NBER WORKING PAPER SERIES

# PANDEMICS, VACCINES AND AN EARNINGS DAMAGE FUNCTION

Harrison Hong
Jeffrey D. Kubik
Neng Wang
Xiao Xu
Jinqiang Yang

Working Paper 27829
http://www.nber.org/papers/w27829

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2020, Revised February 2021

We thank seminar participants at the AFA Annual Meetings, Columbia Business School, Columbia University COVID-19 Symposium, Federal Reserve Bank of Boston, and Shanghai University of Finance and Economics for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Pandemics, Vaccines and an Earnings Damage Function
Harrison Hong, Jeffrey D. Kubik, Neng Wang, Xiao Xu, and Jinqiang Yang
NBER Working Paper No. 27829
September 2020, Revised February 2021
JEL No. G10,G12,G31,G32

## ABSTRACT

We derive a parsimonious model of damage to corporate earnings from COVID-19. Using measures of expected damage from industry-level earnings forecast revisions, we estimate this model with nonlinear least squares and identifying restrictions related to forecast rationality. Forecasts in mid-May 2020 imply an earnings crash and lower earnings growth until a vaccine arrives in 1.48 years (95% CI [0.61, 5.88]). We extend our framework to account for time-varying vaccine arrival rates. Mid-August 2020 forecasts imply a vaccine arrival in 0.61 years (95% CI [0.35, 1.06]), which is due to positive vaccine news as opposed to fiscal or monetary policy news.

Harrison Hong
Department of Economics
Columbia University
1022 International Affairs Building
Mail Code 3308
420 West 118th Street
New York, NY 10027
and NBER
hh2679@columbia.edu

Jeffrey D. Kubik
Maxwell School
Syracuse University
426 Eggers Hall
Syracuse, NY 13244
jdkubik@maxwell.syr.edu

Neng Wang
Columbia Business School
3022 Broadway, Uris Hall 812
New York, NY 10027
and NBER
nw2128@columbia.edu

Xiao Xu
Columbia University
x.xiao@columbia.edu

Jinqiang Yang
Shanghai University of Finance
and Economics
Guoding Rd. 777
Shanghai, 200433
China
yang.jinqiang@mail.sufe.edu.cn

# 1 Introduction

Damage to the economy from a pandemic depends on the arrival of a vaccine and other forms of costly mitigation in the interim. In the vast epidemiology literature, optimal mitigation typically entails quickly implementing vaccination programs (Anderson and May (1992), Bailey et al. (1975)). To the extent that vaccination takes time or is uncertain, other types of costly mitigation such as quarantines or social distancing are used (Wickwire (1977), Behncke (2000)). This perspective is also adopted in recent models of economic damage from COVID-19 (Kruse and Strack (2020), Alvarez, Argente, and Lippi (2020), Acemoglu, Chernozhukov, Werning, and Whinston (2020)).

While this recent work captures more richness in terms of externalities and markets, their economic damage functions are similar to earlier epidemic models. For instance, a social planner (Eichenbaum, Rebelo, and Trabandt (2020)) or firms (Hong, Wang, and Yang (2020)) take into account when a vaccine will arrive in deciding optimal mitigation that comes at the expense of earnings in the interim.[1] When the vaccine arrives, these costs no longer need to be paid and there is an upward jump in earnings.

However, estimating this damage function is challenging for a few reasons. First, it can involve many parameters. Second, it is inherently nonlinear in key parameters such as the expected vaccine arrival. Third, estimating nonlinear models requires more and timelier data of expectations regarding economic damages. That is, estimation using ex-post outcomes on GDP observed annually will be challenging from a power perspective.

To address these issues, we develop a parsimonious continuous-time regime-switching model of firm earnings with the following three key parameters: vaccine arrival rate, jump in earnings (both on pandemic impact and reflation upon vaccine arrival), and the ratio of growth rates across normal (or non-pandemic) versus pandemic regimes. Firm earnings are assumed to follow a log-normal process in the absence of jumps (Black and Scholes (1973), Merton (1974), Leland (1994)), and the arrival of vaccines is assumed to follow a

[1] Andersen, Hansen, Johannesen, and Sheridan (2020) and Farboodi, Jarosch, and Shimer (2020) also point to the importance of voluntary mitigation by households who stop consuming even in advance of government-imposed lockdowns.

time-homogeneous Poisson process.[2]

We derive a tractable expectations formula that relates earnings forecast revisions from just before the pandemic arrival to just after its arrival to these underlying parameters and several independent variables (i.e., a closed-form damage function). Broadly, the vaccine arrival rate moderates the persistence of the COVID-19 shock to earnings. To the extent an effective vaccine is expected to arrive quickly, the shock should be mostly felt in short-term as opposed to medium-term or long-term earnings forecasts. Hence, we can infer from the revision of forecasts of different horizons the parameters of the earnings process taking into account the effects of COVID-19.

We fit our nonlinear model to timely measures of expected damage for firm earnings using revisions of industry-level consensus earnings forecasts made by security analysts. Security analyst forecasts should integrate not only scientific evidence on the development of effective vaccines but also logistical issues surrounding their distribution as well as macroeconomic consequences to consumers and firms. Plentiful timely data on these forecasts allow for precise estimates of these parameters.

The structure of our model points to a natural set of identifying restrictions related to forecast rationality that allow for estimation using nonlinear least squares (NLS). If forecasts are rational, i.e., they take into consideration the key variables of our analysis in their information set (Nordhaus (1987), Keane and Runkle (1998)), then NLS can retrieve consistent parameter estimates. To the extent forecasts are boundedly rational for different reasons (Coibion and Gorodnichenko (2012), Laster, Bennett, and Geoum (1999), Hong and Kubik (2003)), consistent parameter estimates can still be retrieved using NLS as long as the ratio of biases across forecast horizons is uncorrelated with our independent variables. This exclusion restriction is plausible as industry-level forecast revisions following COVID-19 are highly correlated with cross-sectional industry stock price reactions (Landier and Thesmar (2020)) and also vaccine news as we demonstrate below.

[2] A Poisson process is the usual starting point in modeling vaccines in epidemiology or medical literatures (see, e.g., Arnold, Galloway, McNicholas, and O'Hallahan (2011), Lee, Norman, Assi, Chen, Bailey, Rajgopal, Brown, Wiringa, and Burke (2010), Ball and Sirl (2018))

In our empirical work, our main dependent variable is the revision of earnings forecasts after the arrival date of COVID-19 in the US, which we take to be February 20, 2020. To reduce measurement error, we work with industry portfolios by value-weighing median forecasts for stocks at the GICS 8-digit industry classification. To be conservative and to allow forecasts to be fully revised, we use May 2020 as our forecast date.[3]

The main independent variables from our theory are the horizon of the earnings forecasts and the earnings growth rates in the non-pandemic and pandemic regimes. The horizon of earnings forecasts is straightforward to measure. For our baseline specifications, we pool together both industry FY1 (nearest next fiscal year-end), FY2, FY3, FY4 and FY5 (farthest fiscal year-end) forecasts made in May of 2020. We measure the growth rate in the non-pandemic regime using analysts' growth rate forecasts on January of 2020 and also aggregate these to the industry level. That is, our specification assumes that growth rates return to non-pandemic levels after the arrival of a vaccine.

Our model allows us to simultaneously infer not just the vaccine arrival rate but also disentangle jumps in earnings due to mitigation from the growth rate effects in a pandemic regime. We have the following estimates using forecast revisions in May 2020. The vaccine arrival rate $\lambda$ is 0.674 with a 95% bootstrap confidence interval of [0.17, 1.65]. This implies that a vaccine is expected in $1/\lambda = 1.48$ years as of mid-May 2020, with a confidence interval of [0.61 years, 5.88 years]. These estimates are retrieved by value-weighting the NLS regression to further minimize the impact of outliers.

The initial jump in earnings following the arrival of COVID-19 corresponds to costly mitigation measures (e.g. social distancing) meant to keep the virus at bay and is given by $e^{-n}$. The coefficient $n$ is 0.603 with a 95% bootstrap confidence interval of $[0.31, 1.48]$. This statistically significant estimate of $n$ implies around a negative 45.3% jump in earnings level. We estimate that the proportion of the mean growth rates of the pandemic versus the non-pandemic regime is 0.827, which is a deterioration in growth during the pandemic.

Using likelihood ratio tests, we reject the constrained model where $\lambda = 0$ (i.e., there is

[3]There is naturally a lag in analyst revisions and we only begin to see some revisions starting in April and then most of the forecasts have been revised by May of 2020.

no vaccine or a vaccine is expected to arrive in an infinite number of years) in favor of the unconstrained model. Our model assumes that earnings revert to normal when the vaccine arrives. If analysts expect long-run growth rates are damaged, i.e. post-pandemic growth rates are lower than pre-pandemic growth rates, then it would imply counterfactually that we retrieve a small vaccine arrival rate $\lambda$. Our empirical findings in this regard are consistent with households surveys from Giglio, Maggiori, Stroebel, and Utkus (2020) indicating damage to short-run earnings growth but not necessarily long-run damage. We can also reject the constrained model where there is no damage to growth rates during a pandemic in favor of our unconstrained model. Our findings here are consistent with macroeconomic damage documented in Ludvigson, Ma, and Ng (2020) using vector auto-regressions and Gormsen and Koijen (2020) using stock price and dividend futures.

The stock prices of face-to-face (Montenovo, Jiang, Rojas, Schmutte, Simon, Weinberg, and Wing (2020)) industries are particularly hit by COVID-19 (Pagano, Wagner, and Zechner (2020), Favilukis, Lin, Sharifkhani, and Zhao (2020)). To this end, we re-estimate our model using this subsample of industries. The initial jump for this subsample of industries is $n = 2.678$, implying a much larger initial jump for this set of industries than for all the overall sample. This 2.678 point estimate lies well outside the confidence interval of the estimate for the overall sample. The vaccine arrival rate is also higher for these hardest hit industries compared to the overall sample, consistent with speculation that the rollout of the vaccine might be preferentially based on business interest groups that were the most hit.[4] But the point estimate of around one year lies well within the confidence interval of the mid-May 2020 estimate for the overall sample.

We then consider how robust our estimates are to model mis-specifications regarding the vaccine arrival process. The vaccine arrival rate can be time-varying depending on news regarding intermediate stages of development, such as clinical trials, FDA approval, and distribution logistics. Such a time-varying vaccine arrival process involves many more parameters than a simple Poisson arrival process. We simulate earnings data based on such a

[4] J. David Goodman and Luis Ferre-Sadurni, "Big fight breaks out over which interest groups get vaccine first," *NYTIMES*, December 20, 2020.

process for different assumptions regarding the arrival rate of news and how this news changes the ultimate overall vaccine arrival rate. Then we estimate our parsimonious model using this simulated data and find that the expected arrival rate retrieved from our parsimonious model to be close to the true one time implied by time-varying vaccine arrival process.

Finally, we use our parsimonious model to test for time-varying vaccine arrival rates in our data. To implement this test, we estimate vaccine arrival rates for June, July, and August (which is the latest available) forecasts, while holding fixed the mid-May 2020 estimates of the initial jump in earnings and pandemic growth rate estimate. To the extent the vaccine arrival estimates are the same across the months, it would be consistent with a time-homogenous Poisson arrival model. But if they differ, we can reject a time-homogenous Poisson model.

Whereas the estimate arrival rate for mid-May 2020 is 0.674, the estimates are 0.741, 0.815, and 1.636 for June, July and August, respectively. The 1.636 estimate for August stands out and its 95% bootstrap confidence interval for this estimate is [0.94, 2.83]. This translates to an expected arrival time of 0.61 years, which is significantly quicker than 1.48 years. The arrival rate estimate for August of 1.636 essentially lies outside the confidence intervals for May [0.17, 1.65], June [0.45, 1.23], and July [0.5, 1.35] (1.636 is just inside of 1.65). Moreover the estimates for May, June and July also lie outside the confidence interval for August.

Importantly, there is little fiscal or monetary news in July and August of 2020. However, there were two key pieces of news regarding the clinical trials of Moderna and Pfizer that came up in late July and early August of 2020. Our estimates line up with *Good Judgment*'s survey of experts on when the US would vaccinate 25 million people. Their May 2020 forecast was 23 months, while their August forecast was 11 months. Ours are a bit more optimistic (and ex-post more accurate) than these survey forecasts. Hence, this analysis alleviates the concern that our vaccine arrival estimates might be picking up other mitigating factors, particularly expectations regarding fiscal or monetary interventions (Elenev, Landvoigt, and Van Nieuwerburgh (2020)).

We associate a medical intervention that returns the economy to normal as being a vaccine since the bulk of the government funding in the US and Europe have been for its development. For instance, Warp Speed was established in March-April and contributed to the high number of vaccine candidates that came online by May 2020.[5] In other words, whereas vaccine development has previously taken years, the unprecedented government support and new technology such as genetic sequencing were clearly factored into vaccine forecasts by May 2020. Nonetheless, our regime-switching model can be applied to other countries where it might be the arrival of therapeutics or testing that returns these countries to normal. For instance, rigorous testing has played a bigger role in Asian countries.[6]

Our paper proceeds as follows. We present our model of earnings damage function and estimation strategy in Section 2. Section 3 describes the dataset and main variables. Estimates of our model are presented in Section 4. We extend the model to account for vaccine news in Section 5. We conclude in Section 6.

# 2 Model

We assume that the economy can be in one of the two regimes: the normal (or non-pandemic) and pandemic regimes. The economy starts in the normal regime. At stochastic time $t_0$, it unexpectedly enters into the pandemic regime. Afterwards, the pandemic becomes extinct and the economy returns back to the normal regime when a successful vaccine is developed at time $\tau$, which occurs with probability $\lambda$ per unit of time.

[5] According to Bloomberg News article "Trump administration dips into protective gear, CDC funds to fund vaccine push" (September 23, 2020), the Warp Speed budget is as large as $18 billion and almost all of it allocated to vaccine developments (Moderna, Sanofi, GSK, Pfizer, Novavax, J&J and AstraZeneca) and only a small amount toward therapeutics (Regeneron's antibody cocktail).

[6] Another medical scenario that returns the economy to normal is herd immunity. But this possibility does not seem likely given limited evidence on the length of individual immunity.

## 2.1 Normal Regime

We let $\widehat{Y}_t$ denote the earnings (EBITDA) process of the asset in the normal regime. We assume that $\widehat{Y}_t$ follows a commonly used geometric Brownian motion (GBM) process:[7]

$$\frac{d\widehat{Y}_t}{\widehat{Y}_{t-}} = \widehat{g}dt + \rho\phi\, d\mathcal{B}_t + \sqrt{1-\rho^2}\,\phi\, d\mathcal{W}_t\,, \tag{1}$$

where $\mathcal{B}_t$ is the standard Brownian motion driving the "business-as-usual" aggregate risk and $\mathcal{W}_t$ is the standard Brownian motion driving the idiosyncratic earnings risk. By construction, $\mathcal{B}_t$ and $\mathcal{W}_t$ are orthogonal. In Equation (1), $\widehat{g}$ is the expected earnings growth (drift) and $\phi$ is the volatility of earnings growth, which includes the aggregate component $\rho\phi$ and the idiosyncratic component $\sqrt{1-\rho^2}\,\phi$. That is, $\rho$ is the correlation coefficient between the aggregate shock $\mathcal{B}_t$ and the asset's earnings. For simplicity, we let $\widehat{g}$, $\phi$, and $\rho$ all be constant.

## 2.2 Pandemic Regime

Next, we specify the impact of the *unexpected* pandemic arrival and the *anticipated* stochastic vaccine arrival. Let $Y_t$ denote the asset's earnings process during the pandemic regime. Once in the pandemic regime ($t_0 < t < \tau$), the asset's earnings process $Y_t$ follows:

$$\frac{dY_t}{Y_{t-}} = gdt + v\, d\mathcal{Z}_t + \rho\phi\, d\mathcal{B}_t + \sqrt{1-\rho^2}\,\phi\, d\mathcal{W}_t + (e^n - 1)\, d\mathcal{J}_t\,, \tag{2}$$

where $\mathcal{J}_t$ is a pure jump process and $d\mathcal{J}_t = 1$ if and only if the vaccine arrives.

There are four terms in equation (2). First, earnings will jump discretely by a fraction $(e^n - 1)$ at the moment of the vaccine arrival, i.e., when $d\mathcal{J}_t = 1$. This is to capture earnings reflation once the vaccine returns the economy to normal. (Absent vaccine arrival, $d\mathcal{J}_t = 0$). Second, the pandemic arrival changes the expected earnings growth rate from $\widehat{g}$ to $g$ (leaving aside the effect of vaccine arrival.) Third, the pandemic shock $d\mathcal{Z}_t$ directly causes additional earnings growth volatility, $v$. Finally, as in the normal regime, earnings is subject to the

[7]The GBM process is widely used in asset pricing and corporate finance to model corporate earnings, e.g., Gordon growth model, capital structure models in the tradition of Black and Scholes (1973) and Merton (1974) and Leland (1994) models. While earnings is always positive in this formulation, we can generalize this model to allow for negative earnings. By assuming that a firm's earnings at the enterprise level (after we unlever the firm) follows a GBM earnings process, earnings for equity holders can be negative even when earnings for the enterprise is positive.

business-as-usual aggregate shock $d\mathcal{B}_t$ and idiosyncratic shock $d\mathcal{W}_t$ with volatility $\rho\phi$ and $\sqrt{1-\rho^2}\,\phi$, respectively. All shocks are orthogonal to each other.[8] For simplicity, we let $n$ be constant and keep $\widehat{g}$, $\phi$, and $\rho$ the same as in the normal regime.

More generally, the growth rate $g$ and earnings volatility $v$ in the pandemic regime depend on the optimally mitigated infections in the economy. For simplicity, we model these parameters as constants with particular emphasis that $g$ is expected to be less than $\widehat{g}$ due to the adverse direct effect of the pandemic.

## 2.3 Transition from Normal to Pandemic Regime

The arrival of COVID-19 triggers optimal mitigation in the form of foregone earnings. There is both a fixed and variable cost to mitigation that have to be paid out of earnings each period there is a pandemic. This unexpected but optimal corporate mitigation spending decreases its earnings. That is, as the COVID-19 shock unexpectedly hits at $t_0$, the earnings drops by a fixed fraction $\delta$:

$$Y_{t_0} = Y_{t_0-}e^{-\delta}. \tag{3}$$

And at the moment of vaccine arrival, the earnings instantaneously increases by a fraction $n$ from the pre-arrival time since mitigation costs no longer need to be paid as shown in Equation (2):

$$Y_\tau = e^n Y_{\tau-}\,. \tag{4}$$

We further set $\delta = n$. That is, the percentage of earnings increase at the moment of vaccine arrival $\tau$ is equal to the percentage of earnings decrease at the moment of pandemic arrival time $t_0$. Consider the counter-factual case that helps us understand the mechanism: If $\lambda \to \infty$, we have $\tau- = t_0$. For this case, earnings is not impacted at all by the jumps as $Y_\tau = e^n Y_{\tau-} = e^n Y_{t_0} = e^n e^{-n} Y_{t_0-} = Y_{t_0-}$.

[8] The vaccine arrival process $\mathcal{J}_t$ is independent of $[\mathcal{W}_t, \mathcal{B}_t, \mathcal{Z}_t]^\top$, which is a $3\times 1$ standard Brownian motion.

## 2.4 Linking Earnings Forecasts to Pandemics Damage Model

We can now relate earnings forecasts to our model. Recall that $\tau$ denotes the stochastic vaccine arrival time. Assuming that the consensus analyst forecast is being generated by our model, we have for $t$ in the pandemic regime:

$$\frac{1}{Y_t}\mathbb{E}_t[Y_s] = \int_t^s \lambda e^{-\lambda(\tau-t)} e^{g(\tau-t)} e^n e^{\widehat{g}(s-\tau)} d\tau + e^{-\lambda(s-t)} e^{g(s-t)} \tag{5}$$

$$= \frac{\lambda}{\lambda - g + \widehat{g}} \left[ e^{\widehat{g}(s-t)} - e^{(g-\lambda)(s-t)} \right] e^n + e^{(g-\lambda)(s-t)} . \tag{6}$$

Recall that $\widehat{g}$ is the pre-COVID long-term growth (LTG) rate and $g$ is the constant growth conditional on being in the COVID-19 regime. As we assume that there are only two regimes, normal and pandemic, the non-pandemic regime growth rate is the same as the post-pandemic regime growth rate. In a later section, we extend this formula to allow for these two rates to differ.

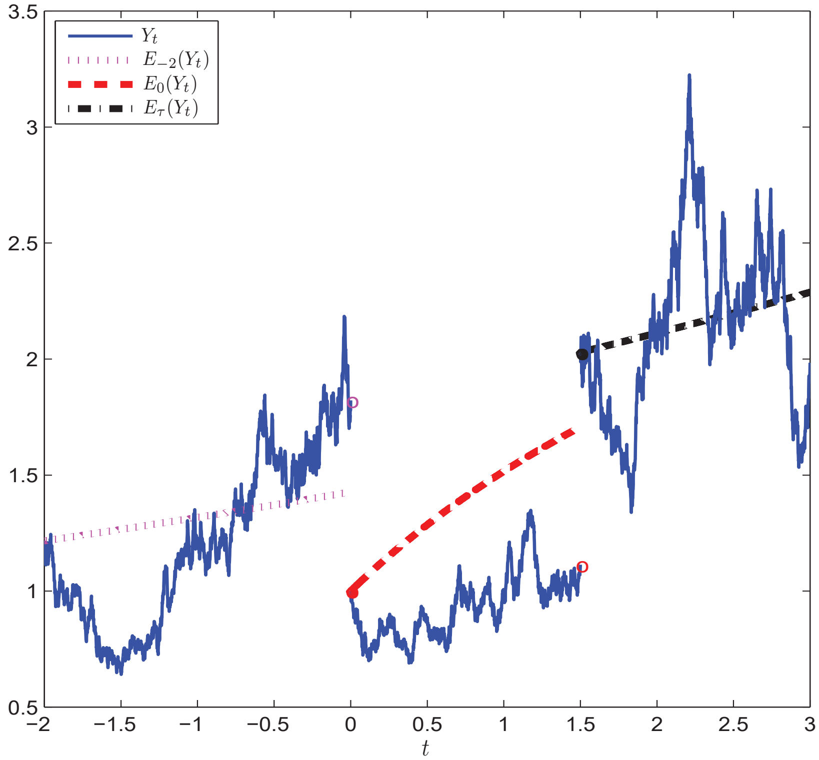
The first term of Equation (5) is conditioned on a vaccine arriving in the interval between $t$ and $s$. Inside the first term, the density of the stochastic vaccine arrival time $\tau$ is $\lambda e^{-\lambda(\tau-t)}$. Before the vaccine arrives (from $t$ to $\tau$) the cumulative (gross) growth is $e^{g(\tau-t)}$. After the vaccine arrives at $\tau$ in this interval $(t, s)$, there is reflation of earnings by a multiple of $e^n$, i.e., $Y_\tau = e^n Y_{\tau-}$, and during the subsequent sub-period $(\tau, s)$, earnings growth reverts to the pre-COVID LTG rate $\widehat{g}$, which gives the cumulative (gross) growth is $e^{\widehat{g}(s-\tau)}$ from $\tau$ to $s$.

As a result, for a given $\tau \in (t, s)$, $\mathbb{E}_t[Y_s] = Y_t e^{g(\tau-t)} e^n e^{\widehat{g}(s-\tau)}$, which explains why the first term is the contribution to $\mathbb{E}_t[Y_s]/Y_t$ conditional on $\tau \in (t, s)$. The probability that a vaccine does not arrive in $(t, s)$ is $e^{-\lambda(s-t)}$. If this is the case, the growth rate in $(t, s)$ is $g$. Therefore, the second term gives the contribution to $\mathbb{E}_t[Y_s]/Y_t$ conditional on $\tau > s$. Adding the two terms together gives $\mathbb{E}_t[Y_s]/Y_t$ for any $t$ in the pandemic regime.

Below in Figure 1, we provide a simulated path of earnings going through the non-pandemic, during-pandemic, and non-pandemic regimes. The plot starts with earnings at 1.22 at $t = -2$. The (continuously compounded) growth rate in the non-pandemic regime is set at $\widehat{g} = 8\%$ per annum. The pandemic unexpectedly arrives at time $t = t_0 = 0$, at which point earnings jumps downward from the magenta dot $Y_{t_0-} = 1.82$ to the red solid

Figure 1: Earnings Path and Expectation Calculations

The parameter values are: $n = \delta = 0.6$, $\widehat{g} = 0.08$, $g = 0.064$, and $\lambda = 0.7$. Parameter values are annualized whenever applicable. $Y_{-2} = 1.22$. At time $t = 0$, earnings jumps from $Y_{t-} = 1.82$ to $Y_t = 1$. And at time $t = 1.5$, earnings jumps from $Y_{t-} = 1.11$ to $Y_t = 2.03$.




dot $Y_{t_0} = 1$ — which we have parameterized as a $\delta = 60\%$ drop. At $t = \tau = 1.5$, the vaccine arrives, earnings $Y_t$ jumps upward by $n = \delta = 60\%$ from $Y_{\tau-} = 1.11$ (the red open dot) to $Y_\tau = 2.03$ (the black solid dot).

We set the vaccine arrival rate at $\lambda = 0.7$ per year (with an implied expected arrival time of around $1/\lambda = 1.43$ years, i.e., $\mathbb{E}_{t_0}(\tau - t_0) = 1.43$) after the unexpected arrival of the pandemic at $t_0$. The (conditional) growth rate in the pandemic regime, $g$, is set to be 0.8 times that of the pandemic regime, $\widehat{g}$, which means $g = \widehat{g} \times 0.8 = 8\% \times 0.8 = 6.4\%$.

In addition to plotting a sample path, we also plot the expected earnings immediately after the pandemic arrival, $\mathbb{E}_0(Y_t)$ given the value of $Y_0 = 1$ at $t = 0$ (see the red dashed line). In contrast, if investors were naive ignoring vaccine arrival and using a constant expected earnings rate $g$ forever, the expected earnings at $t = 0$ is then equal to $Y_0 e^{gt}$. The naive forecasts of $Y_t$ is lower than $\mathbb{E}_0(Y_t)$ due to the assumption that $g \leq \widehat{g}$ and earnings will jump by a fraction $(e^n - 1) > 0$ upon the vaccine arrival.

The magenta dotted line plots the expected earnings at $t = -2$ before the pandemic arrival. As the pandemic is unexpected, we have $\mathbb{E}_{-2}(Y_t) = Y_{-2}e^{\widehat{g}(t+2)} = Y_{-2}e^{0.08\times(t+2)}$. Similarly, the black dash dotted line plots expected earnings $Y_t$ immediately after the arrival of the vaccine at time $\tau$, which is given by $\mathbb{E}_\tau(Y_t) = Y_\tau e^{\widehat{g}(t-\tau)}$. That is, the earnings processes in the normal regimes (both before the pandemic arrival and after the vaccine arrival) are the same. Notice that the growth rate in the non-pandemic regime (the dotted black line) is equal to $\widehat{g}$, which is larger than the growth rate for the dashed red line (the pandemic regime.) Notice that the growth rate (anticipating stochastic vaccine arrival) in the pandemic regime is time-varying and smaller than that in the non-pandemic regime.

Now we calculate the expected earnings from $t_0-$, i.e., the moment that is just prior to the unexpected COVID-19 arrival time $t_0$. Substituting Equation (3), $Y_{t_0}/Y_{t_0-} = e^{-\delta}$, into (6) and with $\delta = n$, we obtain[9]

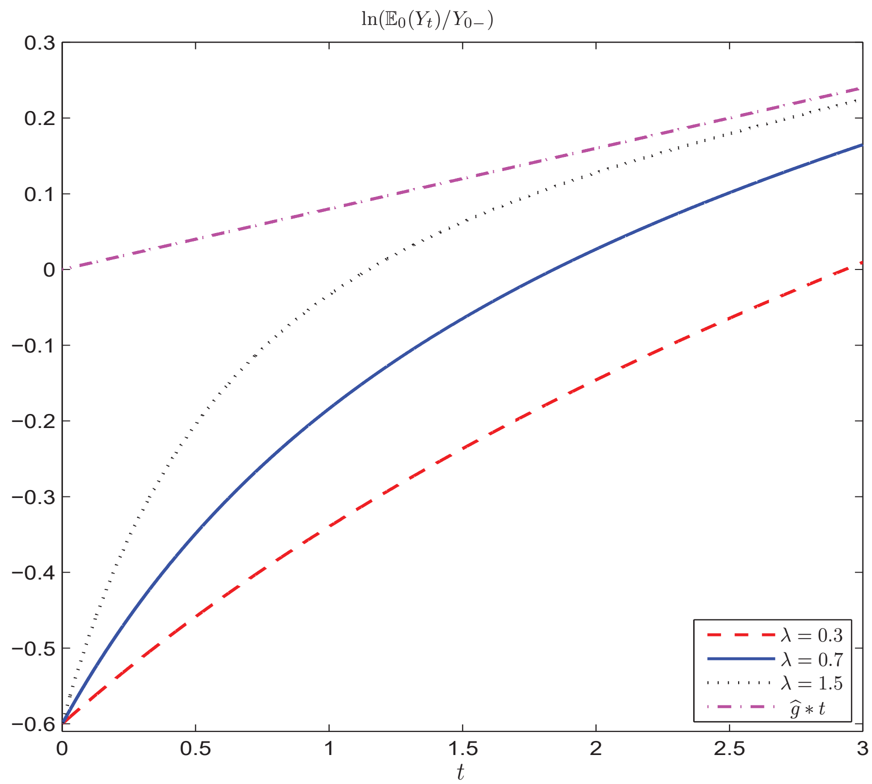
$$\frac{1}{Y_{t_0-}}\mathbb{E}_{t_0}[Y_s] = \frac{Y_{t_0}}{Y_{t_0-}}\frac{1}{Y_{t_0}}\mathbb{E}_{t_0}[Y_s] = \frac{\lambda}{\lambda - g + \widehat{g}}\left[e^{\widehat{g}(s-t_0)} - e^{(g-\lambda)(s-t_0)}\right] + e^{-n}e^{(g-\lambda)(s-t_0)}\,. \tag{7}$$

Figure 2 provides another way to understand the evolution of expectations across the normal and pandemic regimes. In this figure, we examine the effect of the vaccine arrival rate $\lambda$ on $\ln\left[\mathbb{E}_0(Y_t)/Y_{0-}\right]$, the log of forecast revisions between $t = 0-$, the moment just before the pandemic arrives, and any time $t$ subsequently. Compared with the counterfactual that the pandemic did not arrive and the business is then as usual (which means earnings grow at an expected rate of $\widehat{g}$ indefinitely, the earnings responses are naturally negative, meaning that $\mathbb{E}_0(Y_t) < Y_{0-}\, e^{\widehat{g}\,t}$. But because of the anticipated vaccine arrival and the economy eventually reverts to normal, earnings increase over time and approaches the long-run cumulative growth for logarithmic earnings, $\widehat{g}\,t = 0.08t$ (the magenta dash-dotted straight line). For all levels of $\lambda$, the forecast $\ln\left[\mathbb{E}_0(Y_t)/Y_{0-}\right]$ starts at the initial drop $-\delta = -0.6$ at $t = 0$ and then increases over time due to anticipated vaccine arrival and eventually approaches the straight line, $\widehat{g}\,t = 0.08t$.

[9] As COVID-19 is unexpected, we calculate $\mathbb{E}_{t_0}[Y_s]$ from $t_0$, but divide the forecast by $Y_{t_0-}$ for empirical measurement purposes.

Figure 2: The Effect of the Vaccine Arrival Rate $\lambda$ on $\ln[\mathbb{E}_0(Y_t)/Y_{0-}]$

The forecast $\ln[\mathbb{E}_0(Y_t)/Y_{0-}]$ starts at $-\delta = -0.6$ at $t = 0$ and eventually converges to the business-as-usual scenario, depicted by the straight line, $\widehat{g} \cdot t$, as $t \to \infty$. The higher the value of $\lambda$, the faster the convergence. The parameter values are: $n = \delta = 0.6$, $\widehat{g} = 0.08$, and $g = 0.064$.




Intuitively, if an effective vaccine is expected to arrive far out in the future (lower $\lambda$), then forecast revisions will be large for both near term and longer term forecasts (the red dashed line) — that is there is effectively a permanent downward jump in earnings followed by a different pandemic regime growth rate than the one in the non-pandemic regime. In contrast, if we expect a vaccine in a year, then the longer-term forecasts will be revised down much less in comparison to the near-term forecasts.

Figure 3 examines the effect of the size of the jump $n$ and pandemic growth rate $g$ on the term structure of the forecast revision $\ln[\mathbb{E}_0(Y_t)/Y_{0-}]$. Take the blue line as the benchmark case, we implement two experiments to investigate how the shape of the term structure change with respect to different $n$ and $g$.

First, the size of the initial jump is determined by $n$. As we change $n$ from 0.3 to 0.6 (from the blue line to the red dashed line), we see a larger drop in $\ln\left[\mathbb{E}_0(Y_t)/Y_{0-}\right]$ in lower horizon compared to longer horizon. From an identification point of view, this observation implies that data in the shorter horizon are driving the identifiability of $n$ since they are very informative about the initial jump in earnings.

Second, if we further change $g/\widehat{g}$ to 0, we see that there are sizable drops in the level of the $\ln\left[\mathbb{E}_0(Y_t)/Y_{0-}\right]$ (black dotted line compared to the red dashed line) in all horizons. Moreover, the drop is smaller in the short horizon compared to the median and longer horizons. Therefore, the data with longer horizons can help us better identify $g$ given the larger difference generated by the $g/\widehat{g}$ parameter in the median and longer horizon.

## 2.5 Estimation

Using insights from Figure 2 and Figure 3, we take our model to data on analyst forecasts in the following manner. In reality, we do not observe analyst forecasts at $t_0$, which is the immediate moment after the pandemic arrival time. Instead, we observe forecasts at a later time, $t$. As such, we will employ the approximation $Y_t/Y_{t_{0-}} \approx Y_{t_0}/Y_{t_{0-}} = e^{-\delta}$ and assume $\delta = n$ to obtain the following relation:

$$\begin{aligned} \frac{1}{Y_{t_0-}}\mathbb{E}_t[Y_s] = \frac{Y_t}{Y_{t_0-}}\frac{1}{Y_t}\mathbb{E}_t[Y_s] &\approx e^{-\delta}\left[\frac{\lambda}{\lambda-g+\widehat{g}}\left[e^{\widehat{g}(s-t)}-e^{(g-\lambda)(s-t)}\right]e^{n}+e^{(g-\lambda)(s-t)}\right] \\ &= \frac{\lambda}{\lambda-g+\widehat{g}}\left[e^{\widehat{g}(s-t)}-e^{(g-\lambda)(s-t)}\right]+e^{-n}e^{(g-\lambda)(s-t)}. \end{aligned} \tag{8}$$

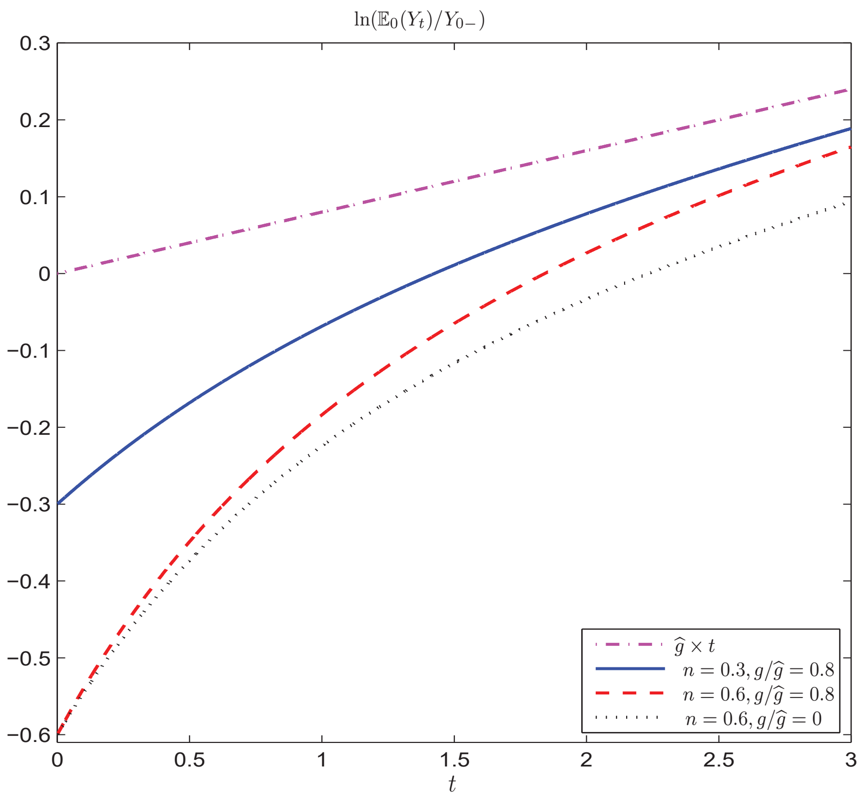
That is, we assume that the jump which in our model occurs over an instant takes place over the period from the end of February 20 to May 14 of 2020.

Moreover, we aggregate corporate earnings forecasts at the firm level up to the industry level, which we denote by $j$. The main dependent variable of interest given by the right side of Equation (8) is constructed in the following manner. As $Y_{j,t_0-}$ is not empirically observable, we measure $Y_{j,t_0-}$ by using the earnings forecast expression before the arrival of COVID-19: $\mathbb{E}_{t_{0-}}[Y_{j,s}] = Y_{j,t_0-}e^{\widehat{g}^{(j)}(s-t_0)}$, where $\widehat{g}^{(j)}$ is the long-run growth rate in the non-pandemic regime, which as we discuss below is observable. Equivalently, we have

$$Y_{j,t_0-} = \exp\left[-\widehat{g}^{(j)}\cdot(s-t_{0-})\right]\cdot\mathbb{E}_{t_{0-}}[Y_{j,s}]\,. \tag{9}$$

Figure 3: The Effect of $n$ and $g$ on $\ln\left[\mathbb{E}_0(Y_t)/Y_{0-}\right]$

This figure presents the forecast $\ln\left[\mathbb{E}_0(Y_t)/Y_{0-}\right]$ based on four sets of parameter values to show the sensitivity of the path to the jump parameter $n$ and the pandemic growth rate parameter $g$. The other parameter values are: $\lambda = 0.7$, $\widehat{g} = 0.08$. The magenta dash-dotted line is the business-as-usual scenario. The blue solid line shows the path when $n = 0.3$ and $g/\widehat{g} = 0.8$. The red dashed line shows the path when $n = 0.6$ and $g/\widehat{g} = 0.8$. The black dotted line presents the path when $n = 0.6$ and $g/\widehat{g} = 0$.



Using equations (8) and (9), and taking natural logs on both sides, we obtain the following relation that we take to data:

$$\ln\left[\frac{\mathbb{E}_t[Y_{j,s}]}{e^{-\widehat{g}^{(j)}(s-t_{0-})}\mathbb{E}_{t_0-}[Y_{j,s}]}\right] = \ln\left[\frac{\lambda}{\lambda - g^{(j)} + \widehat{g}^{(j)}}\left(e^{\widehat{g}^{(j)}(s-t)} - e^{(g^{(j)}-\lambda)(s-t)}\right) + e^{-n^{(j)}}e^{(g^{(j)}-\lambda)(s-t)}\right]. \quad (10)$$

We parameterize the earnings jump parameter $n^{(j)}$ for firms in industry $j$ by

$$n^{(j)} = n_0, \quad (11)$$

The growth rate $g$ for firms in industry $j$ in the pandemic regime, $g^{(j)}$, is parameterized as

$$g^{(j)} = g_0 \cdot \widehat{g}^{(j)}. \quad (12)$$

That is, the growth rate in the pandemic regime $g^{(j)}$ is a multiple of $\widehat{g}^{(j)}$, the growth rate in the non-pandemic regime for firms in industry $j$. The ratio between the two growth rates, $g^{(j)}/\widehat{g}^{(j)}$, captures the average difference in growth rates across the two regimes.

Finally to estimate our model, we need to specify analyst forecast errors. Denote the earnings forecast of industry $j$ at $t$ of horizon $s$ by $f^s_{t,j}$. Suppose

$$f^s_{t,j} = \mathbb{E}_t[Y_{j,s}] \cdot u^s_{t,j}, \quad (13)$$

where $u^s_{t,j}$ is a mean one random variable that is conditionally independent of $\mathbb{E}_t[Y_{j,s}]$. Then, taking logs on both sides of (13) and then using (10) for $\mathbb{E}_t[Y_{j,s}]$, we obtain

$$\begin{aligned}\ln\left[\frac{f^s_{t,j}}{e^{-\widehat{g}^{(j)}(s-t_{0-})}f^s_{0-,j}}\right] \quad &= \quad \ln\left[\frac{\lambda}{\lambda - g^{(j)} + \widehat{g}^{(j)}}(e^{\widehat{g}^{(j)}(s-t)} - e^{(g^{(j)}-\lambda)(s-t)}) + e^{-n^{(j)}}e^{(g^{(j)}-\lambda)(s-t)}\right] \\ &\quad +\ln\left(\frac{u^s_{t,j}}{u^s_{0-,j}}\right) \quad (14)\end{aligned}$$

Therefore, an identifying restriction allowing for estimation of Equation (14) using non-linear least squares (NLS) (Cameron and Trivedi (2005)) is given by:

$$\mathbb{E}\left[\ln\left(\frac{u^s_{t,j}}{u^s_{0-,j}}\right)\middle| s-t, \widehat{g}^{(j)}\right] = 0. \quad (15)$$

Obviously, if forecasts are rational, i.e., forecast errors are white noise, then the exclusion restriction is satisfied. More generally, as long as the log difference of these biases across forecast horizons is uncorrelated with our independent variables, then consistent parameter estimates can still be retrieved using NLS.

This exclusion restriction is plausible. Intuitively, imagine if analyst forecasts were overly optimistic at 5 years out and overly-pessimistic say at 2 years out, then this can bias estimates of the vaccine arrival rate. We know of no obvious research to suggest that our exclusion restriction is unlikely. At the same time, our setting is unique in that we can also compare our estimates implied with forecast to subsequent actual outcomes. Presumably, biased estimates will not be very predictive or subsequent outcomes. But we show below that the vaccine arrival estimate is predictive. Moreover, corporate CEO revenue forecasts also suggest a quick arriving vaccine (Barry, Campello, Graham, and Ma (2021)). While revenue forecasts are not the same as earnings forecasts, they tend to be correlated and hence provide another piece of evidence on the plausibility of our exclusion restriction.

## 2.6 Comments

The upside of our baseline set-up is parsimony. In practice, rather than assuming that a successful vaccine is a silver bullet that instantly brings the economy back to normal upon its arrival as in our baseline model, we may consider a more realistic setting where a successful vaccine development brings the economy back to normal in several stages over time. These stages might correspond to an increasing fraction of the population being vaccinated over time. For example, consider the following setting with $N$ sequentially ordered stages, denoted by $\{S_1, \cdots, S_N\}$, in addition to the pandemic regime, which we denote by $S_0$. We assume that as the stage transitions from stage $S_m$ to stage $S_{m+1}$ at stochastic time $\tau_m$, where $m = 0, \cdots, N-1$, at a constant rate of $\lambda_m$ per unit of time, earnings jumps upward by a constant fraction $\delta_m > 0$. That is, $Y_{\tau_m} = Y_{\tau_m -} e^{\delta_m}$. We can compute the earnings forecast and other key objects in this more general model in closed form, but the model would be less parsimonious.

# 3 Data and Variables

## 3.1 Earnings Forecasts

We obtain the forecasts on earnings per share (EPS) and growth rate forecasts from the monthly IBES summary history files from WRDS. Our data is from January 2020 to May 2020. We keep all stocks that are also in CRSP. We set the starting date of the pandemic regime, $t_0$, to be February 20, 2020. We take the median forecast for each firm in May as the consensus forecast during the pandemic period. We treat the forecasts in January as the most recent non-pandemic period forecast. That is, we link our model notations to our empirical measurement as follows: January 2020 is our $t_0-$, May 2020 is time $t$ for our forecast, and $s$ is the fiscal year end date of the forecasts.

Using February and March of 2020 forecasts is problematic from the point of view of identification since we want timely measures of analyst expectation revisions from just before COVID-19 arrived to after its arrival. February 2020 may capture a bit of information about the pandemic since some analysts might have started revising their forecasts based on infections in other countries such as China. On the other hand, March 2020 might not capture the full extent of the pandemic regime to the extent some analysts might have been slow in revising. As such, we view using January 2020 forecasts as cleanly capturing non-pandemic earnings expectations and either April or May 2020 forecasts as capturing revisions accounting for the pandemic and hence embedding information regarding vaccines. We prefer May 2020 to April 2020 since almost all the analysts have revised their forecasts by then.[10]

We label the EPS annual forecasts based on the time gap between their forecast period end date s (i.e. the fiscal end year end date of the company) and the forecast date t, i.e., the gap $(s-t)$. If the time gap is less than 365 days, we label the annual forecast as $FY1_t$. If the time gap is between 366 days and 730 days, we label the annual forecast as $FY2_t$. We also similarly collect FY3 and FY4 annual forecasts from IBES. In addition, we convert LTG

[10]Moreover, most of the government intervention programs have already been announced and hence ought to be reflected in analyst forecasts as well by then.

forecasts, which are defined as long-run growth rates from the previous announced earnings out to 5 years, to FY5 forecasts.

We use this methodology to label annual forecasts instead of using the classification provided by IBES because their classification is based on when the actual earnings is reported not when the fiscal year ends. For example, IBES will label an April annual forecast of a firm with a fiscal year that ended the previous month as FY1 if the firm has not yet reported the actual earnings for that fiscal year. We want FY1 to reflect future earnings so we use our methodology instead.

In our empirical analysis, FY1 forecasts need to be adjusted for the fact that a certain fraction of the fiscal year has already been realized before the pandemic arrived at $t_0$. Consider a firm in our sample that has a fiscal year ending in October 2020 (time $s$ in our model). In this case, for $FY1_t$, the FY1 earnings forecast for the period from November 2019 to October 2020, made in May 2020 (our $t$), only the sub-period between February 20, 2020 (our $t_0$) to October 2020 is exposed to COVID-19.

Therefore, we need to make adjustments to $FY1_t$ forecasts (e.g. May as our $t$) considering the differential impact of the pandemic on earnings resulting from heterogeneous fiscal year end dates. What enters into our calculation of earnings forecast in Equation (10) at $t$ (May in our empirical analysis) is adjusted as follows:

$$FY1_t^{adj} = FY1_t \cdot \left(\frac{1}{s - t_0}\right) + FY1_{t_0-} \cdot \left(1 - \frac{1}{s - t_0}\right) , \tag{16}$$

where $(s - t_0)$ is the fraction of the fiscal year that is exposed to COVID-19.

For the preceding example, $s - t_0 = (10 - 2)/12$ (the event time $t_0$ is February 2020 and time $s$ in Equation (16) is October 2020.) That is, 8/12 =2/3 of the annual earnings is after the pandemic arrival and the other 4/12=1/3 is non-pandemic. Our adjusted earnings forecast at $t$ (in May for our empirical analysis) is then given by $FY1_t^{adj} = (3/2)FY1_t - (1/2)FY1_{t_0-} = FY1_t + 0.5 \times (FY1_t - FY1_{t_0-})$. That is, the adjusted annual earnings forecast $FY1_t^{adj}$ is equal to the unadjusted FY1 forecast $FY1_t$ plus a term, which accounts for the change of forecasts caused by the pandemic arrival. If pandemic is bad news for the firm, i.e., $FY1_t < FY1_{t_0-}$, this earnings forecast is adjusted downward by $0.5 \times (FY1_t - FY1_{t_0-})$, where the multiple 0.5 reflects the ratio between the non-pandemic 4- month duration and pandemic 8-month duration. In our sample, the non-pandemic forecast $FY1_{t_0-}$ is the FY1 forecasts in January and $FY1_t$ is the unadjusted FY1 forecasts in May.

We merge IBES forecasts with CRSP market capitalization data using historical 8-digit CUSIP identifiers.[11] We then merge in the 8-digit GICS code obtained from Compustat. On each date in our IBES sample, we set the negative values in adjusted FY1 to the lowest positive observation in adjusted FY1 on that date. We also set the negative values of FY2 on each date to the lowest positive FY2 observation on each date. We repeat the same procedure for FY3, FY4 and FY5. We then aggregate the EPS forecasts, pre-pandemic growth rate forecasts, non-pandemic earnings, and time until fiscal year end to the 8-digit GICS industries using the end of 2019 market capitalization from CRSP as the weights. We winsorize these industry $\frac{\mathbb{E}_{t_0}[Y_s]}{Y_{t_0-}}$ and $\widehat{g}$ at the 5% level.

The summary statistics for our dependent variables are presented in Table 1. In Panel A, we report the distribution of $\frac{\mathbb{E}_{t_0}[Y_s]}{Y_{t_0-}}$ for the mid-May 2020 forecasts. The mean is 1.16 and the standard deviation is 0.54. The $\ln\left(\frac{\mathbb{E}_{t_0}[Y_s]}{Y_{t_0-}}\right)$ has a mean of 0.01 with a large standard deviation of 0.61. The mean $(s-t)$ is 2.57 for the May 2020 forecasts.

In Figure 4, we take a closer look at the standard deviation of these forecasts by plotting the industry forecast revisions separately for FY1 to FY5 forecasts. We can see that the FY1 forecast within twelve months before forecast end are significantly revised down: 54% on average for the May 2020 forecasts across the industries in our sample. This is consistent with a significant negative jump on average in our model. But we can also see that the FY2 forecasts farther out are not nearly as impacted.

## 3.2 Leverage, Face-to-Face, and Customer Interaction Measures

We obtain the GICS code and calculate the market leverage of each firm using Compustat. Market Leverage is calculated at the end of 2019 using the following formula: long-term debt (dlttq) plus debt in current liabilities (dlcq) all divided by the sum of market capitalization (prccq $\times$ cshoq) and total assets (atq) net common equity (ceqq).

[11] For the unmatched cases, we obtain additional matching using the official tickers and 6-digit CUSIP.

Table 1: Summary Statistics

This table summarizes the mean, standard deviation, and the quartiles of the key variables used in our main analysis at 8-digit GICS industry level. $\mathbb{E}_t[Y_s]/Y_{t_0-}$ is the earnings forecasts in month $t$ divided by the non-pandemic earnings $Y_{t_0-}$, which is the FY1 forecasts in January 2020 discounted by the I/B/E/S growth rate forecasts in January 2020. $\ln(\mathbb{E}_t[Y_s]/Y_{t_0-})$ is the natural log of $\mathbb{E}_t[Y_s]/Y_{t_0-}$. $s-t$ is the horizon of the earnings forecasts in month $t$, which is the difference between the date of the forecast period end and the I/B/E/S statistical period in month $t$. We include the May sample of I/B/E/S summary statistics in 2020 in our analysis. The sample includes the earnings forecasts with horizons up to 5 years. Panel A presents the summary statistics of $\mathbb{E}_t[Y_s]/Y_{t_0-}$, $\ln(\mathbb{E}_t[Y_s]/Y_{t_0-})$ and $s-t$ in May 2020. Panel B contains the summary statistics of other key variables. Face-to-Face Score is first constructed at the occupation level using O*Net Main database and then aggregated to industry level using the BLS Industry-occupation matrix data (from 2018). Market Leverage is calculated at the end of 2019 using the following formula, (long-term debt+ debt in current liabilities)/(fiscal year end market capitalization + total assets - common equity). $\widehat{g}$ is the I/B/E/S forecasts of growth rates in January 2020. All the firm level variables are aggregated to the industry level using 8-digit GICS code, weighted by the market values of the companies in each industry at the end of 2019. $\mathbb{E}_t[Y_s]/Y_{t_0-}$ is winsorized at 5% level on each date within each horizon. Forecasts horizons are defined by the distance between the forecast end date and the I/B/E/S statistical period. $\widehat{g}$ is also winsorized at 5% level.

(a) Panel A: Distribution of $\mathbb{E}_t[Y_s]/Y_{t_0-}$ and $s-t$ in May 2020

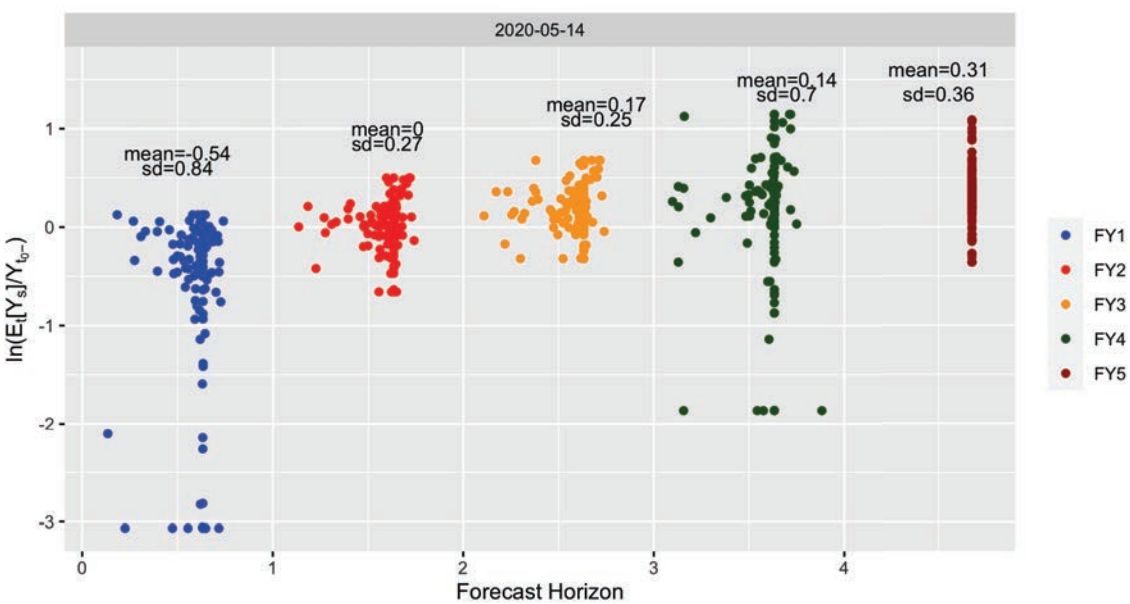
| | Mean | SD | P0 | P25 | P50 | P75 | P100 |
|---|---|---|---|---|---|---|---|
| $\mathbb{E}_t[Y_s]/Y_{t_0-}$ | 1.16 | 0.54 | 0.05 | 0.86 | 1.10 | 1.39 | 3.14 |
| $\ln(\mathbb{E}_t[Y_s]/Y_{t_0-})$ | 0.01 | 0.61 | -3.07 | -0.15 | 0.10 | 0.33 | 1.14 |
| $s-t$ | 2.57 | 1.45 | 0.13 | 1.56 | 2.62 | 3.63 | 4.67 |

(b) Panel B: Distribution of other variables used in analysis

| | Mean | SD | P0 | P25 | P50 | P75 | P100 |
|---|---|---|---|---|---|---|---|
| Market Leverage | 0.20 | 0.10 | 0.03 | 0.13 | 0.19 | 0.25 | 0.72 |
| Face-to-Face Score | 3.94 | 0.14 | 3.59 | 3.85 | 3.90 | 4.01 | 4.33 |
| Customer Score | 3.45 | 0.45 | 2.54 | 3.09 | 3.44 | 3.80 | 4.48 |
| Blinder Score | 2.97 | 0.24 | 2.57 | 2.75 | 2.95 | 3.13 | 3.76 |
| $\widehat{g}$ | 0.10 | 0.09 | -0.05 | 0.06 | 0.08 | 0.13 | 0.35 |

Figure 4: $\ln(\mathbb{E}_t[Y_s]/Y_{t_0-})$ Over Forecast Horizons

This figure plots the natural log of the industry level I/B/E/S earnings forecasts divided by the non-pandemic earnings, $\ln(\mathbb{E}_t[Y_s]/Y_{t_0-})$, against the horizons of the forecasts $(s-t)$. $Y_{t_0-}$, the non-pandemic earnings, is the FY1 forecasts in January 2020 discounted by the I/B/E/S growth rate forecasts in January 2020. The May 2020 cross section is plotted. Forecast horizons are marked with different colors. Forecast are defined by the distance between the forecast end date and the I/B/E/S statistical period.



We then use the O*Net Main database in the U.S. about occupational information to construct the face-to-face exposures of different industries. O*Net collects information on 974 occupations. They are based on the Standard Occupational Classification (SOC), the last update of which was done in 2010. O*Net surveys people in these occupations, asking about the knowledge, skills, and abilities used to perform the activities and tasks of their occupations. Our face-to-face measure is based on Montenovo, Jiang, Rojas, Schmutte, Simon, Weinberg, and Wing (2020).

They use questions taken from the 2019 Work Context module. The questions used in face-to-face measure are: (1) How often do you have face-to-face discussions with individuals or teams in this job? And (2) To what extent does this job require the worker to perform job tasks in close physical proximity to other people? These measures are typically provided on a 1-5 scale, where 1 indicates that a task is performed rarely or is not important to the job, and 5 indicates that the task is performed regularly or is important to the job.

There is also a direct question that asks people to rate how much they work with customers in the O*Net survey. The question is: How important is it to work with external customers or the public in this job? We take the average score for each occupation for this alternative measure. One issue with this customer measure is that it does not necessarily capture face-to-face contact. To this end, we have also constructed a customer measure from Blinder (2009) based on the following questions: (1) establishing and maintaining personal relationships, (2) assisting and caring for others, (3) performing for or working directly with the public, (4) selling or influencing others, and (5) social perceptiveness.

The O*Net provides two ways that people weight how an occupation uses these characteristics: Importance and Level. That is, people in an occupation are asked to rate how important the characteristic is in their job and the level of use of the characteristic in their job. We use the Importance score of each characteristic and take the simple average of the Importance scores to make what we call the Blinder index for each occupation. The social perceptiveness question is in the Social Skills part of the O*Net. The other four measures are in the Work Activities part of the O*Net.

We have occupation-level measures of face-to-face and the two customer measures. We then convert them to an industry-level measure. To do this, we use the BLS Industry-occupation matrix data (from 2018).[12] In the BLS data, for every industry, they measure what percentage of workers work in a given occupation. (They also use the SOC occupation codes just like the O*Net). So we take the O*Net occupation measures and for each industry weight them by the percentage of workers in that industry that work in the occupation. We take a weighted-average to come up with the industry measures. One issue is that the BLS uses NAICS codes for industries. We convert these to 8-digit GICs codes using a crosswalk.[13]

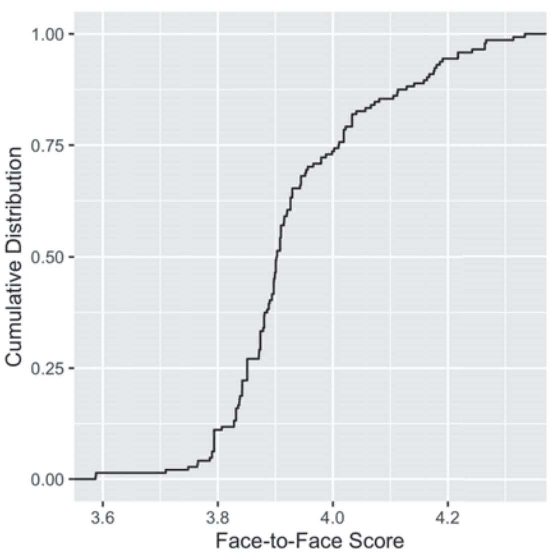
The summary statistics for leverage and these three face-to-face measures are provided in Panel B of Table 1. The mean Market Leverage ratio is 0.2 with a standard deviation of 0.1. The mean Face-to-Face Score is 3.94 with a standard deviation of 0.14. The mean Customer Score is 3.45 with a standard deviation of 0.45, while the Blinder Score has a mean of 2.97 and a standard deviation of 0.24. These measures are correlated (around 0.4 to 0.5 in pairwise correlations). The statistics for $\widehat{g}$ are also displayed — the mean (annual) non-pandemic growth rate is 10% with a standard deviation of 9%.

In our empirical analysis, we will work with percentiles of these measures as opposed to the values themselves. Figure 5 show the empirical cumulative distribution of our Face-to-Face Score and Market Leverage measures, respectively. The correlation at the industry level of face-to-face ranks and leverage ratio ranks is 0.4. There are a number of good economic reasons why these two industry attributes are correlated. Airline and hotels for instance have high Face-to-Face Scores and are also industries that have physical assets such as land or planes that are used for collateralized borrowing. Our goal in this paper is not to disentangle these two effects. Hence we will use both of these measures interchangeably to model latent growth rates in our baseline specifications. We will consider the two customer measures in our robustness exercises.

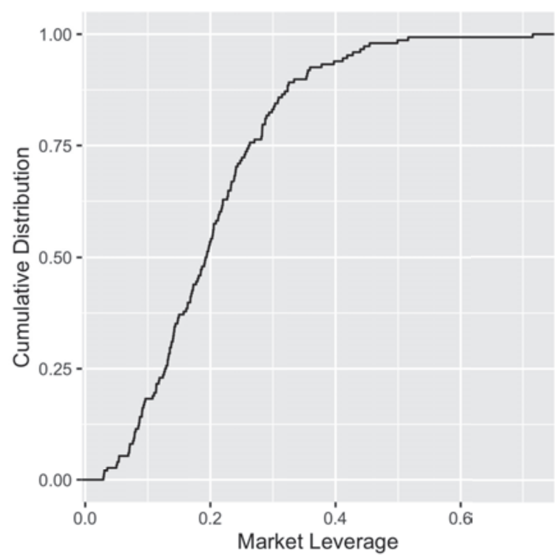
[12]See https://www.bls.gov/emp/tables/industry-occupation-matrix-industry.htm
[13]See https://sites.google.com/site/alisonweingarden/links/industries

Figure 5: The Empirical Distributions of Face-to-Face Scores and Market Leverage

This figure plots the empirical cumulative distributions of Face-to-Face Scores and Market Leverage of industries defined by 8-digit GICS codes. Subfigure (a) is the cumulative distribution of Face-to-Face Scores. Face-to-Face Score is first constructed at the occupation level using O*Net Main database and then aggregated to the industry level using the BLS Industry-occupation matrix data (from 2018). Subfigure (b) is the cumulative distribution of Market Leverage. Market Leverage is calculated at the end of 2019 using the following formula, (long-term debt+ debt in current liabilities)/(market capitalization + total assets - common equity). The variables are from Compustat. In Compustat variable names, the formula is the following, Market Leverage = (dlttq + dlcq)/(atq - ceqq + prccq * cshoq).




(a) The Cumulative Distribution of Face-to-Face Scores




(b) The Cumulative Distribution of Market Leverage

Table 2: NLS Results Using the I/B/E/S Sample in May 2020

This table presents the coefficients and bootstrap confidence intervals from non-linear least square regressions of Equation (14). The regressions are run using I/B/E/S summary statistics in May 2020. $\mathbb{E}_t[Y_s]/Y_{t_0-}$ is the earnings forecasts divided by the non-pandemic earnings. The dependent variable is the natural log of $\mathbb{E}_t[Y_s]/Y_{t_0-}$. $Y_{t_0-}$, the non-pandemic earnings, is the FY1 forecasts in January 2020 discounted by the I/B/E/S growth rate forecasts in January 2020. The explanatory variables include the horizon of the earnings forecasts $s-t$ and the non-pandemic (January 2020) I/B/E/S forecasts of growth rates $\widehat{g}$. $\lambda$ is the vaccine arrival rate. $g_0$ represents the proportional change in the growth rate. $n_0$ governs the size of the jump in earnings. Columns (1)-(3) present the results from three different restrictions on the model parameters. Column (1) contains the results of the unconstrained regression. Column (2) contains the results restricting $\lambda = 0$. Column (3) contains the results restricting $g_0 = 1$. We keep observations with non-missing $\mathbb{E}_t[Y_s]/Y_{t_0-}$ and $\widehat{g}$. The 95% bootstrap confidence intervals are reported in square brackets. We also present the likelihood ratio test statistics for the restricted models. All regressions are run using industry market capitalization as weights.

| | (1) | (2) | (3) |
|---|---|---|---|
| $\lambda$ | 0.674 | | 0.577 |
| | [0.17,1.65] | | [0.2,1.24] |
| $g/\widehat{g}$ | 0.827 | 1.223 | |
| | [-0.61,1.42] | [0.86,1.52] | |
| $n$ | 0.603 | 0.234 | 0.573 |
| | [0.31,1.48] | [0.12,0.35] | [0.3,1.19] |
| Num.Obs. | 677 | 677 | 677 |
| LR. Stat. | | 13.54 | 4.16 |

# 4 Empirical Results

## 4.1 Baseline Specification

In Table 2, we present the coefficients and bootstrap confidence intervals from non-linear least square regressions of Equation (14) using May 2020 earnings forecasts. The dependent variable is the natural log of $\mathbb{E}_t[Y_s]/Y_{t_0-}$, i.e., the revision of forecasts between January and May 2020. The explanatory variables include the (remaining) duration of time-$t$ earnings forecasts $(s-t)$ and the non-pandemic (January 2020) forecasts of the growth rate $\widehat{g}$.

Column (1) contains the results for our baseline and unconstrained model. The estimate of $\lambda$ is 0.674 with 95% bootstrap confident interval of $[0.17, 1.65]$. So the vaccine that returns the earnings to normal is expected in 1/0.674 or 1.48 years.[14] The estimate of $g_0$ is 0.827, indicating that pandemic growth rates are lower than during the non-pandemic periods. The confidence interval is $[-0.61, 1.42]$. Moreover, industries experience a large downward jump in earnings level, captured by $n_0 = 0.603$, which has a confidence interval of $[0.31, 1.48]$. Notice that this parameter also captures the expected reflation in earnings for these industries when the vaccine does arrive. The jump in earnings level is given by $1 - e^{-n}$, which means that there is around a 50% drop in earnings immediately following the arrival of COVID-19.

Recall that the average earnings FY1 forecast revision in the summary statistics is nearly 50% with a fat left-tail. The nonlinear least squares model, which is value weighted, will fit this tail, giving a sizable estimate for the downward jump in earnings $g_0$. There is also a fat-left tail in further out forecasts, which will then impart an attribution of low of negative growth rates in the pandemic regime compared to the non-pandemic regime. Finally, the high $\lambda$ estimate comes from the intuition discussed earlier that there is a sizable disconnect between downward revisions in FY1 forecasts compared to subsequent ones.

In Column (2), we present the estimates for the constrained model where we set $\lambda = 0$, i.e. assuming there is no vaccine. The estimate for $g_0$ is 1.223 with a 95% bootstrap confidence interval of $[0.86, 1.52]$. When $\lambda$ is forced to be zero, the constrained model has to compensate with a positive $g_0$ to account for the higher levels of FY2-FY5 earnings forecasts compared to FY1. Moreover, the initial jump in earnings is $n_0 = 0.234$ with a confidence interval of $[0.12, 0.35]$. This implies a downward jump of $1 - e^{-n_0} = 0.21$ or 21%.

The estimates of the constrained model are nonsensical because they imply higher pandemic growth rates and a small initial jump. Of course, we know from the summary statistics that the FY1 forecast revision for the median industry is nearly 50%. These nonsensical estimates are of course coming from constraining $\lambda$, which is equivalent to an omitted variables

[14]There has been significant attention to the question of when vaccines will arrive and if they will return the economy to normal. For instance, see the *McKinsey Report* (July 29, 2020) "On pins and needles: Will COVID-19 vaccines save the world", and an article in the *Washington Post* (August 2, 2020), entitled "A coronavirus vaccine won't change the world right away". Our estimate of the vaccine arrival rate $\lambda$ as far as we know is the first systematic attempt to speak to this question.

bias where expectations of an imminent vaccine are ignored in estimating damage functions. This lack of fit is reflected in the likelihood ratio test statistic of 13.54 clearly rejecting the constrained model in Column (2) in favor of the unconstrained model in Column (1).

In Column (3), we present the estimates for the constrained model where we set $g_0 = 1$; i.e. there was no damage to growth rates. $\lambda$ is now estimated to be 0.577 with a 95% confidence interval of $[0.2, 1.24]$, and $n_0$ is 0.573 with a 95% confidence interval of $[0.3, 1.19]$. Notice that since we are assuming there is no growth impairment, there is a lower estimated $\lambda$ because the higher growth rate will explain more of the difference between revisions of long-horizon forecasts to short-horizon forecasts. The likelihood ratio statistic comparing Column (1) to (3) is 4.16, rejecting at the 5% level the constrained model in Column (3) in favor of the unconstrained model.

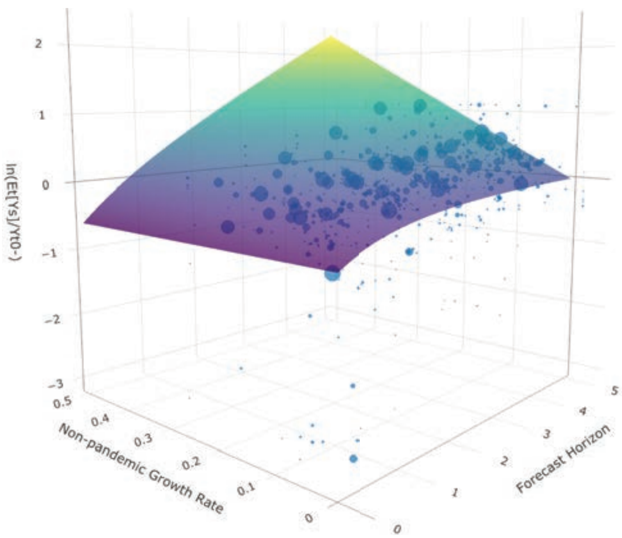
Another way to see that the unconstrained model fits the data is to compare the predicted values as a function of our two main independent variables the forecast horizon $(s - t)$ and non-pandemic industry growth rate $\widehat{g}_0^{(j)}$. These plots are in Figure 6. Panel (a) shows the fitted values for the unconstrained model from Column (1), while panels (b) and (c) show the fitted values from the constrained models in columns (2) and (3), respectively. It is clear from these 3-D plots that only the unconstrained model can fit the data. The constrained models generate poor fits of the data.

## 4.2 Subsamples

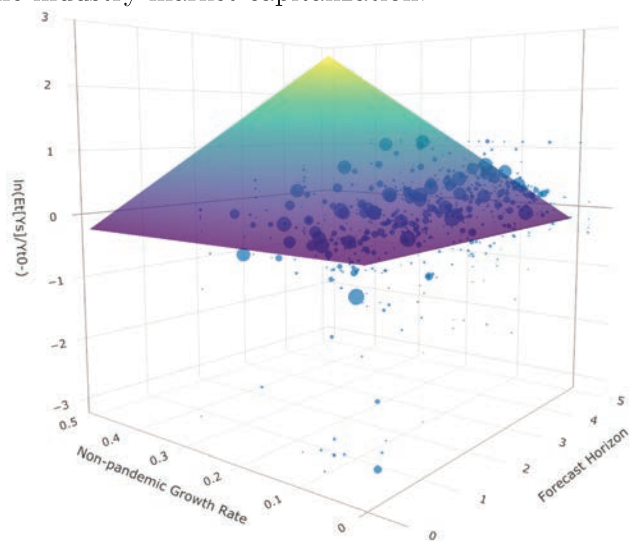
Since earlier work suggests that levered or face-to-face industries are particularly hit by COVID-19 and should be the most informative regarding the damage function, we re-run our model using observations from just these industries. The results are presented in Table 3. In Panel A, we present the results for high face-to-face industries based on our main face-to-face measure. An industry is categorized in the high group if its face-to-face score is in the top tercile of the cross-sectional distribution. $\lambda$ is estimated to be 1.086, higher than our estimate from Column (1) of Table 2. $g_0$ is estimated to be 0.704, which is smaller than the figure from Column (1) of Table 2. However, its confidence interval of $[-1.09, 2.29]$ is quite wide. Hence, there is not a statistical difference across the two sets of estimates.

Figure 6: The Surfaces of the Estimated Models

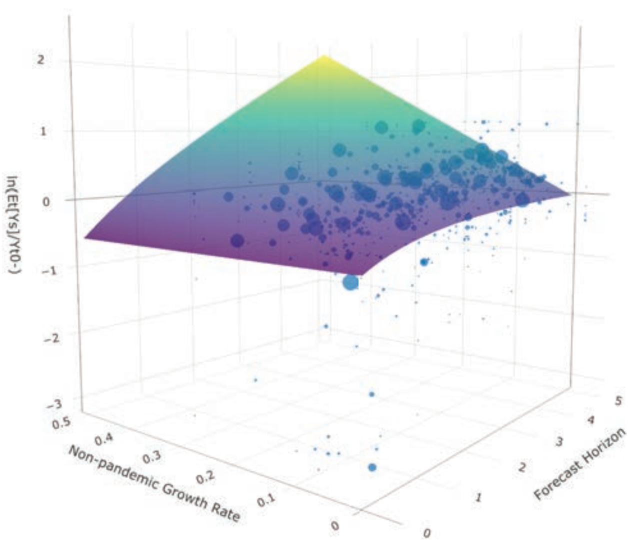
This figure plots the observations and fitted value of $\ln(\mathbb{E}_t[Y_s]/Y_{t_0-})$ using the parameter estimates of Equation (14) on the I/B/E/S sample from May of 2020. All the subfigures plot $\ln(\mathbb{E}_t[Y_s]/Y_{t_0-})$ and the fitted surface against the pre-pandemic growth rate and the horizons of forecasts. Subfigure (a) uses estimates from Column (1) in Table 2. Subfigure (b) uses estimates from Column (2) in Table 2. Subfigure (c) uses estimates from Column (3) in Table 2. The $\ln(\mathbb{E}_t[Y_s]/Y_{t_0-})$ observations are the blue dots. The size of the dot indicates the industry market capitalization.




(a) Unconstrained




(b) Constraint: $\lambda = 0$




(c) Constraint: $g_0 = 1$

Similarly, the coefficient for $n_0$ is 2.678 which is larger than the figure from Table 2. But again, they are not statistically different. In other words, the damage function estimated off of this subsample of firms is quite similar to the overall sample. The same can be said for the constrained models in columns (2) and (3). So overall, our earlier conclusions based on the overall sample continues to hold for this subsample.

In Panel B, we present the results for high leverage industries based on our main leverage measure. An industry is categorized in the high group if its leverage score is in the top tercile of the cross-sectional distribution. Our qualitative conclusions are quite similar to those from Panel A. We have also repeated these exercises by using a net market leverage measure where we deduct corporate cash and short-term investments and by replacing our baseline face-to-face measure with our two customer interaction measures. The conclusions are similar, pointing to the robustness of our damage function estimates.

## 4.3 Placebo Analysis

In Table 4, we then consider a placebo exercise. We run exactly the same empirical procedure but using the forecasts in 2019 far before COVID-19. We report in Table 4 the regressions results with the constraint that $\lambda \geq 0$. Our estimates are zero for both the placebo full sample and the placebo subsamples of high face-to-face and high leverage industries. $g_0$ is 0.746 for the full sample with a tight 95% bootstrap confidence interval of $[0.44, 0.83]$. The point estimates are similar to the placebo subsample of high face-to-face and high leverage industries, though the confidence interval for the placebo high leverage subsample estimate is quite wide. $n_0$ is -0.064 with a tight confidence interval of $[-0.18, -0.02]$ for the full placebo sample. This is quite small in comparison to our earlier estimates. the same conclusions hold for the placebo subsamples. These exercises indicate that our model estimates using the COVID-19 sample are informative.

In Figure 7, we plot the dependent variables, i.e., the forecasts revisions, for the placebo full sample that are analogous to those shown in Figure 4. We can see that the big difference between the COVID-19 period and the other placebo period is that one does not typically see such a large divergence in revisions across FY1 and FY2 forecasts. Understandably, in

Table 3: NLS Results Using Subsamples

This table presents the coefficients and bootstrap confidence intervals from non-linear least square regressions of Equation (14) on subsamples of industries split by the terciles of Face-to-Face Scores and Market Leverage. The top tercile of the Face-to-Face Scores are classified as High Face-to-Face. The top tercile of the Market Leverage are classified as High Market Leverage. The regressions are run using I/B/E/S summary statistics in May 2020. $\mathbb{E}_t[Y_s]/Y_{t_0-}$ is the earnings forecasts divided by the non-pandemic earnings. The dependent variable is the natural log of $\mathbb{E}_t[Y_s]/Y_{t_0-}$. $Y_{t_0-}$, the non-pandemic earnings, is the FY1 forecasts in January 2020 discounted by the I/B/E/S growth rate forecasts in January 2020. The explanatory variables include the horizon of the earnings forecasts $s-t$ and the non-pandemic (January 2020) I/B/E/S forecasts of growth rates $\widehat{g}$. $\lambda$ is the vaccine arrival rate. $g_0$ represents the proportional change in the growth rate. $n_0$ governs the size of the jump in earnings. Panel A and B contain the results using the High Face-to-Face subsample and High Market Leverage subsample correspondingly. In each panel, Columns (1)-(3) present the results from three different restrictions on the model parameters. Column (1) contains the results of the unconstrained regression. Column (2) contains the results restricting $\lambda = 0$. Column (3) contains the results restricting $g_0 = 1$. We keep observations with non-missing $\mathbb{E}_t[Y_s]/Y_{t_0-}$, $\widehat{g}$, Face-to-Face Score, and Market Leverage. The 95% bootstrap confidence intervals are reported in square brackets. We also present the likelihood ratio test statistics for the restricted models. All regressions are run using industry market capitalization as weights.

(a) Panel A: NLS Results Using the High Face-to-Face Subsample

| | (1) | (2) | (3) |
|---|---|---|---|
| $\lambda$ | 1.086 | | 0.993 |
| | [0.08,1.38] | | [0.11,1.43] |
| $g_0$ | 0.704 | 1.308 | |
| | [-1.09,2.29] | [0.18,1.95] | |
| $n_0$ | 2.678 | 0.319 | 2.260 |
| | [0.51,4.41] | [0,0.58] | [0.47,4.53] |
| Num.Obs. | 201 | 201 | 201 |
| LR.Stat. | | 60.19 | 1.80 |

(b) Panel B: NLS Results Using the High Leverage Subsample

| | (1) | (2) | (3) |
|---|---|---|---|
| $\lambda$ | 1.246 | | 0.600 |
| | [0.01,2.31] | | [-0.02,1.6] |
| $g_0$ | -0.222 | 1.018 | |
| | [-3.1,1.56] | [0.27,1.55] | |
| $n_0$ | 1.167 | 0.194 | 0.672 |
| | [0.23,3.56] | [0,0.41] | [0.2,3.32] |
| Num.Obs. | 201 | 201 | 201 |
| LR.Stat. | | 4.03 | 5.93 |

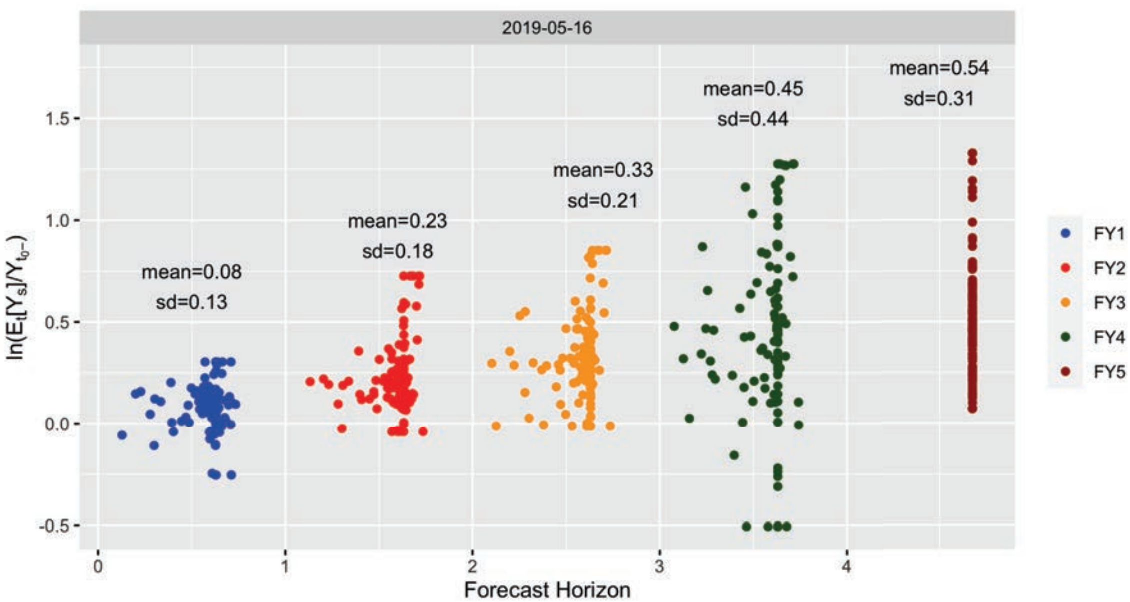
Table 4: Placebo Results Using the I/B/E/S Sample in May 2019

This table presents the coefficients and bootstrap confidence intervals from the placebo non-linear least square regressions of Equation (14) with the constraint that $\lambda \geq 0$. The regressions are run using I/B/E/S summary statistics in May of 2019. The dependent variable is the natural log of $\mathbb{E}_t[Y_s]/Y_{t_0-}$, where $\mathbb{E}_t[Y_s]/Y_{t_0-}$ is the earnings forecasts in May divided by the pseudo non-pandemic earnings. $Y_{t_0-}$, the pseudo non-pandemic earnings, are the FY1 forecasts in January 2019 discounted by the I/B/E/S growth rate forecasts in the same month. The explanatory variables include the horizon of the earnings forecasts $s-t$ and the January I/B/E/S forecasts of growth rate $\widehat{g}$ in 2019. $\lambda$ is the vaccine arrival rate. $g_0$ represents the proportional change in the growth rate. $n_0$ governs the size of the jump in earnings. The first column contains the results using the full sample. The second column ("High Face-to-Face") shows the results using the subsample of industries with Face-to-Face Scores in the top tercile. The last column ("High Leverage") shows the results using the subsample of industries with Market Leverage in the top tercile at the end of 2018. We keep observations with non-missing $\mathbb{E}_t[Y_s]/Y_{t_0-}$ and $\widehat{g}$. The 95% bootstrap confidence intervals are reported in square brackets. All regressions are run using industry market capitalization as weights.

| | Pooled | HiFF | HiLeverage |
|---|---|---|---|
| $\lambda$ | 0.000 | 0.000 | 0.000 |
| | [0,0.78] | [0,3.08] | [0,10.31] |
| $g_0$ | 0.746 | 0.785 | 0.871 |
| | [0.44,0.83] | [-0.21,1.09] | [-1.62,1.1] |
| $n_0$ | -0.064 | -0.091 | -0.015 |
| | [-0.18,-0.02] | [-0.68,-0.01] | [-3.66,0.14] |
| Num.Obs. | 680 | 195 | 227 |

Figure 7: $\ln(\mathbb{E}_t[Y_s]/Y_{t_0-})$ Over Forecast Horizons of the Placebo Sample

This figure plots the natural log of the industry level I/B/E/S earnings forecasts divided by the pseudo non-pandemic earnings, $\ln(\mathbb{E}_t[Y_s]/Y_{t_0-})$, against the horizons of the forecasts $(s-t)$ using I/B/E/S summary statistics in May 2019. The pseudo non-pandemic earnings are the FY1 forecasts in January 2019 discounted by the I/B/E/S growth rate forecasts in the same month. Forecasts horizons are marked with different colors. Forecasts horizons are defined by the distance between the forecast end date and the I/B/E/S statistical period.




most periods, the relationship between FY1 and FY2 revisions should be more synchronized with the growth rate.

But of course, the COVID-19 period data suggest instead that there is a regime switch that might occur between the roughly 1 to 2 year period of forecast horizons. As we said, the alternative is that the growth rates in the pandemic period are just much larger, which is counterfactual. Importantly, this is not an artifact of slow revisions of FY2 since analysts typically revise FY1 and FY2 at the same time and both sets of forecasts experienced significant revisions downward with the arrival of COVID-19.

# 5 Time-Varying Vaccine Arrival Rates

## 5.1 Robustness to vaccine model misspecification

While the earnings process in our model is well understood, the assumption that the vaccine arrives at a constant rate per unit of time (Poisson process), despite its extensive use in the epidemiological literature, needs to be examined further. In reality, the vaccine arrival rate can be time-varying depending on news regarding intermediate stages of development: clinical trials, FDA approval and rollout logistics. In this section, we analyze the implications of model misspecification when it comes to the vaccine arrival process for our inferences.

To this end, we consider a model where an effective vaccine arrival comes after a sequence of informative signals, which we refer to as a multi-signal model. We assume that there are $N$ sequentially ordered signals before the vaccine eventually arrives stochastically at time $\tau^v$. Let $\tau_n$ denote the arrival time of the $n$-th signal, where $n = 1, 2, \cdots, N$. The inter-arrival time between the $n$-th and the $(n+1)$-th signals, $(\tau_{n+1} - \tau_n)$, is independently and identically distributed.[15] Let the distribution for the inter-arrival time $(\tau_{n+1} - \tau_n)$ be exponential with a mean of $1/\lambda^\eta$. That is, at any time $t$ a new signal arrives at a constant rate of $\lambda^\eta$ per unit of time. Let $\lambda^v(t)$ denote the time-$t$ forecast of the vaccine arrival rate. After the arrival of the last signal at $\tau_N$, i.e., $t > \tau_N^\eta$, we assume that a successful and effectively implemented vaccine will arrive at $\tau^v$ at a constant rate of $\lambda^v(t) \equiv \lambda_N^v$ for $\tau_N \leq t < \tau^v$.

What each signal does is to update the time-$t$ perceived vaccine arrival rate $\lambda^v(t)$; i.e., these signals map to news regarding clinical trials, FDA approval and rollout logistics. In between signal arrivals, there is no change of $\lambda^v(t)$. That is, $\lambda^v(t) = \lambda^v(\tau_{n-1})$ for $\tau_{n-1} \leq t < \tau_n$. If news revealed at $\tau_n$ is good, investors increase the perceived vaccine arrival rate from $\lambda^v(\tau_{n-1})$ to $\lambda^v(\tau_n) = u\lambda^v(\tau_{n-1})$ where $u > 1$. If news is bad, investors decrease the perceived vaccine arrival rate to $\lambda^v(\tau_n) = d\lambda^v(\tau_{n-1})$ where $d < 1$. We write the perceived vaccine arrival rate at $t = 0$ as $\lambda_0^v \equiv \lambda^v(0)$. We assume that the likelihood of each signal to be good or bad is independently and identically distributed. Let $\pi^G$ and $\pi^B = 1 - \pi^G$ denote

[15] We start with $t = 0$ and $\tau_1$ is thus the first signal arrive time.

the corresponding likelihood for good ($G$) and bad ($B$) signals.[16]

Suppose that total $n$ signals have arrived by time $t$. Let $\lambda^v(t) = \lambda^v_n$. Then, the expected remaining time that it takes for a vaccine to arrive (including the time it takes to receive the remaining $N-n$ signals) is then given by

$$\begin{aligned}
\mathbb{E}_t(\tau^v - t) &= \mathbb{E}_t(\tau_N - t) + \mathbb{E}_t\left[\mathbb{E}_{\tau_N}(\tau^v - \tau_N)\right] \\
&= \frac{N-n}{\lambda^\eta} + \sum_{j=0}^{N-n} \frac{(N-n)!}{j!(N-n-j)!}\left(\frac{\pi^G}{u}\right)^j \left(\frac{\pi^B}{d}\right)^{N-n-j} \frac{1}{\lambda^v_n} \\
&= \frac{N-n}{\lambda^\eta} + \left(\frac{\pi^G}{u} + \frac{\pi^B}{d}\right)^{N-n} \frac{1}{\lambda^v_n}\,. \qquad (17)
\end{aligned}$$

The first term in (17) is $\mathbb{E}_t(\tau_N - t)$, which is the time-$t$ expected time to receive all remaining $N-n$ signals. The second term is the time-$t$ conditional expected inter-arrival time between the vaccine arrival time $\tau^v$ and the moment that the last signal arrives at $\tau_N$: $\mathbb{E}_t(\tau^v - \tau_N)$.

We use this multi-signal model to generate simulated data that we can then use to check the robustness of our parsimonious model estimates to potential vaccine model misspecification. We assume that analysts understand that the world is described by the multi-signal model while we as econometricians use the simpler baseline model (with constant arrival and no signal). We then check how different our inferences based on the simple model are from parameters of the true model.

We investigate how our estimation procedure performs for different news arrival rates, $\lambda_\eta$. We consider four values of $\lambda_\eta$, $\{4, 12, 24, 48\}$, so that news arrives quarterly, monthly, bi-weekly, and (roughly) weekly, respectively. We set $N = 3$ so that there will be three news signals before the vaccine arrival; i.e., they correspond to the various stages of a vaccine development process that we described above.

We choose our remaining parameter values so that the simulated earnings data correspond to a set of arrival rates that span a large range that includes our mid-May 2020 estimate. We set $\lambda^v_0 = 0.8$, which means that the time-0 estimate of the incremental time (from the arrival of the last signal to the vaccine arrival time), $\mathbb{E}_0(\tau^v - \tau_3)$, is 1.25 years on average. We also

[16] Note that the preceding assumption (similar to those in the recombining binomial tree analysis) implies that there are $N+1$ possible values for the random vaccine arrival rates: $\lambda^v_N = u^m d^{N-m} \lambda^v_0$, where $m$ is the number of good ($G$) signals and the remaining $N-m$ signals are bad ($B$).

set $\pi_G = \pi_B = 0.5$ so that good news and bad news are equally likely and let $u = 1.05$ and $d = 0.95$ so that the vaccine arrival rate will be scaled up by 5% or down by 5% from the previous level when good (bad) news arrives.

For a given value of $\lambda_\eta$ (together with fixed values of $\lambda_0^v = 0.8$, $\pi_G = \pi_B = 0.5$, $u = 1.05$ and $d = 0.95$), we randomly generate $B = 100,000$ paths of our multi-signal model, which includes the stochastic internal-arrive time for each signal $(\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_2)$ and the inter-arrival time $(\tau^v - \tau_3)$ between the vaccine arrival time $(\tau^v)$ and the last signal arrival time $(\tau_3)$. We consider five earnings forecast horizons: $s \in \{0, 1, 2, 3, 4, 5\}$.

We simulate earnings for 140 industries, characterized by a grid of the pre-pandemic growth rates $\widehat{g}^{(j)} \in \{0.0025, 0.005, \cdots, 0.35\}$ per annum based on the data. We set $n = n_0^* = 0.6$, which implies that earnings decrease by 45% upon the pandemic arrival, and $g^{(j)} = g_0^* \, \widehat{g}^{(j)} = 0.8 \, \widehat{g}^{(j)}$ close to our estimate in mid-May 2020. We set $Y_{j,0} = 1$ and then calculate the expected earnings for each path conditional on whether $s > \tau^v$ or $s \leq \tau^v$.

Specifically, for a given path $b$, if the vaccine arrives after the forecast horizon, i.e., $s < \tau^v$, we have $\mathbb{E}_0(Y_{j,s}^b | s < \tau^v) = Y_{j,0} e^{g^{(j)} \times s}$, where $g^{(j)}$ is the expected earnings growth rate for industry $j$ during the pandemic. If the vaccine arrives before the forecast horizon, i.e., $s \geq \tau^v, Y_{j,s}^b$, $\mathbb{E}_0(Y_{j,s}^b | s \geq \tau^v) = Y_{j,0} e^{g^{(j)} \times \tau^v + n + \widehat{g}^{(j)}(s - \tau^v)}$, where $\widehat{g}^{(j)}$ is the expected earnings growth rate for industry $j$ during the normal regime (after the pandemic.) For each industry $j$, we estimate the rational earnings forecast (before any news arrival), $\mathbb{E}_0[Y_{j,s}]$, by using the average of expected earnings at horizon $s$ for this industry across the 100, 000 paths, $\frac{1}{B} \sum_b Y_{j,s}^b$.

Using the preceding procedure, we obtain a set of 840 rational earnings forecasts (140 industries × 6 forecast horizons) under the multi-signal model. Using (8), we then estimate $\frac{\mathbb{E}_0[Y_{j,s}]}{Y_{j,0-}} = \frac{\mathbb{E}_0[Y_{j,s}]}{Y_{j,0}} \cdot e^{-n}$ with $\frac{(\frac{1}{B} \sum_b Y_{j,s}^b)}{Y_{j,0}} \cdot e^{-n}$ for each of the 840 earnings forecasts. Next, we generate a data sample by taking the logarithm of $\frac{\frac{1}{B} \sum_b Y_{j,s}^b}{Y_{j,0}} \cdot e^{-n}$ and then adding an independent random error term, $\epsilon_{0,j}^s$, drawn from a standard normal distribution. That is, the earnings forecast of industry $j$ at horizon $s$ before any news arrival, $\phi_{0,j}^s$, is equal to $\phi_{0,j}^s = \ln \left( \frac{\frac{1}{B} \sum_b Y_{j,s}^b}{Y_{j,0}} \cdot e^{-n} \right) + \epsilon_{0,j}^s$, which corresponds to the sample average of the left-side variable in Equation (14). We then obtain a sample of 840 analyst forecast revisions and

Table 5: Simulation on the Performance of Potentially Misspecified Econometric Model

This table reports the results of our estimation procedure applied to simulated earnings data. The underlying vaccine arrival process follows our multi-signal model. We choose 4 values of $\lambda_\eta$, $\{4, 12, 24, 48\}$, so that news arrives quarterly, monthly, bi-weekly, and weekly respectively. The first column contains the four values of expected news arrival time ($1/\lambda_\eta$). The second column reports the true vaccine arrival time ($1/\lambda^*$) from the multi-signal model. The third column reports the sample mean of our estimates of the vaccine arrival time ($1/\overline{\lambda}$) using Equation (14). The fourth column reports the mean $\overline{g}_0$ of our estimates of $g_0$ using equation (14). The last column reports the mean of our estimates of $n_0$. We set $g_0^* = 0.8$ and $n_0^* = 0.6$ in our simulation.

| $1/\lambda_\eta$ | $1/\lambda^*$ | $1/\overline{\lambda}$ | $\overline{g}_0$ | $\overline{n}_0$ |
|---|---|---|---|---|
| 0.25 | 2.01 | 2.31 | 0.87 | 0.63 |
| 0.08 | 1.51 | 1.43 | 0.79 | 0.60 |
| 0.04 | 1.38 | 1.25 | 0.76 | 0.60 |
| 0.02 | 1.32 | 1.18 | 0.75 | 0.60 |

then repeat this sample generation 10,000 times. We then estimate Equation (14) using this sample.

We report our estimation results in Table 5. The first column lists the four values of the expected inter-arrival time between two signals, $1/\lambda_\eta$. In the second column, we report $\mathbb{E}_0(\tau^v) = 1/\lambda^*$, which is the expected vaccine arrival time (including the time it takes for all signals to arrive) in our multi-signal model:

$$\mathbb{E}_0(\tau^v) = 1/\lambda^* = \left( \frac{N-n}{\lambda^\eta} + \left( \frac{\pi^G}{u} + \frac{\pi^B}{d} \right)^{N-n} \frac{1}{\lambda_n^v} \right). \tag{18}$$

The value of $\mathbb{E}_0(\tau^v)$ decreases from 2.01 (when the news arrival rate is on average one per quarter, $1/\lambda_\eta = 0.25$) to 1.32 (when the news arrival rate is on average (slightly more than) one per week, $1/\lambda_\eta = 1/48$).

In the third column, we present the mean of the estimated expected vaccine arrival time, $1/\overline{\lambda}$, using the simulated data from the multi-signal model, e.g., Equation (14). Notice that our estimate for our parsimonious model is not too far off from the true value of $\mathbb{E}_0(\tau^v)$. In the first row, the true $\mathbb{E}_0(\tau^v)$ is 2.01 while the one estimated using the parsimonious model is 2.31. This is a mild over-estimate. In row 2, the true $\mathbb{E}_0(\tau^v)$ is 1.51 years while the one estimated using the parsimonious model is 1.43 years, which is a mild under-estimate. Across all these four scenarios, our parsimonious model delivers an inference that is not too far off from the actual vaccine arrival rate parameter.

The fourth column shows the mean of the estimated $g_0$ from the parsimonious model. The true underlying parameter is 0.8. We retrieve depending on the arrival rate of news estimates that range from 0.75 to 0.87. The fifth column shows that the mean of the estimated jump in earnings, $\overline{n}_0$, is also barely different from the true underlying parameter value of 0.6. Overall, we conclude that our parsimonious model delivers valuable inference regarding the underlying earnings damage function even with potential misspecification of the vaccine arrival process.

## 5.2 Test of Time-Varying Arrival Rates

Our parsimonious model, in addition to being robust to potential model misspecification of the vaccine arrival process, can also be used to conduct a test of time-varying arrival rates. Using our May 2020 estimates for the jump in earnings and latent growth rates, we re-estimate the vaccine arrival rate in June, July, and August with the latest earnings forecasts. Differences in arrival rates from our baseline May 2020 estimates suggest positive vaccine news.

More specifically, we take our model's predictions for $Y_t$ at time $t$ in the pandemic regime, $Y_t^{pred}$, by using the May 2020 estimates of $n$ and $g$, which we denote by the subscript $may$ (i.e., $n_{may}$ and $g_{may}$). Then, with June, July or August forecasts, we can estimate $\lambda$ with the same expectation formula as in our baseline model:

$$\begin{aligned} \frac{1}{Y_{j,t}^{pred}}\mathbb{E}_t[Y_s] &= \int_t^s \lambda e^{-\lambda(\tau-t)} e^{g_{may}(\tau-t)} e^{n_{may}} e^{\widehat{g}(s-\tau)} d\tau + e^{-\lambda(s-t)} e^{g_{may}(s-t)} \\ &= \frac{\lambda}{\lambda - g_{may} + \widehat{g}} \left[ e^{\widehat{g}(s-t)} - e^{(g_{may}-\lambda)(s-t)} \right] e^{n_{may}} + e^{(g_{may}-\lambda)(s-t)} . \end{aligned}$$

In the pandemic regime at time $t$, conditional on no news arrival, we expect our estimate of $\lambda$ using these other months to be the same as that obtained from the May 2020 forecasts.

We report the results of this estimation in Table 6, where we use the estimated values of $n_0$ and $g_0$ from Column (1) of Table 2 and estimate $\lambda$ using June, July and August forecasts. First, we estimate $\lambda$ to be 0.741 for the June forecasts and 0.815 for the July forecasts,

Table 6: Updated Estimates of the Vaccine Arrival Rate

This table presents the updated estimates of the vaccine arrival rate $\lambda$ using I/B/E/S summary statistics in June, July, and August. The dependent variable is the natural log of $\mathbb{E}_t[Y_s]/Y_t^{pred}$. $Y_t^{pred}$ is the earnings predicted using the estimates in Column (1) of Table 2. The explanatory variables include the horizons of the earnings forecasts $s-t$ and the non-pandemic (January 2020) I/B/E/S forecasts of growth rates $\widehat{g}$. $\lambda$ is the vaccine arrival rate.

| | June | July | August |
|---|---|---|---|
| $\lambda$ | 0.741 | 0.815 | 1.636 |
| | [0.45,1.23] | [0.5,1.35] | [0.94,2.83] |
| Num.Obs. | 676 | 676 | 673 |

Table 7: The Estimates of the Average Time to a Vaccine by the 'Superforecasters'

This table presents the average time to a vaccine implied by the forecasts of the 'Superforecasters' from Good Judgment at the beginning of the months from May 2020 to October 2020.

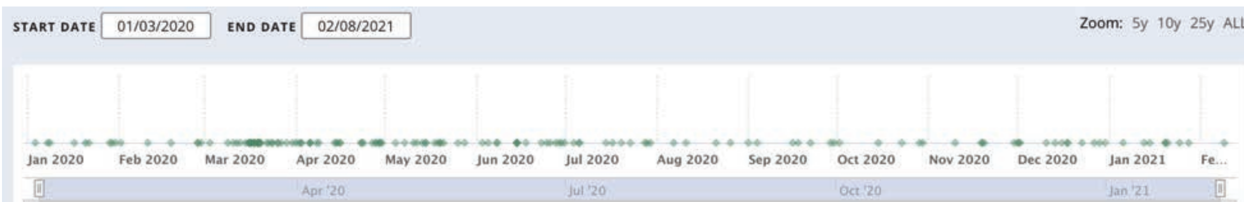
| Date | Implied Months |
|---|---|
| May 1, 2020 | 23 |
| Jun 1, 2020 | 20 |
| Jul 1, 2020 | 15 |
| Aug 1, 2020 | 11 |
| Sep 1, 2020 | 7 |
| Oct 1, 2020 | 7 |

respectively. These estimates are close to the May estimates of 0.674. However, we find that the estimated arrival rate significantly increased when using the August 2020 forecasts: The estimated value of $\lambda$ in August is 1.636 with a 95% confidence interval of $[0.94, 2.83]$. The arrival rate estimate for August of of 1.636 essentially lies outside the confidence intervals for May [0.17, 1.65], June [0.45,1.23], and July [0.5,1.35] (1.636 is just inside of 1.65). Moreover the estimates for May, June and July also lie outside the confidence interval for August.

In summary, we find that the vaccine arrival rate is the same in June and July compared to May but is higher in August. This is evidence of time-varying vaccine arrival rates and hence a time-varying damage function.

Figure 8: Timeline of Events Related to the COVID-19 Pandemic from Federal Reserve

(a) Timeline of Important Events Between January 2020 and February 2021




(b) List of Events During July and August 2020

| Date | Event |
|---|---|
| July 4, 2020 | President Trump signs extension of Paycheck Protection Program (PPP) |
| July 6, 2020 | Main Street Lending Program (MSLP) fully operational |
| July 15, 2020 | Federal Reserve announces rule change to PPP |
| July 17, 2020 | Federal Reserve modifies MSLP to expand access to qualifying nonprofits |
| July 20, 2020 | Congressional Oversight Commission releases third report |
| July 23, 2020 | Federal Reserve expands counterparties in Term Asset-Backed Securities Loan Facility (TALF), Commercial Paper Funding Facility (CPFF) and Secondary Market Corporate Credit Facility (SMCCF) |
| July 28, 2020 | Federal Reserve extends operation of emergency lending programs |
| July 29, 2020 | Federal Reserve announces extension of international US dollar liquidity arrangement |
| July 30, 2020 | Real GDP falls sharply in the second quarter of 2020 |
| August 7, 2020 | Congressional Oversight Commission holds hearing on MSLP |
| August 11, 2020 | Federal Reserve lowers interest rates for Municipal Liquidity Facility |
| August 21, 2020 | Congressional Oversight Commission releases fourth report |
| August 26, 2020 | Agencies issue final rules on modifications to bank leverage ratios |

## 5.3 Vaccine News versus Fiscal or Monetary Policy News

This sequence of estimates alleviates the concern that our vaccine arrival estimate might also be picking up other mitigating factors, particularly expectations regarding fiscal or monetary interventions. To see why, first consider Figure 8. In Panel A of this figure, we report the amount of fiscal and monetary policy news.[17] Notice that most of the news is concentrated in March and April. In Panel B, we report all the news in July and August. If one scans through this list of news, there is no significant fiscal or monetary policy news in these two months. Most of the news have to do with obvious extensions or expansions of existing programs. In other words, there is little important fiscal or monetary news in July and August of 2020.

In contrast, there were two key pieces of vaccine news in late July and early August of 2020. By mid-summer, Moderna and Pfizer established themselves as the leaders in the race to develop a COVID-19 vaccine. Both companies were also the only to take the mRNA vaccine approach, publishing initial Phase I/II clinical trial data on July 14th for Moderna, and on August 12th for Pfizer. Despite small sample sizes, the results demonstrated promising safety measures and antibody production against the spike protein from those who got the vaccine.

The July earnings forecasts were posted by analysts on July 15th and hence information regarding Moderna was unlikely to be incorporated into them. However, the mid-August forecast would have conditioned on the extremely good news from Moderna on July 14th. Even the August 12th Pfizer news might have also been in the forecasts.

Our sequence of estimates also lines up with *Good Judgment*'s survey of experts (known as 'Superforecasters') on when the US would vaccinate 25 million people, show in Table 7. We converted their survey questions into an expected arrival time. Their May 2020 forecast was 23 months, while their August forecast was 11 months. There is a monotonically declining pattern through the time series, similar to our estimates. Overall, our inferred expected arrival times are a bit more optimistic (and ex-post more accurate) than these

[17] This timeline is taken from the Federal Reserve. See https://fraser.stlouisfed.org/timeline/covid-19-pandemic.

survey forecasts but the two time series line up and are highly correlated. If anything, our estimates have been more accurate ex post than these surveys since the US passed 50 million vaccinated by the end of February 2021.

# 6 Conclusion

Despite a large theoretical literature on the inherent nonlinearity of pandemic damage functions, there has been relatively little work in estimating them. To address this challenge, we propose a parsimonious damage function that we take to the data using timely measures of expected damage given by revisions of industry-level earnings forecasts. The structure of our model suggests a natural set of identifying restrictions related to forecast rationality that allow for estimation using nonlinear least squares.

We also extend our framework to account for time-varying arrival rates. Forecasts in mid-May 2020 imply an earnings crash and lower earnings growth until a vaccine arrives in 1.48 years. Mid-August 2020 forecasts imply a much quicker vaccine arrival in 0.61 years, which is due to positive vaccine news as opposed to fiscal or monetary policy news.

Our estimates have implications for a number of policy questions. Moreover, there are several natural inquiries based on our model and estimates. One can consider the stock pricing of vaccine risks, such as in Hong, Wang, and Yang (2020) and Acharya, Johnson, Sundaresan, and Zheng (2020). It would also be valuable to combine these estimates with an asset pricing model to assess the extent to which stock prices, particularly for distressed industries such as airlines or hotels, are efficient. We leave these inquiries for future research.

# References


Acemoglu, D., V. Chernozhukov, I. Werning, and M. D. Whinston, 2020, "A multi-risk SIR model with optimally targeted lockdown," working paper, National Bureau of Economic Research.

Acharya, V. V., T. Johnson, S. Sundaresan, and S. Zheng, 2020, "The value of a cure: An asset pricing perspective," working paper, National Bureau of Economic Research.

Alvarez, F. E., D. Argente, and F. Lippi, 2020, "A simple planning problem for COVID-19 lockdown," Working Paper 26981, National Bureau of Economic Research.

Andersen, A. L., E. T. Hansen, N. Johannesen, and A. Sheridan, 2020, "Pandemic, shutdown and consumer spending: lessons from Scandinavian policy responses to COVID-19," Working paper.

Anderson, R. M., and R. M. May, 1992, *Infectious diseases of humans: dynamics and control.* Oxford university press.

Arnold, R., Y. Galloway, A. McNicholas, and J. O'Hallahan, 2011, "Effectiveness of a vaccination programme for an epidemic of meningococcal B in New Zealand," *Vaccine*, 29(40), 7100–7106.

Bailey, N. T., et al., 1975, *The mathematical theory of infectious diseases and its applications.* Charles Griffin & Company Ltd.

Ball, F., and D. Sirl, 2018, "Evaluation of vaccination strategies for SIR epidemics on random networks incorporating household structure," *Journal of Mathematical Biology*, 76(1-2), 483–530.

Barry, J., M. Campello, J. Graham, and Y. Ma, 2021, "Corporate flexibility in a time of crisis," working paper, National Bureau of Economic Research.

Behncke, H., 2000, "Optimal control of deterministic epidemics," *Optimal control applications and methods*, 21(6), 269–285.

Black, F., and M. Scholes, 1973, "The pricing of options and corporate liabilities," *Journal of political economy*, 81(3), 637–654.

Blinder, A. S., 2009, "How many US jobs might be offshorable?," *World Economics*, 10(2), 41.

Cameron, A. C., and P. K. Trivedi, 2005, *Microeconometrics: methods and applications*. Cambridge University Press.

Coibion, O., and Y. Gorodnichenko, 2012, "What can survey forecasts tell us about information rigidities?," *Journal of Political Economy*, 120(1), 116–159.

Eichenbaum, M. S., S. Rebelo, and M. Trabandt, 2020, "The macroeconomics of epidemics," Working Paper 26882, National Bureau of Economic Research.

Elenev, V., T. Landvoigt, and S. Van Nieuwerburgh, 2020, "Can the COVID bailouts save the economy?," Working Paper 27207, National Bureau of Economic Research.

Farboodi, M., G. Jarosch, and R. Shimer, 2020, "Internal and external effects of social distancing in a pandemic," Working Paper 27059, National Bureau of Economic Research.

Favilukis, J. Y., X. Lin, A. Sharifkhani, and X. Zhao, 2020, "Labor force telework flexibility and asset prices: evidence from the COVID-19 pandemic," Working paper.

Giglio, S., M. Maggiori, J. Stroebel, and S. Utkus, 2020, "Inside the mind of a stock market crash," Working Paper 27272, National Bureau of Economic Research.

Gormsen, N. J., and R. S. J. Koijen, 2020, "Coronavirus: impact on stock prices and growth expectations," Working Paper 27387, National Bureau of Economic Research.

Hong, H., and J. D. Kubik, 2003, "Analyzing the analysts: career concerns and biased earnings forecasts," *The Journal of Finance*, 58(1), 313–351.

Hong, H., N. Wang, and J. Yang, 2020, "Implications of stochastic transmission rates for managing pandemic risks," *Review of Financial Studies*, forthcoming.

Keane, M. P., and D. E. Runkle, 1998, "Are financial analysts' forecasts of corporate profits rational?," *Journal of Political Economy*, 106(4), 768–805.

Kruse, T., and P. Strack, 2020, "Optimal control of an epidemic through social distancing," Working paper.

Landier, A., and D. Thesmar, 2020, "Earnings expectations in the COVID crisis," Working Paper 27160, National Bureau of Economic Research.

Laster, D., P. Bennett, and I. S. Geoum, 1999, "Rational bias in macroeconomic forecasts," *The Quarterly Journal of Economics*, 114(1), 293–318.

Lee, B. Y., B. A. Norman, T.-M. Assi, S.-I. Chen, R. R. Bailey, J. Rajgopal, S. T. Brown, A. E. Wiringa, and D. S. Burke, 2010, "Single versus multi-dose vaccine vials: an economic computational model," *Vaccine*, 28(32), 5292–5300.

Leland, H. E., 1994, "Corporate debt value, bond covenants, and optimal capital structure," *The Journal of Finance*, 49(4), 1213–1252.

Ludvigson, S. C., S. Ma, and S. Ng, 2020, "COVID-19 and the macroeconomic effects of costly disasters," (26987).

Merton, R. C., 1974, "On the pricing of corporate debt: the risk structure of interest rates," *The Journal of Finance*, 29(2), 449–470.

Montenovo, L., X. Jiang, F. L. Rojas, I. M. Schmutte, K. I. Simon, B. A. Weinberg, and C. Wing, 2020, "Determinants of disparities in COVID-19 job losses," Working Paper 27132, National Bureau of Economic Research.

Nordhaus, W. D., 1987, "Forecasting efficiency: concepts and applications," *The Review of Economics and Statistics*, pp. 667–674.

Pagano, M., C. Wagner, and J. Zechner, 2020, "Disaster resilience and asset prices," Working paper.

Wickwire, K., 1977, "Mathematical models for the control of pests and infectious diseases: a survey," *Theoretical Population Biology*, 11(2), 182–238.