

Absent economic mitigation, the virus spreads rapidly, and after about three months 50 percent of the U.S. population has been (or is currently) infected with the virus: the blue line with the never-infected share of the populations drops rapidly below 50 percent. In contrast, under our projection for the current economic mitigation plan, the never infected share declines more slowly and a significantly larger share of the population is never touched by the virus (60 percent rather than 50). That is, aggressive mitigation measures do not just flatten the curve: they also reduce the total number of infections. The logic is that in the SIR class of models, the growth rate of infections depends not just on how many people are infected, but also on the relative shares of susceptible versus recovered individuals in the non-infected population. More aggressive mitigation measures slow the spread of infection, such that infections peak later. But delaying the peak in infections gives time for more people to recover and develop immunity, which slows infection growth. The result is that the economy converges to a steady state in which a larger share of individuals has never been infected relative to the scenario in which the economy open up at Easter.

The left panel translates infections into mortality associated with the virus. In the absence of economic mitigation the death toll of the virus rises rapidly, and by the end of the outbreak 0.18% of the U.S. population is predicted to have lost their lives, which amounts to $590,000$ people. Under the current benchmark economic mitigation policy that number falls to 0.1275% ($418,000$ individuals). The difference in lives lost ($172,000$) comes from two sources. First, there is less hospital overload and excess associated mortality with economic mitigation in place. Second, with mitigation, a smaller cumulative total number of infections means that fewer people ever risk adverse health outcomes and death.

Finally, the figure also anticipates our optimal mitigation policy finding that the optimal mitigation path is somewhere between the current level of economic shut-down and no shut-down at all, with health consequences that lie in between the benchmark and the no economic mitigation scenario. We will discuss this finding in greater detail in the next section.

Figure 7 plots the dynamics of consumption for workers and non-workers through the course of the pandemic. Recall that in this economy all workers enjoy the same consumption level, independent of sector, and the government provides equal consumption via transfers to all

Covid Economics 6, 17 April 2020: 22-64

Covid Economics 6, 17 April 2020: 22-64

non-workers, irrespective of whether they are not working because they are old, sick, or asked to stay home because of economic mitigation. The four panels correspond to four different economies. In the top two panels, we assume use our baseline value for τ, which implies that it is costly for the planner to redistribute from workers to non-workers. In the bottom two panels, we set $\tau = 0$, so that the planner can redistribute freely. In that case, the planner equates consumption between workers and non-workers at each date.[19]

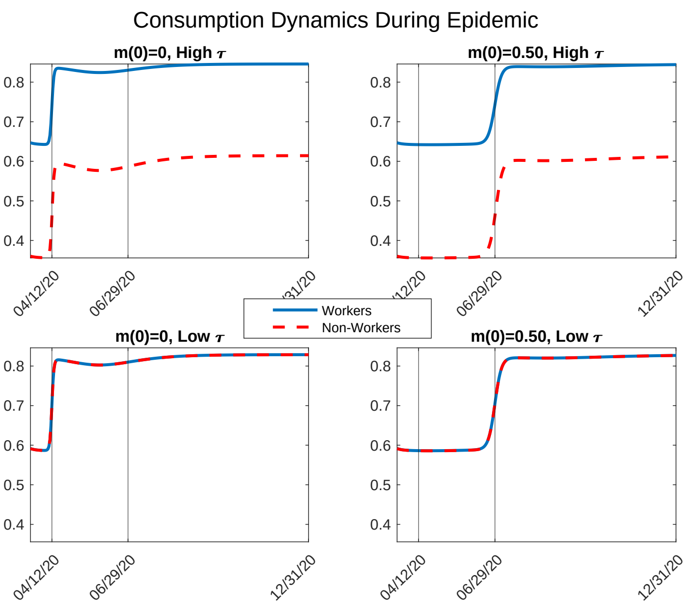


Figure 7: Consumption paths. Top two panels, $\tau = 0.347$. Bottom two panels, $\tau = 0.001$. Left two panels, $m = 0$. Right two panels, $m = 0.5$ for 100 days, then $m = 0$.

[19]Since the levels of consumption chosen by the panner statically do not affect infection rates and thus have no impact on the dynamics of health, the evolution of the population health distribution is independent of the cost of transfers.

Table 4: Welfare Gains (+) or Losses (-) From Mitigation

Mitigated Share	50%		25%		10%	
Young Basic	-0.01%	-0.08%	0.10%	0.06%	0.07%	0.05%
Young Luxury	-0.31%	-0.08%	-0.03%	0.06%	0.03%	0.06%
Old	1.18%	1.72%	1.32%	1.65%	0.73%	0.88%

Comparing across columns, the left two panels display the evolution of consumption when economic mitigation ends on April 12, and the right two panels maintain 50 percent mitigation until the end of June. In the first case, the economy immediately recovers as all healthy workers that were affected by the shut-down in the luxury sector return back to work, increasing output, income and thus aggregate consumption in the economy by about 27.5 percent.[20] The right two panels show that in terms of output and thus consumption a later end to the shut-down simply (and somewhat mechanically) postpones the economic recovery by 2.5 months. Note from the upper right panel of Figure 7 that the cost of economic mitigation is born disproportionately by non-workers: the ratio of non-worker to worker consumption declines (from two thirds to one half) during the mitigation phase. This reflects our assumption that extracting resources to redistribute from workers becomes ever harder the more the planner wants to tax each worker. To avoid very large redistribution costs, the planner optimally chooses to reduce insurance during the mitigation phase and increases it again as the economy recovers.

Next we report the expected welfare gains and losses for each type of individual for various assumptions about the level of economic mitigation and the parameter τ that indexes the cost of redistribution. In particular, we consider three mitigation levels: $m = 0.5$ (our baseline used to construct the previous plots), $m = 0.25$, and $m = 0.1$. In each case we assume mitigation is in place for 100 days from March 21. The welfare calculation asks: what percent of consumption would a person be willing to pay every day for the rest of his life to move from the economy with $m = 0$ to $m = 0.5$ (or $m = 0.25$ or $m = 0.1$) for 100 days. We report results for our baseline value for τ (3.51) and for a case in which redistribution is costless ($\tau = 0$).

The first clear message from Table 4 is that economic mitigation offers significant welfare gains for old, but has much more modest welfare effects on the young. For example, in

[20]Note that we assume that infected people with symptoms stay home rather than go to work, and since the share of infected individuals is endogeneously evolving over time, the increase is not exactly equal to the 27.5% decline in output when economic mitigation was introduced in the first place.

Covid Economics 6, 17 April 2020: 22-64

our baseline case ($m = 0.5$ and $\tau = 3.51$) the old gain 1.18 percent of consumption, while the young basic workers are essentially indifferent relative to no shutdown, and young luxury workers experience welfare losses equivalent to losing 0.31 percent of consumption. The reason the gains are much larger for the old is simply that the old face a much higher likelihood of being killed by the virus, and that strong economic mitigation policies reduce infections in the workplace, which in turn lowers the risk that the old meet infected individuals at home or while shopping.

The second key message is that the cost of redistribution matters. In particular, when redistribution is costless, young luxury workers and young basic workers perceive essentially identical welfare effects from mitigation.[21] However, when redistribution is costly, young luxury workers fare notably worse than young basic workers, because they risk larger expected consumption losses from economic mitigation. The reason is that when mitigation is increased, the planner needs to redistribute from a smaller pool of workers toward a larger pool of non-workers. Given convex costs of extracting additional resources from workers, this induces the planner to reduce insurance, translating into a larger consumption gap between workers and non-workers.

We now briefly discuss a few factors that shape these welfare calculations. First, the overall level of the welfare numbers is sensitive to several choices. A key one is the value of a statistical life: a lower value would make life-saving economic mitigation trivially less attractive. Second, if we assumed lower recovery rates at different stages of an infection, or a higher mortality rate at the hospital stage, agents would perceive a greater risk of death, and be more willing to sacrifice consumption to avoid that risk. Third, in our model, when a shutdown raises non-employment and reduces consumption, there is no upside in households' utility functions from more leisure. In the analysis of optimal shutdowns in Eichenbaum et al. (2020), the fact that households experience reduced disutility from labor supply when economic activity is taxed compensates strongly for the utility cost of reduced consumption. Finally, the attractiveness of shutdowns clearly depends on the share of virus transmission that occurs through different forms of economic activity: the larger is that share, the more powerful are shutdowns as a tool to slow transmission.

[21] In fact, luxury workers are slightly more pro-mitigation, since they benefit more from reduced infection at work.

5.2 Optimal Policy

The mitigation policies we have compared thus far were not chosen optimally. We now turn to exploring the optimal time path for economic mitigation, and the associated statically optimal degree of redistribution, given that path. To start, we optimize over the three parameters in our parametric process for m_t. That is, we choose γ_0, γ_1, and γ_2 in eq. 44 to maximize social welfare as defined in Section 3.2. The choice of these parameters lets the government control the initial size of economic mitigation, when it ends, and how quickly it is phased out. Figure 8 describes the preferred policies within this class.

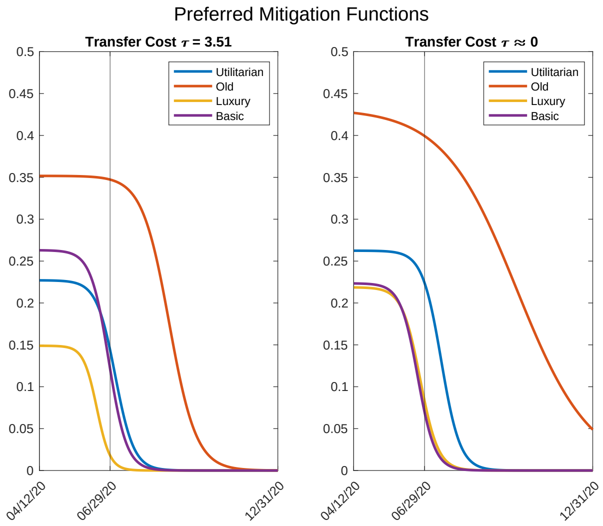


Figure 8: Preferred intensity and duration of lock downs

The left-panel describes optimal policies under our baseline cost for redistribution, with $\tau = 3.51$. The blue line is the policy chosen by a utilitarian planner, who weights expected utility of each type in proportion to their date 0 population shares. The other lines describe the

Covid Economics 6, 17 April 2020: 22-64

policies preferred by each of the three different types (young workers in the basic sector, young workers in the luxury sector, and old individuals, respectively). The right-panel corresponds to a case in which redistribution to soften the economic effects of mitigation is costless ($\tau = 0$).

There are clearly large differences across individual types in terms of what fraction of the economy they would like to see shut down, and for how long. As a point of comparison, recall that up until April 12 the level of mitigation is set at 50 percent of the luxury sector. We first focus on the benchmark calibration with costly transfers (the left panel).

The old (15 percent of the population) would like to see about 35 percent of the luxury sector shut down, and for it to remain shut down for about 110 days, before the shutdown is slowly lifted, and completely wound down by November 2020. In contrast, young luxury workers (close to 50% of the entire population) would prefer a much lower level of mitigation, and for that mitigation to end much earlier. Basic sector workers have a policy preference roughly in the middle of these two extremes, and a utilitarian government adopts a similar policy. Thus a utilitarian government closes about 25 percent of the luxury sector until around the end of June, before gradually opening up over the following month. Note that this policy implies a notably lower level of economic mitigation than the one currently in place. At the same time, some mitigation is certainly called for, and our model suggests it should remain in place for an extended period of time.

When redistribution is costless (right panel of Figure 8), policy preferences remain qualitatively similar, but change quantitatively quite significantly. First, young workers in both sectors now broadly agree on the preferred mitigation policy, which is not surprising since they face identical consumption consequences. Second, the old now prefer even more mitigation, because they do not have to worry about a reduction in relative consumption during a shutdown. The utilitarian policy is more aligned with the preferences of young workers, simply because they constitute the lion's share of the population. Interestingly, the preferred utilitarian mitigation policy is more aggressive when redistribution is (counterfactually) costless, both in terms of level as well as in terms of a longer and more gradual phasing-out. However, even when redistribution is costless, the optimal level of economic mitigation is still below the level we believe to be currently in place.

The next two tables (Tables 5 and 6) describe expected welfare gains, relative to a no economic mitigation baseline, under each of the policies described in Figure 8. The columns of

Covid Economics 6, 17 April 2020: 22-64

Table 5: Welfare Gains (+) or Losses (-) From Preferred Mitigation, $\tau = 3.47$

	Utilitarian	Old	Young Luxury	Young Basic
Young Basic	0.10%	0.03%	0.08%	0.10%
Young Luxury	0.02%	-0.23%	0.04%	0.01%
Old	1.15%	1.54%	0.74%	1.18%

Table 6: Welfare Gains (+) or Losses (-) From Preferred Mitigation, $\tau \approx 0$

	Utilitarian	Old	Young Luxury	Young Basic
Young Basic	0.06%	-0.17%	0.07%	0.07%
Young Luxury	0.07%	-0.17%	0.08%	0.08%
Old	1.66%	2.27%	1.31%	1.26%

each table identify the policy in place. The rows report expected welfare for each type.

Consistent with the results in the previous section, the old experience large welfare gains from any of these policies. Irrespective of the cost of redistribution, the welfare gains or losses for the young are much smaller. Second, and again in line with the previous section, the welfare gains for young luxury workers are always smaller than for young basic workers when redistribution is costly, while the pattern is reversed when redistribution is costless. Third, when redistribution is costly, the policy that is welfare maximizing for the old is actually welfare-reducing (relative to no mitigation) for young luxury workers.

6 Conclusion

We extended a standard epidemiological model of disease progression to include heterogeneity by age, and multiple sources of disease transmission. This disease model was combined with a multi-sector economic model in which workers differ by sector (basic and luxury) as well as by health status. We studied optimal economic mitigation policies and argued that costly redistribution reduces the desire of the government to engage in such policies. Our results also starkly illustrate how unevenly the welfare gains and losses from economic mitigation are likely distributed across different segments of society. The elderly gain much more from extensive reductions in economic activity than the young. Those working in the sector partially shut-

Covid Economics 6, 17 April 2020: 22-64

tered are the most adversely impacted, especially when it is costly to soften the distributional consequences via public transfers.

References

ATKESON, A. (2020): "What Will Be the Economic Impact of COVID-19 in the US? Rough Estimates of Disease Scenarios," Working Paper 26867, National Bureau of Economic Research.

BAYER, C. AND M. KUHN (2020): "Intergenerational ties and case fatality rates: A cross-country analysis," London, Centre for Economic Policy Research DP14519.

BERGER, D., K. HERKENHOFF, AND S. MONGEY (2020): "An SEIR Infectious Disease Model with Testing and Conditional Quarantine," BFI Working Paper NO. 2020-25.

BICK, A. AND A. BLANDIN (2020): "Real Time Labor Market Estimates During the 2020 Coronavirus Outbreak," Unpublished Manuscript, Arizona State University.

EICHENBAUM, M., S. REBELO, AND M. TRABANDT (2020): "The Macroeconomics of Epidemics'. London, Centre for Economic Policy Research," CEPR DP 14520.

FANG, H., L. WANG, AND Y. YANG (2020): "Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-nCoV) in China," Working Paper 26906, National Bureau of Economic Research.

FARIA-E-CASTRO, M. (2020): "Back-of-the-Envelope Estimates of Next Quarter's Unemployment Rate," Saint Louis Fed Blogs:.

FERGUSON, N. M. ET AL. (2020): "Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand," Imperial College COVID-19 Response Team.

FLAXMAN, S., S. MISHRA, A. GANDY, ET AL. (2020): "Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries," Imperial College London (2020) doi.

GREENSTONE, M. AND V. NIGAM (2020): "Does Social Distancing Matter?" (March 30, 2020). University of Chicago, Becker Friedman Institute for Economics Working Paper No. 2020-26. Available at SSRN: https://ssrn.com/abstract=3561244 or http://dx.doi.org/10.2139/ssrn.3561244.

HALL, R. E., C. I. JONES, AND P. J. KLENOW (2020): "Trading Off Consumption and COVID-19 Deaths," Stanford University Working Paper.

Hong, J. H., J. Pijoan-Mas, and J.-V. Rios-Rull (2018): "Health and Inequality," Work in Progress. First Version, 2016.

Krueger, D., H. Uhlig, and T. Xie (2020): "Macroeconomic Dynamics and Reallocation in an Epidemic," Tech. rep., Centre for Economic Policy Research DP 14607.

McAndrew, T. (2020): "COVID19-Survey 8-2020-04-08," Unpublished Paper, University of Massachusetts Amherst.

Moll, B., G. Kaplan, and G. Violante (2020): "Pandemics According to HANK," Virtual Presentation on March 31, 2020.

Mossong, J., N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, et al. (2008): "Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases," *PLoS Med*, 5, https://doi.org/10.1371/journal.pmed.0050074.

Saez, E., J. Slemrod, and S. H. Giertz (2012): "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature*, 50, 3–50.

Toxvaerd, F. (2020): "Equilibrium Social Distancing," Tech. rep., Cambridge-INET Working Paper Series No: 2020/08 Cambridge Working Papers in Economics: 2021.

Tsai, T. C., B. H. Jacobson, and A. K. JhaJack (2020): "American Hospital Capacity And Projected Need for COVID-19 Patient Care," Health Affairs Blog, March 17, 2020. DOI: 10.1377/hblog20200317.457910.

Xu, J., G. Watson, D. Xiong, M. A. Suchard, and C. M. Ramirez (2020): "COVID-19 Infection and Mortality across Workplaces and Intervention Effects," Mimeo, UCLA.

Covid Economics 6, 17 April 2020: 22-64

A The Unrestricted Optimal Policy Problem

The complete characterization of the optimal policy path is the solution to an optimal control problem. In the main text we already have derived the period return function $W(\mathbf{x}, m)$. In addition, the evolution of the state (the distribution of the population by health status $\boldsymbol{x} = (x^{i,j})$) evolves according to the vector-valued equation (summarizing Equations (4) to (22) the paper in a compact form):

$$\dot{\boldsymbol{x}} = G(\boldsymbol{x}, m) \tag{45}$$

To solve for the optimal time path of the scalar mitigation variable is then a straightforward optimal control problem with a multi-dimensional state vector and a one-dimensional control variable. Define the current value Hamiltonian as

$$\mathcal{H}(\boldsymbol{x}, m, \boldsymbol{\mu}) = W(\boldsymbol{x}, m) + \boldsymbol{\mu} G(\boldsymbol{x}, m) \tag{46}$$

where $\boldsymbol{\mu}$ is the vector of co-state variables associated with the population state vector $\boldsymbol{x}$. Necessary conditions at an interior solution for mitigation m are the optimality condition for m

$$\begin{aligned} W_m(\boldsymbol{x}, m) &= -\boldsymbol{\mu} \cdot G_m(\boldsymbol{x}, m) && (47)\\ \dot{\boldsymbol{\mu}} &= \rho\boldsymbol{\mu} - [W_{\boldsymbol{x}}(\boldsymbol{x}, m) + \boldsymbol{\mu} \cdot G_{\boldsymbol{x}}(\boldsymbol{x}, m)] && (48)\\ \dot{\boldsymbol{x}} &= G(\boldsymbol{x}, m) && (49) \end{aligned}$$

The key trade offs with mitigation efforts m discussed in the main text are encoded in equation (47). A marginal increase in m entails static economic costs of $W_m(\boldsymbol{x}, m)$ stemming from the loss of output and thus consumption of all individuals in the economy, as encoded in $y^n(m)$. The dynamic benefit is a better change in the population health distribution, as encoded in the vector $G_m(\boldsymbol{x}, m)$. Concretely, as is clear from equations (4 – 6) an increase in m reduces the outflow of individuals from the susceptible to the asymptomatic state. The value (in units of the objective function) are given by the co-state vector $\boldsymbol{\mu}$.

It should be kept in mind that since $(\boldsymbol{x}, \boldsymbol{\mu})$ are vectors, so are the entities $G_m(\boldsymbol{x}, m) = (G_m^{i,j}(\boldsymbol{x}, m))$ and $W_{\boldsymbol{x}}(\boldsymbol{x}, m) = (W_{x^{i,j}}(\boldsymbol{x}, m))$ and $G_{\boldsymbol{x}}(\boldsymbol{x}, m) = (G_{x^{i,j}}^k(\boldsymbol{x}, m))$ so that equation (47) reads explicitly

$$W_m(\boldsymbol{x}, m) = -\sum_{i,j} \mu^{i,j} G_m^{i,j}(\boldsymbol{x}, m), \tag{50}$$

and a specific row of the vector-valued equation (48) is given by

$$\dot{\mu}^{i,j} = \rho\mu^{i,j} - \left[W_{\boldsymbol{x}^{i,j}}(\boldsymbol{x}, m) + \sum_k \mu^k G_{x^{i,j}}^k(\boldsymbol{x}, m) \right]. \tag{51}$$

CEPR PRESS

Covid Economics **Issue 6, 17 April 2020**

Supply and demand shocks in the COVID-19 pandemic: An industry and occupation perspective[1]

R. Maria del Rio-Chanona,[2] Penny Mealy,[3] Anton Pichler,[4] François Lafond[5] and J. Doyne Farmer[6]

Date submitted: 8 April 2020; **Date accepted:** 10 April 2020

We provide quantitative predictions of first-order supply and demand shocks for the US economy associated with the COVID-19 pandemic at the level of individual occupations and industries. To analyze the supply shock, we classify industries as essential or non-essential and construct a Remote Labor Index, which measures the ability of different occupations to work from home. Demand shocks are based on a study of the likely effect of a severe influenza epidemic developed by the US Congressional Budget Office. Compared to the pre-COVID period, these shocks would threaten around 22% of the US economy's GDP, jeopardise 24% of jobs and reduce total wage income by 17%. At the industry level, sectors such as transport are likely to have output constrained by demand shocks, while sectors relating to manufacturing, mining and services are more likely to be constrained by supply shocks. Entertainment, restaurants and tourism face large supply and demand shocks. At the occupation level, we show that high-wage occupations are relatively immune from adverse supply

1 We would like to thank Eric Beinhocker, Stefania Innocenti, John Muellbauer, Marco Pangallo and David Vines for many comments and discussions. We are also grateful to Andrea Bacilieri and Luca Mungo for their help with the list of essential industries. We thank Baillie Gifford, IARPA, and the Oxford Martin School for the funding that made this possible.

2 Doctoral Student, University of Oxford.
3 Post-doctoral researcher, University of Oxford.
4 Doctoral Student, University of Oxford.
5 Senior Research Associate, University of Oxford.
6 Professor of Mathematics, University of Oxford.

Covid Economics 6, 17 April 2020: 65-103

CEPR PRESS

and demandside shocks, while low-wage occupations are much more vulnerable. We should emphasize that our results are only first-order shocks – we expect them to be substantially amplified by feedback effects in the production network.

Covid Economics 6, 17 April 2020: 65-103

Covid Economics 6, 17 April 2020: 65-103

1 Introduction

The COVID-19 pandemic is having an unprecedented impact on societies around the world. As governments mandate social distancing practices and instruct non-essential businesses to close to slow the spread of the outbreak, there is significant uncertainty about the effect such measures will have on lives and livelihoods. While demand for specific sectors such as healthcare has skyrocketed in recent weeks, other sectors such as air transportation and tourism have seen demand for their services evaporate. At the same time, many sectors are experiencing issues on the supply-side, as governments curtail the activities of non-essential industries and workers are confined to their homes.

Many economists and commentators believe that the economic impact could be dramatic (Baldwin & Weder di Mauro 2020). To give an example based on survey data in an economy under lockdown, the French statistical office estimated on March 26 that the economy is currently at around 65% of its normal level[1]. Bullard (2020) provides an undocumented estimate that around a half of the US economy would be considered either essential, or able to operate without creating risks of diffusing the virus. Inoue & Todo (2020) modeled how shutting down firms in Tokyo would cause a loss of output in other parts of the economy through supply chain linkages, and estimate that after a month, daily output would be 86% lower than pre-shock (i.e. the economy would be operating at only 14% of its capacity!). Using a calibrated extended consumption function, and assuming a labor income shock of 16% and various consumption shocks by expenditure categories, Muellbauer (2020) estimates a fall of quarterly consumption of 20%. Roughly speaking, most of these estimates, like ours, are estimates of instantaneous declines, and would translate to losses of annual GDP if the lockdown lasted for a year.

In a rare study based on aggregating industry-level shocks, the OECD (2020) estimates a drop in immediate GDP of around 25%, which is in line with our results. Another study by Barrot et al. (2020) estimates industry level shocks by considering the list of essential industries, the closure of schools, and an estimate of the ability to work from home (based on ICT use surveys); Using these shocks in a multisector input-output model, they find that six weeks of social distancing would bring GDP down by 5.6%.

In this paper, we aim to provide analytical clarity about the supply and demand shocks caused by public health measures and changes in preferences caused by avoidance of infection. We estimate (i) supply-side reductions due to the closure of non-essential industries and workers not being able to perform their activities at home and (ii) demand-side changes due to peoples' immediate response to the pandemic, such as increased demand for healthcare and reduced demand for goods or services that are likely to place people at risk of infection (e.g. tourism).

It is important to stress that the shocks that we predict here should not be interpreted as the *overall impact* of the COVID-19 pandemic on the economy. Deriving overall impact estimates involves modeling second-order effects, such as the additional reductions in demand as workers who are stood down or laid off experience a reduction in income and additional reductions in supply as potential shortages propagate through supply chains. Further effects, such as cascading firm defaults, which can trigger bank failures and systemic risk in the financial system, could also arise. Understanding these impacts requires a model of the macro-economy and financial sector. We intend to present results from such an economic model in the future, but in the meantime we want to make our estimates of first order impacts available for researchers or governments to build upon or use in their own models.

Several researchers have already provided estimates of the supply shock from labor supply (Dingel & Neiman 2020, Hicks 2020, Koren & Pető 2020). Here we improve on these efforts in three ways: (i) we

[1]https://www.insee.fr/en/statistiques/4473305?sommaire=4473307

Covid Economics 6, 17 April 2020: 65-103

propose a methodology for estimating how much work can be done from home based on work activities, (ii) we identify industries for which working from home is irrelevant because the industries are considered essential, and (iii), we compare our estimated supply shocks to estimates of the demand shock, which in many industries is the more relevant constraint on output.

To see why it is important to compare supply and demand shocks, consider the following thought experiment: Following social distancing measures, suppose industry i is capable of producing only 70% of its pre-crisis output, e.g. because workers can produce only 70% of the output while working from home. If consumers reduce their demand by 90%, the industry will produce only what will be bought, that is, 10%. If instead consumers reduce their demand by 20%, the industry will not be able to satisfy demand but will produce everything it can, that is, 70%. In other words, the experienced first order reduction in output from the immediate shock will be the greater of the supply shock or the demand shock. In other words, most of the first-order impact on the economy will be due to an inability of people to work rather than to consume. However, again, we expect that as wages from work drop, there will be potentially larger second-order negative impacts on demand, and the potential for a self-reinforcing downward spiral in output, employment, income, and demand.

Overall, we find that the supply and demand shocks considered in this paper represent a reduction of around one quarter of the US economy's value added, one fifth of current employment and about 17% of the US total wage income. Supply shocks account for the majority of this reduction. These effects vary substantially across different industries. While we find no negative effects on value added for industries like Legal services, Power generation and distribution or Scientific research, the expected loss of value added reaches up to 80% for Accommodation, Food services and Independent artists.

We show that sectors such as Transports are likely to experience immediate demand-side reductions that are larger than their corresponding supply-side shocks. Other industries such as manufacturing, mining and certain service sectors are likely to experience larger immediate supply-side shocks relative to demand-side shocks. Health unsurprisingly experiences an overall increase in demand for its output. Entertainment, restaurants and hotels experience very large supply and demand shocks, with the demand shock dominating. These results are important because supply and demand shocks might have different degrees of persistence, and industries will react differently to policies depending on the constraints that they face. Overall, however, we find that aggregate effects are dominated by supply shocks, with a large part of manufacturing and services being classified as non-essential while its labor force is unable to work from home.

We also break down our results by occupation and show that there is a strong negative relationship between the overall immediate shock experienced by an occupation and its wage. Relative to the pre-COVID period, 38% of the jobs for workers in the bottom quartile of the wage distribution are predicted to be vulnerable. (And bear in mind that this is only a first-order shock – second order shocks may significantly increase this). In contrast, most high-wage occupations are relatively immune from adverse shocks, with only 6% of the jobs at risk for the 25% of workers working in the highest pay occupations. Absent strong support from governments, most of the economic burden of the pandemic will fall on lower wage workers.

We neglect several effects that, while important, are small compared to those we consider here. First, we have not sought to quantify the reduction in labor supply due to workers contracting COVID-19. A rough estimate suggests that this effect is relatively minor in comparison to the shocks associated with social distancing measures that are being taken in most developed countries.[2] We have also not explicitly

[2] See Appendix D.1 for rough quantitative estimates in support of this argument.

included the effect of school closures. However, in Appendix D.2 we argue that this is not the largest effect and is already partially included in our estimates through indirect channels.

A more serious problem is caused by the need to assume that that within a given occupation, being unable to perform some work activities does not harm the performance of other work activities. Within an industry, we also assume that if workers in a given occupation cannot work, they do not produce output, but this does not prevent other workers in different occupations from producing. In both cases we assume that the effects of labor on production are linear, i.e. that production is proportional to the fraction of workers who can work. In reality however, it is clear that there are important complementarities leading to nonlinear effects. There are many situations where production requires a combination of different occupations, such that if workers in key occupations cannot work at home, production is not possible. For example, while the accountants in a steel plant might be able to work from home, if the steelworkers needed to run the plant cannot come to work, no steel is made. We cannot avoid making linear assumptions because as far as we know there is no detailed understanding of the labor production function and these interdependencies at an industry level. By neglecting nonlinear effects, our work here should consequently be regarded as an approximate lower bound on the size of the first order shocks.

This paper focuses on the United States. We have chosen it as our initial test case because input-output tables are more disaggregated than those of most other countries, and because the O*NET database, which we rely on for information about occupations was developed based on US data. With some additional assumptions it is possible to apply the analysis we perform here to other developed countries.

This paper is structured as follows. In Section 2 we describe our methodology for estimating supply shocks, which involves developing a new Remote-Labor Index (RLI) for occupations and combining it with a list of essential industries. Section 3 discusses likely demand shocks based on estimates developed by the US Congressional Budget Office (2006) to predict the potential economic effects of an influenza pandemic. In Section 4, we show a comparison of the supply and demand shocks across different industries and occupations and identify the extent to which different activities are likely to be constrained by supply or demand. In this section, we also explore which occupations are more exposed to infection and make comparisons to wage and occupation-specific shocks. Finally, in Section 5 we discuss our findings in light of existing research and outline avenues for future work. We also make all of our data available in a continuously updated online repository.

2 Supply shocks

Supply shocks from pandemics are mostly thought of as labor supply shocks. Several pre-COVID-19 studies focused on the direct loss of labor from death and sickness (e.g. McKibbin & Sidorenko (2006), Santos et al. (2013)), although some have also noted the potentially large impact of school closure (Keogh-Brown et al. 2010). McKibbin & Fernando (2020) consider (among other shocks) reduced labor supply due to mortality, morbidity due to infection, and morbidity due to the need to care for affected family members. In countries where social distancing measures are in place, social distancing measures will have a much larger economic effect than the direct effects from mortality and morbidity. This is in part because if social distancing measures work, only a small share of population will be infected and die eventually. Appendix D.1 provides more quantitative estimates of the direct mortality and morbidity effects and argues that the they are likely to be at least an order of magnitude smaller than those due to social distancing measures, especially if the pandemic is contained.

For convenience we neglect mortality and morbidity and assume that the supply shocks are determined

Covid Economics 6, 17 April 2020: 65-103

Covid Economics 6, 17 April 2020: 65-103

only by the amount of labor that is withdrawn due to social distancing. We consider two key factors: (i) the extent to which workers in given occupations can perform their requisite activities at home and (ii) the extent to which workers are likely to be unable to come to work due to being in non-essential industries. We quantify these effects on both industries and occupations. Figure 1 gives a schematic overview of how

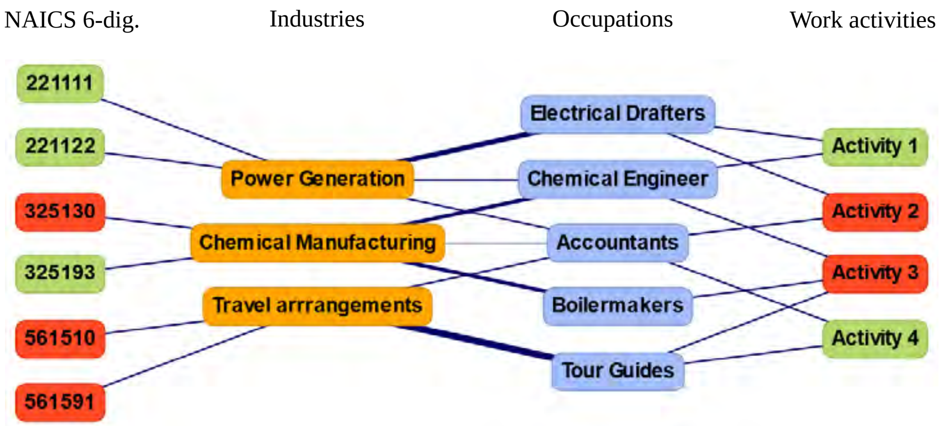


Figure 1: A schematic network representation of supply side shocks. The nodes to the left represent the list of essential industries at the NAICS 6-digit level. A green node indicates essential, a red node non-essential. The orange nodes (center-left) are more aggregate industry categories (e.g. 4-dig. NAICS or the BLS industry categories) for which further economic data is available. These two sets of nodes are connected through industry concordance tables. The blue nodes (center-right) are different occupations. A weighted link connecting an industry category with an occupation represents the number of people of a given occupation employed in each industry. Nodes on the very right are O*NET work activities. Green work activities mean that they can be performed from home, while red means that they cannot. O*NET provides a mapping of work activities to occupations.

2.1 How much work can be performed from home?

One way to assess the degree to which workers are able to work from home during the COVID-19 pandemic is by direct survey. For example, Zhang et al. (2020) conducted a survey of Chinese citizens in late February (one month into the coronavirus-induced lockdown in China) and found that 27% of the labor force continued working at the office, 38% worked from home, and 25% stopped working. Adams-Prassl et al. (2020) surveyed US and UK citizens in late March, and reported that the share of tasks that can be performed from home varies widely between occupations (from around 20 to 70%), and that higher wage occupations tend to be more able to work from home.

Other recent work has instead drawn on occupation-level data from the Occupational Information Network (O*NET) to assess 'work context' characteristics to determine labour shocks due to the COVID-19 pandemic. For example, Hicks (2020) considered the degree to which an occupation is required to 'work with others' or involves 'physical proximity to others' in order to assess which occupations are likely to be most impacted by social distancing. Similarly, to identify which occupations were likely to be able to work from home, Dingel & Neiman (2020) used O*NET data on the extent to which occupations requires daily 'work outdoors' or needed to engage with 'operating vehicles, mechanized devices or equipment'.

CEPR PRESS

We go to a more granular level than 'work context', and instead draw on O*NET's 'intermediate work activity' data, which provides a list of the activities performed by each occupation based on a list of 332 possible work activities. For example, a Nurse undertakes activities such as "maintain health or medical records", "develop patient or client care or treatment plans" and "operate medical equipment', while a Computer Programmer performs activities such as "resolve computer programs", "program computer systems or production equipment" and "document technical designs, producers or activities"[3]. In Fig 1 these work activities are illustrated by the rightmost set of nodes.

Which work activities can be performed from home? Four of us independently assigned a subjective binary rating to each work activity as to whether it could successfully be performed at home. The individual results were in broad agreement. Based on the responses, we assigned an overall consensus rating to each work activity[4]. Ratings for each work activity are available in an online data repository[5]. While O*NET maps each intermediate work activity to 6-digit O*NET occupation codes, employment information from the US Bureau of Labor Statistics (BLS) is available for the 4-digit 2010 Standard Occupation Scheme (SOC) codes, so we mapped O*NET and SOC codes using a crosswalk available from O*NET.[6] Our final sample contains 740 occupations.

[3]In the future we intend to redo this using O*NET's "detailed" work activity data, which involves over 2000 individual activities associated with different occupations. We believe this would somewhat improve our analysis, but we also think that for our purposes here the intermediate activity list provides a good approximation.

[4]An activity was considered to be able to performed at home if three or more respondents rated this as true. We also undertook a robustness analysis where an activity was considered to able to be performed at home based on two or more true ratings. Results remained fairly similar. In post-survey discussion, we agreed that the most contentious point is that some work activities might be done from home or not, depending on the industry in which it is performed.

[5]https://zenodo.org/record/3751068

[6]Available at https://www.onetcenter.org/crosswalks.html.

Covid Economics 6, 17 April 2020: 65-103

Occupation	RLI
Credit Analysts	1.00
Insurance Underwriters	1.00
Tax Preparers	1.00
Mathematical Technicians	1.00
Political Scientists	1.00
Broadcast News Analysts	1.00
Operations Research Analysts	0.92
Eligibility Interviewers, Government Programs	0.92
Social Scientists and Related Workers, All Other	0.92
Technical Writers	0.91
Market Research Analysts and Marketing Specialists	0.90
Editors	0.90
Business Teachers, Postsecondary	0.89
Management Analysts	0.89
Marketing Managers	0.88
Mathematicians	0.88
Astronomers	0.88
Interpreters and Translators	0.88
Mechanical Drafters	0.86
Forestry and Conservation Science Teachers, Postsecondary	0.86
...	...
Bus and Truck Mechanics and Diesel Engine Specialists	0.00
Rail Car Repairers	0.00
Refractory Materials Repairers, Except Brickmasons	0.00
Musical Instrument Repairers and Tuners	0.00
Wind Turbine Service Technicians	0.00
Locksmiths and Safe Repairers	0.00
Signal and Track Switch Repairers	0.00
Meat, Poultry, and Fish Cutters and Trimmers	0.00
Pourers and Casters, Metal	0.00
Foundry Mold and Coremakers	0.00
Extruding and Forming Machine Setters, Operators, and Tenders, Synthetic and Glass Fibers	0.00
Packaging and Filling Machine Operators and Tenders	0.00
Cleaning, Washing, and Metal Pickling Equipment Operators and Tenders	0.00
Cooling and Freezing Equipment Operators and Tenders	0.00
Paper Goods Machine Setters, Operators, and Tenders	0.00
Tire Builders	0.00
Helpers–Production Workers	0.00
Production Workers, All Other	0.00
Machine Feeders and Offbearers	0.00
Packers and Packagers, Hand	0.00

Table 1: **Top and bottom 20 occupations ranked by Remote Labor Index (RLI)**, based on proportion of work activities that can to be performed by home. There are 44 occupations with an RLI of zero; we show only a random sample.

From work activities to occupations. We then created a Remote Labor Index (RLI) for each occupation by calculating the proportion of an occupation's work activities that can be performed at home. An RLI of 1 would indicate that all of the activities associated with an occupation could be undertaken at home, while an RLI of 0 would indicate that none of the occupation's activities could be performed at

Covid Economics 6, 17 April 2020: 65-103

home.[7] The resulting ranking of each of the 740 occupations can be found in the online repository (see footnote 5). While the results are not perfect[8], most of the rankings make sense. For example, in Table 1, we show the top 20 occupations having the highest RLI ranking. Some occupations such as credit analysts, tax preparers and mathematical technician occupations are estimated to be able to perform 100% of their work activities from home. Table 1 also shows a sample of the 43 occupations with an RLI ranking of zero, i.e. those for which there are no activities that are able to be performed at home.

To provide a broader perspective of how the RLI differs across occupation categories, Figure 2 shows a series of box-plots indicating the distribution of RLI for each 4-digit occupation in each 2-digit SOC occupation category. We have ordered 2-digit SOC occupations in accordance with their median values. Occupations with the highest RLI relate to Education, training and library, Computer and Mathematical, and Business and Financial roles, while occupations relating to Production, Farming, Fishing and Forestry, and Construction and Extraction tend to have lower RLI.

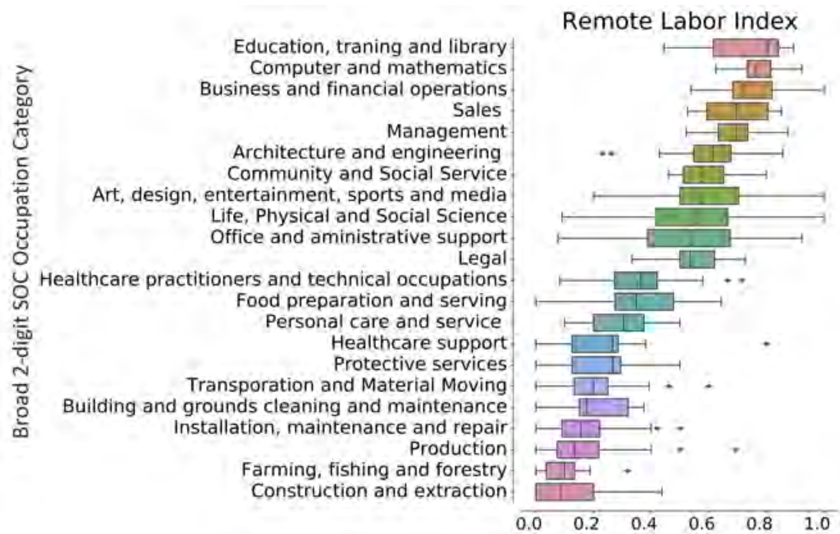


Figure 2: **Distribution of Remote Labor Index across occupations**. We provide boxplots showing distribution of RLI for each 4-digit occupation in each 2-digit SOC occupation category.

From occupations to industries. We next map the RLI to industry categories to quantify industry-specific supply shocks from social distancing measures. We obtain occupational compositions per industry

[7] We omitted ten occupations that had less than five work activities associated with them. These occupations include Insurance Appraisers Auto Damage; Animal Scientists; Court Reporters; Title Examiners, Abstractors, and Searchers; Athletes and Sports Competitors; Shampooers; Models; Fabric Menders, Except Garment; Slaughterers and Meat Packers' and Dredge Operators.

[8] There are a few cases that we believe are misclassified. For example, two occupations with a high RLI that we think cannot be performed remotely are real estate agents (RLI = 0.7) and retail salespersons (RLI = 0.63). However, these are exceptions – in most cases the rankings make sense. The full list can be examined on our online repository at https://zenodo.org/record/3751068. We believe the problems will be fixed when we redo the analysis using fine grained work activities, and we doubt that our results will be qualitatively changed.

Covid Economics 6, 17 April 2020: 65-103

from the BLS, which allows us to match 740 occupations to 277 industries[9].

In Figure 3, we show the RLI distribution for each 4-digit occupation category falling within each broad 2-digit NAICS category. Similar to Figure 2, we have ordered the 2-digit NAICS industry categories in accordance with the median values of each underpinning distribution. As there is a greater variety of different types of occupations within these broader industry categories, distributions tend to be much wider. Industries with the highest median RLI values relate to Information, Finance and Insurance, and Professional, Science and Technical Services, while industries with the lowest median RLI relate to Agriculture, Forestry, Fishing and Hunting and Accommodation and Food Services.

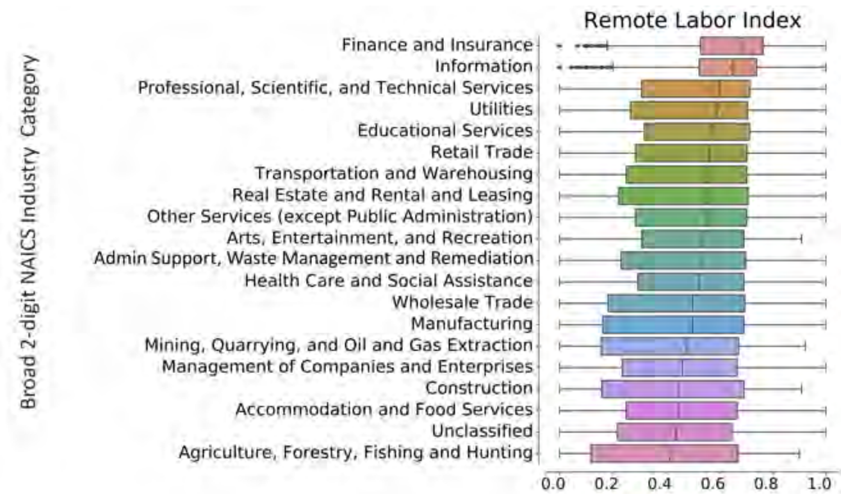


Figure 3: **Distribution of Remote Labor Index across industries**. We provide boxplots showing distribution of RLI for each 4-digit occupation in each 2-digit NAICS Industry category.

In the Appendix, we show industry-specific RLI values for the more detailed 4-digit NAICS industries. To arrive at a single number for each 4-digit industry, we compute the employment-weighted average of occupation-specific RLIs. The resulting industry-specific RLI can be interpreted as a rough estimate of the fraction of jobs which can be performed from home for each industry.

2.2 Which industries are "essential"?

Across the world, many governments have mandated that certain industries deemed 'essential' should remain open over the COVID-19 crisis duration. What constitutes an 'essential' industry has been the subject of significant debate, and it is likely that the endorsed set of essential industries will vary across countries. As the U.S. government has not produced a definitive list, here we draw on the list of essential industries developed by Italy and assume it can be applied, at least as an approximation, to other countries such as the U.S. as well. This list has two key advantages. First, as Italy was one

[9] We use the May 2018 Occupational Employment Statistics (OES) estimates on the level of 4-digit NAICS (North American Industry Classification System), file *nat4d_M2018_dl*, which is available at https://www.bls.gov/oes/tables.htm under *All Data*. Overall, our merged dataset covers 136.8 out of 144 million employed people (95%) initially reported in the OES.

Covid Economics 6, 17 April 2020: 65-103

of the countries affected earliest and most severely, it was one of the first countries to invest significant effort considering which industries should be deemed essential. Second, Italy's list of essential industries includes NACE industrial classification codes, which can be mapped to the NAICS industry classification we use to classify industrial employment in this paper.[10]

Table 2 shows the total numbers of NAICS essential industries at the 6-digit and 4-digit level. More than 50% of 6-digit NAICS industries are considered essential. At the 6-digit level the industries are either classified as essential, and assigned essential score 1, or non-essential and assigned essential score 0. Unfortunately, it is not possible to translate this directly into a labour force proportion as BLS employment data at detailed occupation and industry levels are only available at the NAICS 4-digit level. To derive an estimate at the 4-digit level, we assume that labor in a NAICS 4-digit code is uniformly distributed over its associated 6-digit codes. We then assign an essential 'share' to each 4-digit NAICS industry based on the proportion of its 6-digit NAICS industries that are considered essential. (The distribution of the essential share over 4-digit NAICS industries is shown in the Appendix). Based on this analysis, we estimate that about 89 million (or 64%) of US workers are currently employed in essential industries.

Total 6-digit NAICS industries	1057
Number of essential 6-digit NAICS industries	612
Fraction of essential industries at 6-digit NAICS	0.58
Total 4-digit NAICS industries in our sample	277
Average rating of essential industries at 4-digit NAICS	0.56
Fraction of labor force in essential industries	0.68

Table 2: **Essential industries.** Essential industries at the 6-digit level and essential 'share' at the 4-digit level. Note that 6-digit NAICS industry classifications are binary (0 or 1) whereas 4-digit NAICS industry classifications can take on any value between 0 and 1.

2.3 Supply shock: non-essential industries unable to work from home

Having analyzed both the extent to which jobs in each industry are essential and the likelihood that workers in a given occupation can perform their requisite activities at home, we now combine these to consider the overall first-order effect on labor supply in the US. In Figure 4, we plot the Remote Labor Index of each occupation against the fraction of that occupation employed in an essential industry. Each circle in the scatter plot represents an occupation; the circles are sized proportional to current employment and color coded according to the median wage in each occupation.

Figure 4 indicates the vulnerability of occupations due to supply-side shocks. Occupations in the lower left-hand side of the plot (such as Diswashers, Rock Splitters and Logging Equipment Operators) have lower RLI scores (indicating they are less able to work from home) and are less likely to be employed in an essential industry. If we consider only the immediate supply-side effects of social distancing, workers in these occupations are more likely to face reduced work hours or be at risk of losing their job altogether. In contrast, occupations on the upper right-hand side of the plot, such as Credit Analysis, Political Scientists

[10]Mapping NACE industries to NAICS industries is not straightforward. NACE industry codes at the 4-digit level are internationally defined. However, 6-digit level NACE codes are country specific. Moreover, the list of essential industries developed by Italy involves industries defined by varying levels of aggregation. Most essential industries are defined at the NACE 2-digit and 4-digit level, with a few 6-digit categories thrown in for good measure. As such, much of our industrial mapping methodology involved mapping from one classification to the other by hand. We provide a detailed description of this process in Appendix B.1.

Covid Economics 6, 17 April 2020: 65-103

Covid Economics 6, 17 April 2020: 65-103

and Operations Research Analysts) have higher RLI scores and are more likely to employed in an essential industry. These occupations are less economically vulnerable to the supply-side shocks (though, as we discuss in the next section, they could still face employment risks due to first-order demand-side effects). Occupations in the upper-left hand side of the plot (such as Farmworkers, Healthcare Support Workers and Respiratory Therapists) are less likely to be able to perform their job at home, but since they are more likely to be employed in an essential industry their economic vulnerability from supply-side shocks is lower. Interestingly, there are relatively few occupations on the lower-right hand side of the plot. This indicates that occupations that are predominantly employed in non-essential industries tend to be less able to perform their activities at home.

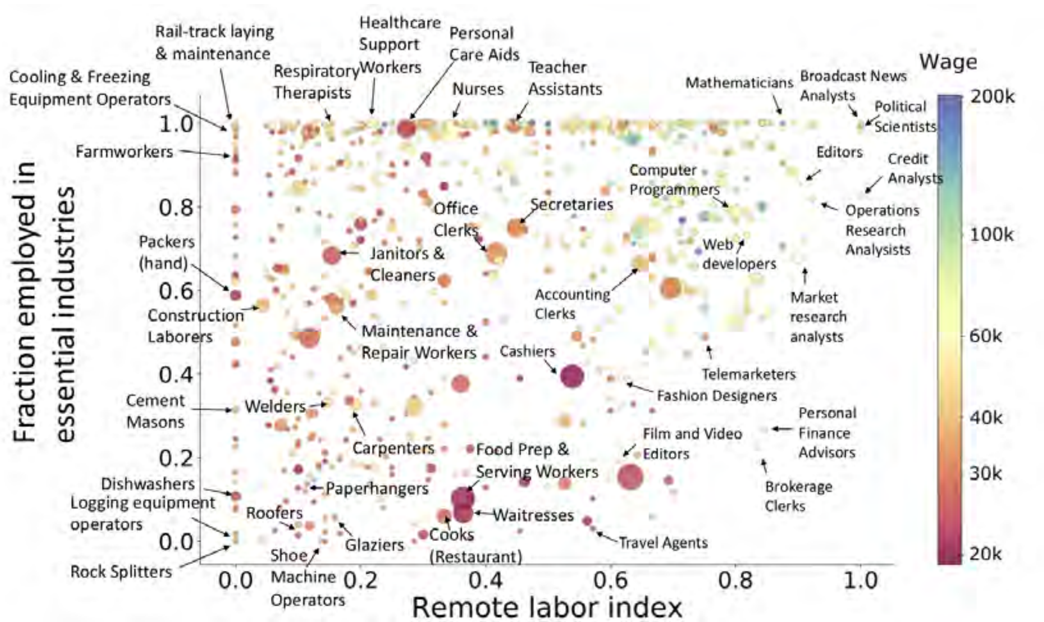


Figure 4: **Fraction employed in an essential industry vs Remote Labor Index for each occupation**. Omitting the effect of demand reduction, the occupations in the lower left corner, with a small proportion of workers in essential industries and a low Remote Labor Index, are the most vulnerable to loss of employment due to social distancing.

CEPR PRESS

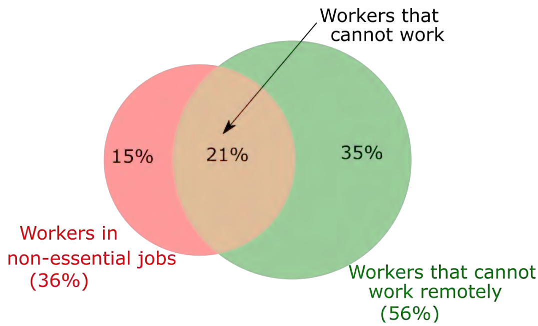


Figure 5: **Workers that cannot work**. On the left is the percentage of workers in a non-essential job (36% in total). On the right is the percentage of workers that cannot work remotely (56% in total). The intersection is the set of workers that cannot work, which is 19% of all workers. A remaining 29% of workers are in essential jobs where they can work remotely.

Covid Economics 6, 17 April 2020: 65-103

To help visualize the problem we provide a summary in the form of a Venn diagram in Figure 5. Before the pandemic, 36% of workers were employed in non-essential jobs. 56% of workers cannot do their job remotely. 21% of workers are in the intersection corresponding to non-essential jobs that cannot be performed remotely. In addition, there are 29% of workers in essential industries that can also work from home.[11]

3 Demand shock

The pre-COVID-19 literature on epidemics and the discussions of the current crisis make it clear that epidemics strongly influence patterns of consumer spending. In addition to increasing their demand for health services, consumers are also likely to seek to reduce their risk of exposure to the virus and decrease demand for products and services that involve close contact with others. In the early days of the outbreak, stockpiling behaviour also drives a direct demand increase in the retail sector (Baker et al. 2020).

Estimates from the CBO. Our estimates of the demand shock are based on expert estimates developed by the US Congressional Budget Office (2006) that attempted to predict the potential impact of an influenza pandemic. Similar to the current COVID-19 pandemic, this analysis assumes that demand is reduced due to the desire to avoid infection. While the analysis is highly relevant to the present COVID-19 situation, it is important to note that the estimates are "extremely rough" and "based loosely on Hong Kong's experience with SARS". The CBO provides two scenarios. We take the severe scenario, which "describes a pandemic that is similar to the 1918-1919 Spanish flu outbreak. It incorporates the assumption that a particularly virulent strain of influenza infects roughly 90 million people in the United States and kills more than 2 million of them".

In this paper, we simply take the CBO estimates as immediate (first-order) demand-side shocks. The CBO lists demand side estimates for broad industry categories, which we mapped to the 2-digit NAICS codes by hand. Table 3 shows the CBO's estimates of the percent decrease in demand by industry, and Table 8 in Appendix E shows the full mapping to 2-digit NAICS.

These estimates, of course, are far from perfect. They are based on expert estimates made more than

[11] In fact we allow for a continuum between the ability to work from home, and an industry can be partially essential.

Broad industry name	Severe Scenario Shock
Agriculture	-10
Mining	-10
Utilities	0
Construction	-10
Manufacturing	-10
Wholesale trade	-10
Retail trade	-10
Transportation and warehousing (including air, rail and transit)	-67
Information (Published, broadcast)	0
Finance	0
Professional and business services	0
Education	0
Healthcare	15
Arts and recreation	-80
Accommodation/food service	-80
Other services except government	-5
Government	0

Table 3: Demand shock by sector according to the Congressional Budget Office (2006)'s severe scenario.

Covid Economics 6, 17 April 2020: 65-103

ten years ago for an hypothetical pandemic scenario. It is not entirely clear if they are for gross output or for final (consumer) demand. However, in Appendix E, we describe three other sources of consumption shocks (Keogh-Brown et al. 2010, Muellbauer 2020, OECD 2020) that provide broadly similar estimates by industry or spending category.

Transitory and permanent shocks. An important question is whether demand reductions are just postponed expenses, and if they are permanent (Mann 2020, Keogh-Brown et al. 2010). Baldwin & Weder di Mauro (2020) also distinguishes between "practical" (the impossibility to shop) and "psychological" demand shocks (the wait-and see attitude adopted by consumers facing strong uncertainty.) We see three possibilities: (i) expenses in a specific good or service are just delayed but will take place later, for instance if households do not go to the restaurant this quarter, but go twice as much as they would normally during the next quarter; (ii) expenses are not incurred this quarter, but will come back to their normal level after the crisis, meaning that restaurants will have a one-quarter loss of sales; and (iii) expenses decrease to a permanently lower level, as household change their preferences in view of the 'new normal'. Appendix E reproduces the scenario adopted by Keogh-Brown et al. (2010), which distinguishes between delay and permanently lost expenses.

Other components of aggregate demand. We do not include direct shocks to investment, net exports, and net inventories. Investment is typically very pro-cyclical and is likely to be strongly affected, with direct factors including cash-flow reductions and high uncertainty (Boone 2020). The impact on trade is likely to be strong and possibly permanent (Baldwin & Weder di Mauro 2020), but would affect exports and imports in a relatively similar way, so the overall effect on net exports is unclear. Finally, it is likely that due to the disruption of supply chains, inventories will be run down so the change in inventories will be negative (Boone 2020).

CEPR PRESS

4 Combining supply and demand shocks

Having described both supply and demand-side shocks, we now compare the two at the industry and occupation level.

4.1 Industry-level supply and demand shocks.

Figure 6 plots the demand shock against the supply shock for each industry. The radius of the circles is proportional to the gross output of the industry[12].

Essential industries have no supply shock and so lie on the horizontal '0' line. Of these industries, sectors such as Utilities and Government experience no demand shock either, since immediate demand for their output is assumed to remain the same. Health experiences an increase in demand and consequently lies below the identity line. Transport on the other hand experiences a reduction in demand and lies well above the identity line. This reflects the current situation, where trains and buses are running because they are deemed essential, but they are mostly empty[13]. Non-essential industries such as Entertainment, Restaurants and Hotels, experience both a demand reduction (due to consumers seeking to avoid infection) and a supply reduction (as many workers are unable to perform their activities at home). Since the demand shock is bigger than the supply shock they lie above the identity line. Other non-essential industries such as manufacturing, mining and retail have supply shocks that are larger than their demand shocks and consequently lie below the identity line.

[12]Since relevant economic variables such as total output per industry are not extensively available on the NAICS 4-digit level, we need to further aggregate the data. We derive industry-specific total output and value added for the year 2018 from the BLS input-output accounts, allowing us to distinguish 169 private industries for we which we can also match the relevant occupation data. The data can be downloaded from https://www.bls.gov/emp/data/input-output-matrix.htm.

[13]Some transport companies such as Transport for London have reduced their traffic, but this is a second-order impact, not a first-order shock. This decrease in output resulting from reduced demand should be modelled as an endogenous effect, not a first-order shock.

Covid Economics 6, 17 April 2020: 65-103

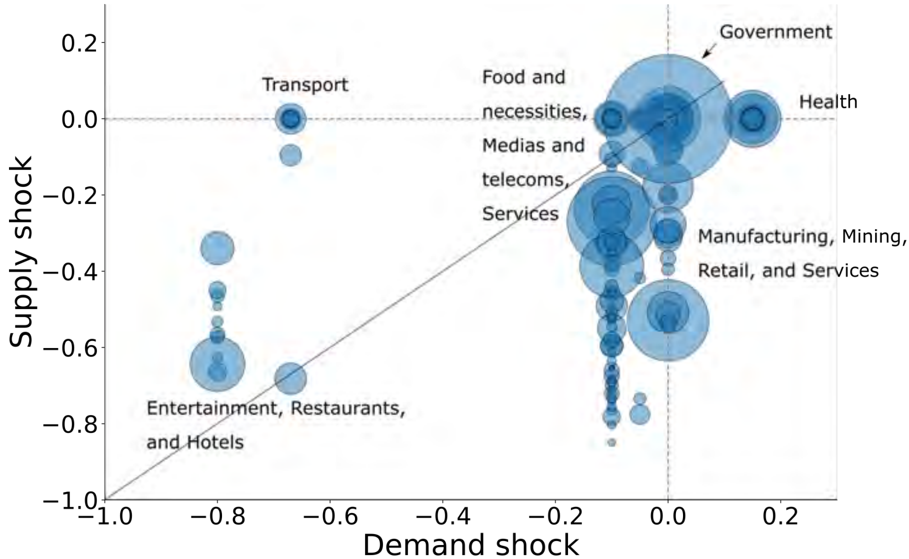


Figure 6: **Supply and demand shocks for industries**. Each circle is an industry, with radius proportional to gross output. Many industries experience exactly the same shock, hence the superposition of some of the circles. Labels correspond to broad classifications of industries.

4.2 Occupation-level supply and demand shocks.

In Figure 7 we show the supply and demand shocks for occupations rather than industries. For each occupation this comparison indicates whether it faces a risk of unemployment due the lack of demand or a lack of supply in its industry.

Several health-related occupations such as Nurses, Medical Equipment Preparers and Healthcare Social Workers are employed in industries experiencing increased demand. Occupations such as Airline Pilots, Lodging Managers and Hotel Desk Clerks face relatively mild supply shocks and strong demand shocks (as consumers reduce their demand for travel and hotel accommodation) and consequently lie above the identity line. Other occupations such as Stonemasons, Rock Splitters, Roofers and Floor Layers face a much stronger supply shock as it is very difficult for these workers to perform their job at home. Finally occupations such as Cooks, Dishwashers and Waiters suffer both adverse demand shocks (since demand for restaurants is reduced) and supply shocks (since they cannot work from home and tend not to work in essential industries).

For the majority of occupations the supply shock is larger than the demand shock. This is not surprising given that we only consider immediate shocks and no feedback-loops in the economy. We expect that once second-order effects are considered the demand shocks are likely to be much larger.

Covid Economics 6, 17 April 2020: 65-103

Covid Economics 6, 17 April 2020: 65-103

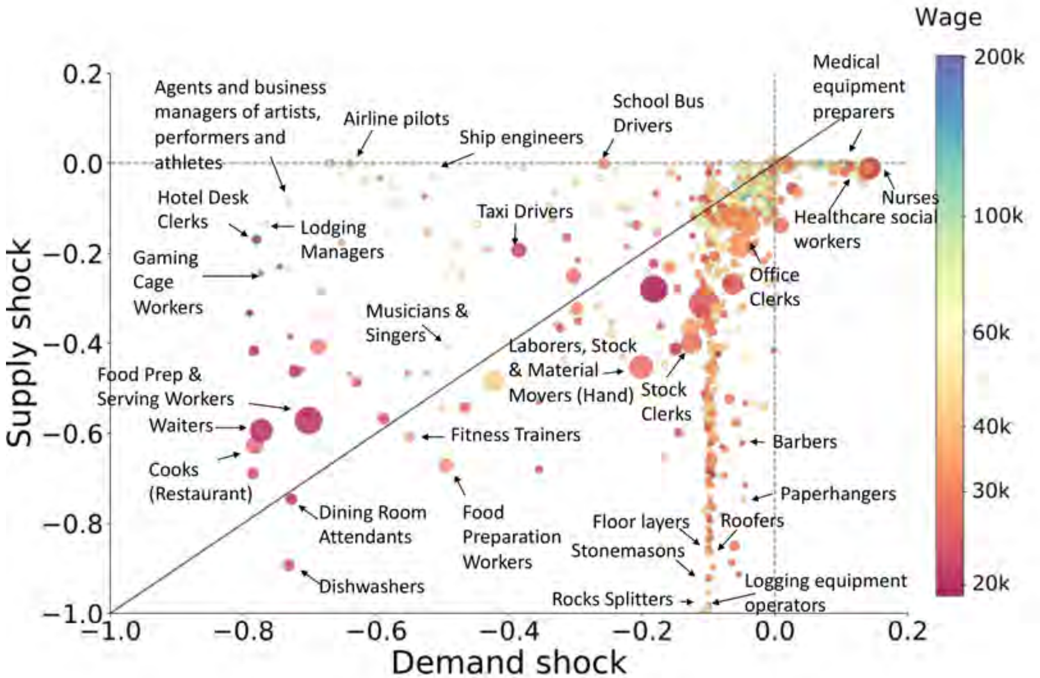


Figure 7: **Supply and demand shocks for occupations.** Each circle is an occupation with radius proportional to employment. Circles are color coded by the log median wage of the occupation. The correlation between wages and demand shocks is 0.26 (p-value = 2.4×10^{-13}) and between wages and supply shocks is 0.40 (p-value = 3.9×10^{-29}).

4.3 Aggregate shocks

We now aggregate shocks to obtain estimates for the whole economy. We assume that, in a given industry, the total shock will be the worse of the supply or demand shocks. The shock on occupations depends on the prevalence of each occupation in each industry (see Appendix A for details). We then aggregate shocks in three different ways.

First we estimate the decline in employment by weighting occupation-level shocks by the number of workers in each occupation. Second, we estimate the decline in total wages paid by weighting occupation-level shocks by the share of occupations in the total wage bill. Finally, we estimate the decline in GDP by weighting industry level shocks by the share of industries in GDP.

Table 4 shows the results. In all cases, by definition, the total shock is larger than both the supply and demand shock, but smaller than the sum. Overall, the supply shock appears to contribute more to the total shock than does the demand shock.

The wage shock is around 17% and is lower than the employment shock (24%). This makes sense, and reflects a fact already well acknowledged in the literature (Office for National Statistics 2020, Adams-

Prassl et al. 2020) that occupations that are most affected tend to have lower wages. We discuss this more below.

The shock on value added (22%) is higher than the shock to the wage bill. This also makes sense; because value added is the sum of the wage bill and payment to capital, this suggests that capital intensive industries are hit harder.

Aggregate Shock	Employment	Wages	Value Added
Supply	-21	-15	-18
Demand	-13	-8	-8
Total	-24	-17	-22

Table 4: **Aggregate shocks to employment, wages and value added**. The size of each shock is shown as a percentage of the pre-pandemic value. Demand shocks include positive values for the health sector. The total shock at the industry level is the minimum of the supply and demand shock, see Appendix A. Note that these are only first order shocks (not total impact), and instantaneous values (not annualized).

For industries and occupations in the health sector, which have experienced an increase in demand, there is no corresponding increase in supply. Table 6 in Appendix A.7 provides the same estimates as Table 4, but now assuming that the increased demand for health will be matched by increased supply. This corresponds to a scenario where the healthcare sector would be immediately able to hire as many workers as necessary and pay them at the normal rate. This assumption does not, however, make a significant difference to the aggregate total shock. In other words, the increase in activity in the health sector is unlikely to be large enough to compensate significantly for the losses from other sectors.

4.4 Shocks by wage level

To understand how the pandemic has affected workers of different income levels differently, we present results for each wage quartile. The results are in Table 4, columns $q_1 \dots q_4$.[14] In Table 5 we show employment shocks by wage quartile. This table shows that workers whose wages are in the lowest quartile (lowest 25%) will bear much higher relative losses than workers whose wages are in the highest quartile. Our results confirm the survey evidence reported by the Office for National Statistics (2020) and Adams-Prassl et al. (2020), showing that low-wage workers are more strongly affected by the COVID crisis in terms of lost employment and lost income. Furthermore, Table 5 shows how the total loss of wages in the economy in split amongst the different quartiles. Even though the lowest quartile have lower salaries, the shock is so high that they bear the highest share of the total loss.

	q_1	q_2	q_3	q_4	Aggregate
Percentage change in employment	-42	-24	-21	-7	-24
Share of total lost wages (%)	30	23	29	18	-17

Table 5: **Total Wages or employment shocks by wage quartile**. We divide workers into wage quartiles based on the average wage of their occupation (q_1 corresponds to the 25% least paid workers). The first row is the number of workers who are vulnerable due to the shock in each quartile divided by the total who are vulnerable. Similarly, the second row is the fraction of whole economy total wages loss that would be lost by vulnerable workers in each quartile. The last column gives the aggregate shocks from Table 4.

[14] As before, Table 6 in Appendix A.7 gives the results assuming positive total shocks for the health sector, but shows that it makes very little difference.

Covid Economics 6, 17 April 2020: 65-103

Next we estimate labor shocks at the occupation level. We define the labor shocks as the declines in employment due to the total shocks in the industries associated with each occupation. We use Eq. (14) (Appendix A.7) to compute the labor shocks, which allows for positive shocks in healthcare workers, to suggest an interpretation in terms of a change in labor demand. Figure 8 plots the relationship between labor shocks and median wage. A strong positive correlation (Pearson $\rho = 0.40$, p-value $= 4.4 \times 10^{-30}$) is clearly evident, with almost no high wage occupations facing a serious shock.

We have also colored occupations by their exposure to disease and infection using an index developed by O*NET[15] (for brevity we refer to this index as "exposure to infection"). As most occupations facing a positive labor shock relate to healthcare[16], it is not surprising to see that they have a much higher risk of being exposed to disease and infection. However, other occupations such as janitors, cleaners, maids and childcare workers also face higher risk of infection. Appendix C explores the relationship between exposure to infection and wage in more detail.

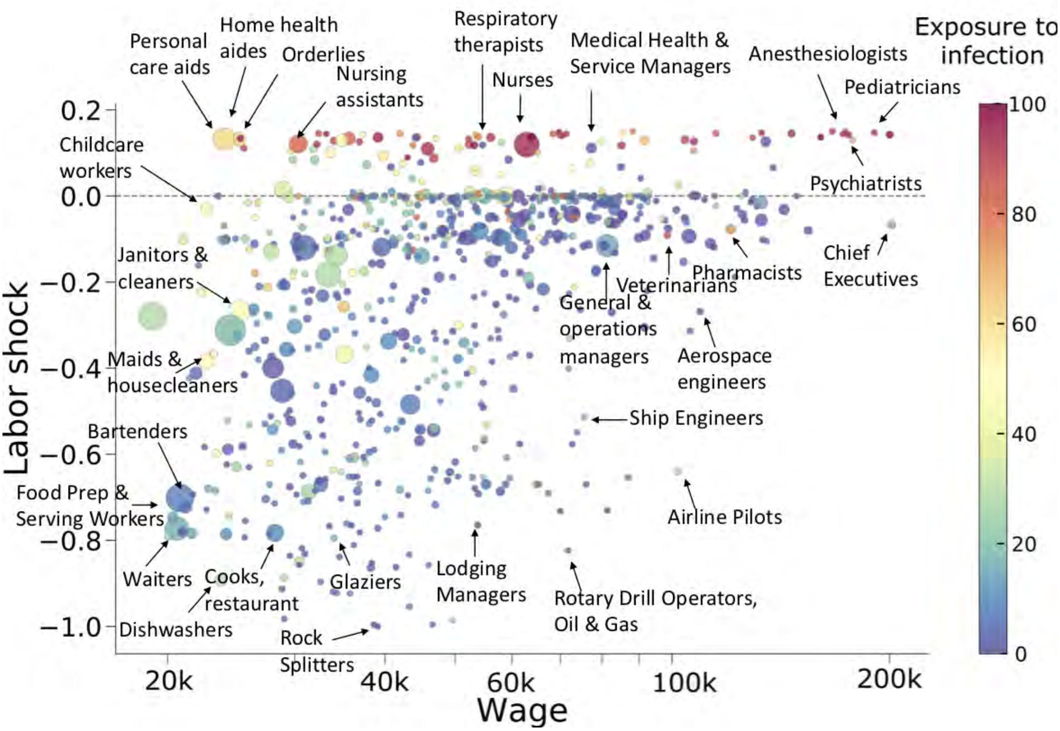


Figure 8: **Labor shock vs. median wage for different occupations.** We color occupations by their exposure to disease and infection. There is a 0.40 correlation between wages and the labor shock (p-value $= 4.4 \times 10^{-40}$). Note the striking lack of high wage occupations with large labor demand shocks.

15 https://www.onetonline.org/find/descriptor/result/4.C.2.c.1.b

16 Our demand shocks do not have an increase in retail, but in the UK supermarkets have been trying to hire several tens of thousands workers (Source: BBC, 21 March, https://www.bbc.co.uk/news/business-51976075). Baker et al. (2020) documents stock piling behavior in the US.

Covid Economics 6, 17 April 2020: 65-103

5 Conclusion

This paper has sought to provide quantitative predictions for the U.S. economy of the supply and demand shocks associated with the COVID-19 pandemic. To characterize supply shocks, we developed a Remote Labor Index to estimate the extent to which workers can perform activities associated with their occupation at home and identified which industries are classified as essential vs. non-essential. We also reported plausible estimates of the demand shocks, in an attempt to acknowledge that some industries will have an immediate reduction in output due to a shortfall in demand, rather due to an impossibility to work. We would like to emphasize that these are *predictions*, not measurements: The estimates of the demand shocks were made in 2006, and the RLI and the list of non-essential industries contain no pandemic-specific information, and could have been made at any time. Putting these predictions together, we estimate that the first order aggregate shock to the economy represents a reduction of roughly a quarter of the economy.

This is the first study seeking to compare supply-side shocks with corresponding demand-side shocks at the occupation and industry level. At the time of writing, the most relevant demand-side estimates available are highly admittedly 'rough' and only available for very aggregate (2-digit) industries. Yet, this suggests that sectors such as Transport are more likely to have output constrained by demand-side shocks, while sectors relating to manufacturing, mining and services are more likely to be constrained by supply-side shocks. Entertainment, restaurants and tourism face both very large supply and demand constraints, with demand shocks dominating in our estimates. By quantifying supply and demand shocks by industry, our paper speaks to the debate on the possibility of inflation after the crisis. Goodhart & Pradhan (2020) argue that the lockdown causes a massive supply shock that will lead to inflation when demand comes back after the crisis. But as Miles & Scott (2020) note, in many sectors it is not obvious that demand will come back immediately after the crisis, and if a gradual reopening of the economy takes place, it may be that supply and demand rise slowly together. However, our paper is the first to raise the fact that because supply and demand shocks are so different by sectors, even a gradual reopening may leave important supply-demand imbalances within industries. Such mismatches could consequently lead to an unusual level of heterogeneity in the inflation for different goods.

When considering total shocks at the occupation level, we find that high-wage occupations are relatively immune from both supply and demand-side shocks, while many low-wage occupations are much more economically vulnerable to both. Interestingly, low-wage occupations that are *not* vulnerable to supply and/or demand-side shocks are nonetheless at higher risk of being exposed to coronavirus (see color code in Figure 8). Such findings suggest that the COVID-19 pandemic is likely to exacerbate income inequality in what is already a highly unequal society.

For policymakers there are three key implications from this study. First, the magnitude of the shocks being experienced by the U.S. economy is very large, with around a quarter of the economy not functioning. As Table 4 shows, even including positive shocks, our estimates of the potential impacts are a drop in employment of 24%, a decline in wages of 17%, and loss in value added of 22%. Bearing in mind the caveats about shocks vs. total impacts, the potential impacts are a multiple of what was experienced during the global financial crisis (e.g. where employment dropped 3.28 percentage points)[17] and comparable only to the Great Depression (e.g. where employment dropped 21.7% 1929-32 (Wallis 1989, Table 2)). Second, as the largest shocks are from the supply-side, strategies for returning people to work as quickly as possible without endangering public health must be a priority. Virus mitigation and containment are

[17]Employment Rate, aged 15-64, all persons for the US (FRED LREM64TTUSM156N) fell from 71.51 in December 2007 to 68.23 in June 2009, the employment peak to trough during the dates of recession as defined by the NBER.

Covid Economics 6, 17 April 2020: 65-103

clearly essential first steps, but strategies such as widespread antibody testing to identify people who are safe to return to work, and rapid testing, tracing, and isolation to minimize future lock-downs, will also be vital until if and when a vaccine is available. Furthermore, aggressive fiscal and monetary policies to minimize first-order shocks cascading into second-order shocks are essential, in particular policies to keep workers in employment and maintain incomes (e.g. the "paycheck protection" schemes announced by several countries), as well as policies to preserve business and financial solvency. Third, and finally, the inequalities highlighted by this study will also require policy responses. Again, higher income knowledge and service workers will likely see relatively little impact, while lower income workers will bear the brunt of the employment, income, and health impacts. In order to ensure that burdens from the crisis are shared as fairly as possible, assistance should be targeted at those most effected, while taxes to support such programs be drawn primarily from those least effected.

To reiterate an important point, our predictions of the shocks are *not* estimates of the overall impact of the COVID-19 on the economy, but are rather estimates of the first-order shocks. Overall impacts can be very different form first order shocks for several reasons: First, shocks to a particular sector propagate and may be amplified as each industry faces a shock and reduces its demand for intermediate goods from other industries. Second, industries with decreased output will stop paying wages of furloughed workers, thereby reducing income and demand. Third, the few industries facing higher demand will increase supply, if they can overcome labor mobility frictions (del Rio-Chanona et al. 2019). We make our predictions of the shocks available here so that other researchers can improve upon them and use them in their own models[18]. We intend to update and use these shocks ourselves in our models in the near future.

We have made a number of strong assumptions and used data from different sources. To recapitulate, we assume that the production function for an industry is linear and that it does not depend on the composition of occupations who are still able to work; we neglect absenteeism due to mortality and morbidity, as well as loss of productivity due to school closures (though we have argued these effects are small – see Appendix D.1). We have constructed our remote labor index based on a subjective rating of work activities and we assumed that all work activities are equally important and they are additive. We have also applied a rating of essential industries for Italy to the U.S. Nonetheless, we believe that the analysis here provides a useful starting point for macroeconomic models attempting to measure the impact of the COVID-19 pandemic on the economy.

As new data becomes available we will be able to test whether our predictions are correct and improve our shock estimate across industries and occupations. Several countries have already started to release survey data. New measurements about the ability to perform work remotely in different occupations are also becoming available. New York and Pennsylvania have released a list of industries that are considered essential[19] (though this is not currently associated with any industrial classification such as NACE or NAICS). As new data becomes available for the mitigation measures different states and countries are taking, we can also refine our analysis to account for different government actions. Thus we hope that the usefulness of methodology we have presented here goes beyond the immediate application, and will provide a useful framework for predicting economic shocks as the pandemic develops.

[18]Our data repository is at https://zenodo.org/record/3751068, where we will post any update.
[19]https://esd.ny.gov/guidance-executive-order-2026

CEPR PRESS

Covid Economics 6, 17 April 2020: 65-103

References

Adams-Prassl, A., Boneva, T., Golin, M. & Rauh, C. (2020), 'The large and unequal impact of covid-19 on workers', *VoxEU.org, 08 April* . https://voxeu.org/article/large-and-unequal-impact-covid-19-workers.

Baker, S. R., Farrokhnia, R., Meyer, S., Pagel, M. & Yannelis, C. (2020), 'How does household spending respond to an epidemic? consumption during the covid-19 pandemic.', *NBER WP 26949* .

Baldwin, R. & Weder di Mauro, B. (2020), Introduction, *in* R. Baldwin & B. Weder di Mauro, eds, 'Economics in the Time of COVID-19', CEPR, London, chapter 1, pp. 1–30.

Barrot, J.-N., Grassi, B. & Sauvagnat, J. (2020), 'Sectoral effects of social distancing', *Covid Economics 3, 10 April 2020: 85-102* .

Boone, L. (2020), Tackling the fallout from covid-19, *in* R. Baldwin & B. Weder di Mauro, eds, 'Economics in the Time of COVID-19', CEPR, London, chapter 2, pp. 37–44.

Bullard, J. (2020), 'Expected us macroeconomic performance during the pandemic adjustment period', *VoxEU.org, 04 April* . https://voxeu.org/article/expected-us-macroeconomic-performance-during-pandemic-adjustment-period.

Chen, W.-C., Huang, A. S., Chuang, J.-H., Chiu, C.-C. & Kuo, H.-S. (2011), 'Social and economic impact of school closure resulting from pandemic influenza a/h1n1', *Journal of Infection* **62**(3), 200–203.

Congressional Budget Office (2006), 'Potential influenza pandemic: Possible macroeconomic effects and policy issues'. https://www.cbo.gov/sites/default/files/109th-congress-2005-2006/reports/12-08-birdflu.pdf.

del Rio-Chanona, R. M., Mealy, P., Beguerisse-Díaz, M., Lafond, F. & Farmer, J. D. (2019), 'Automation and occupational mobility: A data-driven network model', *arXiv:1906.04086* .

Dingel, J. & Neiman, B. (2020), 'How many jobs can be done at home?', *Covid Economics* **1**.

Goodhart, C. & Pradhan, M. (2020), 'Future imperfect after coronavirus', *VoxEU.org, 27 march* . https://voxeu.org/article/future-imperfect-after-coronavirus.

Hicks, Michael J., F. D. D. S. (2020), 'Occupational exposure to social distancing: A preliminary analysis using O'NET data', *Becker Friedman Institute white paper* . https://projects.cberdata.org/170/social-distance.

Inoue, H. & Todo, Y. (2020), 'The propagation of the economic impact through supply chains: The case of a mega-city lockdown against the spread of covid-19', *Covid Economics* **2**.

Keogh-Brown, M. R., Wren-Lewis, S., Edmunds, W. J., Beutels, P. & Smith, R. D. (2010), 'The possible macroeconomic impact on the uk of an influenza pandemic', *Health Economics* **19**(11), 1345–1360.

Koren, M. & Pető, R. (2020), 'Business disruptions from social distancing', *Covid Economics* **2**.

Lempel, H., Epstein, J. M. & Hammond, R. A. (2009), 'Economic cost and health care workforce effects of school closures in the us', *PLoS currents* **1**.

Mann, C. (2020), Real and financial lenses to assess the economic consequences of covid-19, *in* R. Baldwin & B. Weder di Mauro, eds, 'Economics in the Time of COVID-19', CEPR, London, chapter 8, pp. 81–86.

McKibbin, W. J. & Fernando, R. (2020), 'The global macroeconomic impacts of covid-19: Seven scenarios'.

McKibbin, W. J. & Sidorenko, A. A. (2006), 'Global macroeconomic consequences of pandemic influenza', *Lowy Institute for International Policy Sydney, Australia* .

Miles, D. & Scott, A. (2020), 'Will inflation make a comeback after the crisis ends?', *VoxEU.org, 04 April* . https://voxeu.org/article/will-inflation-make-comeback-after-crisis-ends.

Muellbauer, J. (2020), 'The coronavirus pandemic and U.S. consumption', *VoxEU.org, 11 April* . https://voxeu.org/article/coronavirus-pandemic-and-us-consumption.

OECD (2020), 'Evaluating the initial impact of covid-19 containment measures on economic activity', *OECD Technical Report, March 2020* . https://read.oecd-ilibrary.org/view/?ref=126_126496-evgsi2gmqj&title=Evaluating_the_initial_impact_of_COVID-19_containment_measures_on_economic_activity.

Office for National Statistics (2020), 'Coronavirus and homeworking in the UK labour market: 2019', *ONS Technical Report* . https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/coronavirusandhomeworkingintheuklabourmarket/2019.

Santos, J. R., May, L. & Haimar, A. E. (2013), 'Risk-based input-output analysis of influenza epidemic consequences on interdependent workforce sectors', *Risk Analysis* **33**(9), 1620–1635.

Verity, R., Okell, L. C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., Cuomo-Dannenburg, G., Thompson, H., Walker, P. G., Fu, H. et al. (2020), 'Estimates of the severity of coronavirus disease 2019: a model-based analysis', *The Lancet Infectious Diseases* .

Wallis, J. J. (1989), 'Employment in the great depression: New data and hypotheses', *Explorations in Economic History* **26**, 45–72.

Zhang, S. X., Wang, Y., Rauch, A. & Wei, F. (2020), 'Health, distress and life satisfaction of people in china one month into the covid-19 outbreak', *Distress and Life Satisfaction of People in China One Month into the COVID-19 Outbreak (3/12/2020)* .

Appendix

A Derivation of total shocks

A.1 Derivation of supply shocks

As discussed in the main text, we estimate the supply shock by computing an estimate of the share of work that will not be performed, which we compute by estimating the share of work that is not in an essential industry and that cannot be performed from home. We had to use several concordance tables, and make a number of assumptions, which we describe in details here.

Figure 9 illustrates our method. There are four sets of nodes which are connected by three bipartite networks. The first set of nodes are the 6-digit NAICS industries which are classified to be essential or non-essential. This information is encoded in the K-dimensional column vector u which element $u_k = 1$ if NAICS 6-digit industry k is essential and 0 otherwise. Second, there are N different industry categories on which our economic analysis is based. The 6-digit NAICS codes are connected to these industries by the incidence matrix (concordance table) S. The third set of nodes are the J occupations obtained from the BLS and O*NET data. The weighted incidence matrix M couples industries with occupations where the element M_{nj} denotes the number of people in occupation j being employed in industry n. Fourth, we also have a list of I work activities. Each activity was rated whether it can be performed from home. If activity i can be done from home, the i^{th} element of the vector v is equal 1, and otherwise it is equal to 0. The incidence matrix T denotes whether an occupation is associated with any given work activity, i.e $T_{ji} = 1$ if activity i is relevant for occupation j.

Covid Economics 6, 17 April 2020: 65-103

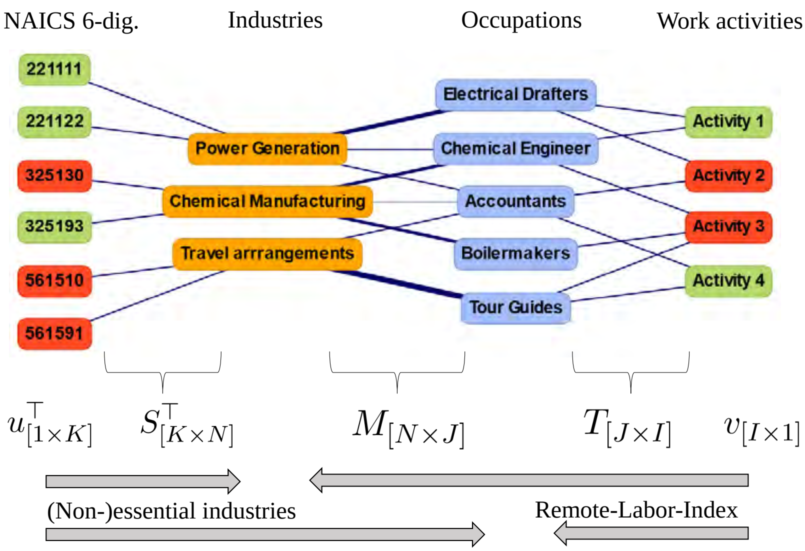


Figure 9: The same schematic network representation of supply side shocks as in the main text, but now also including mathematical notation. The K-dimensional vector u below the NAICS 6-dig. (left nodes) encodes essential industries with binary elements. This set of nodes is connected to relevant industry categories by concordance tables (incidence matrix S). Matrix M connects the N industry categories with J occupations where an element represents the corresponding employment number. The ability to perform work activities (right nodes) from home is represented in vector v, also by binary elements. We use occupation-activity mappings provided from O*NET, represented as incidence matrix T. The grey arrows show the direction of shocks to industries and employment. The shock originating from the list of essential industries is mapped directly onto the broader industry categories, before it can be computed for occupations. Conversely, the remote-labor-index is first mapped onto occupations and then projected onto industries.

The analysis presented here is based on $I = 332$ unique work activities, $J = 740$ occupations and $K = 1,057$ 6-digit NAICS industries. When relating to industry-specific results we use the BLS industry categories of the input-output accounts, leaving us with $N = 169$ industries for which we have reliable data on value added, total output and other key statistics. Employment, occupation and wage statistics are available on a more fine-grained 4-digit NAICS level. We therefore use these $N = 277$ industries for deriving labor-specific results.

A.2 Industry-specific shocks

We can use this simple framework for deriving the supply shocks to industries.

(Non-)essential industries. To estimate the extent to which an industry category is affected by a shutdown of non-essential economic activities, we measure the fraction of its 6-digit NAICS sub-industries which are classified as non-essential. In mathematical terms, the essential-score for every industry is therefore a weighted sum which can be written compactly in matrix notation as

$$e = \tilde{S}u, \tag{1}$$

CEPR PRESS

where $\tilde{S}$ is the row-normalized version of matrix S with elements $\tilde{S}_{nk} = S_{nk} / \sum_h S_{nh}$.

Note that this assumes that the fine-grained NAICS codes contribute uniformly to the more aggregate industry categories. Although this assumption might be violated in several cases, in absence of further information, we use this assumption throughout the text.

Industry Remote-Labor-Index. We can similarly estimate the extent to which the production of occupations or industries can take place by working from home. Since work activities are linked to occupations, but not directly to industries, we need to take two weighted averages to obtain the industry-specific RLI.

For each occupation we first measure the fraction of work activities that can be done from home. We interpret this as the share of work of an occupation that can be performed from home, or 'occupation-level RLI'. This interpretation makes two assumptions: (i) that every work activity contributes equally to an occupation, which is our best guess since we do not have better data, and (ii) that if $z\%$ of activities cannot be done from home, the other $1-z\%$ of activities can still be carried out and and are as productive as before.

For each industry i we then take a weighted average of the occupation-level RLIs, where the weights are the shares of workers employed in each occupation and in industry i. Let $\tilde{T}$ denote the row-normalized version of matrix T, i.e $\tilde{T}_{ji} = T_{ji} / \sum_h T_{jh}$ and similarly let the element of matrix $\tilde{M}$ be $\tilde{M}_{nj} = M_{nj} / \sum_h M_{nh}$. Then the industry-specific remote-labor-index is given by the vector

$$r = \tilde{M}\tilde{T}v. \tag{2}$$

We interpret the remote-labor-index for an industry, r_n, as the fraction of work in an industry n that can be performed from home. As for assumption (ii) above for the occupation-level RLI, this assumes that if $z\%$ of the work of occupations cannot be done, the other $1 - z\%$ of work can still be carried out.

Immediate industry supply shock. To derive industry supply shocks from the scores above, we need to take into account that industries might be exposed to both effects at the same time, but with different magnitudes. For example, consider the illustrative case of *Chemical Manufacturing* in Figure 9. Half of the industry is non-essential (red node '325130') and could therefore be directly affected by an economic shutdown. But different occupations can be found in this industry that are affected heterogeneously. In this simple example, *Chemical Manufacturing* draws heavily on Boilermakers who have only work activities that cannot be done from home. On the other hand, this industry also has a tiny share of accountants and a larger share of Chemical Engineers who are able to do half of their work activities from home.

As stated above, the essential score e_n and the RLI r_n can be interpreted as shares of industry-specific work which can be performed, either thanks to being essential or thanks to being adequately done from home. To compute the share of industry-specific work that can performed due to either effect, we interpret shares as probabilities and assume independence,

$$\mathrm{ISS}_n = -\left(1 - e_n\right)\left(1 - r_n\right), \tag{3}$$

where ISS stands for 'Industry Supply Shock'. We have multiplied the probability by minus one to obtain negative shocks. Although independence is a strong assumption, we have no reason to believe that the work that can be done from home is more or less likely to be judged essential. The empirical correlation

Covid Economics 6, 17 April 2020: 65-103

Covid Economics 6, 17 April 2020: 65-103

coefficient of e and r is 0.03 and is far from being significant (p-value of 0.7). Under the linear assumptions about combining labor we are making here this indicates independence.

When applying these Industry Supply Shocks to value added, we make the implicit assumption that a $z\%$ decrease in labor will cause a $z\%$ decrease in value added.

A.3 Occupation-specific shocks.

We now describe how we compute shocks for specific occupations, rather than specific industries.

Occupations in (non-)essential industries. Occupations are mapped to industries through the weighted incidence matrix M, where an element denotes the number of jobs per occupation and industry. The column-normalized matrix M^* with elements $M^*_{nj} = M_{nj}/\sum_h M_{hj}$ denotes the share of an occupation carried out in a particular industry[20]. The essential-score for occupations is taken as weighted average of the essential score for industries (computed in Eq. 1),

$$x = M^{*\top}\tilde{S}u = M^{*\top}e. \tag{4}$$

Occupation Remote Labor Index. As already indicated in the derivation of the industry-specific RLI, r, in Eq. (2), the occupation-specific RLI, y, is a weighted average of all the corresponding work activities that can be done from home. Formally, the occupation-based RLI is given by

$$y = \tilde{T}v. \tag{5}$$

Total supply-driven occupation shock. Following the same procedure as in Eq. 3, we can get the total immediate shock on occupations from the economy's supply side[21]. The combined immediate shock to occupations is then given as

$$\mathrm{OSS}_j = -\left(1 - x_j\right)\left(1 - y_j\right). \tag{6}$$

Here, the correlation between RLI and the essential-score is larger, $\rho(x, y) = 0.30$ (p-value $= 8.3\times10^{-17}$), and significant which can also be seen from Figure 4. It should therefore be noted that the labor-specific results are expected to be more sensitive with respect to the independence assumption, as it is the case for industry-related results.

A.4 Derivation of demand shocks

Since we have demand shocks only on the 2-digit NAICS level, disaggregating them into the more fine-grained relevant industry categories is straightforward when assuming that the demand shock holds equally for all sub-industries. We let the industry demand shock in percentages for industry n be $-\mathrm{IDS}_n$.

To map the demand shocks onto occupations, we can invoke the same matrix algebra as above. The occupation-specific shock originating from the economy's demand side is then given by the projection

$$\mathrm{ODS} = M^{*\top}\mathrm{IDS}. \tag{7}$$

[20]Note that we column-normalize M to map from industries to occupations and row-normalize when mapping from occupations to industries.

[21]To be clear, this is a product market *supply-side* shock, but this translates into a reduction in labor *demand* in each occupations.

Covid Economics 6, 17 April 2020: 65-103

A.5 Total immediate (first-order) shocks

We now combine supply- and demand-driven shocks to total immediate shocks for occupations and industries.

Let us turn to industries first. As discussed in more depth in the main text, the shock experienced is likely to be the worse of the two (supply and demand) shocks. Since we have expressed shocks as negative if they lead to decrease in output, in more mathematical terms, the *industry total shock* then is

$$\mathrm{ITS}_n = \min(\mathrm{ISS}_n, \mathrm{IDS}_n). \tag{8}$$

and the *occupation total shock* is

$$\mathrm{OTS}_j = \min(\mathrm{OSS}_j, \mathrm{ODS}_j). \tag{9}$$

Note that under these assumptions, the health sector will not experience a positive shock. Our estimates of the shocks suggest that there is an increase in demand, but we have not described positive supply shocks for the health sector. As a result, industries and occupations in the health sector will have a non-positive total shock.

Of course, the actual production of health care is likely to go up. But this is the result of an adjustment to an increase demand, rather than a first-order shock, and is therefore not modelled here. Nevertheless, see Appendix A.7 for an alternative.

A.6 Aggregate total shocks

To provide economy-wide estimate of the shocks, we aggregate industry- or occupation-level shocks. We do this using different sets of weights.

First of all, consider the interpretation that our shocks at the occupation-level represent the share of work that will not be performed. If we assume that if $z\%$ of the work cannot be done, $z\%$ of the workers will become unemployed, we can weight the occupation shocks by the share of employment in each occupation. Using the vector L to denote the share of employed workers that are employed by occupation j, we have

$$\text{Employment total shock} = \mathrm{OTS}^T L \tag{10}$$

The Employment supply (demand) shock is computed similarly but using OSS (ODS) instead of OTS.

Instead of computing how many workers may lose their job, we can compute by how much paid wages will decrease. For each occupation, we compute the total wage bill by multiplying the number of workers by the average wage. We then create a vector w where w_j is the share of occupation j in the total wage bill. Then,

$$\text{Wage total shock} = \mathrm{OTS}^T w, \tag{11}$$

and similarly for the OSS and ODS. Note that we omit three occupations for which we do not have wages (but had employment).

Finally, to get an estimate of the loss of GDP, we can aggregate shocks by industry, weighting by the share of an industry in GDP. Denoting by Y the vector where Y_n is the VA of industry n divided by GDP[22],

$$\text{Value added total shock} = \mathrm{ITS}^T Y, \tag{12}$$

[22]Our estimate of GDP is the sum of VA of industries in our sample.

Covid Economics 6, 17 April 2020: 65-103

and similarly for the industry supply and demand shocks (ISS and IDS). Note that we could have used shares of gross output and compute a shock to gross output rather to GDP.

A.7 Aggregate total shocks with growth of the Health sector

Here we make a different assumption about how to construct the total shock for occupations and industries. For industries, we assume that if they experience a positive demand shock, the industries are able to increase their supply to meet the new demand. Instead of Eqs. (8) we use

$$\mathrm{ITS}_n^h = \begin{cases} \mathrm{ITS}_n, & \text{if } \mathrm{IDS}_n \leq 0 \\ \mathrm{IDS}_n, & \text{if } \mathrm{IDS}_n > 0. \end{cases} \tag{13}$$

Since occupations are employed by different industries, the total shock to an occupation can be influenced by positive demand shocks from the healthcare sector and negative demand shocks from non-essential industries. In Eq. (9) we consider that occupations only experience the negative shocks. An alternative is to consider both the negative shock caused by non-essential industries and the positive shock caused by the health industries. This gives

$$\mathrm{OTS}_j^h = \begin{cases} \mathrm{OTS}_j, & \text{if } \mathrm{ODS}_j \leq 0 \\ \mathrm{ODS}_j + \mathrm{OSS}_j, & \text{if } \mathrm{ODS}_j > 0. \end{cases} \tag{14}$$

In Section 4.4, specifically Figure 8 we use this convention for the y-axis, the Labor Shock. Using Eq. (14) allows us to observe how health related occupations experience a positive shock.

In Table 6 we show the aggregate total shocks when using Eqs.(13) and (14). There is very little difference with the results in the main text. The Health sector and its increase in demand are not large enough to make a big difference to aggregate results.

	Employment					Wages	Value Added
Shock	Aggregate	q_1	q_2	q_3	q_4	Aggregate	Aggregate
Total	-22	-41	-22	-20	-4	-15	-21

Table 6: **Main results allowing for growth in the Health sector**. The results are the same as those presented in Table 4, but assuming that in industries, when demand is positive, the total shock is equal to the demand shock

B Data

In this section we give more details about how we constructed all our variables. We stress that our goal was to produce useful results quickly and transparently, and make them available so that anyone can update and use them. We intend to improve these estimates ourselves in the future, as more information becomes available on the ability to work from home, which industries are essential, and how consumers react to the crisis by shifting their spending patterns.

B.1 Italian list of essential industries

The Italian list of essential industries[23] is based on the Statistical Classification of Economic Activities in the European Community, commonly referred to as NACE. The list of essential industries are listed with

[23] Available at http://www.governo.it/sites/new.governo.it/files/dpcm_20200322.pdf, 22 March.

CEPR PRESS

NACE 2-digit, 4-digit and 6-digit codes. We automatically map industries listed at the 2 or 4-digit NACE level to NAICS 6-digit industries using the crosswalk made available by the European Commission[24]. The 6-digit NACE level classification is country dependent and thus there is no official crosswalk to NAICS codes. We map the 6-digit industries by hand.

In a second step, we looked at the list of resulting list of industries and their essential score and discovered a few implausible cases, resulting from the complex mapping between the various classification systems at different levels. For instance, because Transports are essential, "Scenic and sightseeing transportation, other" was considered essential. In contrast, "Death care services" was classified as non essential. Three of us, as well as two independent colleagues with knowledge of the current situation in Italy, evaluated the list and we proceeded to editing the 4-digit NAICS Essential Scores as follows: From Non-essential to Essential: Grocery stores, Health and personal care stores, Gasoline stations, Death care services. From Essential (sometimes only partly) to non essential: Scenic and sightseeing transportation, Software publishers, Motion picture and video industries, Sound recording industries, and Other amusement and recreation industries. Finally, Federal, State and Local Government were not classified, and we classified as Essential.

B.2 Data for occupations

O*NET has work activities data for 775 occupations, out of which 765 occupations have more than five work activities. We compute the remote labor index for the 765 occupations with more than five work activities. From the May 2018 Occupational Employment Statistics (OES) estimates on the level of 4-digit NAICS (North American Industry Classification System), file *nat4d_M2018_dl*, which is available at `https://www.bls.gov/oes/tables.htm` under *All Data*, we find data for the number of employed workers of 807 occupations in 277 industries. These data covers 144M workers[25]. From the sample of 765 occupations with RLI, and from the sample of 807 occupations with employment data from the BLS, we are able to match 740 occupations, which cover 136.8M workers. Therefore, our final sample has 740 occupations and 136.8M workers.[26]

With the occupation-industry employment data and the essential score of each industry, we estimate the share of essential jobs within each occupation. Additionally, we have wage information for most occupations (i.e. we have median and mean wage data for 737 occupations). We computed all correlations for median wage considering all occupations we had median wage data for. For the 3 occupations for which median wage. data was missing, the color coding of occupations in Figures 4, 7, and the x-axis in Figure 8 corresponds to the average (across all occupations) of the median wage. We used the mean wages and the employment of occupations to define the wage quartiles of our sample. We excluded the 3 occupations for which we did not have mean wage from these calculations.

Finally, we use the O*NET data on exposure to disease and infection of occupations for the color coding in Figure 8. We explain these data further in Appendix C. In the following charts we show the distribution of the remote labor index, exposure to disease and infection, supply, demand and overall shocks across occupations.

[24]`https://ec.europa.eu/eurostat/ramon/relations/index.cfm?TargetUrl=LST_REL&StrLanguageCode=EN&IntCurrentPage=11`

[25](the US economy had 156M workers mid-2018, see `https://fred.stlouisfed.org/series/CE16OV`)

[26]Note that the BLS employment data we use here does not include self-employed workers (which currently accounts for about 16 million people).

Covid Economics 6, 17 April 2020: 65-103

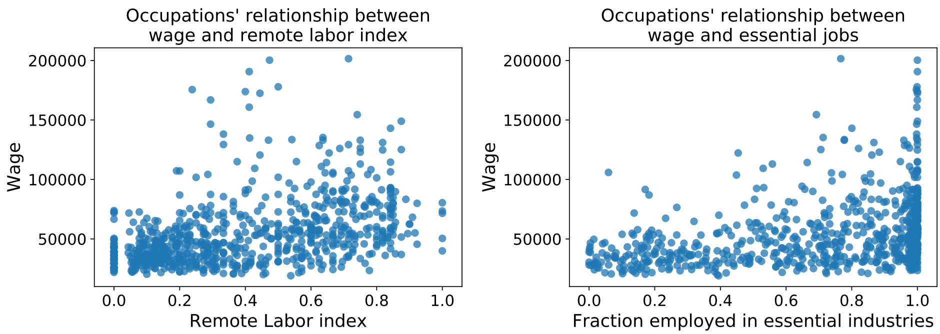


Figure 10: Left. Relationship between remote labor index and median wage. The Pearson correlation is 0.42 (p-value = 7.3×10^{-33}). Right. Relationship between fraction of workers in essential industries and wage. The Pearson correlation is 0.35 (p-value = 1.12×10^{-22}).

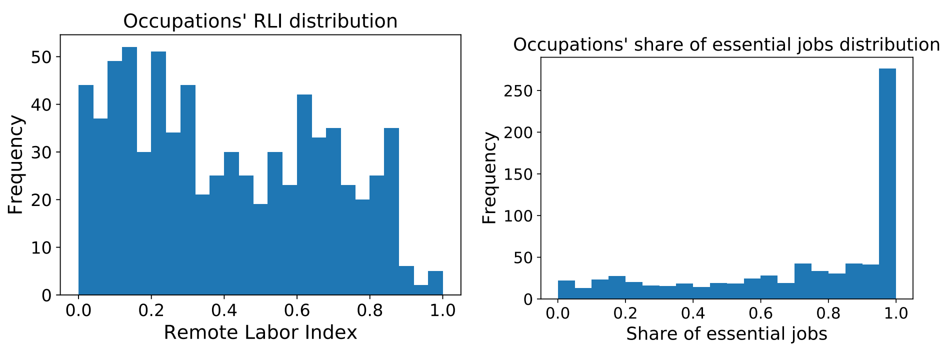


Figure 11: Left. Distribution of the remote labor index for the 660 occupations. Right. Distribution of the share of essential jobs within each of the 660 occupations.

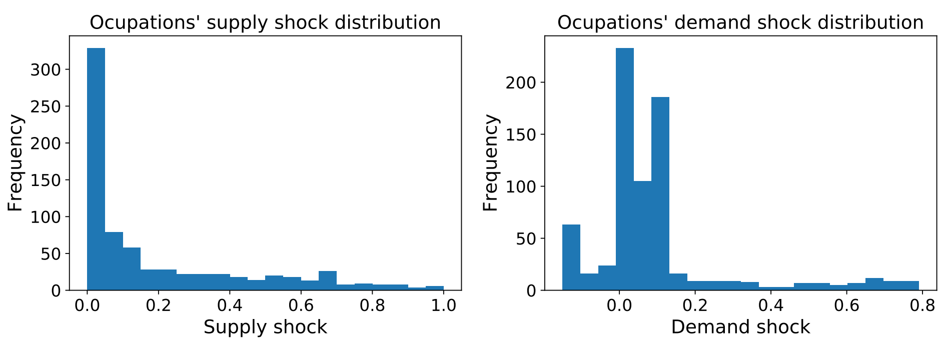


Figure 12: Left. Supply shock distribution across occupation. Right. Demand shock distribution across occupations.

Covid Economics 6, 17 April 2020: 65-103

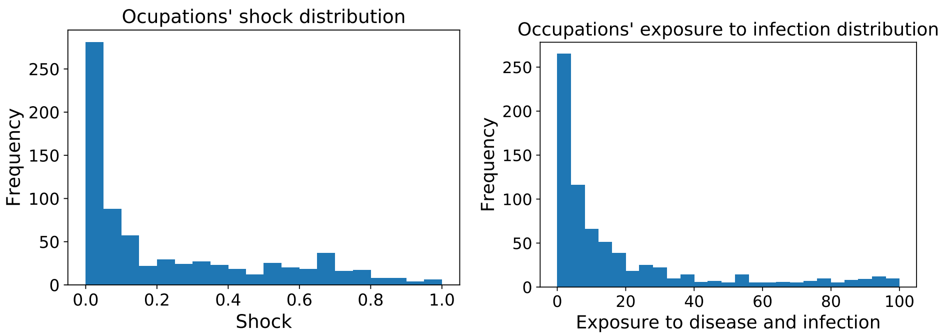


Figure 13: Left. Shock distribution for occupations. Right. Distribution of exposure to disease.

B.3 Data for industries

Matching all data to BLS I-O industries. A key motivation of this paper is to provide relevant economic data which can be used by other researchers and policymakers to model the economic impact of the COVID-19 pandemic. We therefore bring the discussed supply and demand shock data into a format that matches directly to US input-output data.

We use the BLS 2018 input-output account, which allows us to discern 179 private sectors. Moreover, there are the additional industries *Private Households*, NAICS 814 and *Postal Service*, NAICS 491. The data also contains 19 different industries relating to governmental activities. Since these industries are not classified with NAICS codes, we aggregate all governmental industries into a single node *Government*, which can be interpreted as the NAICS 2-digit industry 92. This leaves us with 182 industry categories which are a mixture of 2- to 6-digit NAICS industries.

We are able to match occupational data to 170 out of the 182 industry categories, accounting for 97% of total value added. For this subset we compute industry-specific RLIs, essential scores and supply shocks as spelled out in Appendix A.1, as well as employment-weighted infection exposures.

Since we have demand shocks only on the 2-digit NAICS level, disaggregating them into the more fine-grained BLS input-output data is straightforward when assuming that the demand shock holds equally for all sub-industries.

In the online data repository we also report total wages and total employment per industry. We use the same OES estimates as for the occupational data, but match every industry category according to the corresponding NAICS 2- to 6-digit digit levels.

Figure 14 to Figure 16 show distributions of supply and demand shock-related variables on the industry level. Table 7 summarizes a few key statistics for these industries, when further aggregated to 72 industry categories.

Covid Economics 6, 17 April 2020: 65-103

Covid Economics 6, 17 April 2020: 65-103

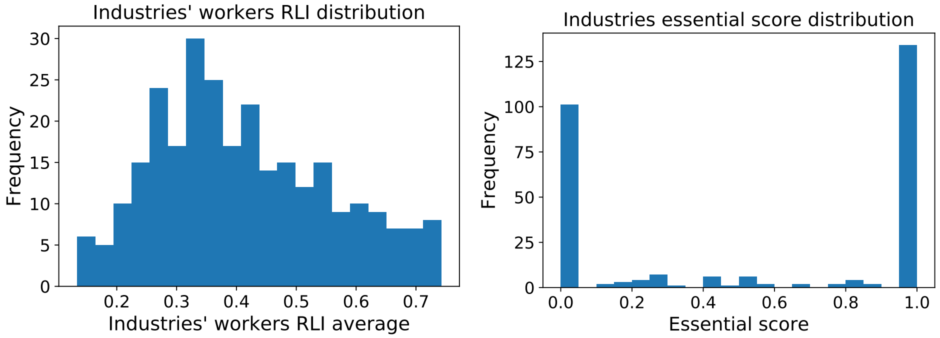


Figure 14: Left. Distribution of the remote labor index, aggregated to 169 industries. Right. Fractions of essential sub-industries per industry category.

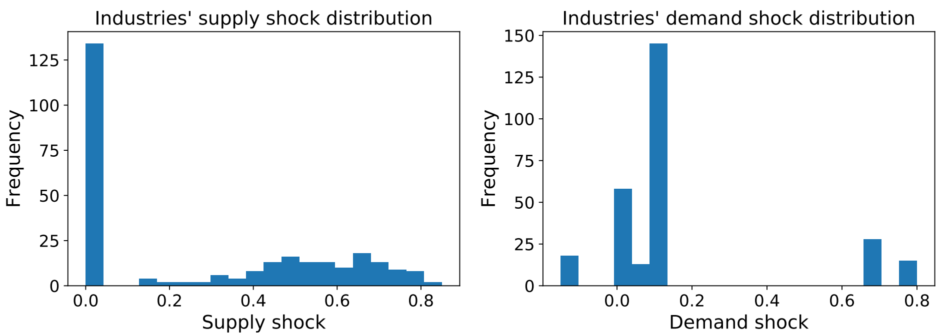


Figure 15: Left. Supply shock distribution across industries. Right. Demand shock distribution across industries.

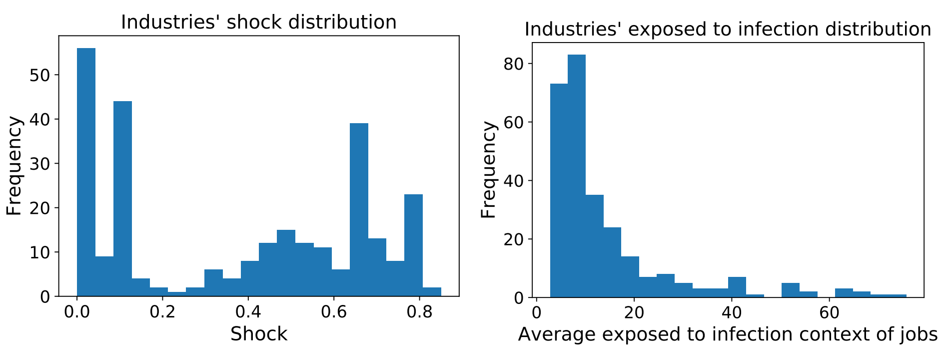


Figure 16: Left. Shock distribution across industries. Right. Distribution of exposure to disease across industries.

NAICS	Title	Outp.	Empl.	Demand	Supply	RLI	Essent.	Expos.
111	Crop Production	209	NA	-10	NA	NA	NA	NA
112	Animal Production and Aquaculture	191	NA	-10	NA	NA	NA	NA
113	Forestry and Logging	19	NA	-10	NA	NA	NA	NA
114	Fishing, Hunting and Trapping	10	NA	-10	NA	NA	NA	NA
115	Support Activities for Agriculture and Forestry	27	378	-10	0	14	100	5
211	Oil and Gas Extraction	332	141	-10	0	47	100	7
212	Mining (except Oil and Gas)	97	190	-10	-54	26	27	8
213	Support Activities for Mining	84	321	-10	-72	28	0	8
221	Utilities	498	554	0	0	42	100	9
23	Construction	1636	7166	-10	-24	31	66	8
311	Food Manufacturing	803	1598	-10	0	21	100	10
312	Beverage and Tobacco Product Manufacturing	192	271	-10	-3	33	96	7
313-4	Wholesale Trade	54	226	-10	-51	26	31	5
315-6	Management of Companies and Enterprises	29	140	-10	-68	25	9	4
321	Wood Product Manufacturing	118	402	-10	-62	26	16	6
322	Paper Manufacturing	189	362	-10	-8	24	89	6
323	Printing and Related Support Activities	80	435	-10	0	38	100	4
324	Petroleum and Coal Products Manufacturing	618	112	-10	-26	36	60	7
325	Chemical Manufacturing	856	828	-10	-2	38	96	9
326	Plastics and Rubber Products Manufacturing	237	722	-10	-8	28	89	7
327	Nonmetallic Mineral Product Manufacturing	140	NA	-10	NA	NA	NA	NA
331	Primary Metal Manufacturing	239	374	-10	-73	27	0	7
332	Fabricated Metal Product Manufacturing	378	1446	-10	-59	33	12	6
333	Machinery Manufacturing	386	1094	-10	-49	42	16	5
334	Computer and Electronic Product Manufacturing	369	1042	-10	-38	58	9	4
335	Electrical Equipment, Appliance, and Component Manufacturing	132	392	-10	-31	45	45	6
336	Transportation Equipment Manufacturing	1087	1671	-10	-58	37	9	5
337	Furniture and Related Product Manufacturing	77	394	-10	-47	35	28	4
339	Miscellaneous Manufacturing	173	601	-10	-16	40	74	11
42	Construction	1980	5798	-10	-27	50	46	7
441	Motor Vehicle and Parts Dealers	334	2006	-10	-23	43	60	11
442-4, 446-8, 451, 453-4	Wholesale Trade	1052	7731	-10	-39	53	17	19
445	Food and Beverage Stores	244	3083	-10	-33	43	43	15
452	General Merchandise Stores	240	3183	-10	-49	51	0	17
481	Air Transportation	210	499	-67	0	29	100	29
482	Rail Transportation	77	233	-67	0	33	100	11
483	Water Transportation	48	64	-67	0	35	100	10
484	Truck Transportation	346	1477	-67	-68	32	0	8
485	Transit and Ground Passenger Transportation	74	495	-67	0	27	100	43
486	Pipeline Transportation	49	49	-67	0	37	100	8
487-8	Management of Companies and Enterprises	146	732	-67	-10	37	85	8
491	Postal Service	58	634	-67	0	35	100	10
492	Couriers and Messengers	94	704	-67	0	37	100	15
493	Warehousing and Storage	141	1146	-67	0	25	100	6
511	Publishing Industries (except Internet)	388	726	0	-16	70	46	4
512	Motion Picture and Sound Recording Industries	155	428	0	-51	49	0	9
515	Broadcasting (except Internet)	196	270	0	0	65	100	6
517	Telecommunications	695	NA	0	NA	NA	NA	NA
518	Data Processing, Hosting, and Related Services	207	319	0	0	70	100	4
519	Other Information Services	192	296	0	-7	71	75	5
521-2	Construction	939	2643	0	0	74	100	11
523, 525	Wholesale Trade	782	945	0	-18	74	32	5
524	Insurance Carriers and Related Activities	1231	2330	0	0	71	100	8
531	Real Estate	1842	1619	0	-53	47	0	19
532	Rental and Leasing Services	163	556	0	-54	46	0	12
533	Lessors of Nonfinancial Intangible Assets (except Copyrighted Works)	182	22	0	-30	70	0	7
541	Professional, Scientific, and Technical Services	2372	9118	0	-2	64	94	9
55	Management of Companies and Enterprises	561	2373	0	0	66	100	8
561	Administrative and Support Services	971	8838	0	-37	35	44	17

Covid Economics 6, 17 April 2020: 65-103

562	Waste Management and Remediation Services	109	427	0	0	30	100	22
611	Educational Services	366	13146	0	0	54	100	29
621	Ambulatory Health Care Services	1120	7399	15	0	37	100	59
622	Hospitals	933	6050	15	0	36	100	65
623	Nursing and Residential Care Facilities	262	3343	15	0	28	100	61
624	Social Assistance	222	3829	15	0	40	100	47
711	Performing Arts, Spectator Sports, and Related Industries	181	505	-80	-51	44	8	13
712	Museums, Historical Sites, and Similar Institutions	20	167	-80	-49	51	0	16
713	Amusement, Gambling, and Recreation Industries	158	1751	-80	-65	35	0	21
721	Accommodation	282	2070	-80	-34	33	50	26
722	Food Services and Drinking Places	832	11802	-80	-64	36	0	13
811	Repair and Maintenance	235	1317	-5	-3	29	96	9
812	Personal and Laundry Services	211	1490	-5	-52	28	28	31
813	Religious, Grantmaking, Civic, Professional, and Similar Organizations	260	1372	-5	0	52	100	19
814	Private Households	20	NA	-5	NA	NA	NA	NA
92	All Public Sector (custom)	3889	9663	0	0	44	100	21

Table 7: Key statistics for different 2- and 3-digit NAICS industries. Column 'Outp.' refers to total output of the industry in current billion USD (2018). 'Emp.' is total employment in thousands. 'Demand' is the immediate severe demand shock in % obtained from the CBO. 'Supply' is the immediate supply shock in % derived from the Remote Labor Index and the list of essential industries. 'RLI' is the industry-specific Remote Labor Index in %. 'Essent.' is the share of sub-industries being classified as essential in %. 'Expos.' denotes the industry-aggregated infection exposure index from O*NET which ranges from 0 to 100 and is explained in Appendix C. We make more disaggregated data with further details available in the corresponding data publication.

C Occupations most at risk of contracting SARS-CoV-2

O*NET make available online work context data for occupations that describe the physical and social factors that influence the nature of work. The "Exposed to disease and infection" work context[27], which we refer to as 'exposure to infection' for short, describes the frequency with which a worker in a given occupation is exposed to disease or infection. It ranges from 0 to 100, where 0 means "never" and 100 "everyday"; an exposed to infection rating of 50 means an exposure of "once a month or more but not every week" and 75 means "Once a week or more but not every day". We have exposure to infection data for 644 of the 740 occupations in our sample. For those occupations for which we did not have the exposure to infection, when colored them as if they had zero exposure to infection.

As we see in Figure 17 there is a U shaped relationship between wages and exposure to infection. There is a correlation of 0.36 (p-value $= 1.7 \times 10^{-21}$) between wages and exposure to infection, but this is misleading[28]. Though many high wage occupations are highly exposed to infection (high paid doctors), there are also many low wage occupations with high probability of infection.

[27] https://www.onetonline.org/find/descriptor/result/4.C.2.c.1.b?s=2

[28] For example, Adams-Prassl et al. (2020), using survey evidence for ~4000 US individuals, found that workers without paid sick leave are more likely to go to work in close proximity to others, which may have suggested a negative correlation between wages and exposures. Note however that our correlation is based on occupations, not individuals, and that wages are not necessarily an excellent predictor of having paid sick leave or not.

Covid Economics 6, 17 April 2020: 65-103

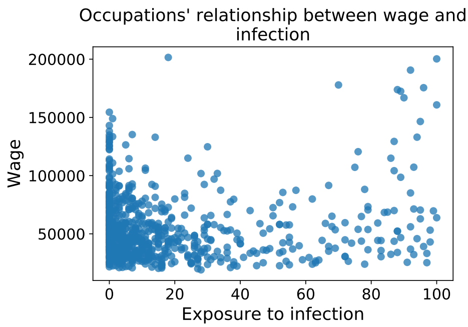


Figure 17: Relationship between wage and probability of infection in linear and log scale. The Pearson correlation is 0.36 (p-value = 1.7×10^{-21}). However we this correlation is mostly driven by high salaries in the health sector, but there are many low-wage occupations with a significant exposure to infection.

D Discussion of labor supply shocks which we do not include

D.1 Labor supply shocks from mortality and morbidity

Typical estimates. McKibbin & Fernando (2020) consider attack rates (share of population who become sick) in the range 1-30% and case-fatality rates (share of those infected who die) in the range 2-3%. From attack rates and case fatality rates, they compute mortality rates. They also assume that sick people stay out of work for 14 days. A third effect they assume is that workers would be care givers to family members.

For their severe scenario of an influenza pandemic, Congressional Budget Office (2006) assumed that 30 percent of the workers in each sector (except for Farms, which is 10%) would become ill and would lose 3 weeks or work, at best, or die (2.5% case fatality rate).

Best guess for current effect of COVID-19. In the case of COVID-19, estimating a labor supply shock is made difficult by several uncertainties. First of all, at the time of writing there are very large uncertainties on the ascertainment rate (the share of infected people who are registered as confirmed cases), making it difficult to know the actual death rate.

We report the result from a recent and careful study by Verity et al. (2020), who estimated an infection fatality ratio of 0.145% (0.08-0.32) for people younger than 60, and 3.28% (1.62-6.18) for people aged 60 or more. The age bracket 60-69, which in many countries will still be part of the labor force, was reported as 1.93% (1.11–3.89).

Taking the infection fatality ratio for granted, the next question is the attack rate. In Verity et al. (2020), the infection fatality ratios are roughly one fourth of the case fatality ratios, suggesting that 3/4 of the cases are undetected. For the sake of the argument, consider Italy, a country that has been strongly affected and appears to have reached a peak (at least of a first wave). There are at the time of writing 132,547 cases in Italy[29] In 2018 the population of Italy[30] was 60,431,283. If we assume that Italy is at the peak today and the curve is symmetric, the total number of cases will be double the current number,

[29] https://coronavirus.jhu.edu/map.html

[30] https://data.worldbank.org/indicator/SP.POP.TOTL?locations=IT

that is 265,094, which is 0.44% of the population. If we assume that the true number of cases if 4 times higher, the attack rate is, roughly speaking, 1.76%. These numbers are more than an order of magnitude smaller than the number who cannot work due to social distancing.

Thus, while it is clear that the virus is causing deep pain and suffering throughout Italy, the actual decrease in labor supply, which is massive, is unlikely to be mostly caused by people being sick, and is much more a result of social distancing measures.

Uncontrolled epidemic. Now, it may be informative to consider the case of an uncontrolled epidemic. If we assume that the uncontrolled epidemic has an attack rate of 80% (a number quoted in Verity et al. (2020)), an infection fatality ratio for people in the labor force of 1% (an arbitrary number between 0.145% for people younger than 60, and 1.93% for the 60-69 age bracket) implies an 0.8% permanent decrease of the labor force. If we assume that those who do not die are out of work for 3 weeks, on an annual basis of 48 worked weeks, we have (3/48)*(0.80-0.01)=4.94% decrease of the labor supply.

Overall, this exercise suggests that left uncontrolled, the epidemic can have a serious effect on labor supply. However, in the current context, the effect on the economy is vastly more a result of social distancing than direct sickness and death.

D.2 Labor supply shocks from school closure

School closures are a major disruption to the functioning of the economy as parents can no longer count on the school system to care for their children during the day.

Chen et al. (2011) surveyed households following a school closure in Taiwan during the H1N1 outbreak, and found that 27% reported workplace absenteeism. Lempel et al. (2009) attempted to estimate the cost of school closure in the US in case of an influenza pandemic. They note that 23% of all civilian workers live in households with a child under 16 and no stay-at-home adults. Their baseline scenario assumes that around half of these workers will miss some work leading to a loss 10% of all labor hours in the civilian U.S. economy, for as long as the school closure lasts.

Some of these effects would already be accounted for in our shocks. For instance, some workers are made redundant because of a supply or demand shock, so while they have to stay at home to care for their children, this is as much a result of labor and supply shock as a result of school closure. For those working from home, we might expect a decline in productivity. Finally, for those in essential industries, it is likely that schools are not close. For instance, in the UK, schools are opened for children of essential workers. Our list of essential industries from Italy includes Education.

Overall, school closure indeed have large effects, but in the current context these may already be accounted for by supply and demand shocks, or non-existent because schools are not fully closed. The loss of productivity from parents working from home remains an open question.

E Additional estimates of demand or consumption shocks

In this appendix we provide additional data on the demand shock. Table 8 shows our crosswalk between the Industry classification of the Congressional Budget Office (2006) and NAICS 2-digit industry codes, and, in addition to the "severe" shocks used here, shows the CBO's "mild" shocks. We have created this concordance table ourselves, by reading the titles of the categories and making a judgement. Whenever NAICS was more detailed, we reported the CBO's numbers in each more fine-grained NAICS.

Covid Economics 6, 17 April 2020: 65-103

We also provide two sources of consumption shocks (in principle, these estimates are meant to reflect actual decreases in consumption rather than shifts of the demand curve). Table 9 shows the consumption shocks used by Keogh-Brown et al. (2010) to model the impact of potential severe influenza outbreak in the UK. Table 10 shows the consumption shocks used by Muellbauer (2020) to model the impact of the COVID-19 on quarterly US consumption. OECD (2020) provided two other sources, both reported in Table 10. The first one is based on assumptions of shocks at the industry level, while Table shows assumptions of shocks by expenditure categories (COICOP: Classification of individual consumption by purpose).

Covid Economics 6, 17 April 2020: 65-103

Covid Economics 6, 17 April 2020: 65-103

NAICS	NAICS	CBO	Severe	Mild
11	Agriculture, Forestry, Fishing and Hunting	Agriculture	-10	-3
21	Mining, Quarrying, and Oil and Gas Extraction	Mining	-10	-3
22	Utilities	Utilities	0	0
23	Construction	Construction	-10	-3
31	Manufacturing	Manufacturing	-10	-3
32	Manufacturing	Manufacturing	-10	-3
33	Manufacturing	Manufacturing	-10	-3
42	Wholesale Trade	Wholesale trade	-10	-3
44	Retail Trade	Retail trade	-10	-3
45	Retail Trade	Retail trade	-10	-3
48	Transportation and Warehousing	Transportation and warehousing (including air, rail and transit)	-67	-17
49	Transportation and Warehousing	Transportation and warehousing (including air, rail and transit)	-67	-17
51	Information	Information (Published, broadcast)	0	0
52	Finance and Insurance	Finance	0	0
53	Real Estate and Rental and Leasing	NA	0	0
54	Professional, Scientific, and Technical Services	Professional and business services	0	0
55	Management of Companies and Enterprises	NA	0	0
56	Administrative and Support and Waste Management and Remediation Services	NA	0	0
61	Educational Services	Education	0	0
62	Health Care and Social Assistance	Healthcare	15	4
71	Arts, Entertainment, and Recreation	Arts and recreation	-80	-20
72	Accommodation and Food Services	Accommodation/food service	-80	-20
81	Other Services (except Public Administration)	Other services except government	-5	-1
92	Public Administration (not covered in economic census)	Government	0	0

Table 8: Mapping of CBO shocks to NAICS 2-digits

Industry	Consumption shock	Only postponed?
Food, drink, alcohol and tobacco	0	NA
Clothing and footwear	-50	yes
Housing, heating, etc.	0	NA
Goods and services (furniture, etc.)	-80	yes
Transport - buying cars	-100	yes
Transport services and car use	-50	no
Recreation and culture - durables	-100	yes
Recreation and culture - games and pets	0	NA
Recreation and culture - sport and culture	-100	no
Recreation and culture - newspapers and books	0	NA
Restaurants, hotels and net tourism	-100	no
Miscellaneous (incl health, communication education)	0	NA

Table 9: Demand shock from Keogh-Brown et al. (2010). The first column gives the percentage decrease, while the second column gives the percentage of the first column which will be recouped in future quarters.

Category	Shock (%)
ISIC.Rev4 shock from OECD (2020)	
Manufacturing of transport equipment (29-30)	-100
Construction (VF)	-50
Wholesale and retail trade (VG)	-75
Air transport (V51)	-75
Accommodation and food services (VI)	-75
Real estate services excluding imputed rent (VL-V68A)	-75
Professional service activities (VM)	-50
Arts, entertainment and recreation (VR)	-75
Other service activities (VS)	-100
COICOP shock from OECD (2020)	
Clothing and footwear (3)	-100
Furnishings and household equipment (5)	-100
Vehicle purchases (7.1)	-100
Operation of private vehicles (7.2)	-50
Transport services (7.3)	-50
Recreation and culture excluding package holidays (9.1-9.5)	-75
Package holidays (9.6)	-100
Hotels and restaurants (11)	-75
Personal care services (12.1)	-100
Consumption shocks from Muellbauer (2020)	
Restaurants and Hotels	-71
Transport services	-70
Recreation services	-63
Food at home	43
Healthcare	18

Table 10: Estimates of consumption shocks from various sources.

Covid Economics 6, 17 April 2020: 65-103

CEPR PRESS

Covid Economics **Issue 6, 17 April 2020**

Who can work from home in developing countries?

Fernando Saltiel[1]

Date submitted: 10 April 2020; **Date accepted:** 11 April 2020

Covid Economics 6, 17 April 2020: 104-118

In this paper, I examine the feasibility of working from home in developing countries. I take advantage of worker-level data from the STEP survey, which collects comparable information on employment outcomes across ten countries. I use information on workers' tasks to define the feasibility of working from home following Dingel and Neiman (2020). I extend the nascent literature on this topic by providing comparable cross-country evidence on the feasibility of telework. Only 13% of workers in STEP countries could work from home, yet this share ranges from 5.5% in Ghana to 23% in Yunnan (China). The feasibility of working from home is positively correlated with high-paying occupations. Educational attainment, formal employment status and household wealth are positively associated with the possibility of working from home, reflecting the vulnerability of various groups of workers. These relationships remain significant within narrowly defined occupations, yet exhibit heterogeneity across countries. I remark on the importance of rapidly identifying vulnerable workers to design adequate policies to combat the negative employment impacts of COVID-19.

1 Postdoctoral Associate, Department of Economics, Duke University. I thank Alejandro Graziano and Mateo Uribe-Castro for useful comments and suggestions.

Covid Economics 6, 17 April 2020: 104-118

1 Introduction

Governments across the world have implemented social distancing and stay-at-home policies to stop the spread of COVID-19. A critical consideration for understanding the negative labor market impacts arising from the virus is whether workers can plausibly work from home. Two recent papers have contributed important evidence in this dimension in the United States by taking advantage of task-content information in occupational dictionaries such as the O*NET (Dingel and Neiman, 2020; Mongey and Weinberg, 2020). However, their analysis cannot be directly extrapolated to developing countries, as the task content of occupations may vary significantly across contexts (Dicarlo et al., 2016; Lo Bello et al., 2019; Saltiel, 2019). In fact, the limited availability of occupational dictionaries in these countries exacerbates the challenge of correctly measuring the share and the types of jobs which can be done from home.

In this paper, I examine the share of jobs which can be done from home in developing countries, as well as the worker characteristics associated with such jobs. I take advantage of worker-level data on task content from the Skills Toward Employability and Productivity (STEP) survey, which follows workers in urban areas across ten low- and middle-income countries, including Armenia, Bolivia, Yunnan Province in China, Colombia, Georgia, Ghana, Kenya, Laos, Macedonia and Vietnam. STEP covers comparable information on employment outcomes across countries, including details on workers' occupations, formal employment status, and tasks performed at work. As a result, I can examine within-occupation heterogeneity in workers' ability to work from home (Autor and Handel, 2013). Furthermore, STEP includes detailed information on observable characteristics, including educational attainment, gender and an asset index, allowing me to consider which workers are less likely to be able to work from home.

As a first approximation to the feasibility of working from home, I consider workers who report using a computer at work. While 29.5% of workers in the sample use a computer at work, there are important cross-country differences, reaching 58% of workers in Macedonia, compared to just 14% of their counterparts in Laos. My preferred work-from-home definition follows Dingel and Neiman (2020). It rules out working from home if workers either (not) use a computer at work, lift heavy objects, repair electronic equipment, operate heavy machinery or report that customer interaction is very important. 12.9% of workers in the full STEP sample are deemed to be able to work from home, which is far below the corresponding share in the U.S. (which equals 34% according to Dingel and Neiman (2020)).[1] I remark important cross-country differences in this measure, as well, as the share of individuals who may work from home reaches 23% in Yunnan compared to just 5.5% of workers in Ghana. Lastly, across all countries, the likelihood of working from home is positively correlated with high-paying occupations, such as managers and professionals, yet a high share of workers in clerical jobs may be able to do so, too.

I further examine the characteristics of workers who may be able to work from home. Across all STEP countries, workers' educational attainment, household wealth, gender and age strongly

[1] The analysis is restricted to urban areas in these countries, which likely overstates the share of jobs which can be done from home in the ten STEP countries.

Covid Economics 6, 17 April 2020: 104-118

predict their ability to work from home. For instance, the sample average indicates that just 4% of high school dropouts can work from home, compared to 24% of their more educated peers. I further document an important wealth gradient, as only 2.8% of households in the within-country bottom wealth quintile can work from home, compared to 25% of their counterparts in the top quintile, thus highlighting the challenge faced by households with limited access to self-insurance. Informal and self-employed workers are more vulnerable to the consequences from COVID-19 across all countries in STEP, as well. These relationships remain significant in a regression of worker-level work-from-home measures against all observed characteristics. To examine whether these patterns remain within narrowly-defined (three digit) occupations, I include occupation fixed effects and find that while occupations are important determinants of the feasibility of working from home in developing countries, vulnerable workers are less likely to do so even within occupations. Since the estimated patterns vary across countries, I remark the importance of rapidly identifying vulnerable workers and designing country-specific policies to limit the negative labor market impacts arising from the spread of COVID-19.

This paper makes various contributions to the nascent literature on workers' ability to work from home. To the best of my knowledge, this is the first paper to provide directly comparable evidence across countries on the feasibility of working from home. Moreover, by using worker-level data on task content, I can examine the within-occupation determinants of telework. This work thus fits in with two recent papers on the feasibility of working from home in the United States, which measure the extent of low work-from-home occupations in the United States along with worker characteristics associated with these jobs (Dingel and Neiman, 2020; Mongey and Weinberg, 2020).[2] By taking advantage of worker-level data on task content, this paper contributes to the existing literature highlighting within-occupation differences in the importance of tasks (Autor and Handel, 2013; Stinebrickner et al., 2018, 2019). I lastly contribute to recent work which presents comparable evidence on the importance of tasks across countries (Dicarlo et al., 2016; Lo Bello et al., 2019; Lewandowski et al., 2019).

The rest of the paper proceeds as follows. Section 2 describes data sources and presents summary statistics. Section 3 presents evidence on the share of jobs that can be done from home, along with the worker characteristics associated with the possibility of home-based work. In Section 4, I discuss the results and conclude.

2 Data Sources and Summary Statistics

I take advantage of data from the Skills Toward Employment and Productivity (STEP) household survey, conducted in urban areas in developing countries by the World Bank. I use information from the first and second survey rounds, which covered workers in ten developing countries, including Armenia, Bolivia, Yunnan Province in China, Colombia, Georgia, Ghana, Kenya, Laos, Macedonia

[2] Leibovici et al. (2020) also provide important evidence on the extent of contact-intensive tasks in the United States by occupation. More broadly, Mas and Pallais (2017) present evidence on the the prevalence of flexible work arrangements.

Covid Economics 6, 17 April 2020: 104-118

and Vietnam.[3] STEP surveys are representative of the working age population in urban areas in these countries, yet collect demographic information on all individuals in the household along with detailed information on employment outcomes for a randomly selected 15-64 year old household member. Since STEP collects detailed information on dwelling characteristics and household assets to construct a wealth-based asset index in each country (Pierre et al., 2014), I consider their quintile ranking in the asset index distribution as a measure of their capacity to cope with the shock.

I observe the main respondent's age, gender and educational attainment, along with detailed information on their employment outcomes, including whether they are currently employed and/or have been in the past twelve months. Respondents report whether they last worked as employees, in self-employment or as unpaid workers in the family business and whether they are employed in a formal or informal job, as defined by the presence of pension benefits.[4] Lastly, STEP includes information on workers' occupations at the one-digit level, encompassing managers, plant workers, among others, and the three-digit level under the harmonized ISCO-08 classification. I restrict the analysis to workers who have been employed in the past twelve months and drop those in unpaid family work or in the armed forces, given the lack of information on their occupational outcomes.

To measure the feasibility of working from home, I take advantage of worker-level data on the tasks performed at work (Dingel and Neiman, 2020; Mongey and Weinberg, 2020). As remarked by Dicarlo et al. (2016), Lo Bello et al. (2019) and Saltiel (2019), an important advantage of the STEP survey is the availability of worker-level task content, which allows me to examine within-occupation heterogeneity in workers' capacity to work from home. Moreover, since all STEP surveys include the same task content questions, the analysis is directly comparable across countries. The first measure of the feasibility of working from home follows directly from a binary response to whether workers use a computer at work. My preferred definition follows Dingel and Neiman (2020) and rules out working from home if workers report performing *either* of the following tasks at work: not using a computer, lifting anything heavier than 50 pounds, repairing/maintaining electronic equipment, operating heavy machinery or industrial equipment, or reporting that contact with customers is very important. Dingel and Neiman (2020) define an occupation as one that cannot be done from home if one of eighteen conditions holds true. Moreover, they assume that jobs cannot be performed from home unless they use e-mail, whereas my definition for developing countries focuses on computer use. As such, my definition could be considered as an upper bound on the share of jobs which can be done from home in these countries, yet it is possible that certain home-based jobs may not require a computer; such as workers in call centers. Since this paper is a first approximation on the feasibility of working from home in developing countries, I remark the need for further work in refining this definition.

[3] STEP also conducted surveys in Sri Lanka and Ukraine. I exclude Sri Lanka from the analysis for comparability, as sample largely covered workers in rural areas. Meanwhile, the Ukraine sample does not include information on workers' three-digit occupations. Throughout the paper, the discussion of results in China is limited to the Yunnan Province.

[4] I observe formal employment status only for workers employed at the time of the survey.

Summary Statistics. Table 1 presents summary statistics for the sample used in the paper, which includes upwards of 17,000 workers in the full sample. I use sample weights to make the samples representative of the working-age populations in each country.[5] On average, respondents in the STEP survey have completed 10.2 years of education, yet there are important cross-country differences which positively correlate with levels of economic development. In terms of employment outcomes, there is a high prevalence of self-employment and informality in the full sample — reaching 43% and 66% of all workers, respectively — and these measures are negatively correlated with countries' GDP per capita.[6] The task measures indicate that a small share of workers in the full sample either repair electronic equipment or operate heavy machinery, reaching 7% in the full sample However, almost 40% of workers lift heavy items at work and 27% report having frequent interactions with their customers. The prevalence of these tasks also varies across countries, as only 16.7% of Ghanaian workers frequently contact customers compared to 38% of their Colombian counterparts. Lastly, 29.5% of workers in the sample use a computer at work, yet are important differences across countries, which I further explore in the next section.

Table 1: Summary Statistics

	All (1)	Armenia (2)	Bolivia (3)	China (4)	Colombia (5)	Georgia (6)	Ghana (7)	Kenya (8)	Laos (9)	Macedonia (10)	Vietnam (11)
Observables											
Years of Education	10.2	13.787	11.158	12.634	10.052	15.237	8.343	9.161	8.379	13.561	11.092
Male	0.467	0.412	0.461	0.522	0.46	0.395	0.438	0.55	0.418	0.522	0.432
Age	36.111	40.283	34.137	39.506	35.135	39.761	36.414	32.324	38.31	41.413	39.479
Employment Outcomes											
Self-Employed	0.427	0.104	0.419	0.137	0.45	0.139	0.612	0.396	0.636	0.147	0.398
Informal	0.661	0.173	0.771	0.456	0.633	0.611	0.836	0.743	0.855	0.142	0.58
Tasks											
Repair Items	0.073	0.257	0.063	0.184	0.053	0.063	0.061	0.067	0.023	0.09	0.086
Operate Machinery	0.074	0.06	0.079	0.054	0.098	0.064	0.054	0.071	0.041	0.084	0.053
Lift Heavy	0.384	0.268	0.45	0.24	0.393	0.275	0.477	0.374	0.532	0.254	0.324
Contact Others	0.273	0.353	0.223	0.3	0.38	0.361	0.167	0.319	0.109	0.407	0.087
Computer Use	0.295	0.431	0.316	0.545	0.324	0.44	0.112	0.225	0.14	0.577	0.34
Observations	17,616	1,092	1,848	1,298	1,930	1,076	2,223	2,529	1,420	1,923	2,277

Source: Skills Toward Employability and Productivity (STEP) Survey.
Note: Table 1 presents summary statistics for the full sample and for each country using sample weights to represent the working-age population of 15-64 year olds. The 'Tasks' row indicates the share of workers in each country who perform the corresponding job activities at work.

[5] Summary statistics in the full sample are weighted by each country's population of 15-64 year olds. As such, these results give greater weight to countries with larger populations.

[6] While there is significant cross-country heterogeneity in these two measures, I remark that part of the difference in the informality share may be driven by its definition on the lack of pension benefits, which may vary across countries due to government policies (Maloney, 2004).

Covid Economics 6, 17 April 2020: 104-118

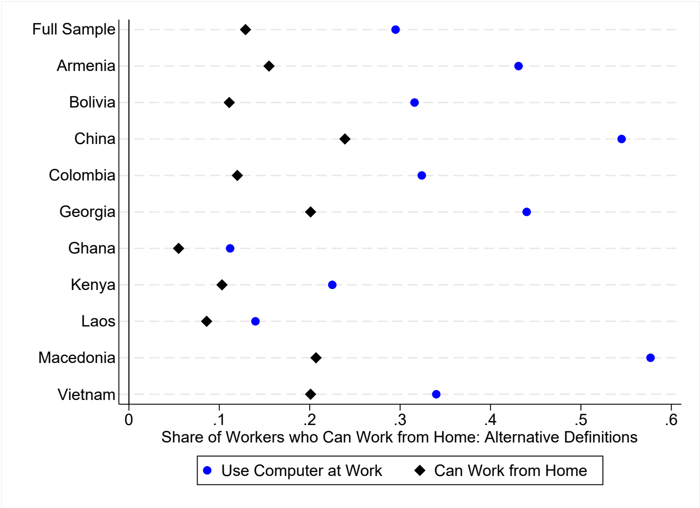
CEPR PRESS

3 Empirical Evidence

3.1 Descriptive Analysis

Figure 1 presents graphical evidence on the share of workers who use a computer at work and on those who can work from home across STEP countries. There is substantial heterogeneity in computer usage in the sample, as 58% of Macedonian workers report using a computer at work, compared to just 11% and 14% of workers in Ghana and Laos, respectively. The share of workers who could possibly work from home equals 12.9% in the full sample, which is far lower than the share of workers using computers due to additional task-based restrictions defined in Section 2. I similarly find important differences in the feasibility of working from home across countries, as 23% of workers in the Yunnan Province in China may do so, compared to just 5.5% of their counterparts in Ghana. All in all, the feasibility of working from home is strongly correlated with GDP per capita, as the correlation equals 0.684. As such, despite the conservative work-from-home definition adopted in this paper, the share of workers who may work from home in the STEP countries is far behind the corresponding share in the United States, which equals 34% (Dingel and Neiman, 2020). I further remark that since the STEP survey is representative of urban areas, the share of work which can be done from home in these countries is likely substantially lower, as agricultural employment is far more prevalent in rural areas.

Figure 1: Share of Jobs Which Can be Done from Home, by Country



Source: Skills Toward Employability and Productivity (STEP) Survey.
Note: Figure 1 presents the share of workers who use a computer at work as well as those who are defined to be able to work from home (Section 2) by country. Results are weighted using sample weights to represent the working-age population of 15-64 year olds.

In Table 2, I extend the analysis to consider heterogeneity in the share of individuals who

Covid Economics 6, 17 April 2020: 104-118

Covid Economics 6, 17 April 2020: 104-118

can work from home across one-digit occupations.[7] Unsurprisingly, the feasibility of working from home is positively correlated with occupation-level wages. In the full sample, only 0.5% of plant and machine operators could potentially work from home, compared to 34.4% of workers in 'Professional' jobs. The feasibility of working from home for managers is not higher than for those in professional occupations due to the prevalence of interactive tasks for managers. Interestingly, 41% of clerical occupations may be done from home, which follows from the lack of physical and interactive tasks involved in these jobs. Meanwhile, just 6.5% of workers employed in services and sales occupations may work from home.

There are important differences in the feasibility of working from home across countries in high-paying occupations. For instance, 14% of managers in Bolivia may do so, compared to 60% of their peers in Vietnam. Similarly, just 18% of professionals in Armenia can work from home, relative to 39% of their counterparts in Laos. Nonetheless, these differences are largely absent in lower-paying jobs, as at most 2.4% of machine operators are deemed to be able to work from home in STEP countries. Similar patterns emerge for workers in services/sales, crafts and trades and those in elementary occupations. Remarkably, work-from-home patterns at the one-digit occupation level are similar across all STEP countries. Dingel and Neiman (2020) document similar patterns in the United States, albeit with important differences in the prevalence of working from home. For instance, the authors document that 84% of jobs in management occupations can be performed at home in the U.S., compared to just 34% in the full STEP sample. As a result, this analysis further highlights that the economic costs from the pandemic in developing countries may far exceed those in the developed world.

Table 2: Share of Work-from-Home by One-Digit Occupation and Country

	All	Armenia	Bolivia	China	Colombia	Georgia	Ghana	Kenya	Laos	Macedonia	Vietnam
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Manager	0.34	0.316	0.142	0.237	0.338	0.339	0.281	0.349	0.298	0.388	0.604
Professional	0.344	0.177	0.283	0.374	0.325	0.271	0.223	0.317	0.39	0.334	0.523
Technician	0.274	0.228	0.271	0.378	0.132	0.303	0.272	0.357	0.372	0.314	0.463
Clerical	0.418	0.461	0.438	0.494	0.376	0.474	0.335	0.465	0.611	0.369	0.424
Services/Sales	0.064	0.042	0.044	0.128	0.103	0.092	0.008	0.023	0.029	0.131	0.086
Agricultural	0.001	0	0	0	0.	0	0	0	0.005	0.019	0
Craft/Trades	0.033	0.052	0.026	0.072	0.056	0.02	0.009	0.01	0	0.05	0.022
Machine Operators	0.005	0.024	0.001	0.005	0.006	0	0.005	0	0	0.011	0.005
Elementary Occupations	0.023	0.013	0.021	0.154	0.02	0.017	0	0.004	0	0.021	0.024

Source: Skills Toward Employability and Productivity (STEP) Survey.
Note: Table 2 presents evidence on the share of workers who can work from home by one-digit occupation and country. Results are weighted using sample weights to represent the working-age population of 15-64 year olds.

3.2 Characteristics of Work-from-Home Workers

Despite the low prevalence of work-from-home jobs in developing countries, examining the characteristics of individuals who cannot work from home can help governments to target policies towards

[7] I present graphical evidence across three one-digit occupations in Figure A1.

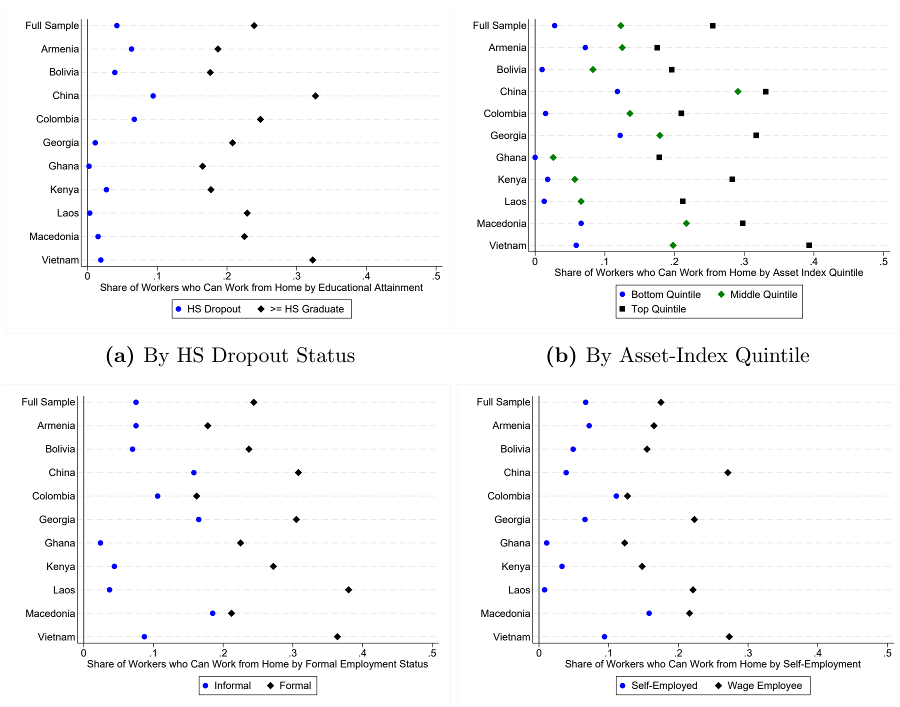
particularly affected workers. In the first panel of Figure 2, I present evidence by educational attainment. In the full STEP sample, just 4.2% of high school dropouts are able to work from home, compared to almost 24% of their peers who at least completed a high school degree. Mongey and Weinberg (2020) similarly show that workers with less than a college degree in the U.S. are 33 percentage points more likely to work in a low work-from-home occupation vis-a-vis their more educated peers. Educational differences in this dimension are present in all countries in the STEP sample — while 9.4% of high school dropouts in Yunnan can work from home, this share is far behind that of their more educated counterparts (exceeding 32%). In the second panel, I present heterogeneity across the within-country asset index distribution. The sample average shows that just 2.8% of households in the bottom quintile can work from home, far trailing their wealthier peers in the top quintile at 25.5%. While the top-bottom quintile asset gap in this measure exceeds 30 percentage points in Vietnam and equals 10 percentage points in Armenia, it is present in all countries in the sample.[8] As such, I remark the challenge facing governments in developing policies for helping households with limited access to self-insurance through their current assets.[9]

Covid Economics 6, 17 April 2020: 104-118

[8]While Mongey and Weinberg (2020) do not directly observe information on households' assets, they find that workers below the median income are far more likely to work in low work-from-home occupations in the U.S.

[9]In Figure A2, I further show that males are less likely to be able to work from home than females, reaching 10.9% and 14.6% of the sample considered in the paper, respectively. Similarly, the feasibility of working from home for workers over 40 is lower (10.7%) than for their younger counterparts in STEP countries.

Figure 2: Characteristics of Work-From-Home Workers, by Country



(a) By HS Dropout Status

(b) By Asset-Index Quintile

(c) By Formal Employment Status

(d) By Self-Employment

Source: Skills Toward Employability and Productivity (STEP) Survey.
Note: Figure 2 presents the share of jobs which can be done from home in the full sample and across STEP countries by workers' high school dropout status (Panel A), asset index quintile (Panel B), formal employment status (Panel C) and self-employment status (Panel D). Results are weighted using sample weights to represent the working-age population of 15-64 year olds.

I additionally consider differences by employment outcomes. In Panel C, I show that just 7.5% of informal sector workers may carry on with their jobs at home, far below the corresponding share (24.4%) for their formal sector counterparts. This result holds across all countries, though the extent of the differences varies in magnitude. While the differences across self-employment status are smaller in magnitude (Panel D), the work-from-home gap is also present in this dimension across all STEP countries. To further understand which observed characteristics drive the relationships presented in Figure 2, I estimate the following OLS regression:

$$NWFH_{ij} = \beta_0 + \beta_1 \boldsymbol{X_i} + v_{ij} \tag{1}$$

where $NWFH_{ij}$ is a binary variable which equals 1 if worker i in occupation j **cannot** work from home and $\boldsymbol{X_i}$ includes binary variables measuring whether worker i is a high school dropout,

Covid Economics 6, 17 April 2020: 104-118

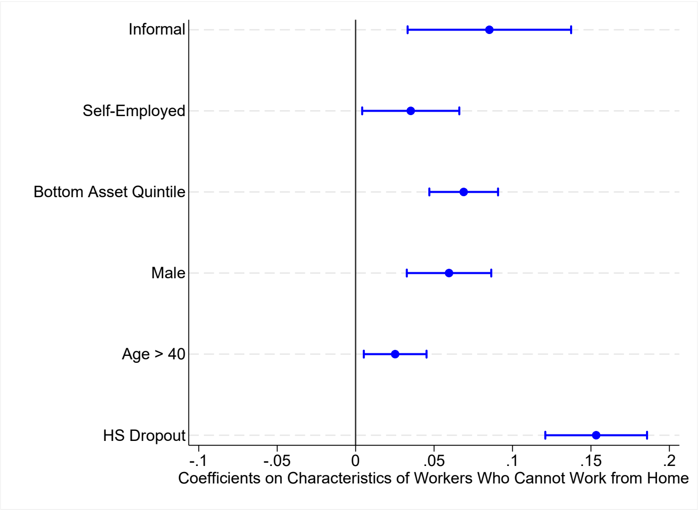
CEPR PRESS

Covid Economics 6, 17 April 2020: 104-118

male, older than 40, in the bottom asset quintile, self-employed, and/or in informal employment. I estimate equation (1) separately by country and also for the full sample, which includes country fixed effects.

In Figure 3, I present the results from equation (1). Fitting in with the evidence presented in Figure 2, the full sample results indicate that high school dropouts are 15.3 percentage points less likely to be able to work from home vis-a-vis their more educated peers, conditional on observed characteristics. Similarly, informality is associated with a lower likelihood of working from home by 8.5 percentage points, yet the estimated gap is about half as in the results from the bivariate regressions presented above. I further find that vulnerable groups are less likely to work from home, including workers older than 40, those in the bottom quintile of the asset distribution and self-employed individuals. I lastly find significant gender gaps in the feasibility of working from home, driven by the prevalence of physical tasks for males in developing countries. Mongey and Weinberg (2020) document similar gender differences in the United States.

Figure 3: Worker Characteristics Associated with **Not** Working from Home, Full Sample



Source: Skills Toward Employability and Productivity (STEP) Survey.
Note: Figure 3 presents the estimated coefficients from equation (1) for the full sample, including country fixed effects. Results are weighted using sample weights to represent the working-age population of 15-64 year olds. Standard errors clustered at the country level. 90% confidence intervals reported in solid lines.

In Table 3, I present results from equation (1) separately by country. The sign of the estimated coefficients largely follow the full sample results presented above, yet there are important cross-country differences. For instance, conditional on other observables, high school dropouts are 22 percentage points less likely to work from home than their more educated peers in Vietnam, yet the corresponding difference in Kenya equals 8 percentage points. The estimated work-from-home asset gap remains across all countries, yet it is not significant in Ghana and Laos. Similarly,

Covid Economics 6, 17 April 2020: 104-118

while the gap across formal employment status remains in all countries (except for Macedonia), the difference is significant in just six countries in the sample. The cross-country heterogeneity in workers' capacity to work from home across observed characteristics highlights the importance of developing country-specific evidence regarding on the most vulnerable workers, to correctly design policies aimed at these workers.

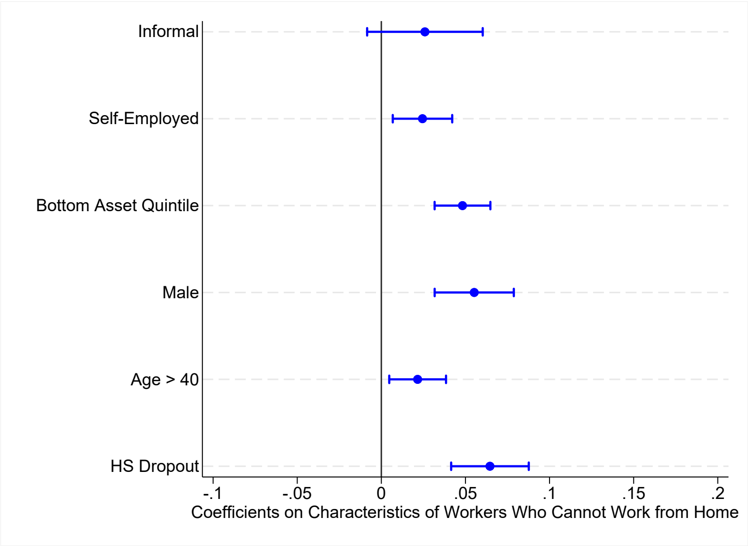
Table 3: Worker Characteristics Associated with **Not** Working from Home, By Country

	Armenia (1)	Bolivia (2)	China (3)	Colombia (4)	Georgia (5)	Ghana (6)	Kenya (7)	Laos (8)	Macedonia (9)	Vietnam (10)
HS Dropout	0.103***	0.095***	0.162***	0.156***	0.182**	0.113***	0.083***	0.121***	0.170***	0.220***
	(0.028)	(0.015)	(0.027)	(0.018)	(0.074)	(0.012)	(0.013)	(0.016)	(0.035)	(0.018)
Age > 40	0.089***	-0.001	-0.004	0.044***	0.058**	0.005	-0.008	0.040***	0.017	0.067***
	(0.023)	(0.016)	(0.024)	(0.016)	(0.027)	(0.010)	(0.014)	(0.014)	(0.019)	(0.016)
Male	0.034	0.020	0.096***	0.078***	0.073***	0.006	0.015	0.020	0.072***	0.106***
	(0.023)	(0.015)	(0.023)	(0.016)	(0.028)	(0.010)	(0.012)	(0.014)	(0.019)	(0.016)
Bottom Asset Quintile	0.086***	0.083***	0.071**	0.097***	0.088**	0.013	0.064***	0.009	0.126***	0.085***
	(0.030)	(0.018)	(0.031)	(0.020)	(0.035)	(0.012)	(0.015)	(0.016)	(0.029)	(0.020)
Self-Employed	0.023	0.063***	0.172***	-0.009	0.103**	0.031**	0.046***	0.081***	0.059**	0.060***
	(0.045)	(0.017)	(0.035)	(0.020)	(0.041)	(0.012)	(0.013)	(0.017)	(0.030)	(0.019)
Informal	0.053	0.094***	0.036	0.034	0.097***	0.117***	0.170***	0.210***	-0.036	0.134***
	(0.038)	(0.020)	(0.026)	(0.021)	(0.029)	(0.016)	(0.016)	(0.023)	(0.031)	(0.020)
Observations	995	1691	1239	1695	927	2074	2327	1390	1806	2176
R^2	0.053	0.099	0.104	0.096	0.063	0.158	0.138	0.255	0.043	0.220

Source: Skills Toward Employability and Productivity (STEP) Survey.
Note: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Table 3 presents the estimated coefficients from equation (1) separately for each country in the sample. Results are weighted using sample weights to represent the working-age population of 15-64 year olds.

As discussed above, by observing worker-level task information, I can further examine differences in workers' capacity to work from home *within* occupations in each country. To understand the role of occupations in determining the feasibility of working from home in developing countries, I re-estimate equation (1) including three-digit occupation fixed effects, which encompass 140 occupations. I present the estimated results for the full sample in Figure 4. The estimated signs on workers' observed characteristics directly follow the results presented in Figure 3, which indicate that high school dropouts, those in less wealthy households, males, older workers and self-employed workers are less likely to be able to work from home in the full sample. Nonetheless, the estimated coefficients are muted relative to the results presented above. For instance, within three-digit occupations, high school dropouts are 6.5 percentage points less likely to work from home vis-a-vis their more educated peers, yet the estimated coefficient is 60% lower than in the results from equation (1). Similarly, the estimated informality-work-from-home gap falls to 2.6 percentage points and is no longer statistically significant at the 10% level. As such, occupations partly explain the feasibility of working from home in developing countries, remarking their importance for structuring employment outcomes (Autor and Handel, 2013). Nonetheless, the remaining variation unaccounted by occupational characteristics highlights the importance of considering worker-level information in this context.

Figure 4: Within-Occupation Worker Characteristics Associated with **Not** Working from Home, Full Sample



Source: Skills Toward Employability and Productivity (STEP) Survey.
Note: Figure 4 presents the estimated coefficients from equation (1) for the full sample, including country fixed effects and three-digit occupation fixed effects. Results are weighted using sample weights to represent the working-age population of 15-64 year olds. Standard errors clustered at the country and three-digit occupation levels. 90% confidence intervals reported in solid lines.

In Table A1, I present the estimates from equation (1) with occupation fixed effects separately for each country. As with the results for the full sample, the estimated coefficients on workers' observed characteristics are significantly smaller and of varying statistical significance. Across all countries, high school dropouts are less likely to be able to work from home, and the coefficient remains significant in all but three countries in the sample. The relationship between household assets and work-from-home status remains negative across all countries even within occupations, yet the result is not significant in Georgia, Ghana and Laos. Lastly, within-occupation differences across formal employment status are similarly narrowed, as the coefficient is positive and significant in just two countries in the sample. All in all, these results indicate that more vulnerable workers are far less likely to continue working from home, thus likely suffering the worst labor market impacts from the pandemic. Since the estimated relationships vary across developing countries, identifying vulnerable workers and developing policies for correctly reaching them should be a priority for governments in these countries.

Covid Economics 6, 17 April 2020: 104-118

Covid Economics 6, 17 April 2020: 104-118

4 Conclusion

Social distancing and stay-at-home policies will play a critical role in stopping the spread of COVID-19. In this context, the negative employment impacts arising from the virus may be muted if workers are able to perform their jobs at home. In this paper, I have considered the feasibility of working from home in developing countries. As discussed above, measuring the number of jobs that can be done at home in these countries is challenging due to the lack of occupational dictionaries which map tasks to occupations. To overcome this limitation, I have relied on data from the World Bank's STEP survey, which has the advantage of measuring task content at the worker level across ten developing countries. The results presented in this paper indicate that a small share of workers in these countries may feasibly carry out their jobs from home, ranging from 6% in Ghana to 23% in Yunnan (China). Moreover, various vulnerable groups are more likely to suffer the negative consequences of the virus as their jobs cannot be done from home, including high school dropouts, informal workers and those in low-asset households. On the other hand, the extent of these relationships varies across countries, remarking the need for governments to identify vulnerable workers and create policies aimed at lessening the negative effects arising from COVID-19.

References

Autor, D. H. and M. J. Handel (2013). Putting tasks to the test: Human capital, job tasks, and wages. *Journal of Labor Economics 31*(S1), S59–S96.

Dicarlo, E., S. L. Bello, S. Monroy-Taborda, A. M. Oviedo, M. L. Sanchez-Puerta, and I. Santos (2016). The Skill Content of Occupations Across Low and Middle Income Countries: Evidence from Harmonized Data. *IZA DP No. 10224*.

Dingel, J. and B. Neiman (2020). How Many Jobs Can be Done at Home? *BFI White Paper, Becker Friedman Institute*.

Leibovici, F., A. M. Santacrue, and M. Famiglietti (2020). Social Distancing and Contact-Intensive Occupations. *Technical Report, St. Louis Federal Reserve Bank. On the Economy Blog*.

Lewandowski, P., A. Park, W. Hardy, and Y. Du (2019). Technology, skills, and globalization: Explaining international differences in routine and nonroutine work using survey data. *IZA DP No. 12339*.

Lo Bello, S., M. L. Sanchez Puerta, and H. Winkler (2019). *From Ghana to America: The Skill Content of Jobs and Economic Development.*

Maloney, W. F. (2004). Informality Revisited. *World Development 32*(7), 1159–1178.

Mas, A. and A. Pallais (2017). Valuing Alternative Work Arrangements. *American Economic Review 107*(12), 3722–59.

Mongey, S. and A. Weinberg (2020). Characteristics of Workers in Low Work-From-Home and High Personal-Proximity Occupations. *BFI White Paper, Becker Friedman Institute*.

Pierre, G., M. L. Sanchez Puerta, A. Valerio, and T. Rajadel (2014). STEP Skills Measurement Surveys: Innovative Tools for Assessing Skills.

Saltiel, F. (2019). Comparative Evidence on the Returns of Tasks in Developing Countries. *mimeo*.

Stinebrickner, R., T. Stinebrickner, and P. Sullivan (2019). Job Tasks, Time Allocation, and Wages. *Journal of Labor Economics 37*(2), 399–433.

Stinebrickner, T. R., R. Stinebrickner, and P. J. Sullivan (2018). Job Tasks and the Gender Wage Gap among College Graduates. Technical report, National Bureau of Economic Research.

Appendices

A Tables and Figures

Table A1: Within-Occupation Worker Characteristics Associated with **Not** Working from Home, By Country

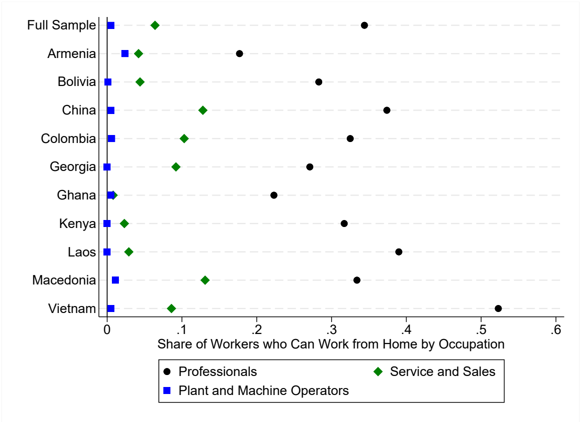
	Armenia (1)	Bolivia (2)	China (3)	Colombia (4)	Georgia (5)	Ghana (6)	Kenya (7)	Laos (8)	Macedonia (9)	Vietnam (10)
HS Dropout	0.005	0.010	0.102***	0.072***	0.092**	0.031**	0.018	0.043**	0.040**	0.085***
	(0.027)	(0.015)	(0.034)	(0.023)	(0.042)	(0.012)	(0.015)	(0.018)	(0.018)	(0.016)
Age > 40	0.026	-0.003	-0.000	0.041*	0.040	0.002	-0.009	0.031	0.023	0.046***
	(0.023)	(0.020)	(0.027)	(0.024)	(0.037)	(0.009)	(0.012)	(0.022)	(0.021)	(0.018)
Male	0.009	0.020	0.080**	0.062***	0.110**	0.008	0.009	0.026	0.075**	0.076***
	(0.038)	(0.021)	(0.034)	(0.023)	(0.044)	(0.019)	(0.015)	(0.027)	(0.031)	(0.020)
Bottom Asset Quintile	0.087***	0.059***	0.070***	0.069***	0.051	0.009	0.032***	0.004	0.036**	0.040***
	(0.027)	(0.019)	(0.022)	(0.018)	(0.032)	(0.008)	(0.011)	(0.011)	(0.017)	(0.014)
Self-Employed	0.020	0.015	0.115***	0.009	0.071	0.030*	0.022*	0.084*	0.098**	-0.005
	(0.063)	(0.025)	(0.027)	(0.025)	(0.045)	(0.017)	(0.013)	(0.048)	(0.042)	(0.036)
Informal	0.025	-0.007	0.028	-0.014	0.047	0.047	0.080***	0.093	-0.085**	0.049*
	(0.045)	(0.033)	(0.020)	(0.026)	(0.039)	(0.042)	(0.030)	(0.075)	(0.038)	(0.028)
Observations	995	1691	1239	1695	927	2074	2327	1390	1806	2176
R^2	0.280	0.347	0.267	0.300	0.313	0.445	0.448	0.473	0.316	0.425

Source: Skills Toward Employability and Productivity (STEP) Survey.
Note: Standard errors in parentheses, clustered at the three-digit occupation level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Table A1 presents the estimated coefficients from equation (1) with three-digit occupation fixed effects separately for each country in the sample. Results are weighted using sample weights to represent the working-age population of 15-64 year olds.

Covid Economics 6, 17 April 2020: 104-118

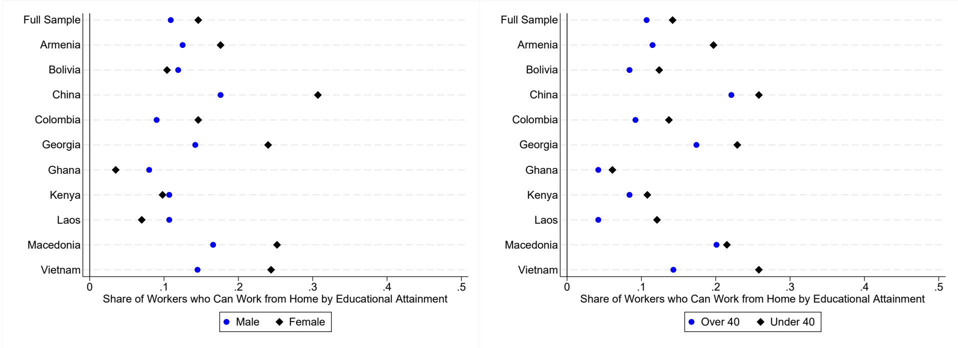
Covid Economics 6, 17 April 2020: 104-118

Figure A1: Share of Jobs Which Can be Done from Home by Occupation and Country



Source: Skills Toward Employability and Productivity (STEP) Survey.
Note: Figure A1 presents the share of workers in Professional, Service/Sales and Plant/Machine Operator occupations who are defined to be able to work from home by country. Results are weighted using sample weights to represent the working-age population of 15-64 year olds.

Figure A2: Characteristics of Work-From-Home Workers, by Country



(a) By Gender **(b)** By Age

Source: Skills Toward Employability and Productivity (STEP) Survey.
Note: Figure A2 presents the share of jobs which can be done from home in the full sample and across STEP countries by workers' gender (Panel A) and age (Panel B). Results are weighted using sample weights to represent the working-age population of 15-64 year olds.

CEPR PRESS

Covid Economics **Issue 6, 17 April 2020**

The invisible risk: Pandemics and the financial markets

Jordan Schoenfeld[1]

Date submitted: 8 April 2020; **Date accepted:** 11 April 2020

Are pandemics systemically important to modern-day financial markets? This study uses the COVID-19 pandemic as a natural experiment for testing how large-scale pandemics affect the financial markets. Using hand-collected data at the firm level, I find that managers systematically underestimated their exposure to pandemics in their SEC-mandated risk factors, and the vast majority of firms decreased in value at the pandemic's onset. I also find that the pandemic triggered unprecedented changes in U.S. employment levels and the values of bonds, commodities, and currencies. These types of findings suggest that pandemics are systemically important to the financial markets. Overall, this study provides some of the first large-scale evidence on how pandemics affect the financial markets.

Covid Economics 6, 17 April 2020: 119-136

1 Associate Professor, Tuck School of Business, Dartmouth College.

Covid Economics 6, 17 April 2020: 119-136

1 Introduction

The novel coronavirus (COVID-19) has disrupted supply chains, consumer spending, and social mores. As a result, governments and central banks have drastically stepped up their operations, and many economists are forecasting a deep coronavirus-induced recession (e.g., Schwartz, 2020). However, we lack answers to some of the most basic questions about how pandemics affect the financial markets. Foremost, are pandemics systemically important to the financial markets? If so, in what ways do pandemics pose a serious risk to the financial system when they materialize? Indeed, Reinhart (2020) argues that there is no "historical episode that can provide any insight as to the likely economic consequences of the unfolding global coronavirus crisis." This study therefore performs one of the first large-scale empirical analyses of the pandemic's impact on the financial markets.

Specifically, I ask whether managers anticipated their exposure to the pandemic, and whether (and to what extent) the pandemic affected U.S. employment levels and the value of important assets such as stocks, bonds, commodities, and currencies. Theoretically, a pandemic potentially represents a significant demand and supply shock, and the nature of a pandemic's financial impact is multifaceted and ambiguous. A pandemic may affect supply chains, production processes, workers, and consumers. In contrast, some firms may see heightened demand for their products during a pandemic. Some managers may also hedge or insure against their pandemic exposure through operational or financial strategies. Governments may also be able to mitigate any economic effect of a pandemic. The COVID-19 pandemic is an ideal laboratory in which to examine these issues.[1]

I begin by examining whether managers anticipated their exposure to pandemics. The Securities and Exchange Commission (SEC) mandates that public firms include key business risk factors in "Item 1A" of their annual reports. To examine the extent to which managers anticipated their exposure to pandemics, I hand collect and read the annual reports for all

[1]Prior disease outbreaks were much less widespread than COVID-19. For example, as of early March, 2020, COVID-19 had infected more than fifty times as many people as the SARS outbreak, and the situation is expected to worsen (Abadi et al., 2020).

firms in the S&P 500 as of January, 2020.[2] I find that about 46 percent of these 500 firms included pandemics (or related risks such as diseases or health crises) in their business risk factors leading up to the COVID-19's onset. To illustrate, Norwegian Cruise Lines includes the following in its business risk factors:

> ***Epidemics and viral outbreaks could have an adverse effect on our business, financial condition and results of operations.***
>
> Public perception about the safety of travel and adverse publicity related to passenger or crew illness, such as incidents of viral illnesses, stomach flu or other contagious diseases, may impact demand for cruises and result in cruise cancellations and employee absenteeism. If any wide-ranging health scare should occur, our business, financial condition and results of operations would likely be adversely affected.

Given that 46 percent of the firms in S&P 500 include pandemics in their business risk factors, any financial impact from a pandemic should, in theory, be largely contained to these exposed firms. However, about 95 percent of the S&P 500 firms decreased in value from January to March, 2020. In fact, the 54 percent of firms that did *not* include pandemics in their risk factors exhibited a mean decrease in value of 32 percent from January to March, 2020. The magnitude and sign of this result almost exactly equal the 32 percent mean decrease in value exhibited by firms that *did* include pandemics in their risk factors. On aggregate, these findings translate to an economic loss of $18 billion per firm on average, or about $9 trillion in total for S&P 500 firms alone.[3]

The above findings suggest that managers systematically underestimated their exposure to pandemics, and that with respect to pandemics, firms' business risk factors had a significant blind spot. In addition, whether a firm included pandemics as a business risk factor does not significantly correlate with (or predict) the signed magnitude of its change in value

[2]Focusing on these firms curbs the labor required to collect and read firms' annual reports, which is considerable. Firms in the S&P 500 accounted for about 83 percent of total market capitalization as of January, 2020. See Section 2 for more detail on the data-collection procedure and why textual analysis techniques are problematic in this setting.

[3]As Section 2 explains, I use buy and hold returns from January 2 to March 20, 2020, to mitigate any effect of subsequent fiscal policy and other government reactions to pandemics. The results are robust to several alternative buy and hold periods (see Section 3 for these results).

Covid Economics 6, 17 April 2020: 119-136

Covid Economics 6, 17 April 2020: 119-136

at the onset of the COVID-19 pandemic. This null result also obtains for return volatility at the pandemic's onset, and for measures akin to return volatility such as unsigned changes in firm value, and after controlling for a firm's industry, size, and other attributes. Thus, investors could not have deduced which stocks would be exposed to the COVID-19 pandemic based on firms' business risk factors.

At the industry level, I observe the largest decreases in firm value for petroleum and natural gas firms, apparel firms, restaurant and hotel firms, automobile firms, transportation firms, machinery firms, and aircraft, ship, and railroad firms. By contrast, I observe smaller decreases in firm value for food product firms, healthcare firms, utility firms, and business services firms. The firms with the largest decreases in value at the pandemic's onset include Norwegian Cruise Lines (−86.0 percent change in value), Noble Energy (−83.7 percent), Royal Caribbean Cruises (−83.4 percent), Halliburton (−80.6 percent), and Carnival (−80.5 percent). A small set of firms also increased in value at the pandemic's onset, suggesting that investors might be expecting increased demand for these firms' goods and services. This group of firms includes Walmart (+0.4 percent), General Mills (+3.0 percent), Netflix (+0.6 percent), Clorox (+26.0 percent), and Regeneron Pharmaceuticals (+31.2 percent).

Other asset classes were also affected by the pandemic. Government bonds increased in value as central banks cut benchmark interest rates and unveiled large bond-buying programs (e.g., Smialek and Irwin, 2020). Although the U.S. Federal Reserve also committed to purchasing corporate bonds, the S&P corporate bond index decreased in value, presumably because of heightened credit risk among investors. Many commodities also decreased in value, particularly oil and natural gas. Gold, by contrast, increased in value. The U.S. dollar also strengthened relative to several foreign currencies.

The pandemic also affected the U.S. labor market. According to the Department of Labor (DOL), for the week ending March 21, 2020, seasonally adjusted initial unemployment claims were 3,307,000, an increase of 3,025,000 from the previous week's level of 282,000. This was the highest level ever for such claims, and that number was surpassed in the week

Covid Economics 6, 17 April 2020: 119-136

ending March 28, 2020, when it reached 6,480,000. California, Washington, and Nevada experienced some of the largest numbers of claims, and many states directly attributed the increase in their unemployment claims to COVID-19-related causes. In a Goldman Sachs national survey of small business owners, more than half said they would be forced out of business in less than three months under current conditions.[4]

Overall, this study is one of the first to examine empirically how the COVID-19 pandemic affected modern-day financial markets. Using hand-collected data, I find that at the firm level, managers systematically underestimated their exposure to the pandemic, and firm value changes at the pandemic's onset were virtually impossible to predict using firms' business risk factors from their annual reports. The COVID-19 pandemic also triggered significant changes to U.S. employment levels and the values of bonds, commodities, and currencies. These findings supplement prior research that examines the health and economic implications of pandemics at the individual level (e.g., Almond and Mazumder, 2005; Almond, 2006; Karlsson et al., 2014).[5]

The remainder of this study is organized as follows. Section 2 describes the data. Section 3 provides the empirical results. Section 4 concludes.

2 Overview of data

This study focuses mainly on the period of January to March, 2020—the onset of when COVID-19 spread to Europe and North America and the World Health Organization declared the virus a global pandemic. Financial data, stock price quotes, and buy-and-hold returns from commonly used data sources such as the Center for Research in Security Prices (CRSP) are typically not made available in real-time. As a result, I cannot use these datasets and instead hand collect financial data for all the S&P 500 firms as of early 2020. To determine which firms include pandemics as a business risk factor, I hand collect the most recent annual

[4]See `https://www.goldmansachs.com/citizenship/10000-small-businesses/US/no-time-to-waste/index.html`.

[5]See Acemoglu et al. (2003) for a perspective on this topic.

CEPR PRESS

reports for each of these firms as of January 1, 2020. Recall that as of January, 2020, firms in the S&P 500 accounted for about 83 percent of total market capitalization. This index is a good representation of large publicly traded U.S. firms and covers all the Fama-French 30 industries. Focusing on these firms curbs the labor required to collect and read firms' annual reports, which is considerable. After eliminating foreign firms and firms with missing data, the final sample consists of 492 firms.[6]

The above approach of collecting annual reports filed before January 1, 2020, ensures that I am not capturing managers' decision to include pandemics as a business risk factor in response to the COVID-19 pandemic. Most firms will file their 2019 annual report in February and March, 2020, which is after the COVID-19 virus began spreading. In addition, the SEC recently gave some firms extra time to file their 2019 annual report and other documents due to COVID-19 (Kiernan et al., 2020).

The SEC has long mandated that firms list their key business risk factors in "Item 1A" of their annual reports. To identify which firms include pandemic in their business risk factors, I cannot simply use textual analysis to search for terms such as "disease" or "pandemic." This technique would be confounded by pharmaceutical firms that often list business risk factors such as the possibility that a drug trial fails to effectively treat a disease or health issue. In addition, there is no widely agreed upon term for exposure to pandemics and related risks. Some firms denote such risks as "health crises" or "health scares," while others use different terms. I therefore read the risk factors for all the S&P 500 firms whose annual reports I collected, and I then code an indicator variable that equals one for firms that include pandemics (or a related issue) as a business risk factor (zero otherwise).

The week of March 22, 2020, is about when many governments and central banks began introducing and passing large economic stimulus plans in response to the pandemic. Although these plans are ostensibly meant to target the pandemic, the reality is that they also contain many significant and politically charged programs unrelated to the pandemic

[6]Since I require a firm to file an annual report in 2019, firms in the S&P 500 as of early 2020 that did not file an annual report in 2019 are excluded from the analysis.

Covid Economics 6, 17 April 2020: 119-136

(e.g., provisions for real-estate taxes). Thus, for the buy and hold returns analysis, I use asset prices from January 2 to March 20, 2020, and source data from Bloomberg, Calcbench, COMEX, the Intercontinental Exchange, and NYMEX. This approach helps to ensure that my analyses isolate the effect of the pandemic and not subsequent government reactions (Section 3.2 performs robustness checks using several alternative buy and hold windows).

3 Empirical results

3.1 Employment results

According to data from the Department of Labor (DOL), initial unemployment claims in the U.S. were relatively stable in the early weeks of 2020, and had steadily decreased over the last decade of economic expansion. One of the first signs that the pandemic was negatively impacting the labor market occurred in the week ending March 14, 2020. Table 1 shows that seasonally adjusted initial unemployment claims were 282,000 for the week ending March 14, 2020, a 71,000 increase from the previous week. More strikingly, Table 1 shows that the situation continued to worsen, with initial unemployment claims reaching 3,307,000 in the week ending March 21, 2020, the highest level ever recorded and a 3,025,000 increase from the prior week. This number was surpassed in the week ending March 28, 2020, when it reached 6,480,000. These unprecedented trends are evident in both the seasonally adjusted numbers and the raw numbers.

Additional data from the DOL show that the findings in Table 1 are being driven by large increases in unemployment claims in California, Illinois, Massachusetts, Nevada, Pennsylvania, Texas, and Washington. The DOL also notes that Washington's unemployment claims stem from "Layoffs in the transportation and warehousing, real estate rental and leasing, arts, entertainment and recreation, accommodation and food services, and service industries." In addition, Nevada's unemployment claims stem from layoffs "due to the COVID-19 virus," and many other states also attributed their increase in unemployment claims to

Covid Economics 6, 17 April 2020: 119-136

CEPR PRESS

Covid Economics 6, 17 April 2020: 119-136

COVID-19-related causes.

3.2 Results for key asset classes

Recall that theoretically, the COVID-19 pandemic potentially represents a significant demand and supply shock, and the nature of the pandemic's financial impact is multifaceted and ambiguous. I therefore next examine how different asset classes have reacted to the COVID-19 pandemic using buy and hold returns.

Recall that the week of March 22, 2020 is about when many governments began introducing and passing large economic stimulus plans in response to the pandemic. Although these plans are ostensibly meant to target the pandemic, the reality is that they also contain many significant and politically charged programs unrelated to the pandemic (e.g., provisions for real-estate taxes). To mitigate the concern that I am capturing how investors perceive subsequent fiscal policy and other government reactions to the pandemic, I focus the following analysis on the buy and hold return for January 2 to March 20, 2020 (hereafter referred to as "the onset of the pandemic").

Table 2 shows that the S&P 500 stocks decreased in value by 28.6 percent at the onset of the pandemic. This finding translates to an economic loss of $18 billion per firm on average, or about $9 trillion in total for S&P 500 firms alone.[7] The Dow 30 index, Nasdaq index, and Russell 2000 index similarly declined in value at the onset of the pandemic.

The firms with the largest declines in value are Norwegian Cruise Lines, Noble Energy, Royal Caribbean Cruises, Halliburton, and Carnival. The firms with the largest increases in value are Regeneron Pharmaceuticals, Clorox, Gilead Sciences, Kroger, and Citrix Systems. These findings provide initial evidence that the transportation, leisure, and energy industries were hurt the most at the pandemic's onset. Additional industry analyses are performed below.

[7]Note that the S&P 500 return in Table 2 differs from the mean return in Table 3 because the former is value-weighted while the latter is equal-weighted.

CEPR PRESS

Table 2 also shows that commodities experienced dramatic price changes at the pandemic's onset. The S&P 500 commodity index decreased in value by 39.2 percent, while crude oil decreased even further at about 60 percent. Natural gas, corn, and wheat also decreased in value. Gold, by contrast, increased in value. Table 2 also shows that the U.S. dollar increased in value relative to the Euro, GBP, and Japanese Yen.

Table 3 provides descriptive statistics for the firm-level analyses. The mean (median) firm has $65.8 billion ($19.7 billion) in assets and a market capitalization of $56.1 billion ($23.7 billion) as of January, 2020. About 46.3 percent of firms include pandemics as a business risk factor. The average firm has a dividend yield of 2.3 percent, a current ratio of 2.93, a PPE (property, plant, and equipment) ratio of 27.0 percent, and a cash ratio of 7.6 percent. For firms that include pandemics as a risk factor versus those that do not, the buy and hold returns at the pandemic's onset are nearly identical at −32.5 percent. This finding provides initial evidence that a firm's risk factors were a misleading indicator of its true exposure to pandemics.

Table 4 examines the industry characteristics of the firm-level buy and hold returns at the onset of the pandemic. I regress buy and hold returns on all the Fama-French 30 industry factors, inserting each industry factor one at a time, which lets the baseline equal the mean of the buy and hold returns variable of about −32.0 percent. I also control for a firm's size as measured by the log of total assets. Table 4 includes the industry factors that are statistically significant at a p-value of less than ten percent. Petroleum and natural gas firms; apparel firms; restaurant and hotel firms; automobile firms; transportation firms; aircraft, ship, and railroad firms; and machinery firms exhibit the largest decreases in value at the pandemic's outset, with magnitudes ranging from −9.0 percent for machinery to −31.6 percent for petroleum and natural gas in excess of the −32.0 percent mean return. In contrast, food product firms, healthcare firms, utility firms, and business services firms exhibit smaller decreases in value at the pandemic's outset, with magnitudes ranging from +8.1 percent for business services to +16.3 percent for food products relative to the −32.0

Covid Economics 6, 17 April 2020: 119-136

CEPR PRESS

Covid Economics 6, 17 April 2020: 119-136

percent mean return.

I next address the question of whether managers systematically anticipated their exposure to the pandemic; that is, whether a firm's business risk factors would have helped investors predict that firm's exposure to the pandemic. If this is the case, I would expect to find little or no change in firm value for firms that did *not* include pandemics as a risk factor. In addition, and assuming that the net effect of the pandemic on firm value is negative for exposed firms, I would expect to find a negative association between a firm's buy and hold return at the pandemic's onset and the indicator variable for whether that firm includes pandemics as a business risk factor. That is, more exposed firms would likely decrease more in value at the pandemic's onset. If the net effect of the pandemic on firm value is positive, I would expect the prior association to be positive.[8]

Table 3 shows that firms that did not include pandemics as a business risk factor decreased in value significantly at the pandemic's onset, indicating significant pandemic exposure for firms that did not consider themselves exposed. Table 5, column 1 further shows that a firm's buy and hold return at the pandemic's onset is not significantly associated ($p > 0.1$) with the indicator variable for whether that firm includes pandemics as a risk factor. I also fail to reject the null for this test after controlling for industry-fixed effects in Table 5, column 2. Nonetheless, the r-squared in Table 5, column 2 is 0.31, which suggests that investors weighted industry factors quite heavily at the pandemic's onset. Table 5, columns 3 and 4 control for firm size as measured by the log of total assets, and again fail to reject the null for the risk factor indicator variable.

I next examine other relevant firm-level attributes in Table 5. First, in Table 5, column 5, I find that a firm's buy and hold return at the pandemic's onset is significantly negatively associated ($p < 0.01$) with its dividend yield. This finding suggests that investors might be expecting these firms to cut their dividend payout to conserve cash during the pandemic.

[8]To illustrate, if the probability of a pandemic is 30 percent and a firm's estimated loss due to a pandemic is \$100, that firm's value will initially drop by \$30 on the day of the risk factor announcement (assuming investors are risk neutral and ignoring discounting). If a pandemic does materialize, its value falls by an additional \$70 (or whatever the true loss is less \$30).

Covid Economics 6, 17 April 2020: 119-136

Indeed, there are several recent reports of firms reducing their dividend payout as a result of the pandemic (e.g., McFarlane and Katz, 2020). Given this finding, I examine whether a firm's liquidity (i.e., its ability to pay short-term obligations such as debt and dividends) attenuated its decline in value at the pandemic's onset. Specifically, I use a firm's current ratio (current assets ÷ current liabilities) and cash ratio (cash holdings ÷ total assets). Table 5, columns 6 and 7 show that a firm's buy and hold return at the pandemic's onset is not significantly associated ($p > 0.1$) with its liquidity condition. I also examine whether a firm's change in value at the pandemic's onset is associated with having high fixed costs (i.e., costs that are difficult to scale down during demand and supply shocks), as measured by its PPE ratio (PPE ÷ total assets). Table 5, column 8 finds no significant effect for this ($p > 0.1$).[9]

The failure to reject the null for the business risk factor indicator variable in Table 5, columns 1 through 4, could be due to the possibility that some exposed firms benefit from the pandemic while other firms are harmed by it, causing gains in the former firms to offset losses in the latter firms (and leading to a relative net change of zero in buy and hold returns). An appropriate way to address this issue is to use the *unsigned* magnitudes of buy and hold returns and return volatility as dependent variables, both of which would increase in the presence of gains *or* losses and are thus largely immune to the prior concern. Table 5, column 9 therefore regresses *unsigned* buy and hold returns on the indicator variable for whether a firm includes pandemics as a business risk factor, and this test also fails to reject the null for that coefficient. I also find a similar null result in Table 5, column 10, which uses a firm's return volatility (standard deviation of daily returns) from January 2 to March 20, 2020 as the dependent variable. Overall, the evidence suggests that managers systematically underestimated their exposure to pandemics, and that investors could not have deduced which stocks would be exposed to the pandemic based on firms' business risk factors.

As additional robustness checks, I re-run all the analyses in Tables 2 through 5 using the following buy and hold return periods: January 2 to March 13, 2020; February 3 to March

[9] I also perform several tests using interaction effects of these variables (e.g., the interaction of dividend yield and the liquidity measures) and fail to reject the null for these coefficients ($p > 0.1$).

20, 2020; February 3 to March 13, 2020; and February 10 to March 20, 2020. Although the magnitudes of the results change somewhat for the different holding periods, the inferences and statistical significance of the findings are consistent with the results above.

4 Conclusion

The COVID-19 pandemic has caused businesses to close, manufacturing plants to idle, and layoffs to spike. In response, governments and central banks have drastically stepped up their operations, stimulating the economy while prohibiting travel and social gatherings. This study uses the COVID-19 pandemic as a natural experiment for testing how large-scale pandemics affect modern-day financial markets. I find that at the onset of the pandemic, most public firms decreased in value, yet the majority of firms indicated no exposure to pandemics in their SEC-mandated business risk factors. I also find that the pandemic triggered unprecedented changes in U.S. employment levels and the values of bonds, commodities, and currencies. These types of findings suggest that pandemics are systemically important to the financial markets.

Overall, this study provides some of the first systematic evidence on how the COVID-19 pandemic affected the financial markets. Future research can continue exploring how the pandemic affected investors, managers, and other entities.

Covid Economics 6, 17 April 2020: 119-136

Covid Economics 6, 17 April 2020: 119-136

References

Abadi, M., Cooper, H., Teckman-Fullard, M., 2020. How the coronavirus compares to SARS, swine flu, Zika, and other epidemics. Business Insider, March 9, 2020.

Acemoglu, D., Johnson, S., Robinson, J., 2003. Disease and Development in Historical Perspective. Journal of the European Economic Association 1, 397–405.

Almond, D., 2006. Is the 1918 Influenza Pandemic Over? Long-Term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population. Journal of Political Economy 114, 672–712.

Almond, D., Mazumder, B., 2005. The 1918 Influenza Pandemic and Subsequent Health Outcomes: An Analysis of SIPP Data. American Economic Review 95, 258–262.

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

Kiernan, P., Maurer, M., Michaels, D., 2020. SEC Gives Extra Time to File Reports as Firms Assess Coronavirus Impact. The Wall Street Journal, March 27, 2020.

McFarlane, S., Katz, B., 2020. Companies Race for Cash in Coronavirus Crisis. The Wall Street Journal, March 23, 2020.

Reinhart, C. M., 2020. This Time Truly Is Different. Project Syndicate, March 23, 2020.

Schwartz, N. D., 2020. Coronavirus Recession Looms, Its Course ‘Unrecognizable’. The New York Times, March 21, 2020.

Smialek, J., Irwin, N., 2020. Fed Slashes Rates to Near-Zero and Unveils Sweeping Program to Aid Economy. The New York Times, March 15, 2020.

CEPR PRESS

Table 1: The Pandemic and U.S. Unemployment
"SA" stands for seasonally adjusted. "NSA" stands for not seasonablly adjusted. Data source: U.S. Department of Labor.

	2020					2019	
Week Ending	March 28	March 21	March 14	March 7	February 29	March 28	March 21
Initial Unemployment Claims (SA)	6,648,000	3,307,000	282,000	211,000	217,000	211,000	215,000
Initial Unemployment Claims (NSA)	5,823,917	2,920,160	251,416	200,375	216,982	183,775	190,023

Covid Economics 6, 17 April 2020: 119-136

Table 2: **Pandemic Buy and Hold Returns for Key Asset Classes**
Commodity returns are computed using COMEX, NYMEX, or ICE data for April, May, or June, 2020 futures contracts depending on the commodity.

Asset	Buy and Hold Return from January 2 to March 20, 2020
Stocks Indices	
S&P 500 Index	−28.6%
Dow 30 Index	−32.3%
Nasdaq Index	−23.3%
Russell 2000 Index	−39.2%
Fixed Income	
S&P 500 Corporate Bond Index	−8.6%
S&P U.S. Treasury Bond Index	+5.7%
Individual Stocks (10 largest losses)	
Norwegian Cruise Line Holdings	−86.0%
Noble Energy	−83.7%
Royal Caribbean Cruises	−83.4%
Halliburton	−80.6%
Carnival	−80.5%
Apache	−80.5%
Diamondback Energy	−79.8%
Alliance Data Systems	−77.4%
MGM Resorts International	−77.1%
United Airlines Holdings	−76.3%
Individual Stocks (10 largest gains)	
Regeneron Pharmaceuticals	+31.2%
Clorox	+25.7%
Gilead Sciences	+20.4%
Kroger	+19.3%
Citrix Systems	+12.2%
Rollins	+6.8%
J.M. Smucker	+6.7%
Hormel Foods	+6.6%
Digital Realty Trust	+6.2%
Costco Wholesale	+4.7%
Commodities	
S&P GSCI (commodity index)	−39.2%
WTI Crude Oil	−60.9%
Brent Crude	−58.6%
Natural Gas	−24.3%
Gold	+9.6%
Corn	−10.9%
Wheat	−3.5%
Currencies	
EUR-USD	−4.8%
GBP-USD	−12.3%
JPY-USD	−2.1%

Covid Economics 6, 17 April 2020: 119-136

CEPR PRESS

Table 3: Descriptive Statistics for S&P 500 Firms as of March, 2020

Index i represents each firm in the sample. Note that the mean return in this table differs from the S&P 500 return in Table 2 because the former is equal-weighted and the latter is value-weighted. "Pandemic Included as Risk Factor" equals one if a firm includes pandemics as a business risk factor in its most recent annual report, zero otherwise. "Dividend Yield" equals a firm's dividends paid to common stockholders ÷ market value. "Current Ratio" equals a firm's current assets ÷ current liabilities. "PPE Ratio" equals a firm's property, plant, and equipment ÷ total assets. "Cash Ratio" equals a firm's cash holdings ÷ total assets. Financial statement data are taken from a firm's most recent annual report as of January 1, 2020. Observations vary based on data availability. The "Diff." column provides the t-statistic from a two-tailed t-test of the difference in means between firms that include pandemics as a risk factor (Mean_2) and firms that do not (Mean_3).

	Full Sample$_1$			Firms w/ Risk Factor$_2$		Firms w/o Risk Factor$_3$		
Variable	N_1	Mean_1	σ_1	N_2	Mean_2	N_3	Mean_3	Diff.
Pandemic Included as Risk Factor$_i$	492	0.46	0.50	228	1.00	264	0.00	(.)
Log(Assets)$_i$	492	23.81	1.31	228	23.99	264	23.66	(2.81)
Log(Market Capitalization as of January, 2020)$_i$	490	24.05	1.07	226	24.22	264	23.91	(3.25)
Buy and Hold Return from January 2 to March 20, 2020$_i$	492	-0.32	0.19	228	-0.32	264	-0.33	(0.45)
Dividend Yield$_i$	490	0.02	0.08	226	0.02	264	0.03	(1.01)
Current Ratio$_i$	477	2.93	12.00	222	3.48	255	2.45	(0.93)
PPE Ratio$_i$	477	0.27	0.28	218	0.22	259	0.31	(3.76)
Cash Ratio$_i$	489	0.08	0.08	227	0.08	262	0.07	(2.09)

Covid Economics 6, 17 April 2020: 119-136

CEPR PRESS

Covid Economics 6, 17 April 2020: 119-136

Table 4: Pandemic Buy and Hold Returns by Industry for S&P 500 Firms
Index i represents each firm in the sample. The industry-fixed effects represent the Fama-French 30 industries. Standard errors are in parentheses and robust to heteroscedasticity.

	Dependent Variable: Buy and Hold Return from January 2 to March 20, 2020$_i$										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Log(Assets)$_i$	-0.020	-0.022	-0.019	-0.021	-0.020	-0.020	-0.017	-0.022	-0.017	-0.020	-0.021
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)	(0.006)	(0.006)
Food Products$_i$	0.163										
	(0.049)										
Apparel$_i$		-0.210									
		(0.072)									
Healthcare$_i$			0.123								
			(0.030)								
Machinery$_i$				-0.090							
				(0.048)							
Automobiles$_i$					-0.153						
					(0.085)						
Aircraft, Ships, Railroads$_i$						-0.117					
						(0.067)					
Petroleum, Natural Gas$_i$							-0.316				
							(0.040)				
Utilities$_i$								0.107			
								(0.036)			
Business Services$_i$									0.081		
									(0.027)		
Transportation$_i$										-0.137	
										(0.043)	
Resaurants, Hotels$_i$											-0.173
											(0.057)
Observations	492	492	492	492	492	492	492	492	492	492	492
R^2	0.04	0.04	0.05	0.03	0.03	0.03	0.13	0.04	0.04	0.04	0.04

CEPR PRESS

Table 5: Pandemic Buy and Hold Returns and Firm-Level Attributes for S&P 500 Firms
Index i represents each firm in the sample. The industry-fixed effects represent the Fama-French 30 industries. Standard errors are in parentheses and robust to heteroscedasticity. Note that column (9) of this table uses *unsigned* buy and hold returns as the dependent variable, and column (10) uses the standard deviation of daily returns from January 2 to March 20, 2020 as the dependent variable (see Section 3.2 for more detail). Observations vary based on data availability.

	Dependent Variable: Buy and Hold Return from January 2 to March 20, 2020$_i$									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(Assets)$_i$			-0.021	-0.014	-0.015	-0.013	-0.014	-0.010	0.013	-0.003
			(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.003)
Pandemic Included as Risk Factor$_i$	0.008	-0.025	0.015	-0.019					0.022	0.010
	(0.017)	(0.016)	(0.018)	(0.016)					(0.015)	(0.007)
Dividend Yield$_i$					-0.276					
					(0.059)					
Current Ratio$_i$						0.000				
						(0.000)				
Cash Ratio$_i$							0.124			
							(0.095)			
PPE Ratio$_i$								0.053		
								(0.038)		
Industry-Fixed Effects	N	Y	N	Y	Y	Y	Y	Y	Y	Y
Observations	492	492	492	492	490	477	489	477	492	492
R^2	0.00	0.31	0.02	0.31	0.32	0.32	0.31	0.31	0.32	0.25

Covid Economics 6, 17 April 2020: 119-136

COVID ECONOMICS
VETTED AND REAL-TIME PAPERS

Covid Economics Issue 6, 17 April 2020

The 1918 epidemic and a V-shaped recession: Evidence from municipal income data

Christian Møller Dahl,[1] Casper Worm Hansen[2] and Peter Sandholt Jensen[3]

Date submitted: 8 April 2020; **Date accepted:** 9 April 2020

We combine high-quality vital statistics data with annual income data at the municipality level to study the economic aftermath of the 1918-inuenza epidemic in Denmark. Controlling for pre-epidemic trends, we find that more severely affected municipalities experienced short-run declines in income, suggesting that the epidemic led to a V-shaped recession, with relatively moderate, negative effects and a full recovery after 2-3 years. Month-by-industry unemployment data shows that unemployment rates were high during the epidemic, but decreased again only a couple of months after it receded. This evidence also indicates that part of the economic downturn in 1918 predates the epidemic.

Covid Economics 6, 17 April 2020: 137-162

1 Professor of Econometrics, University of Southern Denmark.
2 Associate Professor, University of Copenhagen.
3 Professor, Department of Business and Economics, University of Southern Denmark.

Covid Economics 6, 17 April 2020: 137-162

1 Introduction

The Covid-19 Pandemic and the associated policy responses have very rapidly disrupted the economies of most countries in the World. As economies shut down to prevent the virus from spreading, millions of people lose their jobs and unemployment rates have been predicted to rise to levels that the world has not witnessed since the great depression. How much is the pandemic going to depress the economy and for how many years? Such questions are naturally important, but difficult to answer, since such major health shocks are rare and the current crisis seems almost unprecedented.

This paper leverages evidence from the 1918-influenza epidemic to help answering these questions. We combine high quality vital statistics with annual income data for 76 Danish municipalities in a differences-in-differences (DiD) analysis, which exploits excess influenza deaths in 1918 across municipalities to measure the local severity of the epidemic. We find that pre-epidemic local economic activities is predictive of the severity and one should accordingly be cautious when interpreting the economic effects of the epidemic. Nevertheless, controlling for pre-epidemic trends, our analysis indicates that the epidemic at most led to a V-shaped recession with no significant medium or long-run effects. Panel C of Figure 2 illustrates this V-shaped effect on the economy. We find an average decline in income per worker of around five percent from 1917 to 1918 and this decline is larger in more severely affected municipalities.[1] We also provide evidence from monthly unemployment data for 61 industries at the national level as well as annual level municipality banking data. While the unemployment evidence indicates that an economic downturn was underway the last half of 1917—almost one year before the influenza arrived in Denmark—we see that unemployment rates were high during the epidemic months, but bounced back to more normal levels already in the summer of 1919. We do not find that banks located in more severely affected municipalities performed worse after the epidemic.

The severity of the Covid-19 pandemic has naturally spurred renewed interest in the short and medium run effects of pandemics and the public health response on economic activity. Given some similarities with the current pandemic, economists have recently presented different estimates of the economy wide impact of the 1918 influenza pandemic. Using U.S. city evidence, the study by Correia et al. (2020) shows that the epidemic had negative effects on economic activities in the manufacturing and banking sectors, while cities rolling out more aggressively non-pharmaceutical interventions (NPIs) performed better in these sectors afterwards. Barro et al. (2020) exploit international evidence from 43 countries to show negative effects on GDP per capita growth. They also report evidence suggesting

[1] This simple quantification does not take into account the pre-epidemic trends and should be viewed as an upper bound average effect.

that the pandemic had strong short-run negative effect. We contribute to this research by documenting a V-shaped effect of the epidemic using—by historical standards—high-quality within-country data. We also digitized data on NPIs, but do not find any robust evidence indicating that municipalities implementing more NPIs experienced less or more income growth after the epidemic.[2] In addition, we provide, to our knowledge, new evidence on the severity of the epidemic by showing that the short-run pre-epidemic economic activity is associated with excess influenza deaths in 1918 and that the severity of the first wave in 1918 is predictive of the second wave in 1920.[3]

There is also some less recent papers studying the economic effects of the 1918-influenza pandemic. The paper by Brainerd and Siegler (2003) finds a positive effect of the epidemic on GDP per capita using U.S. state evidence. In a similar vein, Garrett (2009) finds positive effects on U.S. Wages. The work by Karlsson et al. (2014) finds that the 1918-epidemic increased long-run poverty, decreased capital returns, but had no effect on earnings in Sweden. There is also a larger related literature studying the effects of health/mortality on different long-run economic outcomes. See for example the papers by Acemoglu and Johnson (2007), Bleakley (2007), Weil (2007), Cervellati and Sunde (2011), Hansen (2014) or the review in Weil (2014). We contribute to this literature by demonstrating that the 1918-epidemic health shock only had short-run income effect and no measurable effects in the longer run.

2 Background

In this section, we provide background information on the 1918-influenza epidemic and on how World War I affected Denmark since the onset of the epidemic overlapped with the end of World War I.[4]

2.1 The 1918-Influenza epidemic

The first wave of the influenza reached Denmark in the summer of 1918 and it is generally believed that marines, patrolling the Sound between Denmark and Sweden, were the first to being reported infected, which dates the first cases to July 8th. Unlike many other countries, the number of influenza cases is known for because a reporting system had been in existence since 1803. The number of influenza cases

[2]We are somewhat cautious when interpreting these non-findings as our NPIs are not coded with the same detail as those used in Correia et al. (2020).

[3]The finding that the epidemic was facilitated by economic activities is consistent with the evidence reported in Adda (2016), who shows that the spread of infectious diseases, such as influenza, is generally related to economic activities.

[4]This section draws on the work by Kolte et al. (2008) and Heisz (2018) for the subsection on the 1918-epidemic. For the subsection on World War I, we draw on Christiansen et al. (1988), Hansen and Henriksen (1984) and Gram-Skjoldager (2019).

Covid Economics 6, 17 April 2020: 137-162

was 34,877 in 1917 and then rose sharply to 496,755 in 1918. In 1919 and 1920, the numbers were 236,217 and 174,736 respectively. For 1918, the total number of influenza deaths were about 4,150 in the urban municipalities for which reliable data exist. These municipalities had a population of circa 1,2 million people or 41 percent of the total population in the census year of 1916. We show how epidemic severity varied between these municipalities in Section 4.

The incidence of the disease was highest among the age group 5 to 15 years, whereas mortality was highest among infants and 15 to 65 years old in 1918. More detailed data on mortality for 1918 reveals that most of the mortality was concentrated in the age group 15 to 44 years with the peak being in the age group 20 to 34 years. These patterns were similar in 1919 and 1920. Danish authorities responded to the epidemic by extending the school holiday until September 2nd, and the capital, Copenhagen closed its schools in the beginning of October 1918. Moreover, theaters and cinemas were closed by the end of the same month. Many, but not all, Danish municipalities followed the example of Copenhagen. For example, what is now the second largest city, Aarhus, did not close its schools during the epidemic. Danish newspapers of the times show that many municipalities had introduced some non-pharmaceutical interventions (NPIs) by the beginning of November.

Historical national accounts reveal that GDP contracted by 3 percent in 1918 (Hansen 1974), though it should be kept in mind that annual GDP data for this period should be interpreted with caution as many of the underlying series needed to compute GDP were not available at the annual level. Yet, there is circumstantial evidence suggesting that both the epidemic and the introduction of NPIs had some adverse economic effects. For example, newspaper articles suggest that absenteeism was a problem in some businesses. In line with this, Trier (2018) mentions this was true for the railway, the postal service and telephone companies. Yet, there are also signs that the disease led to a lack of labor in some municipalities. Moreover, parts of Copenhagen are described as being a ghost town during the epidemic. Newspapers at the time also write about unemployment in the entertainment business due to theater closures.

2.2 World War I

In August 1914 when World War I broke out, the Danish government issued a declaration of neutrality and there was never any battles on Danish soil. Denmark's main trading partners were Britain and Germany. Denmark's main export was agriculture, while it imported coal and fertilizer from Britain and coal and other industrial products from Germany. The warring parties attempted to block the Danish trade, but Denmark managed to maintain its trade and experienced a boom-like situation from

Covid Economics 6, 17 April 2020: 137-162

CEPR PRESS

1915 to 1916.

In February 1917, Germany declared unrestricted submarine warfare, which implied that all ships bound for England sailing within a certain range from the French and Italian coasts would be torpedoed regardless of whether they were neutral or not. This meant that the Danish ships almost stopped sailing routes via the North Sea because of this danger. The result was that important parts of foreign trade came to a halt: Exports to England and imports of coal and fuels. The U.S. joined the war in April 1917 and immediately banned exports of grain and fodder to neutral countries, which exacerbated the situation further. The result was that Denmark oriented its trade more to Germany, though this could not compensate from the loss of overseas deliveries of raw materials, fodder and fuel. According to the historical narrative, the situation led to lower production in 1917-18 and increasing unemployment as we demonstrate below.[5]

3 Data

We combine different data sources to construct a municipality level data set annually from 1904 to 1929, containing mortality counts for different causes of death (all causes, influenza, pneumonia, and influenza), population size, total taxable income, number of tax payers, along with other various municipality characteristics, which we describe as they are introduced. All the data sources are reported in Appendix Table A1. We are able to include 76 municipalities, which constituted the Danish so-called market towns in 1918.[6]

All mortality rates (by cause) are constructed by scaling the mortality counts by municipality population size and multiplied by 1,000. There are different reasons to believe that the quality of the mortality statistics is high for historical standards. First, Egedesø et al. (2020) observe that disease registration on pre-printed forms had been in place in municipalities since 1856 and the cause of death had to be verified by a medical doctor. Second, Lindhardt (1939) states that the Danish historical mortality statistics were viewed as being in the very front rank in terms of quality by foreign and Danish investigators.

Our main economic outcome variable is the average taxable income per tax payer. Denmark introduced income taxation in 1903 (Philip 1955), for which reason taxable income is available from 1904 onwards. Generally, the principles for income taxation remained the same throughout our study

[5]According to Jensen (2020), Russia was Denmark's third or fourth largest trading partner at the time, and trading relations were disrupted by the October revolution.

[6]The number of Danish market towns increased to 87 after the reunion between Southern Jutland and Denmark in 1920.

Covid Economics 6, 17 April 2020: 137-162

period. There was an exemption of 800 Danish kroner in Copenhagen and of 700 Danish kroner in the market towns. The income tax was progressive and rates were initially very low but increased in 1912, 1915, and 1919. For most income-tax payers, the rates were in the range of 3 to 8 percent in 1919.

To supplement our data on income, we have also digitized monthly unemployment rates for the period 1915 to 1919. The data were collected by the unions and processed by Statistics Denmark. Coverage is of 61 industries. Most of the data cover blue collar workers, but also white collar professions, e.g. office workers. Further to obtain data on NPIs, we have consulted newspapers from 1918-1920, various books and archival sources. We have collected data on whether municipalities closed schools, cinemas and theaters during the 1918 autumn part of the epidemic. Finally, we also employ data on total bank assets collected by Statistics Denmark and digitized by Abildgren (2018).

Covid Economics 6, 17 April 2020: 137-162

4 Research strategy

This section starts by describing how we measure the impact (or severity) of the 1918-influenza epidemic. We next investigate how different pre-epidemic municipality characteristics correlate with severity, which is important in its own right in terms of understanding the spread of the disease, but such balancing tests are also crucial in our attempt to study the causal economic impact of the shock. Here we also provide preliminary evidence of how the epidemic influenced outcomes. The final subsection describes our main estimation approach.

4.1 Measuring epidemic severity

We measure the impact of the epidemic using mortality data rather than case data, since mortality data are generally believed to be more reliable. Most other papers, on this particular topic, follow a similar strategy (Brainerd and Siegler 2003; Hatchett et al. 2007, for example). While the recent papers by Barro et al. (2020) and Correia et al. (2020) use influenza (and pneumonia) mortality rates, we calculate the local intensity of the epidemic by excess influenza deaths as:

$$Epidemic_{t,c} = M_{t,c} - \bar{M}_{1904-16,c}, \tag{1}$$

where $M_{t,c}$ is the influenza mortality rate in year $t \in (1918, 1919, 1920)$ per 1,000 people in municipality c and $\bar{M}_{1904-16,c}$ is the unweighted average influenza mortality rate from 1904 to 1916. In this way, we measure the local severity by the excess influenza mortality rate in year t of the epidemic. Besides this number being more realistic in terms of the actual mortality penalty associated with the epidemic, using the *excess* rate, instead of just the annual rate, also has the advantage of taking into account that some areas might just be more unhealthy than others for unrelated reasons.

The epidemic hit most areas in Denmark during 1918, but some more remote areas were not affected before the first months of 1919. In addition, the influenza swept the country in a second wave in 1920. Appendix Figure A.1 shows the excess mortality rate for the three epidemic years. The left (right) panel shows the rates for the most (less) populated municipalities sorted from the bottom up. As expected, the excess influenza mortality rate was highest in 1918. The most affected municipality had around 10 excess influenza deaths per 1,000 people. The largest municipality and capital, Copenhagen, had around 3 excess deaths per 1,000 people, which is close to the municipality average (unweighted = 3.16; population weighted = 3.26). In an international perspective, Denmark was not that severely affected: According to Barro et al. (2020) one of the worst hit countries in the World was India with 41 influenza deaths per 1,000 people in 1918, while their corresponding U.S. number is 3.90.[7] The US-city epidemic intensity in Correia et al. (2020) is substantially higher with an average of around 6.86, however, they also include pneumonia deaths and do not consider *excess* rates. The distribution in epidemic severity appears more similar with a one standard deviation, in the Danish municipalities, being equal to 1.8, while the U.S. city number is 2.0.

4.2 Descriptive evidence

We now focus on the excess influenza mortality rate in 1918—since this epidemic year constitutes our baseline intensity measure going forward—and study how it correlates with different pre-epidemic municipality characteristics. We next show the development of the main outcomes by epidemic severity, which is going to provide us with the first evidence on the impacts epidemic.

Despite that it has been argued that the geographic distribution of the epidemic severity has a component of randomness (Brainerd and Siegler 2003; Almond 2006), research by Clay et al. (2019) for example documents that excess U.S. city mortality in 1918 is positively correlated with factors such as illiteracy, infant mortality, and pollution. We provide new insights to this by showing that short-run economic activity is an important explanation of epidemic severity.

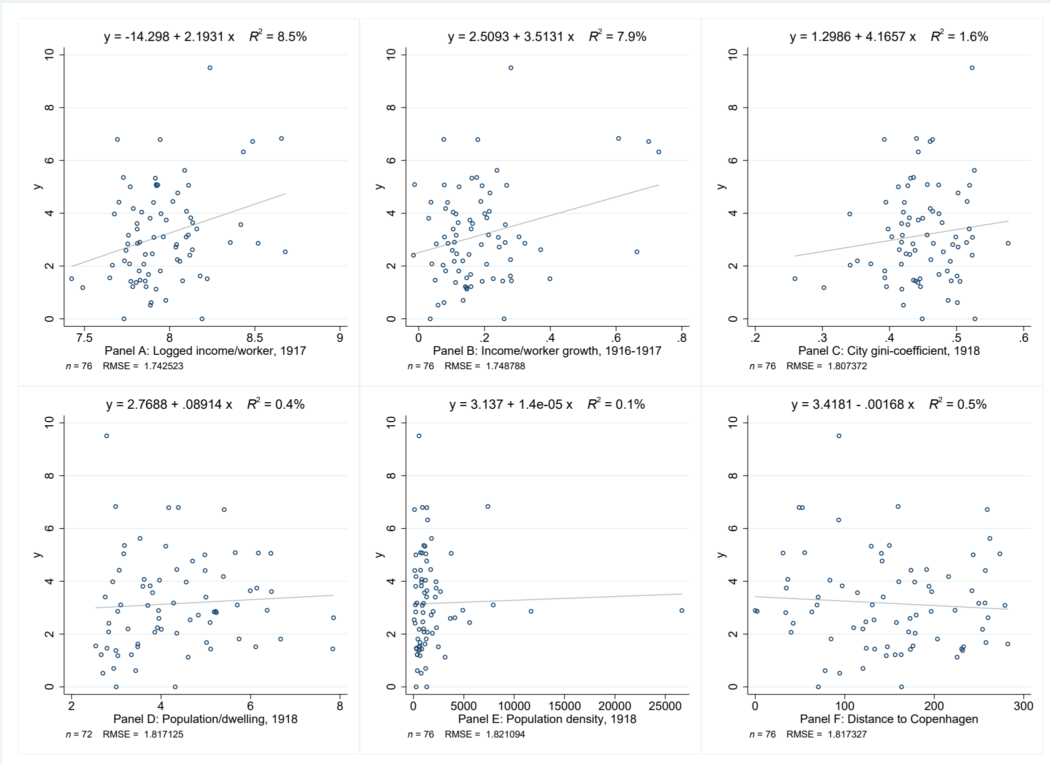
Specifically, Panel A and B of Figure 1 report the correlations between epidemic intensity and economic activity during the war years, as measured by logged income per worker in 1917 and income per worker growth from 1916 to 1917, respectively. We find positive and statistically significant corre-

[7] In Barro et al. (2020), Denmark had a 1918 influenza mortality rate of 1.7, which is substantially smaller than our municipality weighted average, suggesting that the cities were more affected by the virus compared to the countryside.

Covid Economics 6, 17 April 2020: 137-162

Covid Economics 6, 17 April 2020: 137-162

Figure 1: Pre-epidemic municipality characteristics of the 1918-influenza epidemic

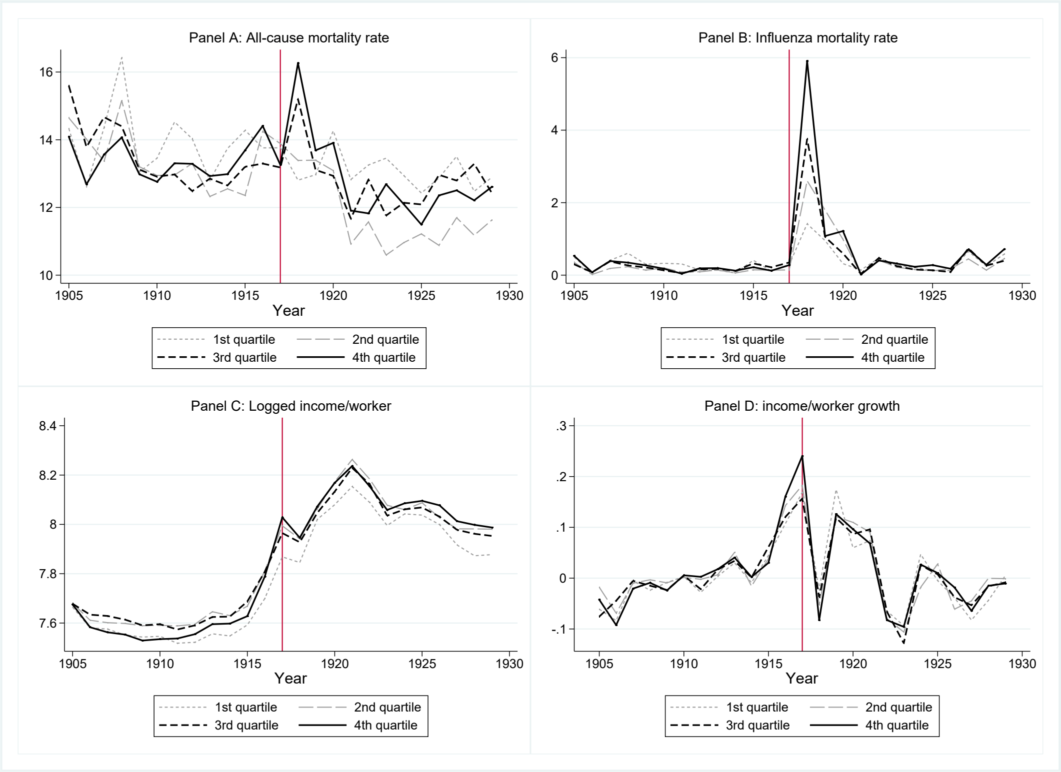


Notes: This figure shows different municipality level correlates of the 1918-influenza epidemic. Panel A: logged income/worker, 1917; Panel B: income/worker growth, 1916-197; Panel C: Gini-coefficient, 1918; Panel D: population per dwelling in 1918; Panel E: population per km2 in 1918; Panel F: distance to Copenhagen in kms. The outlier in Panel E is Copenhagen and removing this observation does not change the reported relationship.

lations with relative high R-squared values. As will also become apparent in our later event-studies, these correlations are *not* driven by longer run differences in income levels across municipalities. For example, logged income per worker four or more years before the epidemic is not a good predictor of the severity. We only have within city inequality data staring in 1918, but for this year we find a positive correlation with severity (Panel C), albeit the estimate is insignificant and the R-squared value is substantially lower. Panels D and E of Figure 1 show that the number of people per dwelling (in 1918) and population density are not very good predictors of severity. Finally, Panel F documents that severity is unrelated to distance to Copenhagen.

Next, we display the municipality average development of the all-cause mortality rate, the influenza mortality rate, logged income per worker, and annual income per worker growth rates by quartiles of $Epidemic_{1918,c}$ in Figure 2. While this exercise is mainly meant to show the broader development of

Figure 2: Mortality and income by epidemic severity in 1918, 1905-1929



Notes: This figure reports the annual average of all-cause mortality rate, influenza mortality rate, logged income per worker, and the annual change in logged income per worker (or income/worker growth). We compute these averages for 4 groups according to the epidemic severity in 1918, where the most/least affected municipalities would be placed in the 4th/1st quartile. The vertical red lines (in 1917) separate the pre and post epidemic periods.

the outcomes over the sample period, these group-averages (by treatment intensity) provide us with the first evidence on how the epidemic influenced the economy. The municipalities belonging to the 4th quartile of the epidemic severity are the most affected ("highly-treated group"), while municipalities belonging to the 1st quartile are the least affected ("control group").

Panel A shows that the all-cause mortality rate is trending downwards in all four groups with sharp increases during the epidemic years. As expected, the most-affected municipalities experience the largest increases in 1918. The influenza mortality rate, in Panel B, varies from year to year with no clear time trend and increases more than ten-fold in 1918 for the most-affected municipalities. We note that the all-cause mortality rate is increasing by less than the influenza mortality rate, indicating evidence of so-called harvesting effects. However, we do not see such differences in our event-study framework below, implying that our strategy of controlling for municipality fixed effects takes this

Covid Economics 6, 17 April 2020: 137-162

Covid Economics 6, 17 April 2020: 137-162

matter into account.

Panel C of Figure 2 displays the development of logged taxable income per tax payer, which we refer to as income per worker. We see that income per worker is generally increasing over the sample period with a pronounced downturn from 1917 to 1918, coinciding with the timing of the epidemic. The average decline, across all municipalities, in income per worker was around 5 percent. Interesting for our analysis, we see that the most/least affected municipalities experienced the largest/smallest income increases in the intermediate years before the epidemic, followed by the largest/smallest declines from 1917 to 1918. Finally, Panel D shows the corresponding development for the annual growth rate in income per worker. Again, we see a significant drop in the annual growth rate between 1917 and 1918 and this drop is more pronounced for more severely affected municipalities. Our event-study estimates below shed more light into these dynamics and reveal to what extent these differences are statistically significant. However, these preliminary patterns are indicative for the epidemic having negative effects on income growth, while at the same time suggesting that one should be careful when interpreting this evidence due to pre-epidemic trends.

4.3 Estimation approach

Our estimation strategy compares outcomes before and after the epidemic in 1918 between less and more severely affected municipalities, as measured by $Epidemic_{1918,c}$ given in eq.1. This type of strategy is often referred to as differences-in-differences (DiD) with a continuous measure of treatment intensity. We consider both DiD and event-study estimates. Our event-study specification takes on the following form:

$$y_{ct} = \sum_{k=1906}^{1929} \beta_k \left(Epidemic_{1918,c} \times 1\left[\tau = k\right] \right) + \gamma_c + \eta_t + \varepsilon_{ct}, \tag{2}$$

where y_{ct} is the outcome (mortality rate by cause, logged income/worker, or the annual growth rate of income per worker) in municipality c in year t. The 1918-epidemic severity measure is interacted with a full set of year fixed effects, where the omitted year of comparison is 1905, γ_c and η_t are municipality and year fixed effects.[8]

We use 1918-influenza severity as treatment intensity and not the 1919 or 1920 measures (or some average/sum of those) because of this event-study specification, which we utilize to think about reverse causality (in the form of pre-epidemic trends) and dynamic treatment effects. For example, one could

[8] In the DiD specification, the 1918-epidemic severity measure is interacted with a post 1918 indicator instead of year fixed effects, and we include additional controls.

argue that the severity in 1919/1920 is endogenous to the initial epidemic and taking some average/sum of the three years could prevent us from seeing pre-epidemic trends clearly and separating out treatment dynamics.

5 Main Results

Our baseline event-study estimates are shown in Figure 3. We start by explaining how the epidemic affected the all-cause mortality rate and the influenza mortality rate (Panels A and B). These results are interesting for different reasons. First, they allow us to check if the epidemic was harvesting within our empirical setup as the descriptive evidence indicated. Second, we can study if the epidemic was related to the changes in the mortality environment both before and after the epidemic; can one really think of the shock being unanticipated and temporary? Third, we can study if initially hard-affected municipalities were less/more severely affected during the final wave of the influenza epidemic in 1920.

Panel A demonstrates that the 1918-severity is unrelated to the all-cause mortality rate prior to the epidemic, so particular unhealthy municipalities were not harder hit by the epidemic. We find that $\beta_{1918} = 0.95$ meaning that one additional influenza death translates into a one-to-one increase in the all-cause mortality rate, suggesting that our empirical design is not capture harvesting effects. Panel B shows two spikes for the influenza mortality rate—one in 1918, which is almost there by construction, and another one in 1920. Thus, municipalities harder hit in the first wave were also more severely affected in the final wave of 1920. Indeed, we also find that $Epidemic_{1918,c}$ is a robust significant positive predictor of $Epidemic_{1920,c}$.[9]

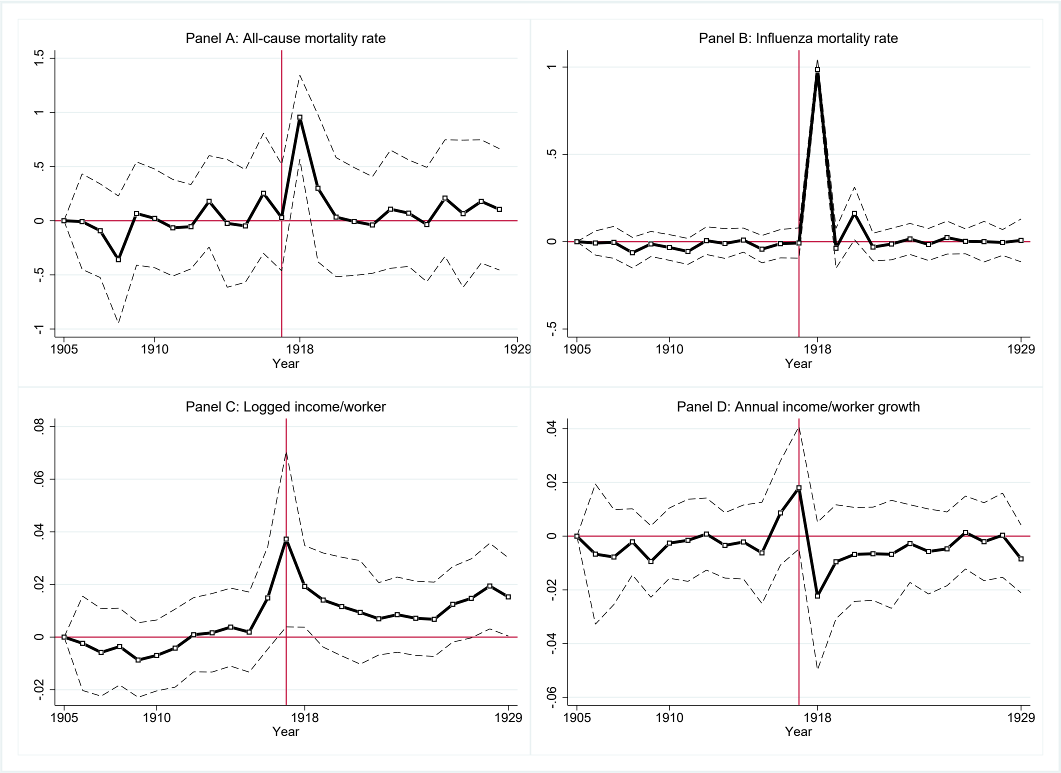
Next, we discuss our finding on how the epidemic influenced municipality economic activities. In the remaining two panels of Figure 3, we report the event-study estimates for logged income per worker (Panel C) and annual income per worker growth rates (Panel D). We find that more-affected municipalities were developing similarly to less-affected municipalities up until around World War I after which the former group experienced larger income increases. This pattern suggest that reverse causality is likely to be an issue when quantifying the effect of the epidemic on income. The interpretation could be relative straightforward, albeit to our knowledge this has not been stressed before in the context of the 1918-influenza epidemic: short-run pre-epidemic economic activities somehow facilitated the spread of the influenza or increased its severity. At the onset of the epidemic, we see that more-affected mu-

[9]In Appendix Figure A.2, we document that the pneumonia and TB mortality rates did not respond to the epidemic. One might interpret this as suggesting little co-mortality or competing risk with these diseases and/or that most deaths during the epidemic were classified as being influenza deaths (and not pneumonia, for example).

Covid Economics 6, 17 April 2020: 137-162

CEPR PRESS

Figure 3: Event-study estimates for mortality and income



Notes: This figure shows the event-study estimates from estimating eq. 2 for the all-cause mortality rate (Panel A), influenza mortality rate (Panel B), logged income per worker (Panel C) and the annual change in logged income per worker (Panel D). The vertical red line (in 1917) separates the pre- and post-epidemic periods.

Covid Economics 6, 17 April 2020: 137-162

nicipalities show sharp declines in income (Panel C) or more negative growth rates (Panel D). These effects become numerically smaller and statistically insignificant the subsequent years, implying that the epidemic only had relatively short-lived effects on municipality-level economic activities.[10]

Finally, we quantify the average effect of the epidemic on annual income growth by reporting DiD estimates in Table 1, paying careful attention to the pre-epidemic income trends. In all specifications, we use pre-epidemic data starting in 1910, but similar results are obtained when starting in 1905. Column 1 (Panel A) reports a negative and statistically significant baseline estimates, using the full period from 1910 to 1929, implying that a one-standard deviation increase in epidemic intensity is associated with a decrease in annual income growth of 1.25 percentage points. This is a dramatic negative growth effect, also considering that the post-epidemic period goes up until 1929. In column 2, we attempt to mitigate the issue of reverse causality by including logged income per worker in 1917 interacted with the post indicator. This reduces the magnitude of DiD estimate substantially and becomes statistically insignificant.

The structure in columns (3) and (4) of Panel A is the same, but here we only use a shorter post-epidemic period from 1918 to 1921 and keep the pre-epidemic period unchanged (1910-1917). Therefore, these estimates give the short-run impact of the epidemic on the economy. We see that the estimates become larger in numerical magnitude from this. However, the specification in which we control for pre-epidemic economic activities shows that the our main effect is statistically insignificant (column 4). Ignoring this fact, the magnitude implies that a one-standard deviation increase in intensity is associated with a decrease in the annual growth rate of 0.47 percentage points.

The remaining columns in Panel A report the medium and long-run effects by using the post-epidemic periods 1922-1925 and 1926-1929, which therefore excludes the short-run negative effects. Basically, we compare growth rates between more and less affected municipalities before the epidemic to growth rates 4-10 years after. We find some non-robust negative effects in the medium run (columns 5 and 6), while there is not much evidence suggesting that the epidemic changed growth rates in the longer run (columns 7 and 8).

Finally, we follow an alternative approach in dealing with the fact that pre-epidemic income-growth predicts of the severity of the epidemic by excluding the war years from the pre-epidemic period and controlling for "convergence" by including logged income per worker in 1914 interacted with the post indicator. These results are reported in Panel B of Table 1. We first note that the DiD estimates are numerically smaller than compared to the estimates in Panel B when *not* controlling for convergence

[10]Appendix Figure A.3 shows a similar patterns if we look at growth rates in total taxable income.

Covid Economics 6, 17 April 2020: 137-162

Covid Economics 6, 17 April 2020: 137-162

Table 1: DiD estimates

	Depedent variable: annual growth rate of income per worker							
	Panel A: Baseline sample							
	All years		Short run		Medium run		Long run	
	Pre-period: 1910-1917 Post-period: 1918-1929		Pre-period: 1910-1917 Post-period:1918-1921		Pre-period: 1910-1917 Post-period: 1922-1925		Pre-period: 1910-1917 Post-period 1926-1929	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Epidemic x Post	-0.00735**	-0.00113	-0.0128**	-0.00266	-0.00637*	-0.00120	-0.00368	0.000197
	(0.00329)	(0.00124)	(0.00554)	(0.00239)	(0.00325)	(0.00220)	(0.00231)	(0.00116)
Initial income x Post		-0.158***		-0.259***		-0.132***		-0.0991***
		(0.0138)		(0.0218)		(0.0167)		(0.0107)
Observations	1,592	1,592	908	908	908	908	984	984
R-squared	0.459	0.486	0.401	0.465	0.535	0.553	0.507	0.521
	Panel B: Alternative sample without war years:							
	All years		Short run		Medium run		Long run	
	Pre-period:1910-1914 Post-period:1918-1929		Pre-period: 1910-1914 Post-period: 1918-1921		Pre-period: 1910-1914 Post-period: 1922-1925		Pre-period 1910-1914 Post period: 1926-1929	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Epidemic x Post	-0.00418**	-0.00364**	-0.00980**	-0.00887**	-0.00318	-0.00271	-0.000488	-0.000227
	(0.00169)	(0.00176)	(0.00379)	(0.00405)	(0.00247)	(0.00241)	(0.00108)	(0.00109)
Initial income x Post		-0.0981***		-0.172***		-0.0867***		-0.0480***
		(0.0212)		(0.0411)		(0.0294)		(0.0171)
Observations	1,364	1,364	680	680	680	680	756	756
R-squared	0.403	0.407	0.384	0.398	0.382	0.388	0.237	0.240

This table report our baseline DiD estimates. The outcome is the annual change in logged income per worker, which is approximately equal to the annual growth rate. The variable Epidemic is explained in eq (1). Post is an indicator variable equal to one after 1917. Initial income is in Panel A logged income per worker in 1917 and in Panel B logged income per worker in 1914. All specifications include municipality and year fixed effects. Cluster robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

(i.e., the odd-numbered columns).[11] However, in Panel B we do find negative significant effects even when controlling for convergence, but these are relatively short lived. In particular, all estimates from the medium and long-run specifications are statistically insignificant (columns 5-8).

Thus, one important takeaway from Table 1 is that the epidemic at most had short-run negative effects: Controlling for pre-epidemic trends in income, we find that—in the short run—one additional influenza death per 1000 people in 1918 reduced income growth by 0.2 to 0.8 percentage points (based on estimates reported in column 4 of Panels A and B in Table 1).

6 Further evidence

6.1 Month-industry unemployment rates

This subsection documents that unemployment rates were high during the epidemic months, bounced back shortly after the epidemic receded and that part of the economic downturn in 1918 potentially predates the epidemic, using national monthly unemployment data by industry from 1915 to 1919. We have unemployment data for 61 industries, providing us with close to 3,000 observations.

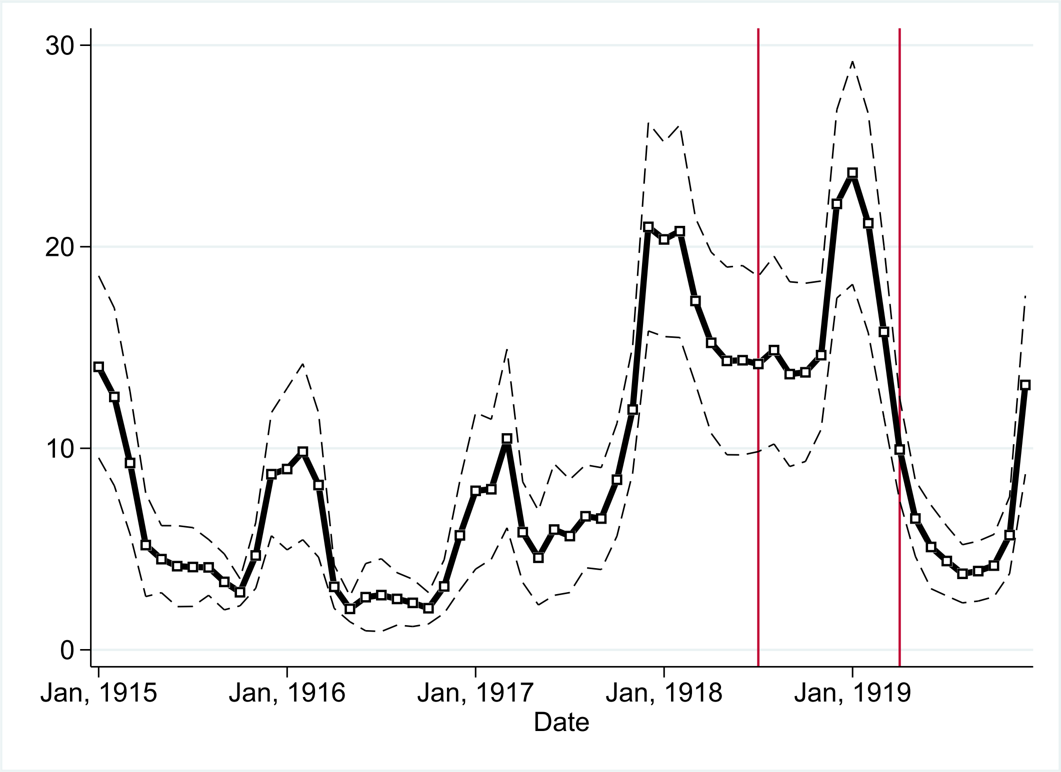
In Figure 4, we display the cross-industry average unemployment number from January 1915 to December 1919. Consistent with our income data, we see that average unemployment rates were low during the war years from the January 1915 to around June/July of 1917. The average unemployment rate for this period is around 6.8 percent. The unemployment rate starts to increase more than what can be explained by seasonal variation in the last half of 1917, which is almost one year before the epidemic in Denmark. Thus, we can argue that the economic downturn in 1918 predates the epidemic with some confidence (or at least that it is rooted before the epidemic). The vertical red lines indicate the epidemic period in Denmark. We do see that unemployment rates were high during November and December of 1918 when the epidemic peaked in Denmark, but not much higher than compared to the same months the previous year. Finally, we see that unemployment was back to its pre June-1917 levels already in the summer of 1919, and so if there was any effect of the epidemic on unemployment, it was short lived, consistent with the income evidence.

Of course, this conclusion is only based on time series evidence and the reported averages might mask shifts in unemployment across industries caused by the epidemic. This issue could have been addressed with unemployment data at the municipality level, however, such data do not exists to our

[11]Thus, one might conclude from this that reverse causality is leading to a numerical upward bias in the estimated reported in Panel A.

Covid Economics 6, 17 April 2020: 137-162

Figure 4: Monthly unemployment rates, 1915-1919



Notes: This figures show the unweighted average unemployment rates (in percent) across 61 industries with 95 percent confidence bands from January 1915 to December 1919. The vertical red lines indicate the epidemic period. Number of unemployment observations are 2,997.

knowledge.

6.2 NPIs and epidemic effects on banks

This subsection studies how municipality NPIs affected excess influenza mortality rates, and short-run economic income growth, followed by an analysis of how the bank sector was influenced by epidemic severity.

The main NPIs, which was implemented at the municipality levels, included the closing of schools, cinemas, and theaters. We were able to find evidence of NPIs for 73 out of the 76 municipalities used in the baseline analysis. We know when these NPIs were introduced, but we do not have information on the first outbreak of the epidemic at the municipality level, so we cannot construct similar intervention variables as constructed by Hatchett et al. (2007) and used in the analysis by Correia et al. (2020). 59 municipalities closed schools, 18 municipalities closed theaters, and 26 municipalities closed cinemas.

Covid Economics 6, 17 April 2020: 137-162

We construct indicators out of these data as well as an NPI index, which is equal to the sum of the different NPIs. For example, if a municipality closed schools, theaters as well as cinemas, the index takes on the maximum value of three. Appendix Table 2 reports the NPI results. We find that the closing of schools and cinemas reduced the severity of the 1918 epidemic, although the coefficients are not significant at any conventional levels. The closing of theaters is associated with higher epidemic severity (see column 1). It is important to stress that this evidence is only suggestive and higher frequency mortality data are required to assess obvious endogeneity issues (e.g., municipality might introduce NPIs because of many influenza deaths).

Using the short-run specification for annual income growth without the war years, the remaining columns show limited effects on income. If anything, closing school is associated with less growth, though this finding is not robust to specification choice (in terms of significance; compare columns 3 and 4). This conclusion is the same if we include the war years of if the post period is extended (i.e., the medium and long-run effects)

Appendix Table 3 reports the banking results. We consider total bank assets as the outcome in the first two columns and the growth rate of assets in the remaining. Since, in some municipalities, there are more than one banks operating, these specifications also include bank fixed effects, besides the usual municipality and year fixed effects. Regardless, all four estimates are small and statistically insignificant, suggesting that the 1918-influenza epidemic had little impact on the banking sector. Event-study analysis reveals that these non-findings are not driven by differential pre-epidemic trends (available upon request).

7 Lessons and perspectives

This paper has provided evidence showing that the 1918-influenza epidemic led to a V-shaped recession in Denmark with moderate short-run effects on economic activities as measured by income and unemployment rates. One might wonder if this pattern is specific to Denmark and what lessons can drawn for the current Covid-19 pandemic. First, while Denmark was not among the worst-affected countries according to Barro et al. (2020), some Danish municipalities had 1918-influenza mortality rates comparable to countries such as the U.S. and the U.K. In addition, some municipalities but not all implemented NPIs like in those countries. Therefore, our result for the 1918-epidemic is not necessarily specific to Denmark.

Second, while the Danish 1918-influenza experience gives hope that the economic aftermath of

Covid Economics 6, 17 April 2020: 137-162

the Covid-19 pandemic is going to be short lived, it is important to stress the limitations in using the experience of the 1918-influenza pandemic to inform us about the future effects of the Covid-19 pandemic. For one thing, while the Danish municipalities implemented NPIs, they were far less restrictive than the lockdowns observed in many countries today. In addition, the 1918-influenza had a very different age-profile, mainly killing people of working ages. Finally and importantly, the 1918-influenza epidemic hit during a period of time in human history, where most deaths were due to infectious diseases. For example, the leading causes of deaths were pneumonia and tuberculosis before 1940. Thus, being sick from an infectious disease was not by any means something unusual, while today this is very different in most developed countries and one might speculate that the Covid-19 pandemic is going to change long-run behavior for this very reason.

Covid Economics 6, 17 April 2020: 137-162

References

Abildgren, Kim, "A Collection of Danish Historical Micro Data," 2018.

Acemoglu, Daron and Simon Johnson, "Disease and development: the effect of life expectancy on economic growth," *Journal of political Economy*, 2007, *115* (6), 925–985.

Adda, Jérôme, "Economic activity and the spread of viral diseases: Evidence from high frequency data," *The Quarterly Journal of Economics*, 2016, *131* (2), 891–941.

Almond, Douglas, "Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population," *Journal of political Economy*, 2006, *114* (4), 672–712.

Barro, Robert J, Jose F Ursua, and Joanna Weng, "The Coronavirus and the Great Influenza Pandemic: Lessons from the Spanish Flu for the Coronavirus Potential Effects on Mortality and Economic Activity," Technical Report, National Bureau of Economic Research 2020.

Bleakley, Hoyt, "Disease and development: evidence from hookworm eradication in the American South," *The quarterly journal of economics*, 2007, *122* (1), 73–117.

Brainerd, Elizabeth and Mark V Siegler, "The economic effects of the 1918 influenza epidemic," 2003.

Cervellati, Matteo and Uwe Sunde, "Life expectancy and economic growth: the role of the demographic transition," *Journal of economic growth*, 2011, *16* (2), 99–133.

Christiansen, Niels Finn, Karl Christian Lammers, and Henrik S Nissen, *Danmarks historie: Tiden 1914-1945*, Vol. 7, Gyldendal, 1988.

Clay, Karen, Joshua Lewis, and Edson Severnini, "What explains cross-city variation in mortality during the 1918 influenza pandemic? Evidence from 438 US cities," *Economics & Human Biology*, 2019, *35*, 42–50.

Correia, Sergio, Stephan Luck, and Emil Verner, "Pandemics Depress the Economy, Public Health Interventions Do Not: Evidence from the 1918 Flu," 2020.

Egedesø, Peter Juul, Casper Worm Hansen, and Peter Sandholt Jensen, "Preventing the White Death: Tuberculosis Dispensaries," *The Economic Journal*, 2020.

Garrett, Thomas A, "War and pestilence as labor market shocks: US manufacturing wage growth 1914–1919," *Economic Inquiry*, 2009, *47* (4), 711–725.

Gram-Skjoldager, Karen, "Denmark during the First World War: Neutral policy, economy and culture," *Journal of Modern European History*, 2019, *17* (2), 234–250.

Hansen, Casper Worm, "Cause of death and development in the US," *Journal of Development Economics*, 2014, *109*, 143–153.

Hansen, Svend A, *Økonomisk vækst i Danmark. 2. 1914-1970*, Gads, 1974.

Hansen, Svend Aage and Ingrid Henriksen, "Dansk Social Historie 1914-39," 1984.

Hatchett, Richard J, Carter E Mecher, and Marc Lipsitch, "Public health interventions and epidemic intensity during the 1918 influenza pandemic," *Proceedings of the National Academy of Sciences*, 2007, *104* (18), 7582–7587.

Heisz, T, *Den spanske syge. Da historiens mest dødbringende epidemi kom til Danmark*, Politikens forlag, Copenhagen, Denmark, 2018.

Covid Economics 6, 17 April 2020: 137-162

Jensen, Bent, *Danmark og det russiske spørgsmål 1917-1924. Dansk Ruslandspolitik fra bolsjevikkernes magterobring til anerkendelsen af det bolsjevikkiske regime de*, Lindhardt og Ringhof, 2020.

Karlsson, Martin, Therese Nilsson, and Stefan Pichler, "The impact of the 1918 Spanish flu epidemic on economic performance in Sweden: An investigation into the consequences of an extraordinary mortality shock," *Journal of health economics*, 2014, *36*, 1–19.

Kolte, Victoria Ida, Peter Skinhøj, Niels Keiding, and Elsebeth Lynge, "The Spanish flu in Denmark," *Scandinavian journal of infectious diseases*, 2008, *40* (6-7), 538–546.

Lindhardt, Marie, *The statistics of pulmonary tuberculosis in Denmark 1925-1934*, Ejnar Munksgaard, 1939.

Philip, K, *Skattepolitik*, Gyldendal, Copenhagen, Denmark, 1955.

Trier, Hans, *Angst og Engle. Den spanske syge i Danmark.*, Gads Forlag, København, 2018, 2018.

Weil, David N, "Accounting for the effect of health on economic growth," *The quarterly journal of economics*, 2007, *122* (3), 1265–1306.

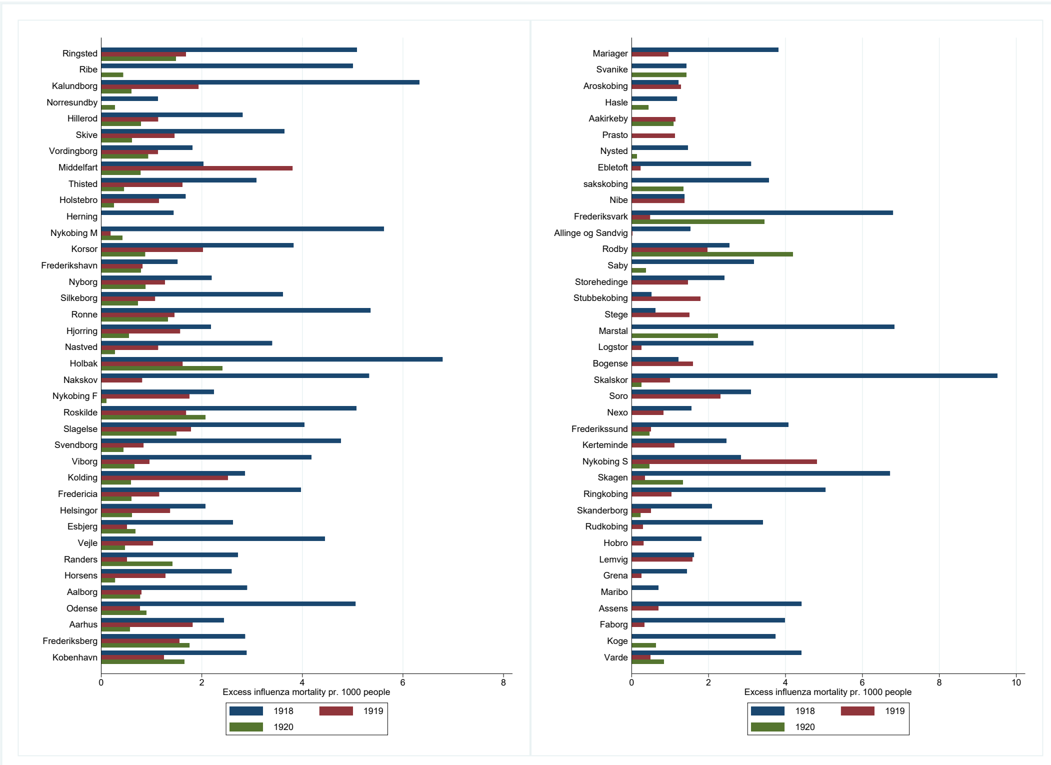
_ , "Health and economic growth," in "Handbook of economic growth," Vol. 2, Elsevier, 2014, pp. 623–682.

Covid Economics 6, 17 April 2020: 137-162

A Online Appendix

A.1 Additional tables and figures

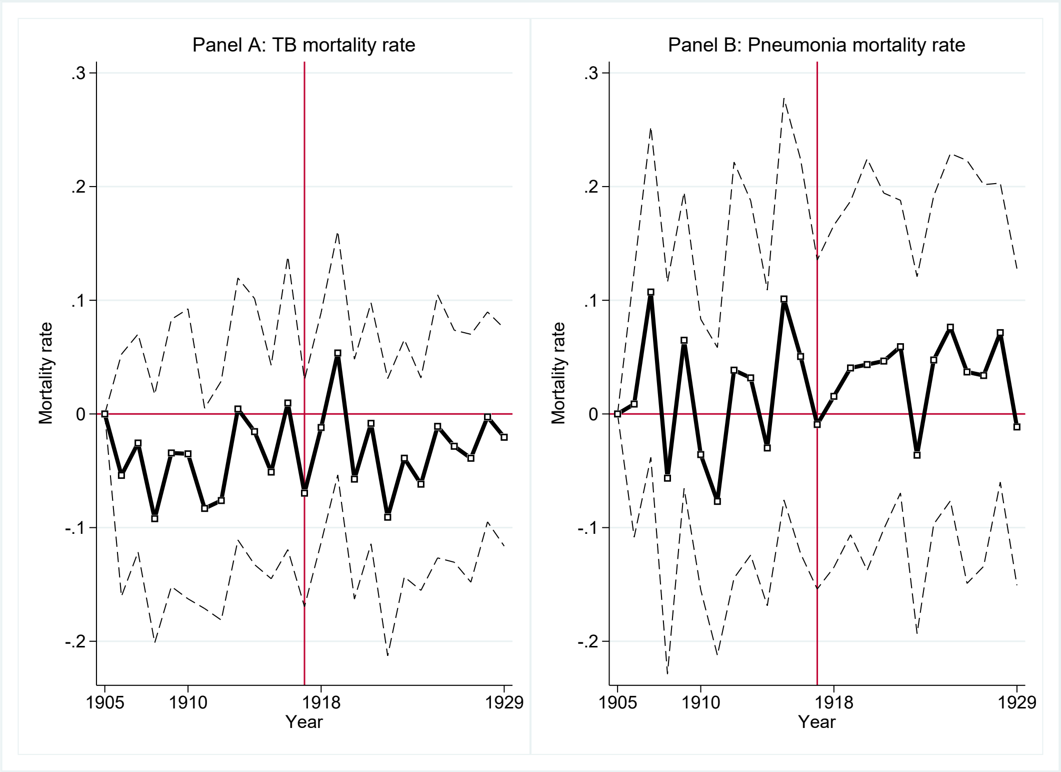
Figure A.1: Epidemic severity by municipality and year



Notes: This figure shows excess influenza mortality rate per 1,000 people as calculated in eq.1.

Covid Economics 6, 17 April 2020: 137-162

Figure A.2: Event-study estimates: Tuberculosis and pneumonia mortality rates



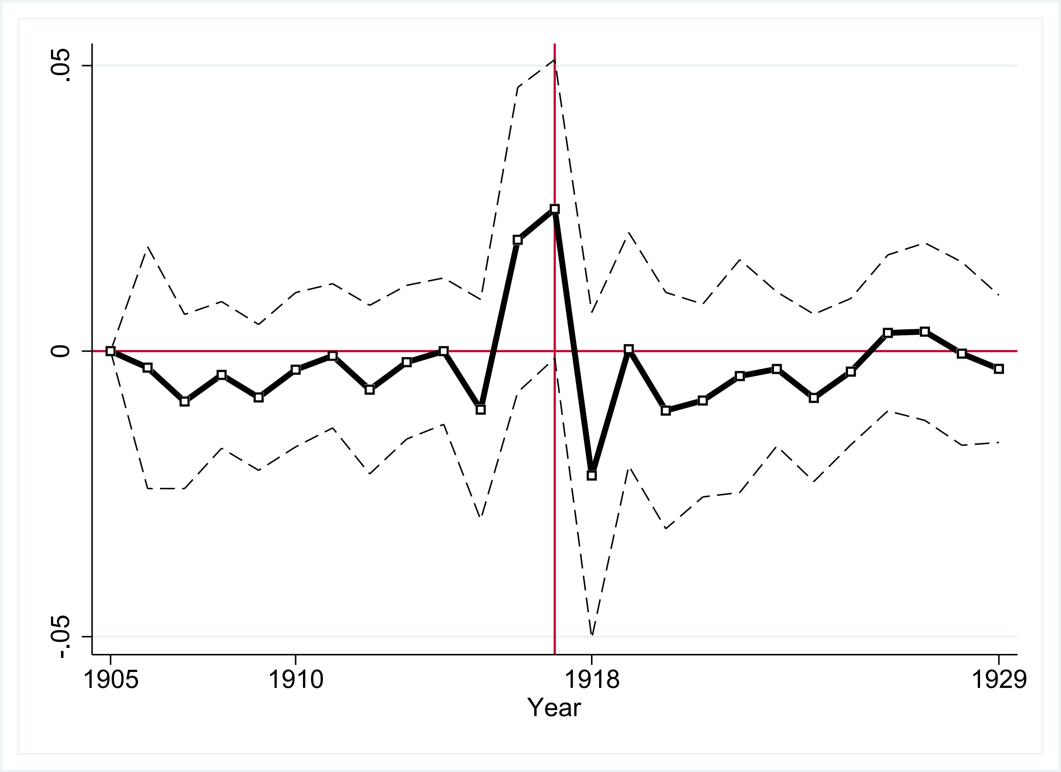
Notes: This figure shows the event-study estimates from estimating eq. 2 for the tuberculosis (TB) mortality rate, and the pneumonia mortality rate. The vertical red line (in 1917) separates the pre- and post-epidemic periods.

Covid Economics 6, 17 April 2020: 137-162

CEPR PRESS

Covid Economics 6, 17 April 2020: 137-162

Figure A.3: Event-study estimates: Annual total taxable income growth



Notes: This figure shows the event-study estimates from estimating eq. 2 for growth rates in total taxable municipality income. The vertical red line (in 1917) separates the pre- and post-epidemic periods.

CEPR PRESS

Appendix Table 1: Data explanation

Variable:	Explanation and source:
Income:	Total taxable income of the inhabitants liable to pay taxes. Source: Statistiske meddelelser (1905-1929).
Taxpayers:	Number of inhabitants liable to pay taxes. Source: Statistiske meddelelser (1905-1929).
Death rate:	Number of deaths excluding stillbirths per 1,000 people, as stillbirths are only available from 1901. Source: Cause of Death Statistics (1905-1929).
Influenza rate:	Number of deaths from influenza per 1,000 people. Source: Cause of Death Statistics (1905-1929).
Pneumonia rate:	Number of deaths from any form of pneumonia per 1,000 people. Source: Cause of Death Statistics (1905-1929).
TB rate:	Number of deaths from any form of tuberculosis per 1,000 people. Source: Cause of Death Statistics (1905-1929).
Population:	Number of inhabitants. Source: Cause of Death Statistics (1905-1929).
Population density:	Number of inhabitants in 1890 per 1890 acreage of the city. Source: Cause of Death Statistics (1918) and DigDag.
Gini:	The Gini coefficient for the inhabitants liable to pay taxes calculated using the lowest point in the intervals of the income distribution. Source: Statistiske meddelelser (1918).
Dwellings:	Number of dwellings in 1918. Source: Statistiske meddelelser (1919).
Total assets:	Total assets of banks, 1900-1920. Source: Statistiske undersøgelser (1969) and Abildgren (2018).
Unemployment rate:	Average unemployment rate across occupations (1915-1919). Source: Statistiske meddelelser (1919).
Non Pharmaceutial Interventions (NPIs):	NPIs are measured by whether a municipality closed schools, cinemas or theathers. Sources: Newspapers, archives and various books.

Notes: This table describes the main variables used in the analysis.
DigDag is a geographic database of Denmark's historic administrative division, see `http://www.digdag.dk/`.

Covid Economics 6, 17 April 2020: 137-162

CEPR PRESS
Covid Economics 6, 17 April 2020: 137-162

Appendix Table 2: Effects of NPIs

	Dependent variable:					
	Epidemic 1918		Income per worker growth rates			
	(1)	(2)	(3)	(4)	(5)	(6)
NPI: Theater	1.361*					
	(0.686)					
NPI: School	-1.070					
	(0.716)					
NPI: Cinema	-0.877					
	(0.633)					
NPI index		-0.0732				
		(0.248)				
NPI: Theater x Post			-0.00872	0.00626		
			(0.0145)	(0.0149)		
NPI: School x Post			-0.0111	-0.0234*		
			(0.0154)	(0.0140)		
NPI: Cinema x Post			0.0178	0.00836		
			(0.0145)	(0.0136)		
Epidemic x Post				-0.0115***		-0.00988**
				(0.00382)		(0.00377)
NPI index x Post					0.00263	0.00201
					(0.00591)	(0.00535)
Observations	73	73	653	653	653	653
R-squared	0.111	0.001	0.381	0.391	0.380	0.389

This table report effects of NPI on the epidemic intensity in 1918 (columns 1 and 2) and on income growth (columns 3-6). School is an indicator equal to one if the municipality closed public schools. Cinema is an indicator equal to one of the municipality closed cinemas. Theater is an indicator equal to one if the municipality closed theaters. Post is an indicator variable equal to one after 1917. The specifications in columns 3-6 include municipality and year fixed effects. In these specifications, the pre-period is 1910-1914 and the post period is 1918-1921 Cluster robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Covid Economics 6, 17 April 2020: 137-162

Appendix Table 3: Effects on bank assets

	Depedent variable:			
	Logged total assets		Annual growth rate in assets	
	(1)	(2)	(3)	(4)
Epidemic x Post	0.0108	0.00662	-0.000993	-0.000778
	(0.0185)	(0.0168)	(0.00573)	(0.00579)
Year FE	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes
Population controls	No	Yes	No	Yes
Observations	1,134	1,134	1,100	1,100
R-squared	0.977	0.978	0.484	0.485

This tables report DiD estimates for the banking sector. The outcome variable is logged total bank assets (columns 1 and 2) and logged annual differences in total bank assets (columns 3 and 4). The variable Epidemic is explained in eq (1). Post is an indicator variable equal to one after 1917. All specifications include municipality, year, and bank fixed effects. The equal-numbered columns include control for municipality population size. Cluster robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1