Federal Reserve Bank of Chicago

A Model of Economic Activity in San Francisco During the 1918 Influenza Epidemic

François R. Velde

February 15, 2022

WP 2022-04

https://doi.org/10.21033/wp-2022-04

A Model of Economic Activity in San Francisco During the 1918 Influenza Epidemic

François R. Velde*

Federal Reserve Bank of Chicago

February 15, 2022

Abstract

I jointly use daily data on deaths and public transportation ridership in San Francisco in 1918–19 to estimate a model in which agents choose their level of economic activity based on perceived infection risk, modeled as a function of current and lagged infections or deaths. Agents' choices in turn affect the dynamics of the epidemic by reducing contacts in an otherwise standard SEIR model. Non-pharmaceutical interventions restrict agents' activity either as a tax or a bound. I estimate the parameters by maximum likelihood and use the best-fitting model to compute counterfactuals. San Francisco's intervention reduced deaths by a few percent only, and it was away from the Pareto frontier: an earlier and milder intervention would have done better. The behavioral feedback narrows the room for intervention compared to a model with unresponsive agents, and ill-timed interventions can worsen outcomes. Masks also had an effect on transmission rates.

Keywords: 1918 influenza epidemic, San Francisco, public transportation, non-pharmaceutical interventions, SIR macro model, policy evaluation, counterfactuals (JEL H12, I18, I19, N12, R40).

*I thank my colleagues at the Chicago Fed, particularly Gadi Barlevy, François Gourio, and Martì Mestieri; as well as participants in seminars at Manchester University and Università di Roma La Sapienza. The views expressed do not necessarily represent those of the Federal Reserve System or the Federal Reserve bank of Chicago.

Introduction

The Covid-19 pandemic has sparked interest in earlier pandemics, in particular the 1918–19 influenza pandemic (Beach, Clay, and Saavedra 2021). During that event, almost all large US cities (and many small ones as well) used non-pharmaceutical interventions (NPIs) to contain the disease. The nature of the restrictions was broadly similar: closings of places of amusement, public gatherings, schools and churches. The timing and duration of the interventions varied, and in some places restrictions were reimposed during the second wave.

This variation across a relatively homogenous set has provided ways to evaluate NPIs. An older epidemiological literature (Markel et al. 2007; Bootsma and Ferguson 2007) evaluates the impact on the dynamics of epidemics. With the Covid-19 pandemic, economists have used the same data to make arguments about the tradeoffs (or absence thereof) between epidemic containment and foregone economic activity (Correia, Luck, and Verner 2020).

I also want to understand the choices that were and could have been made in 1918. The contribution is twofold: (1) I use mobility data in addition to deaths data, and (2) I add optimizing agents to a SEIR model in order to make inferences from the mobility and death data jointly, taking mobility data as a proxy for economic activity as in Fernández-Villaverde and Jones (2020). Korolev (2020) has pointed out identification problems in the basic SIR model and called for using additional data such as mobility. The structural approach allows me to estimate the parameters of the model and run counterfactuals. This paper centers on the city of San Francisco because I have found data at the daily frequency for both deaths and mobility. The SEIR model is standard (Bootsma and Ferguson 2007) and the economic component taken from Eichenbaum, Rebelo, and Trabandt (2021).

The ridership data and anecdotal evidence show that the use of public transportation began to decline before the imposition of restrictions. If individuals were limiting interactions even in the absence of public policy, how much would the transmission rate of the disease have been reduced anyway? Empirically, the crucial element is the behavioral feedback, or endogenous social distancing. Epidemiologists have long recognized that the rate at which agents infect each other need not be constant as the epidemic explodes, and during the Covid-19 pandemic economists have made the point repeatedly (Atkeson 2020, e.g.,) and devised various techniques to estimate this rate as a time-varying parameter (see Arias et al. 2021, and the literature cited therein). I derive the rate from agents' perceptions of risk, which nest rational expectations or full information (Toxvaerd 2021, as in), but may depend on lags of infections or deaths, rather than depend on current infections only (Jones, Philippon, and Venkateswaran 2021, as in). This behavioral feedback through agents' precautionary reductions in mobility and economic activity, not present in Acemoğlu et al. (2020) and Glover et al. (2020, e.g.,) is important, not only to match the data, but also to understand the counterfactuals and evaluate the policies.

The model does a good job of matching the mobility data and rejects rational expectations. Parameter estimates indicate strong persistence in the feedback: current and lagged

deaths affect current behavior. The counterfactual implies that the reduction in deaths achieved by the interventions in San Francisco was not negligible but relatively modest: a few percent of total deaths, or two hundred lives. Moreover the intervention was well away from the Pareto frontier: authorities could have saved more lives with an earlier, shorter, and milder intervention. The reason is that the behavioral channel moderates the epidemic considerably, compared to an epidemic of similar characteristics but with completely unresponsive agents. There is less overshooting than in the SIR model with constant transmission, but the herd immunity level puts the same upper bound on reduction in deaths: hence, the room for maneuver is limited. Interestingly, the Pareto frontier has a sharp kink: most reduction in deaths could be achieved with very mild interventions, and further gains were very limited and required much longer interventions.

Finally, the model allows me to estimate the impact of masks separately from other restrictions, because during San Francisco's second wave masks were required again but there were no closings. I find that masks in the second wave may have been as effective on their own as masks and closings during the first wave.

1 The epidemic in San Francisco

The purpose of this section is to provide some context for the data and model.[1] The focus is on the development of the epidemic as seen by contemporaries, and the reaction of the city authorities. As in many other places in the US, measures were taken at the city level, either by the Board of Health of the city and county, or by the Board of Supervisors. Dr. William C. Hassler, as head of the Health Department or Health Officer from 1915 to his death in 1931, was the most prominent official and led the city's response.

The influenza pandemic of 1918 (aside from a little-noticed herald wave in the spring of 1918) arrived in the US in late August 1918 and hit the East Coast first. It spread quickly to other parts of the country, and by mid-September authorities in California were well aware of the danger, but the timing of its arrival on the West Coast was uncertain and they initially thought they had time. As of September 20, George Ebright, the president of the State Board of Health, thought that there was no evidence of influenza in California, but that it would arrive within two to three weeks, following the lines of railroad travel. In the meantime health officials focused on educating the public and making recommendations. Hassler described the state of knowledge about the disease, its propagation from person to person through nose and mouth, the rapid progress and frequently fatal pneumonia that ensued, and the susceptibility of "the most robust members of the community." He recommended avoiding public places and contacts with infected persons. There was no talk of restrictions or closings yet (the first such measures were taken on September 20 in Milford, MA).

When the first case was reported in the city on September 24, Hassler was optimistic that the epidemic would be contained, but within a few days he admitted that the disease

[1] The main source of this section is contemporary news articles from the *San Francisco Examiner*, the *San Francisco Chronicle*, and the *Recorder*; see also Crosby (2003, 91–116).

was well distributed and a serious epidemic was to be feared, and on October 4 he conceded that "the spread of the disease here now will be very rapid [. . .] an exceedingly large roll of cases can be expected." From the end of September large cities on the East coast resorted to closings of places of amusement and schools, and the US Surgeon General Rupert Blue stated on October 4 that "the only way to stop the spread of Spanish influenza is to close churches, schools, theatres and public institutions in every community where the epidemic has developed."

Yet in San Francisco authorities remained ambivalent. While nearby communities adopted such measures (Redwood City on October 10, San Jose the next day), the city's Board of Health decided on October 10 that there was no need to close schools or theaters, and Hassler opined that reports from eastern cities suggested that closing theaters did little to check the epidemic. As late as October 19, the State Board of Health stated:

> The closing of the public schools is a measure that the State Board of Health does not favor, provided that the pupils are inspected daily by teacher or nurse, those who show signs of illness being immediately sent home. The State Board of Health has not issued any general order nor made any drastic moves in prohibiting public meetings or in closing places of amusement. The Board believes that it is much better to supply citizens with full information as to the ways in which the disease may be contracted.

These ways included acts of personal hygiene, walking to work and avoiding crowded places, self-isolation in case of symptoms.

Case numbers continued to climb: by October 13 hospitals were overwhelmed, the San Francisco hospital was devoted wholly to caring for influenza cases and hospitals were "requested not to take in any cases which do not absolutely demand hospital care." Notably, attendance at places of amusement was already declining: "all the theatres are feeling the effects of the campaign against Spanish influenza, and even on O'Farrell Street, though the audiences are still large, there has been a falling off in patronage" (*Examiner*, Oct 14, p. 11). Finally, the city Board of Health held a special meeting on the evening of October 17, attended by the mayor, various officials of the US military, US Public Health Service, the Red Cross, and representatives of the entertainment industry, and voted to close "all places of amusement, including theaters, moving-picture theaters, concert halls, dance halls and dances in all cabarets, cafes and hotels, and all form of entertainment in any or all of them". The order also applied to lodge and fraternal meetings, public amusement places (penny arcades, merry-go-rounds), private dances, halls, social gatherings, Sunday school classes, church services, public and private schools and kindergartens. Any public meeting required a permit. Representatives of entertainment houses were strongly in favor because they were already suffering financially and hoped that the disease would be eradicated quickly (*Examiner*, Oct 18).

Hassler strongly believed in the effectiveness of masks and urged their use on the public street. Dr. Woods Hutchinson, a health writer and strong promoter of masks, argued that "if one half or even one third of the population would wear masks the number of

attacks and deaths due to Spanish influenza would be cut from one-half to two-thirds" (*Recorder*, Oct 23, p. 8) and Hassler said that they gave "90% immunity" (*Examiner*, Oct. 22, p. 13). At Hassler's urging the Health Board made masks compulsory for clerks, druggists, tellers, hotel employees, barbers on October 19. Then, on October 24, the Board of Supervisors passed a penal ordinance making the wearing of masks mandatory in all places except the home, under penalty of a fine between $5 and $100, up to ten days imprisonment, or both (*Municipal Record*, 1918, 352). Newspapers reported enforcement actions against "mask slackers" (250 arrested on November 4), but Hassler claimed that the ordinance was widely abided. The State Board of Health, which ultimately concluded that masks had been useless, agreed that "contrary to expectation, the masks were worn cheerfully and universally" (Kellogg 1919a, 39).[2]

Reported cases peaked in late October. Hassler attributed this to masks but expected that the city would remain masked for two months to avoid a second wave (*Chronicle*, Nov 6, p. 9) and quashed early talk of reopening theaters. Although "fully aware of the losses incurred under the present rigid regime," he wanted closings to continue until December: "I can almost deplore the necessity for printing daily the number of new cases, for I fear it will tend to lull the public into a false sense of security, as was the case in Los Angeles, where , with all the amusement houses permitted to remain open, the disease passed beyond the control of the authorities" (*Chronicle*, Nov 9, p. 9). But after four weeks of closure, "been the maximum time for the closing of amusement places in other cities," the entertainment sector increased pressure on the Board of Health and on the mayor. The closing order was lifted: theaters were allowed to reopen on the afternoon of November 16,[3] churches on November 17, and schools on November 18 (although the Board of Education deferred the reopening for a week). After waiting for a few days to see the impact of reopening, Hassler proclaimed that "Spanish influenza has been eradicated from San Francisco" and announced that the mask mandate would be lifted on November 21 (*Examiner*, Nov 18, p. 3).

The epidemic dropped from headlines and newspapers ceased to report cases and deaths for a while, but in early December new cases were reported again. At first Hassler denied any indication of a flare-up (*Chronicle*, Dec 4, p. 8) but the next day he required store clerks to wear masks again and warned that a reinstatement of the ban would be necessary if cases continued to rise. Within a few days he urged the mayor to consider reinstating the mask mandate. This time he met with serious opposition: at a conference on December 9 with Hassler and the Board of Supervisors, the entertainment and retail sectors disputed the case numbers and pushed back against masks, despite Hassler's prediction that the second wave would be worse than the first "as has been the case in many other cities." After being widely feted, Hassler became the object of strong feelings, to the point that a bomb was mailed to him (*Examiner*, December 18, p. 1). The Supervisors

[2]Another measure that Hassler thought important, despite widespread skepticism, was the administration of a vaccine developed by Dr, Timothy J. Leary of Tufts Medical College, sent from Boston with great fanfare and administered for free, but with disappointing take-up (Crosby 2003, 100).

[3]The Mission and North Beach areas remained closed a little longer, and large dancing remained banned until November 30.

postponed a decision on masks for several days, took it up again on December 17 for 4 and a half hours with no conclusion, and finally rejected the mask mandate by a 9–7 vote on December 19. Cases continued to grow, albeit more slowly than in October. Hassler judged the situation "bad" on December 25 and "alarming" on Dec 28, but was still reluctant to go back to the Supervisors. The Board of Health stated on January 2 that it had done all it could and it was up to the legislative body of the city to act. At its urging the Board of Education required masks in public schools on January 4. Finally, enough supervisors were willing to change their minds: the Health committee approved a mask ordinance just as cases passed 600 (compared with a peak of over 2000 on October 25) and on January 11 the ordinance was announced, although it would not take effect until January 17. Immediately cases began a precipitous drop, which Hassler attributed to the masks. Within a week cases had fallen by a factor of ten, hospitals were relieved, and Hassler declared the situation "bright beyond our highest expectations." Late January the pressure to repeal the ordinance increased, as compliance fell, and at a special meeting held on a Sunday morning, February 1, the Board of Health relented and the mayor promptly nullified the ordinance at 11am.

Hassler's views apparently shifted from the first to the second wave. He was initially optimistic that the city would be spared, and only slowly came to recommend restrictive measures. He proclaimed the disease eradicated after the first wave, and dismissed the first signs of the second wave, but then strongly pushed for another mask mandate. It's unclear what model or data he had in mind, although he explicitly referred to the experience of other cities during the 1918 pandemic, and also the pandemic of 1889–91 which was the main reference point at the time (California State Board of Health 1918).

In summary, San Francisco authorities saw the influenza epidemic arrive by September 24, and within a week or two had the examples of other cities' closings, but no NPI was undertaken until October 18, when schools, churches, and places of amusement were closed. On October 24 a mask mandate was also put in place and enforced, and seems to have been widely obeyed. The closing of places of amusement lasted until November 16 (this is the end date I will use in the estimation) while schools remained closed until November 24 and the mask mandate was lifted on November 21. Health officials would have wanted to extend the closings but strong pressure led the city authorities to lift the ban. When the second wave arrived in early December Hassler urged a reinstatement of the mask mandate but this time met with opposition and did not succeed until January 17, at which point the wave dropped off quickly and the mandate was lifted on February 1. There were no closings during the second wave.

2 The Data

This section describes the two data series on mortality and streetcar ridership.

Mortality data

As mentioned in the previous section cases and deaths were reported regularly at news conferences by Dr. Hassler and relayed in the local newspapers. Although they tell us what information was available to contemporaries (see section 3 below), these series are not good measures of actual cases and deaths, as was understood even then (Kellogg 1919a).

I have collected all deaths in San Francisco on a daily basis from September 1, 1918 to April 15, 1919 using `ancestry.com`.[4] The data includes first and last names, year of birth, and day of death. They appear to have been manually transcribed from indices of death registration records and are not free of error, but aggregated at the monthly frequency the totals differ by no more than 3% from the official monthly totals for San Francisco (Bureau of the Census 1913–23). I also collected daily deaths for the same months in the years 1912 to 1916 in order to establish a baseline of daily expected deaths (adjusting for population growth, Department of Commerce, Bureau of the Census 1923). This gives me excess deaths on a daily basis, which I assume to be entirely due to the epidemic.

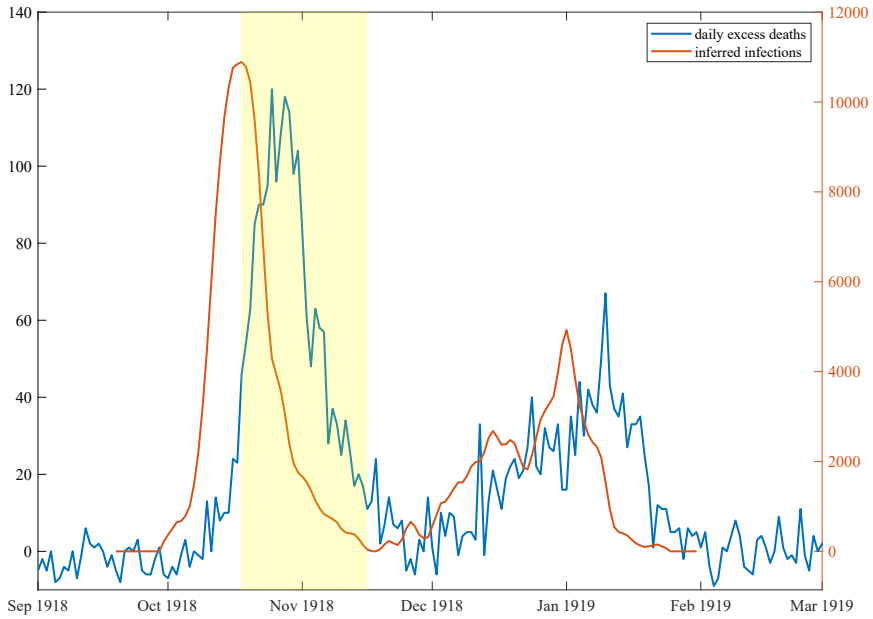

Figure 1: Excess deaths in San Francisco, daily (left axis) and estimated infections based on a deconvolution of deaths (right axis). The yellow area indicates the period when places of amusement were closed. Sources: `ancestry.com`.

Using the delay from exposure to death[5] it is possible to deconvolute the daily death series and construct an estimated infection series (Goldstein et al. 2009). The result is shown in Figure 1 and will be used later.

[4]Ideally one would want to collect all deaths ascribed to pneumonia and influenza in death certificates, but they are not available digitally for California in that time period.

[5]The delay from exposure to symptoms is taken from Ferguson et al. (2005) and the delay from symptoms to death is from Keeton and Cushman (1918, chart 2), based on 603 deaths observed in Cook County Hospital.

Ridership data

The ridership data pertains to the San Francisco Municipal Railways (the "Muni"), one of three street railway systems in San Francisco (O'Shaughnessy 1921). The largest system was the United Railroads, which operated 257 miles; the California Street Cable Railroad Company operated 10.5 miles of cable car. The Muni, city-owned as its name indicates, began operating a line on Geary Street in 1912 and by June 1918 it operated 57.4 miles.[6] Figure 2 shows the network in 1920, which covered the central business district (Market Street) and connected to the main residential areas of the city. In 1920 the Muni accounted for around 22% of the total ridership, which was on average 511,060 per day for a city population of 511,300 (O'Shaughnessy 1921, 13, 122). The Muni ridership data is thus likely to be a good measure of the use of public transportation.

I collected weekly reports of daily receipts of the Muni from 1917 to 1922 in local newspapers.[7] Newspapers reported gross revenues, but passenger revenues represent 99.1% of gross revenues. Revenues are also an excellent measurement of (paying) passengers: the fare remained unchanged at 5 cents throughout the period, the average fare per revenue passenger was 4.92 cents (Wilcox 1921, 185), and the correlation of monthly passengers to monthly gross revenues is 0.998 for 1917–19 (San Francisco Municipal Railway 1915–22).

Figure 3 shows the data (smoothed using a 7-day centered moving average) and two methods of deriving an index of ridership. The first method is to regress the data over the full sample (June 1918 to December 1922) on a linear time trend, week-of-year, and day-of-week dummies.[8] The second method is to take the ratio of a day's revenues to the revenues 364 days later, and then scale the resulting series by the average annual growth rate estimated in the previous regression. The two methods show little difference; I use the first in the estimation.

Is ridership a good proxy for economic activity?

One goal of this paper is to use the ridership data to validate the epidemiological model, and in particular to assess the nature and importance of any behavioral feedback. To do this I will use in the next section a simple model in which agents decrease their economic activity in response to the epidemic as well as to restrictions imposed by authorities. But is ridership a good proxy?

The impact of the epidemic on electric railway ridership was widely noted. The journal of the American Electric Railway Association noted that nationwide the epidemic "caused an immense falling off in the number of passengers carried" (AERA, 1918, 7(6):575). In San Francisco, the *Examiner* reported (28 Nov 1918, p. 6):

[6] Mileage increased by 2.4mi on Feb 21 1919 and 3.0mi on Apr 12, 1919.

[7] The newspaper are the *San Francisco Examiner*, *San Francisco Chronicle*, and *Daily Recorder*. It's not clear why newspapers published these reports, but they begin in August 1917, during a strike that affected the private competitors from August 12 to November 23, 1917 (*Commercial and Financial Chronicle*, 106(2744):119). They become regular in October 1917 but data is occasionally missing (from Jan 6 to Feb 9 and Feb 17 to Mar 2, 1918; weeks ending Apr 13, Jul 13, 1918, May 3, 1919, Dec 16, 1920; from Dec 24 to Jan 15, 1921; weeks ending Jul 9, Aug 6, 1921; Jan 7, Aug 5 1922).

[8] I only use dummies for Saturday and Sunday because they are the only ones significant. I interact these day-of-week dummies with calendar year dummies.

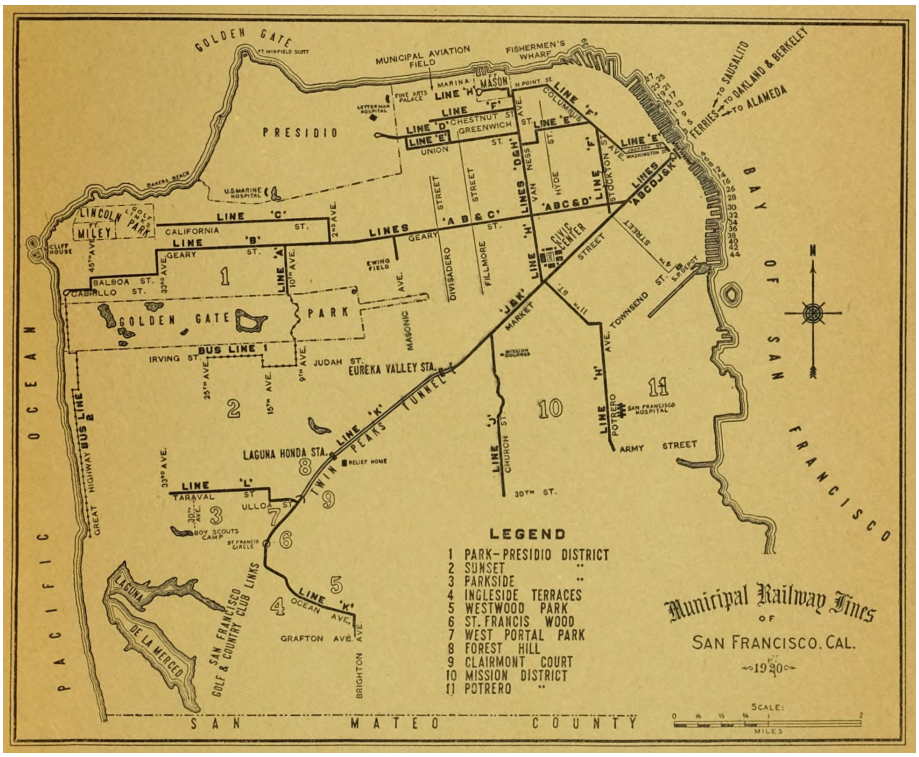

Figure 2: Map of the San Francisco Municipal network in 1920 (San Francisco Municipal Railway 1915–22, 1920, 15).

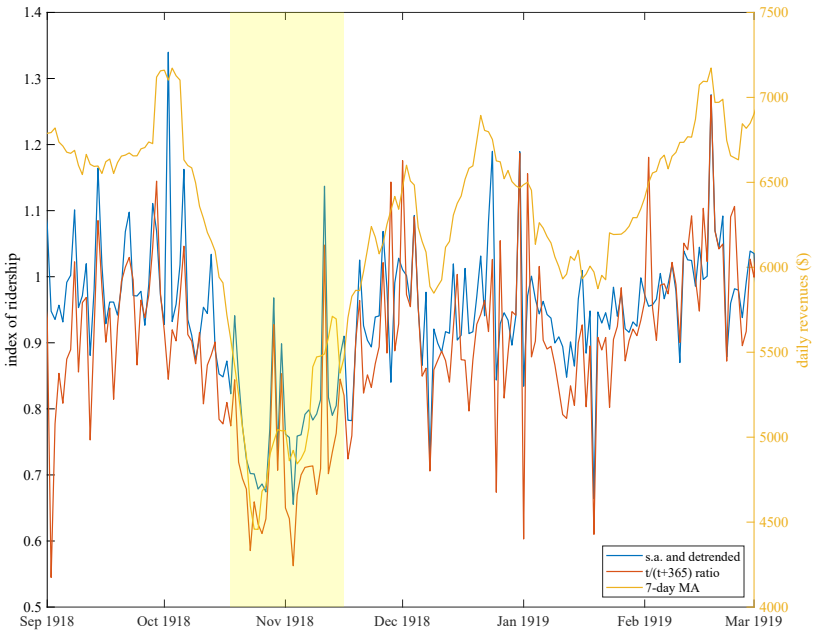

Figure 3: Daily revenues on the San Francisco Municipal Railways, 7-day moving average (right axis), and two indices (left axis): one detrends and seasonally adjusts, the other takes the ratio of t to $t + 365$ and adjusts for 1-year average growth. The yellow area indicates the period when places of amusement were closed. Sources: *San Francisco Examiner*, *San Francisco Chronicle*, *Daily Recorder* (various issues).

> During the epidemic all street railway earnings were cut down enormously. The United Railroads suffered in proportion to the Municipal Railway lines. The deficit has amounted to more than $1,000 a day on the Municipal lines ever since the ban has been raised. Traffic has been abnormally light since the influenza epidemic broke out, whereas it should be the heaviest during the rainy season.

Fewer riders does not translate one to one into less economic activity. Not all workers and shoppers used public transportation, and those that usually did could choose to walk during the epidemic, albeit at a cost in terms of time. But workers who usually rode the Muni and stayed home would reduce measured ridership.

The maximum decrease in ridership in Figure 3 is about 30%. The only quantitative evidence I have found for the effect of the epidemic on activity is a statement from the entertainment industry on November 15 that $400,000 was "weekly diverted from the normal channels." This translates into about $1 per person or $4 per household per week, compared with GDP per capita per week of $12.50 or per household of $60, this means a reduction of 8%. Using median hourly wages of $0.50 for men, $0.30 for women, with only half of women in the labor force compared to men, and a 50-hour workweek would give a higher number of 12%. It seems likely, then, that the 30% figure is an over-estimate.

Beyond the historian's defense that this is all the data that I have, mobility data arguably provides an upper bound on the reduction in economic activity. Moreover, if one assumes a proportional (though less than one-for-one) relation between the two, mobility data allows for ordinal comparisons in outcomes: a counterfactual with more or less ridership corresponds to a counterfactual with more or less economic activity. This will be important when thinking about the Pareto frontier in section 5.

3 Model

The model, adapted from Eichenbaum, Rebelo, and Trabandt (2021), has an economic and an epidemiological side.

The epidemiological component

The epidemiological side is based on the classic compartmental model (Bootsma and Ferguson 2007). The population is partitioned into S_t susceptibles, E exposed, I infected, and R removed:

$$S_t + E_t + I_t + R_t = 1.$$

The categories of interest evolve according to the following laws of motion:

$$\dot{S}_t = -\lambda_t I_t S_t \tag{1}$$

$$\dot{E}_t = \lambda_t S_t I_t - \alpha E_t \tag{2}$$

$$\dot{I}_t = \alpha E_t - \nu I_t \tag{3}$$

$$\dot{R}_t = \nu I_t \tag{4}$$

In addition, deaths evolve according to

$$D_t = \mu \int_0^\infty f(s)\lambda(t-s)S(t-s)I(t-s)ds$$

where the delay function $f(s)$ is the distribution of time from exposure to death.

The literature has long recognized that the force of infection λ_t need not be constant over time, and can depend on current and past values of the epidemic's variables. Several ad-hoc forms have been used.

Let $f(t)$ be some function of the epidemic's history $\{S_s, E_s, I_s, D_s\}_{s\leq t}$ satisfying $f(0) = 0$. One common specification is that λ_t is a Hill function of $f(t)$:

$$\frac{\lambda_t}{\lambda_0} = \frac{1}{1+f(t)}. \tag{5}$$

Yu et al. (2017) compare the Hill function with alternatives such as $e^{-\kappa f(t)}$ and $(1-f(t))^\kappa$. Bootsma and Ferguson (2007) use the Hill function and assume that $f(t)$ a function of observed current and lagged deaths over a window of length T_m: $f(t) = \int_0^{T_m} D(t-s)ds$ while Yu et al. (2017) use exponential weighting of lagged deaths: $\dot{f}(t) = D_t - \tau f(t)$.

The economic component of the model will provide a foundation for the functional form, and link λ_t to measurements of economic activity.

The economic component

The agent's preferences over consumption c and hours worked n

$$\sum_t \beta^t \left(\frac{c^{1-\sigma}}{1-\sigma} - v\frac{n^{1+\epsilon}}{1+\epsilon} \right)$$

are maximized subject to the budget constraint $w_t n_t = c_t$. The parameters σ and ϵ are the inverses of the intertemporal elasticity of substitution and of the Frisch elasticity. A linear technology converts hours worked into consumption good at a constant rate, so that in equilibrium the wage rate is constant.

The force of infection λ_t becomes endogenous in the following way. Eichenbaum, Rebelo, and Trabandt (2021) assume that incidence (new infections per period) $\lambda_t S_t I_t$ arises from interactions between the susceptible and the infected, and these interactions are chosen by agents taking account the contagion risk. I also want to allow for economic activity to be affected directly through the labor supply, so I assume that a fraction ϕ of the infected are asymptomatic, and make the same economic choices as the susceptible, while the fraction $1-\phi$ is out of the labor force.

Specifically, new infections are modeled as

$$-(S_{t+1} - S_t) = \lambda_t S_t I_t = \pi_1 (S_t c_t^s)(\phi I_t c_t^i) + \pi_2 (S_t n_t^s)(\phi I_t n_t^i) + \pi_3 S_t I_t. \tag{6}$$

The first two terms on the right-hand side measure infections of susceptibles during the purchase of consumption and the production of goods respectively, while the third term measures infections that are not affected by economic choices and would occur even if the economy shut down completely.

A susceptible individual understands that her probability of infection depends on her consumption and labor decisions. The true probability derived from (6)

$$[\phi(\pi_1 w^2 + \pi_2) n_t^i n + \pi_3] I_t \tag{7}$$

is an affine function of her choice n, given the prevalence I_t and the choices n_t^i and $c_t^i = w n_t^i$ of the infected. Rational expectations would require that the agent use (7). I will be more flexible and only require that the agent models sees the mapping from n to the probability of not being infected as an affine function $a_t - b_t n$ in which a_t and b_t depend on the current and lagged epidemic variables. This assumption nests rational expectations as we shall see in the next section.

To simplify I will assume that agents maximize

$$\begin{aligned} \max_{n_t, c_t} \quad & (a_t - b_t n_t) W + \frac{c_t^{1-\sigma}}{1-\sigma} - v \frac{n_t^{1+\epsilon}}{1+\epsilon} \\ \text{s.t.} \quad & c_t = w n_t \end{aligned} \tag{8}$$

where W is the continuation value, or value of life and w is the (exogenous) wage. The first-order condition

$$w^{1-\sigma} = b_t W n_t^{\sigma} + v n_t^{\epsilon+\sigma} \equiv F(n_t, b_t) \tag{9}$$

maps b_t, the marginal riskiness of working, into the choice of work n_t. The function $F(\cdot, b)$ is increasing from $F(0, b) = 0$ and has a unique solution $n(b)$, interior for low enough w. Denote $n^* = v^{-\frac{1}{\epsilon+\sigma}} w^{\frac{1-\sigma}{\epsilon+\sigma}}$ the solution to $F(n^*, 0) = 0$, and denote $c^* = w n^*$ the corresponding choice of consumption. The allocation (c^*, n^*) represents the economy before the epidemic, at $t = 0$.

Eichenbaum, Rebelo, and Trabandt (2021) assume $\sigma = \epsilon = 1$, in which case $n^* = v^{-\frac{1}{2}}$ and the agent's choice of hours worked is given by

$$\frac{n}{n^*} = \sqrt{1 + \alpha_t^2} - \alpha_t \tag{10a}$$

$$\alpha_t = \frac{1}{2} W n^* b_t. \tag{10b}$$

Assuming that the infected not out of the labor force behave like the susceptibles, the factor λ_t in (6) and (1) becomes

$$\begin{aligned} \lambda_t &= \pi_1 c_t^s c_t^i \phi + \pi_2 n_t^s n_t^i \phi + \pi_3 \\ &= (\pi_1 w^2 + \pi_2) \phi n^{*2} \left(\frac{n}{n^*}\right)^2 + \pi_3 \\ &= (\lambda_0 - \pi_3) \left(\frac{n}{n^*}\right)^2 + \pi_3 \end{aligned}$$

with $\lambda_0 = (\pi_1 w^2 + \pi_2) \phi n^{*2} + \pi_3$. If we further assume $\pi_3 = 0$ this reduces to

$$\frac{\lambda_t}{\lambda_0} = \left(\frac{n_t}{n^*}\right)^2.$$

The force of infection is the square of agents's choices of economic activity.

Aggregate economic activity y_t is the sum of individual activities. The recovered choose n^*, while the susceptible, exposed, and asymptomatic infected choose n_t. The

joint behavior of the epidemic and economic activity (relative to the pre-epidemic level $y^* = n^*$) is described by

$$\frac{y_t}{y^*} = R_t + (S_t + E_t + \phi I_t)\left(\frac{n_t}{n^*}\right) \tag{11a}$$

$$\frac{\lambda_t}{\lambda_0} = \left(\frac{n_t}{n^*}\right)^2 \tag{11b}$$

$$\frac{n_t}{n^*} = \sqrt{1+{\alpha_t}^2} - \alpha_t \tag{11c}$$

$$\alpha_t = \frac{1}{2} W n^* b_t. \tag{11d}$$

At this point I haven't yet specified the form taken by b_t, which is the marginal reduction in infection risk from reducing economic activity.

Rational expectations equilibrium

Under rational expectations the probability perceived by the agent must equal the true probability (7) when $n_t^i = n_t$ satisfies (10a):

$$b_t = \phi(\pi_1 w^2 + \pi_2) n_t^i I_t = n^* \phi(\pi_1 w^2 + \pi_2)(\sqrt{1+{\alpha_t}^2} - \alpha_t) I_t.$$

Substituting $b_t = \frac{2}{W n^*}\alpha_t$ from (10b) leads to

$$\alpha_t = \frac{1}{2}\frac{\kappa I_t}{\sqrt{1+\kappa I_t}} \tag{12a}$$

$$\frac{n}{n^*} = \frac{1}{\sqrt{1+\kappa I_t}} \tag{12b}$$

$$\frac{\lambda_t}{\lambda_0} = \frac{1}{1+\kappa I_t}. \tag{12c}$$

with $\kappa = W\lambda_0$.

Note that (12c coincides with the Hill function assumed by epidemiologists, with $f(t) = I_t$ and the κ parameter has a structural interpretation as the value of life.

Under rational expectations the dynamics of an epidemic with and without behavioral feedback are easy to compare in a phase diagram, shown in Figure 4.[9] The process starts at $S = 1$, when the whole population is susceptible. Since S cannot increase the path can be read from right to left as a time series.

When there is no behavioral channel the vertical line at $S^* = \frac{\nu}{\lambda_0} = 1/R_0$ (herd immunity) separates two regions with I increasing to the right and decreasing to the left. The epidemic starts at $S = 0$, peaks at $S = S^*$, and ends at S_∞. The best outcome would be a path that ends at $(S^*, 0)$, the herd immunity level. At that point, by (2) and (3) $\dot{E} + \dot{I} < 0$ even if $\lambda_t = \lambda_0$, in the absence of any behavioral reaction. Without intervention the epidemic overshoots herd immunity.

With the behavioral channel the red dotted line separates the two regions, hence the epidemic peaks earlier (for a higher S), but there is only one wave. Hence behavioral feedback, of itself, is not sufficient to generate multiple waves; nor does it change the herd immunity level or the best outcome.

[9] For simplicity I consider here a simpler SIR model in which the compartment E is eliminated, and (2–3) are replaced by $\dot{I}_t = \lambda S_t I_t - \nu I_t$.

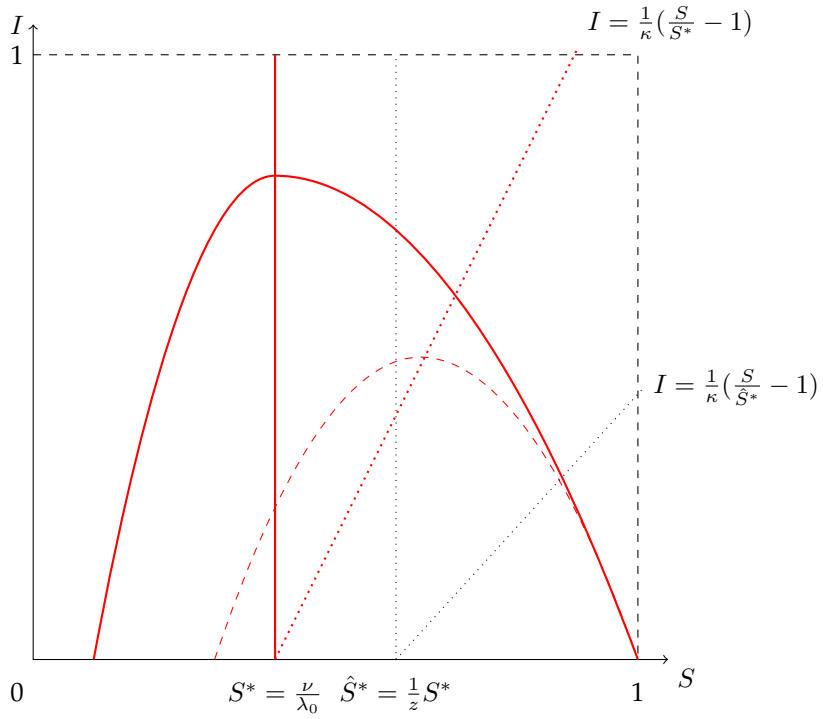

Figure 4: Phase diagram with and without behavioral channel.

One way to generate waves is to introduce exogenous time variation in λ, that is, lockdowns. The dashed black lines outline the phase diagram when transmissibility has been reduced by a factor z. If z is large enough ($z > 1/S^*$) the vertical line is shifted to the right of the current position (S_t, I_t) and infections begin decreasing immediately upon imposition, no matter what the current value of I is. But if restrictions are removed before S_t has fallen below S^*, infections rise again and produce a second wave (multiple lockdowns can generate multiple waves).

In terms of optimal policy, a well-timed lockdown can bring path of the epidemic to end at $(S, I) = (S^*, 0)$ at which point the epidemic cannot restart. One such policy is to let the epidemic progress and, as soon as S reaches S^*, impose an absolute lockdown ($z = 0$) that halts new infections: $\dot{I}$ immediately, S remains constant at S^* and I decreases to 0. This eliminates the epidemic's overshoot.

Ad hoc behavioral models

The assumption of rational expectations raises two questions: What information did agents have in real time? Did behavior seem to respond to current infections only?

On the first question, contemporary newspapers suggest that there was a strong demand for information about the epidemic. The city's Health Commissioner, William C. Hassler, held briefings at least once a day on the state of the epidemic and relayed numbers of cases and deaths. But the numbers were problematic. Determining whether a death was due to influenza or opportunistic pneumonia was not obvious. Influenza had not been made a reportable disease until the beginning of the epidemic (September 27),

so physicians were not in the habit of reporting it. In the midst of an epidemic, many were too busy to provide information in a timely manner, and sometimes filed their reports very late and by mail. It was often unclear if cases were dated by the date on which the report was sent or on which it was recorded by the city, and Hassler himself was increasingly unclear on this point.

Such as they are, the numbers of cases and deaths were reported in newspapers and may have been used by people as a measure of the epidemic. I have collected these reports; in addition, Chowell, Nishiura, and Bettencourt (2006) provide a shorter but apparently official series of cases.[10]

The top panel of Figure 5 shows that reported cases do not line up well with the infections that can be inferred from excess deaths and the known delay between infection and deaths. In addition, the difference in scale suggests that only about 20% of actual cases were reported. The bottom panel shows that death reports were delayed by a few days but otherwise close to the truth, and that reported cases match the time pattern (but not the scale) of deaths.

Thus, although rational expectations would impose that b_t is a function of current infections I_t only, I allow b_t to depend on deaths, possibly with lags or cumulation.

The model I estimate consists of (11a–11b), rewritten as

$$\frac{y_t}{y^*} = R_t + (S_t + E_t + \phi I_t)\sqrt{K(t)} \tag{13a}$$

$$\frac{\lambda_t}{\lambda_0} = K(t) \tag{13b}$$

and several variants, namely:

$$K(t) = \begin{cases} \frac{1}{1+\kappa f(t)} & \texttt{(Hill)} \\ (1-f(t))^{\kappa} & \texttt{(Power)} \\ e^{-\kappa f(t)} & \texttt{(Exp)} \\ 1 + 2{\alpha_t}^2 - 2\alpha_t\sqrt{1+{\alpha_t}^2},\ \alpha_t = \frac{1}{2}\kappa f(t) & \texttt{(model)} \end{cases}$$

and

$$f(t) = \int_0^{T_m} X_{t-s} ds \qquad \text{(default)}$$

$$\dot{f}(t) = X_t - \tau f(t) \qquad \texttt{(Alt)}$$

and $X_t = M_t$ (default) or $X_t = I_t$ (`Inf`).[11]

The different functional forms $K(t)$ as functions of $f(t)$ coincide to the first order:

$$\begin{aligned} \texttt{Hill} &: 1 - \kappa x + \kappa^2 x^2 - \kappa^3 x^3 + O(x^4) \\ \texttt{Power} &: 1 - \kappa x + \frac{1}{2}\kappa(\kappa-1)x^2 - \frac{1}{6}\kappa(\kappa-1)(\kappa-2)x^3 + O(x^4) \\ \texttt{Exp} &: 1 - \kappa x + \frac{1}{2}\kappa^2 x^2 - \frac{1}{6}\kappa^3 x^3 + O(x^4) \\ \texttt{model} &: 1 - \kappa x + \frac{1}{2}\kappa^2 x^2 - \frac{1}{8}\kappa^3 x^3 + O(x^5) \end{aligned}$$

[10] They cite a Japanese publication from the 1920s, whose ultimate source is unclear.

[11] Deaths are normalized per 10^5 and infections per 10^8.

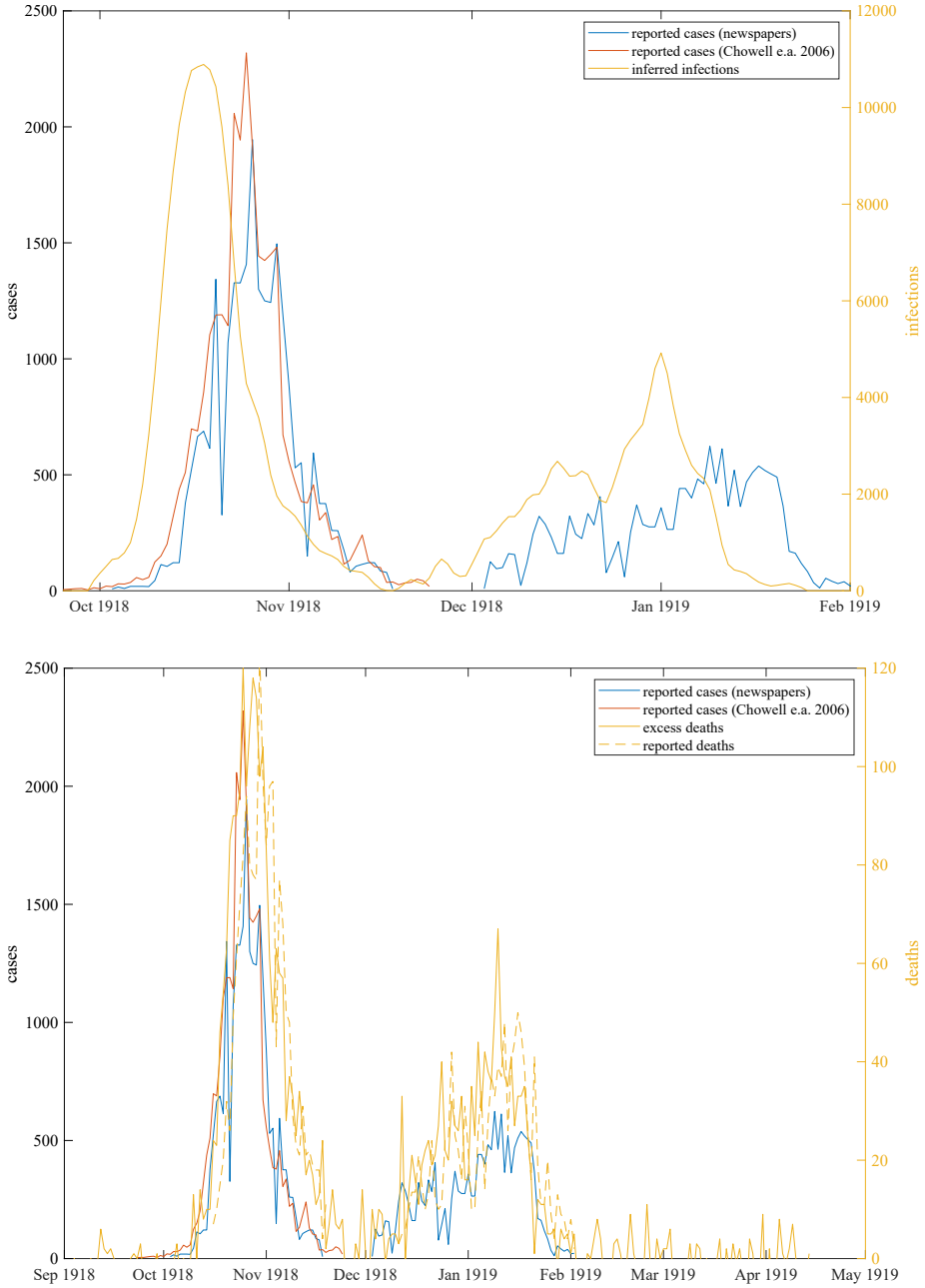

Figure 5: Reported cases from newspapers and Chowell, Nishiura, and Bettencourt (2006) (left axis) compared to inferred infections from Figure 1 in the top figure, to reported deaths from newspapers and excess deaths from Figure 1 in the bottom figure.

Note that a Hill function with $f(t) = \int_0^{T_m} I(t-s)ds$ nests the rational expectation model for small $T_m = 1$ ($\tau = 1$ for the `Alt` model).

Introducing a lockdown

There are various ways one can introduce a lockdown, or restrictions on economic activity. Bootsma and Ferguson (2007) assume that the effect is multiplicative, that is, lockdowns act as a tax (Eichenbaum, Rebelo, and Trabandt 2021). Alternatively, they may act as an upper bound on n. With a lockdown of intensity $0 \leq p_{ct} \leq 1$, (10a) becomes either of

$$\frac{n_t}{n^*} = \begin{cases} (1-p_{ct})(\sqrt{1+\alpha_t^2} - \alpha_t) & \text{(default)} \\ \min\{1-p_{ct}, \sqrt{1+\alpha_t^2} - \alpha_t\} & \texttt{(Min)} \end{cases}$$

and $K(t)$ is replaced in (13) by either of

$$\tilde{K}(t) = \begin{cases} (1-p_{ct})^2 K(t) \\ \min\{(1-p_{ct})^2, K(t)\}. \end{cases}$$

A suite of models

I will consider a suite of models. The naming convention will indicate the choices for:

1. the argument of the behavioral function b_t: deaths (the default) or infections (`Inf`)
2. the functional form for memory: default or modified (`Alt`)
3. the functional form for λ_t: `Hill`, `Exp`, `Power`, or that implied by (13) (`model`)
4. the way in which lockdowns interact with the dynamics: multiplicative (default) or `Min`.

4 Estimation

The estimation procedure can be summarized as follows. Given parameter values (including the epidemic's starting date) I solve the deterministic model and compute predicted for deaths and activity.

I assume that the each day's deaths follow a Poisson process whose parameter is the model's predicted value for that day, while measured economic activity follows a Gaussian process whose mean is the model's predicted activity for that day and whose variance is σ.[12] The data consists of daily deaths D and economic activity y (proxied by the index of ridership).

The log likelihood is proportional to

$$\mathcal{L}(D, y, \theta) \propto \sum_t \log(\hat{D}_t(\theta))D_t - \hat{D}_t(\theta) - \frac{y_t - \hat{y}_t(\theta)}{2\sigma} - \frac{\log(\sigma)}{2}$$

[12] This treats the noise term in activity as measurement error. An alternative would be to model the transmission rate as stochastic and let the daily fluctuations in ridership affect the epidemic.

model (LL)	μ (%)	R_0	κ	τ (%)	T_m	p_c	deaths (%)
Hill	1.22	2.25	320.1		17	0.41	5
9211.3	[1.17,1.27]	[2.24,2.27]	[284,357]		[16.2,18.3]	[0.39,0.43]	[3,7]
HillAlt	1.17	2.42	638.5	1		0.42	2
9208.7	[1.13,1.21]	[2.41,2.43]	[563,719]	[0.93,1.10]		[0.40,0.43]	[-1,4]
HillMin	1.22	2.39	369.3		22	0.61	9
9233.8	[1.18,1.27]	[2.38,2.4]	[333,406]		[21.3,23.1]	[0.61,0.62]	[7,10]
HillAltMin	1.15	2.66	660.8	0.76		0.66	7
9221.9	[1.1,1.19]	[2.65,2.67]	[586,741]	[0.69,0.83]		[0.65,0.67]	[5,9]
HillAltInf	1.10	2.86	213.3	0.4		0.46	-2
9186.5	[1.06,1.14]	[2.83,2.88]	[187,241]	[0.36,0.44]		[0.44,0.49]	[-4,1]
HillAltInfMin	1.08	2.99	223.9	0.39		0.70	-2
9187.1	[1.04,1.12]	[2.97,3.01]	[196,253]	[0.34,0.44]		[0.69,0.71]	[-4,2]
Power	1.22	2.20	259.7		16	0.40	1
9203.2	[1.17,1.27]	[2.19,2.22]	[231,289]		[15.3,17.5]	[0.38,0.42]	[-2,4]
PowerAlt	1.18	2.36	515.0	1		0.39	-4
9203.9	[1.13,1.22]	[2.34,2.37]	[456,577]	[0.96,1.15]		[0.37,0.42]	[-6,-1]
PowerAltMin	1.15	2.65	533.7	0.73		0.66	3
9226.1	[1.1,1.19]	[2.64,2.67]	[477,594]	[0.67,0.80]		[0.65,0.66]	[-0,5]
Exp	1.22	2.20	259.8		16	0.40	1
9203.2	[1.17,1.27]	[2.19,2.22]	[232,289]		[15.3,17.5]	[0.38,0.42]	[-2,4]
ExpAlt	1.18	2.36	515.3	1		0.39	-4
9203.9	[1.13,1.22]	[2.34,2.37]	[456,577]	[0.96,1.15]		[0.37,0.42]	[-6,-1]
ExpAltMin	1.15	2.65	534.0	0.73		0.66	3
9226.1	[1.1,1.19]	[2.64,2.67]	[478,593]	[0.67,0.80]		[0.65,0.66]	[-0,5]
model	1.22	2.20	261.3		17	0.40	1
9203.9	[1.18,1.27]	[2.19,2.22]	[233,291]		[15.4,17.6]	[0.38,0.42]	[-1,4]
modelAlt	1.18	2.36	519.2	1		0.39	-4
9204.6	[1.13,1.22]	[2.34,2.37]	[460,582]	[0.96,1.14]		[0.37,0.42]	[-6,-1]
modelMin	1.22	2.39	312.0		21	0.61	5
9235.3	[1.17,1.26]	[2.37,2.4]	[285,340]		[20.5,22.4]	[0.60,0.62]	[3,7]
modelAltMin	1.15	2.65	537.9	0.73		0.66	3
9225.9	[1.1,1.19]	[2.64,2.67]	[481,598]	[0.67,0.80]		[0.65,0.66]	[1,6]

Table 1: Median estimates without the ridership data (95% confidence intervals between brackets).

I calibrate the mean latency period to $1/\alpha = 1.5$ days and the mean infectious period to $1/\nu = 1.8$ days as in Bootsma and Ferguson (2007). I set the fraction of asymptomatics ϕ at 0.7. I estimate the rest, namely σ (the variance of ridership data) and $\theta = (\mu, R_0, \kappa, T_m, p_c, d_0)$, where $R_0 = \lambda/\nu$ and d_0 is the first day of the epidemic. The mappings $D_t(\theta)$ and $y_t(\theta)$ is clearly nonlinear: for a given θ I simulate the epidemic with a seed infection of 1 at d_0 and a time scale of 0.1 day.[13] I use MCMC methods with a flat prior to estimate the vector of parameters.[14]

Tables 1 and 2 report the results with and without the ridership data.

There is little difference between Tables 1 to 2: the additional ridership data does not change the parameter estimates. This may be disappointing (the ridership data contain no information) or comforting (the model's behavioral channel is validated by observations on mobility). I tend to the latter view. Validating the behavioral channel postulated by epidemiologists with actual mobility data has important consequences for optimal policy as we shall see.

The models that use infections rather than deaths as input to the behavioral function do poorly. In addition, rational expectations (which imposes a dependence on current infections only, hence τ close to 1) is rejected.

The models that assume that lockdowns are an upper bound (`Min`) rather than a tax do systematically better. The functional forms `Power`, `Exp`, and `Hill` do reasonably well, but the form `model` does at least as well, and it has the additional advantage that it maps into an economic interpretation.

[13] The seed makes some difference for the parameter estimates, notably R_0 and the starting date which is later for a larger seed. A time unit smaller than 0.1 does not alter the estimates.

[14] The chains are 500,000 long with a 20% burn-in.

model (LL)	μ (%)	R_0	κ	τ (%)	T_m	p_c	deaths (%)
Hill	1.21	2.26	312.4		17	0.41	6
9581.6	[1.17,1.26]	[2.24,2.27]	[277,349]		[16.0,18.1]	[0.39,0.43]	[4,8]
HillAlt	1.16	2.42	635.0	1		0.41	2
9580.2	[1.12,1.21]	[2.41,2.44]	[558,716]	[0.96,1.13]		[0.39,0.43]	[-1,4]
HillMin	1.21	2.39	355.3		22	0.61	9
9603.9	[1.17,1.26]	[2.38,2.4]	[320,392]		[21.1,22.9]	[0.60,0.62]	[8,10]
HillAltMin	1.15	2.59	630.7	0.81		0.65	6
9585.8	[1.1,1.19]	[2.58,2.61]	[556,710]	[0.74,0.88]		[0.64,0.66]	[4,8]
HillAltInf	1.09	2.78	202.7	0.44		0.45	-2
9552.4	[1.05,1.13]	[2.76,2.8]	[177,229]	[0.39,0.48]		[0.43,0.47]	[-4,1]
HillAltInfMin	1.09	2.81	208.0	0.45		0.67	-3
9548.4	[1.05,1.12]	[2.79,2.83]	[179,238]	[0.39,0.50]		[0.67,0.68]	[-5,0]
Power	1.21	2.21	253.9		16	0.40	1
9572.7	[1.17,1.26]	[2.19,2.22]	[226,283]		[15.1,17.3]	[0.38,0.42]	[-1,4]
PowerAlt	1.17	2.36	513.6	1.1		0.39	-4
9575.8	[1.13,1.21]	[2.34,2.37]	[454,577]	[0.98,1.18]		[0.37,0.41]	[-6,-1]
PowerAltMin	1.15	2.59	522.3	0.78		0.64	1
9591.4	[1.11,1.19]	[2.57,2.6]	[465,582]	[0.72,0.85]		[0.64,0.65]	[-2,4]
Exp	1.21	2.21	253.9		16	0.40	1
9572.8	[1.17,1.26]	[2.19,2.22]	[225,284]		[15.1,17.3]	[0.38,0.42]	[-1,4]
ExpAlt	1.17	2.36	512.9	1.1		0.39	-4
9575.8	[1.13,1.21]	[2.34,2.37]	[453,577]	[0.98,1.18]		[0.37,0.41]	[-6,-1]
ExpAltMin	1.15	2.59	521.8	0.78		0.64	1
9591.4	[1.11,1.19]	[2.57,2.6]	[465,582]	[0.72,0.85]		[0.64,0.65]	[-2,4]
model	1.21	2.21	256.3		16	0.40	2
9573.6	[1.17,1.26]	[2.19,2.22]	[228,286]		[15.2,17.4]	[0.38,0.42]	[-1,4]
modelAlt	1.17	2.36	517.9	1.1		0.39	-3
9576.5	[1.13,1.21]	[2.35,2.37]	[457,583]	[0.98,1.17]		[0.37,0.41]	[-6,-1]
modelMin	1.21	2.39	302.7		21	0.60	6
9606.5	[1.16,1.25]	[2.38,2.4]	[276,331]		[20.4,22.3]	[0.60,0.61]	[4,7]
modelAltMin	1.15	2.59	524.9	0.78		0.64	2
9591.1	[1.11,1.19]	[2.57,2.6]	[468,586]	[0.72,0.85]		[0.64,0.65]	[-1,5]

Table 2: Median estimates with the ridership data (95% confidence intervals between brackets).

The best model is (`modelMin`), both without and with ridership data. The values of μ and R_0 are fairly consistent across models, and with the epidemiological literature (Mills, Robins, and Lipsitch 2004; Biggerstaff et al. 2014). The magnitudes of κ can be gauged as follows. The level of f_t (in deaths per 100,000) that halves the transmission rate is N/κ for `Hill`, $N/\sqrt{2}\kappa$ for `model`, $N(1-2^{1/\kappa})$ for `Power`, $N\log(2)/\kappa$ for `Exp` (the population of San Francisco can be estimated at 494,350 in September 1918 Department of Commerce, Bureau of the Census 1923). For the estimates of κ these values are all in the range of 1300–1800. By comparison the peak of $\int_0^{T_m} D(t-s)ds$ (for $T_m = 20$) was 1680 on November 7. This suggests a fairly strong, but (given the value of T_m) sluggish response. The values of p_c indicate a substantial reduction in activity (in the `Min` model, the upper bound is $\sqrt{1-p_c} \sim 60\%$).

A final remark on the results: the last column of both tables gives the reduction in deaths, compared to a counterfactual of no intervention. In terms of magnitudes, a reduction of 6% corresponds to about 200 lives saved during the six months of the epidemic. The numbers are relatively small and even negative for some models, for reasons that will become apparent in the next section.

Figures 6 and 7 compare the simulations of the best-fitting model with the data for deaths and ridership (Figures 12 to 13 in the Appendix show the same graphs for all models).

In Figure 6 the model's fit can be compared with the data. The counterfactual assumes no restrictions and results in a single wave that peaks at 230 deaths per day instead of 110. Figure 7 plots predicted ridership, and also the ridership at each point in time that would result if restrictions were lifted. Comparing the two lines shows when the restrictions

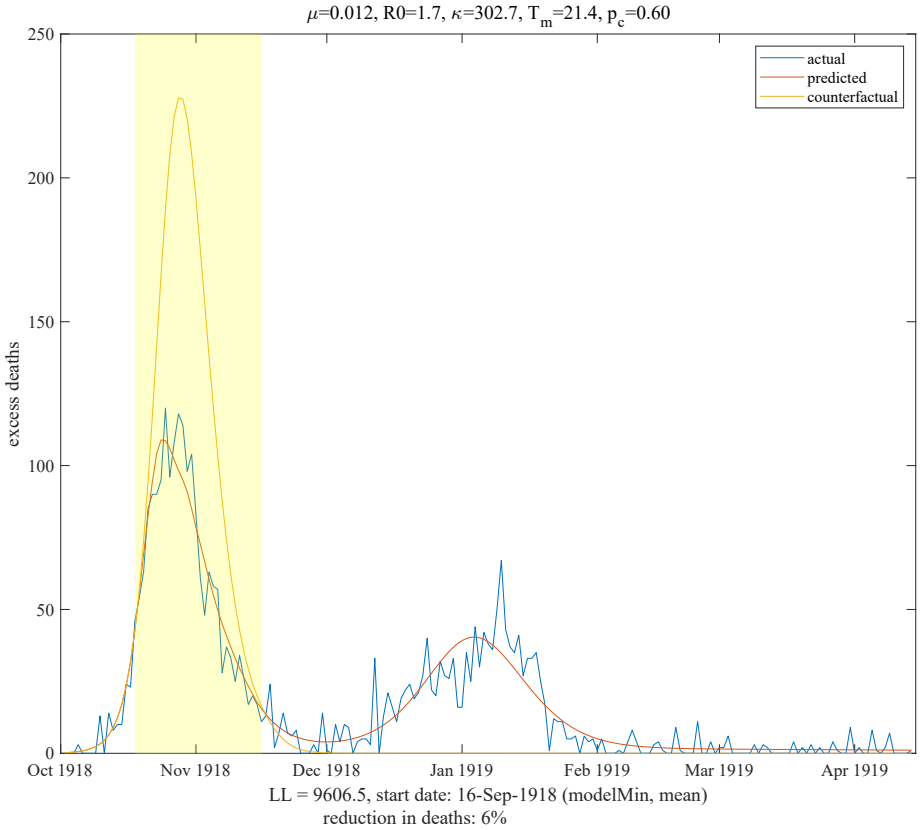

Figure 6: Deaths, actual (estimated excess deaths), fitted, and counterfactual (with no lockdown), best model. The yellow area indicates the period when places of amusement were closed.

were binding: most of the time, but not always. The restrictions were mostly binding in the early part of the restriction period (the yellow area in the Figure), when cumulated deaths hadn't risen enough to prompt an equal-sized reduction in activity through the behavioral channel.

5 Optimal policy

The previous section showed that the reduction in deaths was small. The phase diagram, constructed with the estimated parameters and shown in Figure 8 shows why.[15]

The red line shows the path traced by the counterfactual without lockdown. The blue line shows the simulated path with lockdown. The light grey lines are not properly speaking tracing the phase diagram, which is difficult to do with the lags incorporated in the behavioral function. Instead they trace paths without lockdowns and starting from values of $S < 1$. The dark blue line represents the points along those paths where $\dot{E} + \dot{I} = 0$, the equivalent of the vertical line at $1/R_0$ in Figure 4. Finally, the outer green line plots a counterfactual with no behavioral channel.

If one knew the parameters but ignored the behavioral channel, one would take the green line to represent the result of no intervention: the overshoot is massive and interven-

[15]The simulations done for this section were computed for 1200 days, to make sure that the epidemic is not merely postponed beyond the simulation horizon.

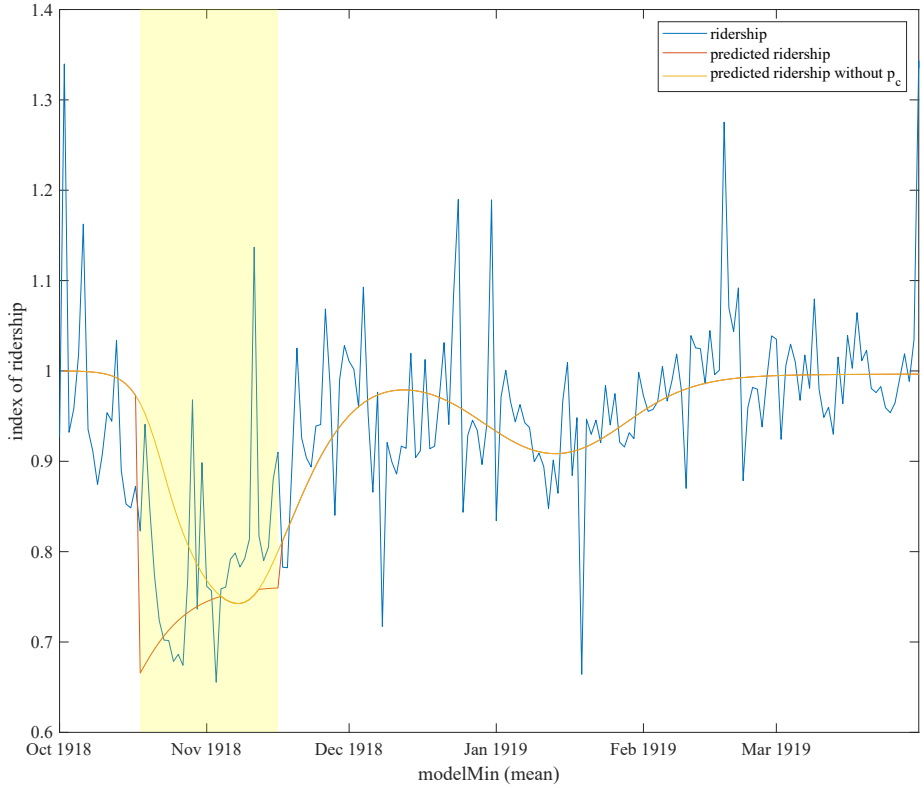

Figure 7: Economic activity (as proxied by ridership), actual, predicted, and predicted if restrictions were lifted. The yellow area indicates the period when places of amusement were closed.

tion could reduce deaths considerably. In reality there is little margin between the herd immunity threshold and the final outcome without intervention. The behavioral channel is to some degree a substitute for intervention.

Note that an ill-timed intervention can do worse than no intervention at all, which is not possible in the model without lags in the behavioral function. The reason is that there are various paths leading to the same point in the (S, I) plane. A point on the counterfactual (no-lockdown) path corresponds to a particular prior history of deaths, and hence current value of $M(t)$. An intervention can arrive at the same point but with a lower value of $M(t)$, leading to less precautionary behavior and higher transmission rates going forward. This is the reason why some models estimate negative death reductions in Tables 1 and 2.

To simplify the optimal policy problem I assume that authorities can institute only one lockdown of constant intensity. There are thus three control variables: the start date, the end date, and the intensity of the intervention.[16] By varying the control variables one can construct the Pareto frontier tracing the trade-off between deaths and economic activity.

Figure 9 plots that frontier, with economic activity measured in annualized consumption equivalent relative to no pandemic. The no-lockdown point represents the maximum

[16]In the computation the start date is constrained between 0, day of the first case, and 50. The actual intervention ran from 32 to 61. The duration is bounded above by 180. I let simulations run for 1200 days to make sure that the epidemic does not restart: for that purpose, an infection of 1e-3 individual is the same as 0. The longer simulation length means that some results will be slightly different from those presented in the previous section, where simulations lasted seven months.

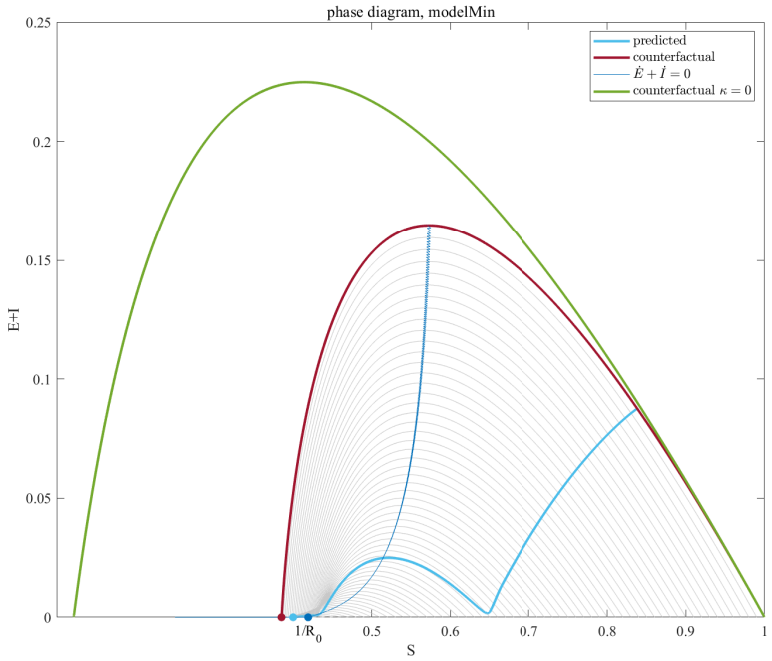

Figure 8: Phase diagram for the best-fitting model.

of deaths and minimum reduction in economic activity. The actual policy followed appears to be clearly interior to the frontier. Authorities in San Francisco (had they known the model) could have done better on both dimensions. The very steep slopes of the frontier indicate that, starting from no lockdown, important gains could be made at little economic cost up to a point, after which further gains would abruptly become costly.

Arguably it might have been difficult for authorities to control the intensity of the lockdown, but the start and end dates were clear choices. The inner line in Figure 9 constrains the intensity to be its estimated (actual) value. The results are the same.

The actual shape of the Pareto frontier effectively obviates the need to choose terms for the tradeoff. The upper-right point is nearly a kink, suggesting that for a wide range of relative weights on deaths and economic activity, the answer would be the same.

Underlying the Pareto frontier are three choices for the intervention: start date, end date, and intensity. Figure 9 shows the values of these choices, with intensity on the left axis, start and end date on the right axis. The actual values of the control variables are also shown. Going from left to right traces the Pareto frontier from no-lockdown. The steeply vertical part of the Pareto frontier corresponds the left part of the graph, with short but early lockdowns of increasing but moderate intensity. The kink of the Pareto frontier is passed when the lockdown suddenly becomes much longer. Past that point further reductions in deaths require an increasingly stringent lockdown.

Along the Pareto frontier the relation between controls (length, intensity, earliness of intervention) and death outcomes is monotonic, but very nonlinear. Along the whole frontier earliness does not change much; length is nearly constant up to the kink and then after. Only the relation between intensity and reduction in deaths seems roughly

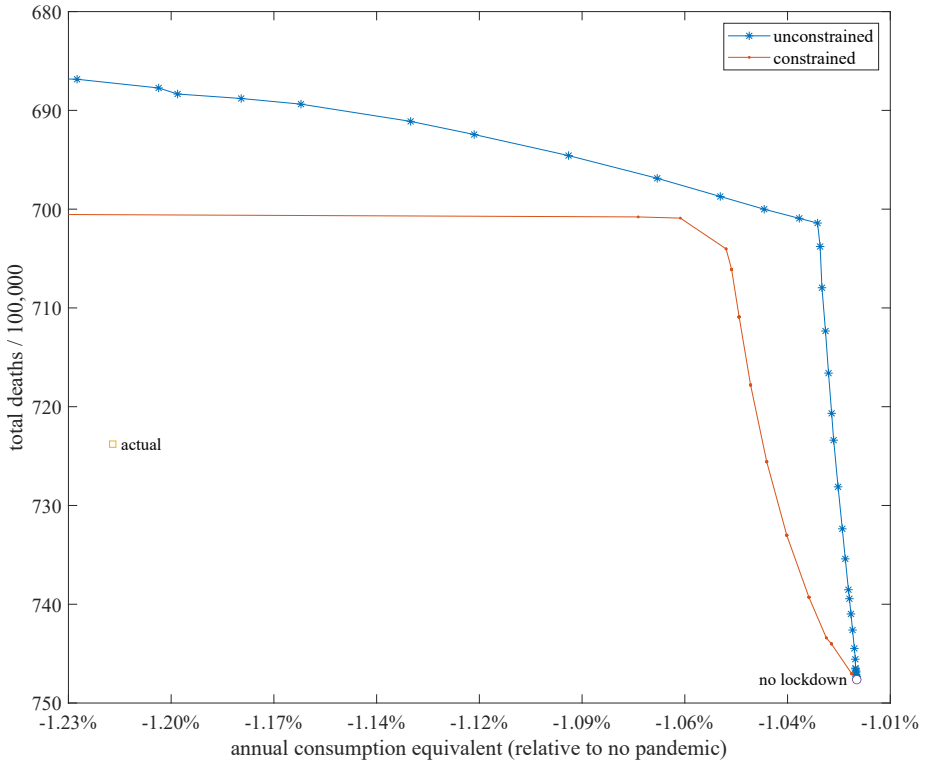

Figure 9: Pareto frontier between economic activity and deaths. The lockdown intensity is fixed at its actual value in the constrained case.

linear. Away from the frontier (as the actual policy is) there is no telling how controls and outcomes would relate—as noted earlier, interventions can actually worsen outcomes.

The next figure compares the dynamics of the epidemic under a few policies of interest. Figure 11 plots four paths in the $(S, E + I)$ space. One is the actual (or fitted) path, the three others correspond to the end points and the kink of the frontier: maximum activity (no lockdown), the kink which I label Pareto optimum, and minimum deaths on the far left of the Pareto curve.

The maximum activity (no intervention) obviously does the worst in terms of deaths, ending at the lowest value of S. The Pareto optimum corresponds to an early but very brief intervention that bends the curve slightly at its start and lets the behavioral response take over for the rest of the epidemic. The death-minimizing intervention lets infections rise as high as the no-lockdown, but then imposes a strong and long intervention to bring down infections definitively. The actual policy starts earlier but is milder and shorter. Once it ends a second wave rises again, moderated by the behavioral channel, but not enough to prevent a trailing end of deaths that extends beyond what the death-minimizing path achieves.

Compared with the (fitted) actual policy, an earlier, shorter, and milder closing (what I call the Pareto optimum), would have reduced incidence by 1.7%, or 8,400 cases, and deaths (using the estimated rate of 1.2%) by 100. Compared to no lockdowns, the actual policy avoided 10,900 cases and 130 deaths. The most stringent policy would have saved an additional 270 lives.[17] This is another way of stating that, while mitigation policies

[17]This number is slightly smaller than the 200 deaths avoided in the previous section, because the horizon

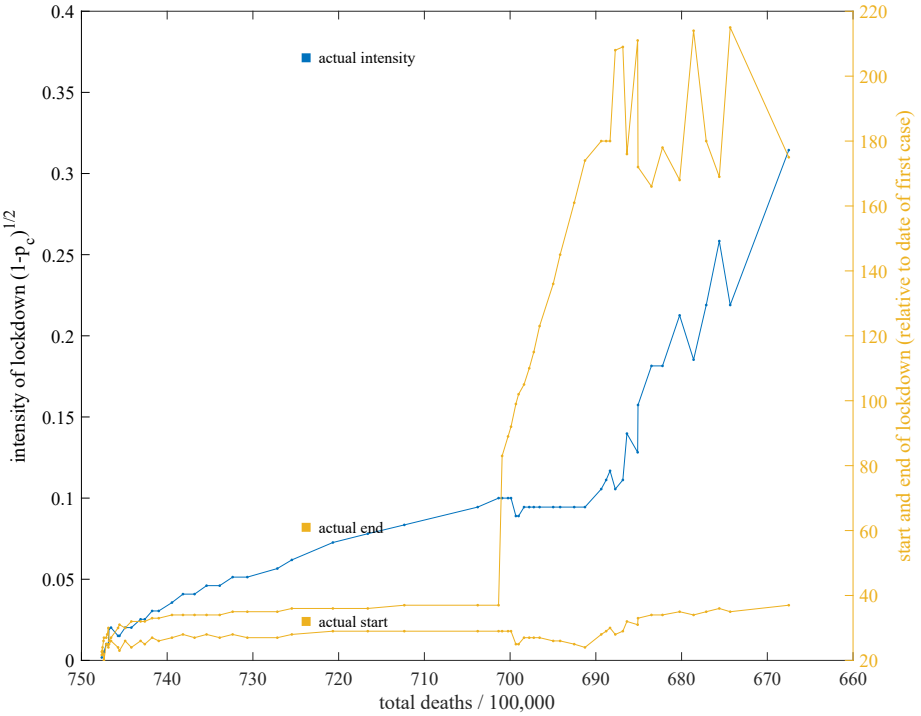

Figure 10: Policies that implement the Pareto frontier, indexed by total deaths. The left axis plots the intensity of lockdown, the right axis plots the start and end dates of lockdown.

could save lives, the room for action was limited.

6 The effect of masks

The final exercise consists in estimating the effectiveness of masks. Recall that a mask mandate was in effect from January 17 to February 1, but was not accompanied by any other restrictions. The assumption is that masks did not affect mobility during the second wave. I allow them to affect the transmission rate λ_t, with an intensity to be estimated, but not economic activity. Table 3 presents the results. The last two columns show the overall reduction in deaths compared to no intervention, and the reduction from the January mask mandate alone compared to no mask mandate (the denominator is therefore deaths from January 1919 onwards).

The parameters p_c and p_{c2} measure the force of the first (masks and closings in November) and second (masks in January) interventions. With one additional parameters, the models do at least as well as in Table 2. The ordering changes, so that the best fit is now achieved by `HillAlt` (Hill function with exponential discounting of deaths). Estimated death rates (μ) are higher and basic reproduction numbers (R_0) are lower. Most models (including the best fitting) see little difference between p_c and p_{c2}, although the estimates are imprecise. If so, this suggests that masks alone would have done as well in November, without economic costs. However the best model of the previous section (`modelMin`)

of the simulation is 1200 days in this section; as incidence drags on at a low level for a long period of time in the model, the differences between policies narrow.

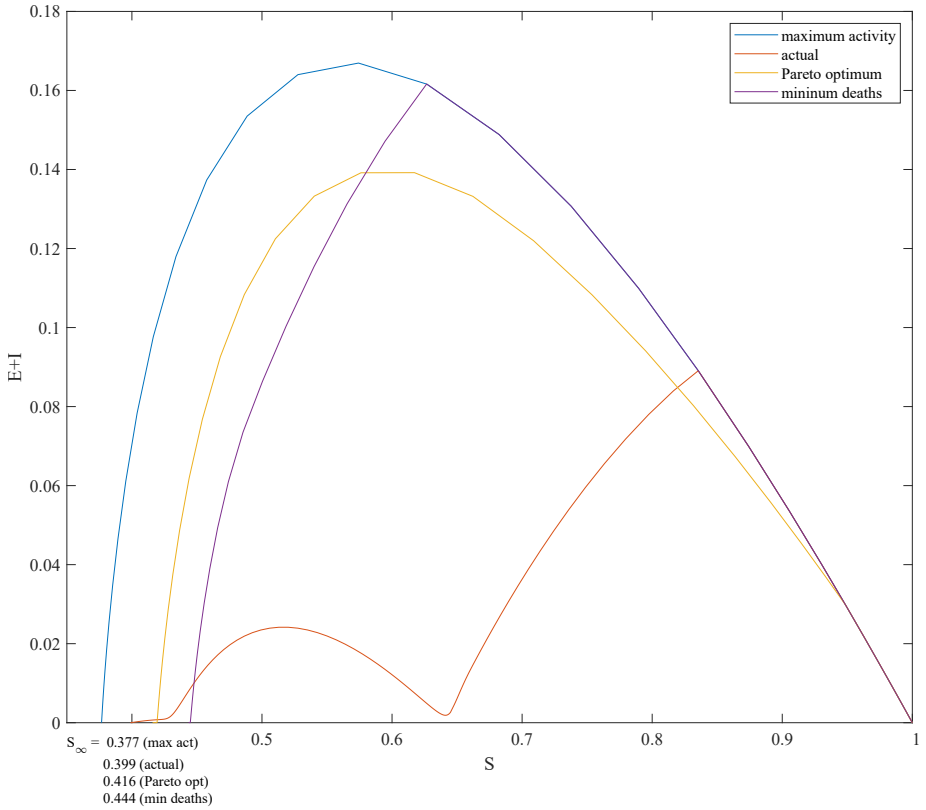

Figure 11: Paths for selected policies along the Pareto frontier.

rejects any role for masks, while another (`HillMin`) finds an effect of masks but the additional parameter does not increase the log-likelihood significantly.

Hassler claimed that wearing masks in early December when he first requested it would have saved two or three hundred lives (*Chronicle*, 26 January 1919, p. 9). Conversely F. Holmes Smith, Health Officer of San Mateo County, expressed doubts about the value of reimposing a mask mandate in January 1919. He pointed to the fact that in his county, where masks had not been required, the same phenomenal decline in new cases has been reported as in San Francisco: "I think the decline of the epidemic can be explained by the statement that masks were required at the time when the epidemic had reached its peak, and under normal conditions would have declined anyway" (*Chronicle*, 27 January 1919, p. 3). In the model Smith is partly correct: the second wave had peaked, but the mask mandate still slowed the epidemic and reduced overshooting. However Hassler's claim is not supported by the model: a counterfactual with a mask mandate starting on December 5 saves only about thirty lives rather than three hundred.

Interestingly there was interest in evaluating the effectiveness of masks, using data (but without the benefit of a model). Bases on the similarities in curves across cities that required or didn't require masks, the State Board of Health concluded that gauze masks, while sound in principle, were useless because of their defective quality and the difficulty of requiring their use in close quarters (Kellogg 1919b). My results suggest less skepticism.

model (LL)	μ (%)	R_0	κ	τ (%)	T_m	p_c	p_{c2}	deaths (%)	
Hill	1.35	2.13	339.0		16	0.36	0.28	4	32
9608.2	[1.29,1.41]	[2.12,2.14]	[298,382]		[15.0,17.0]	[0.33,0.38]	[0.22,0.35]	[2,6]	[25,38]
HillAlt	1.32	2.22	655.3	1.1		0.35	0.31	2	33
9612.8	[1.27,1.38]	[2.21,2.23]	[571,745]	[1.01,1.19]		[0.33,0.38]	[0.25,0.38]	[1,3]	[27,39]
HillMin	1.25	2.34	353.7		22	0.60	0.37	8	12
9605.1	[1.2,1.3]	[2.33,2.35]	[313,395]		[20.5,22.6]	[0.59,0.61]	[0.01,0.44]	[7,10]	[0,21]
HillAltMin	1.23	2.46	640.6	0.84		0.63	0.44	5	24
9601.4	[1.18,1.28]	[2.44,2.47]	[560,726]	[0.76,0.92]		[0.62,0.64]	[0.38,0.50]	[2,7]	[18,30]
HillAltInf	1.23	2.54	219.8	0.43		0.41	0.33	1	33
9594.7	[1.18,1.28]	[2.52,2.56]	[191,249]	[0.39,0.47]		[0.38,0.43]	[0.27,0.40]	[-0,3]	[28,39]
HillAltInfMin	1.20	2.64	229.9	0.44		0.65	0.48	0	31
9584.1	[1.15,1.25]	[2.62,2.66]	[198,263]	[0.38,0.49]		[0.65,0.66]	[0.42,0.54]	[-1,3]	[26,37]
Power	1.36	2.09	290.3		15	0.33	0.29	-1	34
9604.0	[1.3,1.42]	[2.07,2.1]	[256,326]		[13.8,15.9]	[0.30,0.36]	[0.23,0.35]	[-2,1]	[27,40]
PowerAlt	1.33	2.17	559.7	1.2		0.32	0.31	2	34
9611.5	[1.28,1.39]	[2.15,2.18]	[491,633]	[1.06,1.26]		[0.28,0.35]	[0.25,0.37]	[-2,6]	[28,40]
PowerAltMin	1.25	2.40	526.0	0.84		0.62	0.44	-2	25
9604.4	[1.2,1.3]	[2.38,2.41]	[463,595]	[0.77,0.92]		[0.61,0.63]	[0.38,0.49]	[-4,1]	[18,32]
Exp	1.36	2.09	290.5		15	0.33	0.29	-1	34
9604.0	[1.3,1.42]	[2.07,2.1]	[256,328]		[13.8,15.8]	[0.30,0.36]	[0.23,0.35]	[-2,2]	[27,40]
ExpAlt	1.33	2.17	559.8	1.2		0.32	0.31	2	34
9611.5	[1.28,1.39]	[2.16,2.18]	[491,633]	[1.07,1.26]		[0.28,0.35]	[0.25,0.37]	[-2,5]	[28,40]
ExpAltMin	1.25	2.40	526.4	0.84		0.62	0.44	-2	25
9604.4	[1.2,1.3]	[2.38,2.41]	[464,595]	[0.77,0.92]		[0.61,0.63]	[0.38,0.49]	[-4,1]	[19,32]
model	1.36	2.09	292.3		15	0.33	0.29	-0	34
9604.6	[1.31,1.42]	[2.08,2.1]	[258,329]		[13.9,15.9]	[0.30,0.36]	[0.23,0.35]	[-1,2]	[28,40]
modelAlt	1.34	2.17	562.8	1.2		0.32	0.31	2	34
9611.9	[1.28,1.39]	[2.16,2.18]	[494,636]	[1.06,1.25]		[0.29,0.35]	[0.25,0.37]	[-2,5]	[28,40]
modelMin	1.23	2.33	293.0		22	0.60	0.00	5	0
9606.7	[1.18,1.27]	[2.32,2.35]	[267,320]		[20.6,22.5]	[0.59,0.61]	[0.00,0.04]	[3,7]	[0,0]
modelAltMin	1.25	2.40	528.3	0.84		0.62	0.44	-1	25
9604.3	[1.2,1.3]	[2.38,2.41]	[465,594]	[0.77,0.92]		[0.61,0.63]	[0.38,0.49]	[-3,2]	[19,32]

Table 3: Median estimates with the ridership data and allowing for masks in January.

7 Conclusion

Even before the second wave hit, Hassler claimed that masks and other measures taken had avoided 20,000 cases and saved 1,500 lives (*Chronicle*, 22 Nov, p. 4). Was he anywhere near the truth?

Policy evaluation requires counterfactuals, which in turn require a model. The one I use is structural. Daily observations of ridership on the San Francisco Muni, interpreted through the lens of a simple model of activity under risk of infection, confirms that a standard SEIR model captures the dynamics of the 1918 Spanish influenza epidemic. Substantial behavioral response incorporating long lags is needed, and this has implications for optimal policy. To a large degree the behavioral response substitutes for policy intervention, so that the room for actual improvement (in terms of deaths) is much narrower than in a model that ignores the response. My estimate of the effect of the intervention in San Francisco on deaths is an order of magnitude lower than Hassler's.

The Pareto frontier between economic activity and deaths is steep, implying limited trade-offs. Actual policy was away from the frontier, and the optimal policy would have achieved a better outcome with a shorter intervention. Reducing deaths further was feasible, but at the cost of a six-month long intervention that seems unrealistic given the pressures I documented in the historical section. Masks may well have been effective.

The fact that San Francisco was away from the frontier suggests that other cities may have been. The dynamics of the SEIR model with behavioral feedback indicate that interventions can be counterproductive away from the frontier. Even along the frontier the relation between intervention and outcome is non linear. This suggests caution when using reduced form, linear models that rely on cross-sectional variation to evaluate the effectiveness or long-term impact of interventions.

Bibliography

Acemoğlu, Daron, Victor Chernozhukov, Iván Werning, and Michael Whinston. 2020. *Optimal Targeted Lockdowns in a Multi-Group SIR Model.* Working paper 27102. NBER, May.

Arias, Jonas, Jesús Fernández-Villaverde, Juan Rubio Ramírez, and Minchul Shin. 2021. *Bayesian Estimation of Epidemiological Models: Methods, Causality, and Policy Trade-Offs.* Technical report. March.

Atkeson, Andrew. 2020. *How Deadly Is COVID-19? Understanding The Difficulties With Estimation Of Its Fatality Rate.* Technical report. April.

Beach, Brian, Karen Clay, and Martin Saavedra. 2021. "The 1918 Inuenza Pandemic and its Lessons for COVID-19." *Journal of Economic Literature* forthcoming.

Biggerstaff, Matthew, Simon Cauchemez, Carrie Reed, Manoj Gambhir, and Lyn Finelli. 2014. "Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature." *BMC Infectious Diseases* 14, no. 1 (September).

Bootsma, Martin C. J., and Neil M. Ferguson. 2007. "The effect of public health measures on the 1918 influenza pandemic in U.S. cities." *Proceedings of the National Academy of Sciences* 104, no. 18 (April): 7588–93.

Bureau of the Census. 1913–23. *Mortality Statistics.* Government Printing Office.

California State Board of Health. 1918. "The Influenza Outbreak of 1889–91 in California." *California State Board of Health Bulletin* 14, nos. 5–6 (January): 165–69.

Chowell, Gerardo, Hiroshi Nishiura, and Luís M.A Bettencourt. 2006. "Comparative estimation of the reproduction number for pandemic influenza from daily case notification data." *Journal of The Royal Society Interface* 4, no. 12 (October): 155–66.

Correia, Sergio, Stephan Luck, and Emil Verner. 2020. *Pandemics Depress the Economy, Public Health Interventions Do Not: Evidence from the 1918 Flu.* Technical report. SSRN.

Crosby, Alfred W. 2003. *America's Forgotten Pandemic.* Cambridge University Press.

Department of Commerce, Bureau of the Census. 1923. *Estimates of Population of the United States by States and Cities, 1910 to 1923.* Washington, DC: Government Printing Office.

Eichenbaum, Martin S., Sergio Rebelo, and Mathias Trabandt. 2021. "The Macroeconomics of Epidemics." *The Review of Financial Studies* 34, no. 11 (April): 5149–5187.

Ferguson, Neil M., Derek A.T. Cummings, Simon Cauchemez, Christophe Fraser, Steven Riley, Aronrag Meeyai, Sopon Iamsirithaworn, and Donald S. Burke. 2005. "Strategies for containing an emerging influenza pandemic in Southeast Asia." *Nature* 437, no. 7056 (August): 209–14.

Fernández-Villaverde, Jesús, and Charles I. Jones. 2020. "Macroeconomic Outcomes and COVID-19: A Progress Report." *Brookings Papers on Economic Activity* 2020 (3): 111–166.

Glover, Andrew, Jonathan Heathcote, Dirk Krueger, and José-Víctor Ríos-Rull. 2020. *Health versus Wealth: On the Distributional Effects of Controlling a Pandemic.* Technical report. April.

Goldstein, E., J. Dushoff, J. Ma, J. B. Plotkin, D. J. D. Earn, and M. Lipsitch. 2009. "Reconstructing influenza incidence by deconvolution of daily mortality time series." *Proceedings of the National Academy of Sciences* 106, no. 51 (December): 21825–29.

Jones, Callum, Thomas Philippon, and Venky Venkateswaran. 2021. "Optimal Mitigation Policies in a Pandemic: Social Distancing and Working from Home." Edited by Ralph Koijen. *The Review of Financial Studies* 34, no. 11 (September): 5188–5223.

Keeton, Robert W., and A. Beulah Cushman. 1918. "The Influenza Epidemic in Chicago." *Journal of the American Medical Association* 71 (24): 1963–67.

Kellogg, W. H. 1919a. "A Consideration of the Methods Used in the Control of Influenza." *California State Journal of Medicine* 17, no. 7 (July): 228–31.

———. 1919b. "The Influenza Epidemic, with Summary of Conclusions Reached as a Result of the Study of the Control Measures Adopted." *California State Board of Health Bulletin* 14, no. 7 (January): 226–41.

Korolev, Ivan. 2020. "Challenges with Estimating the Effect of Anti-Contagion Policies on the COVID-19 Pandemic." *SSRN Electronic Journal.*

Markel, Howard, Harvey B. Lipman, J. Alexander Navarro, Alexandra Sloan, Joseph R. Michalsen, Alexandra Minna Stern, and Martin S. Cetron. 2007. "Nonpharmaceutical Interventions Implemented by US Cities During the 1918-1919 Influenza Pandemic." *JAMA* 298, no. 6 (August): 644.

Mills, Christina E., James M. Robins, and Marc Lipsitch. 2004. "Transmissibility of 1918 pandemic influenza." *Nature* 432, no. 7019 (December): 904–06.

O'Shaughnessy, M. M. 1921. *The Municipal Railway of San Francisco.* San Francisco: J.A. Prud'homme Composition Co.

San Francisco Municipal Railway. 1915–22. *Financial report of the Municipal Railway of San Francisco.* San Francisco: The Recorder Company.

Toxvaerd, Flavio. 2021. *Equilibrium Social Distancing*. Cambridge-INET Working Paper Series 2020/08. University of Cambridge , Faculty of Economics.

Wilcox, Delos F. 1921. *Analysis of the electric railway problem.* New York, NY: The Author.

Yu, Duo, Qianying Lin, Alice PY Chiu, and Daihai He. 2017. "Effects of reactive social distancing on the 1918 influenza pandemic." Edited by Gerardo Chowell. *PLOS ONE* 12, no. 7 (July): e0180545.

Appendix

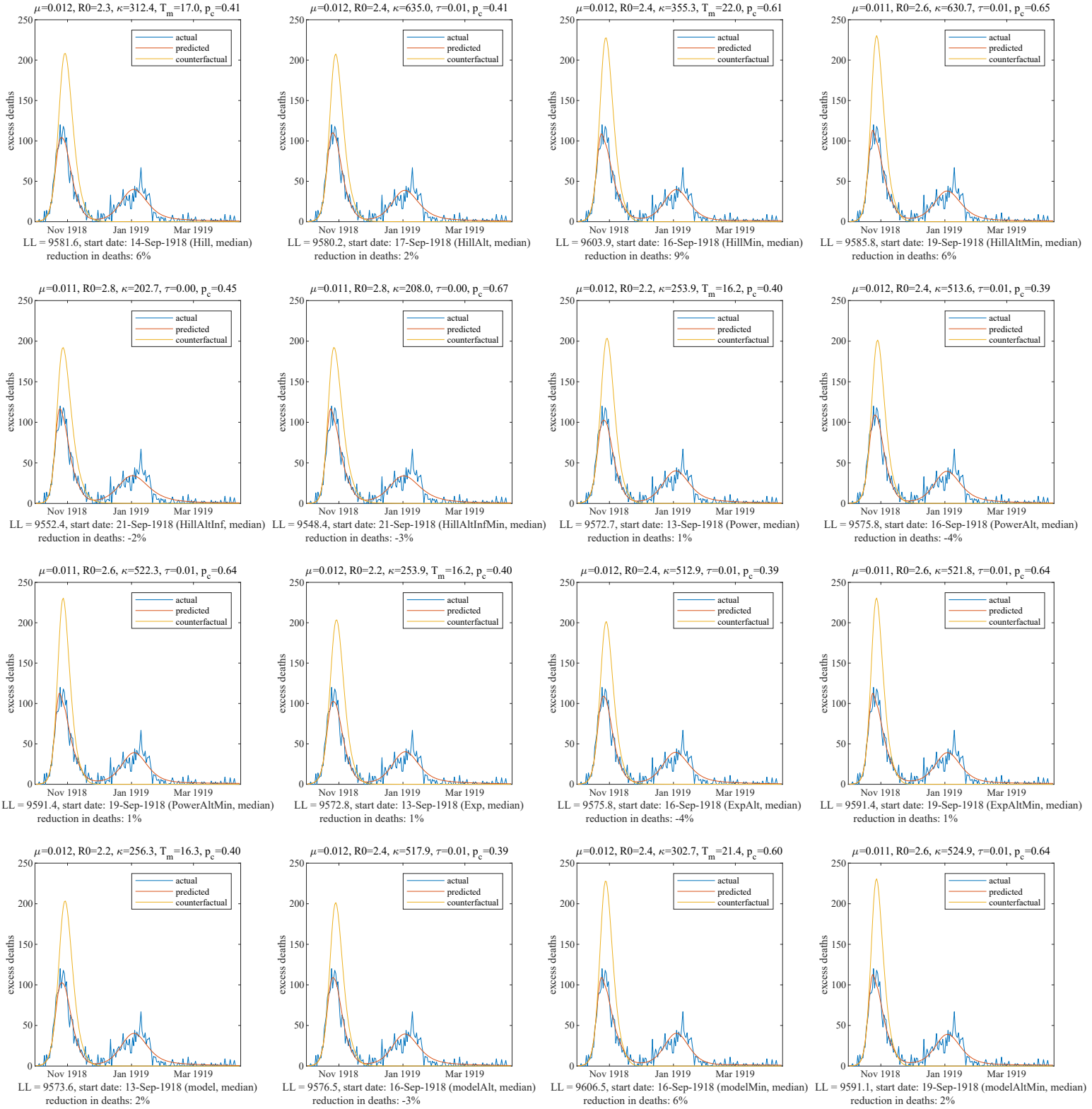

Figure 12: Deaths, actual (estimated excess deaths), fitted, and counterfactual (with no lockdown).

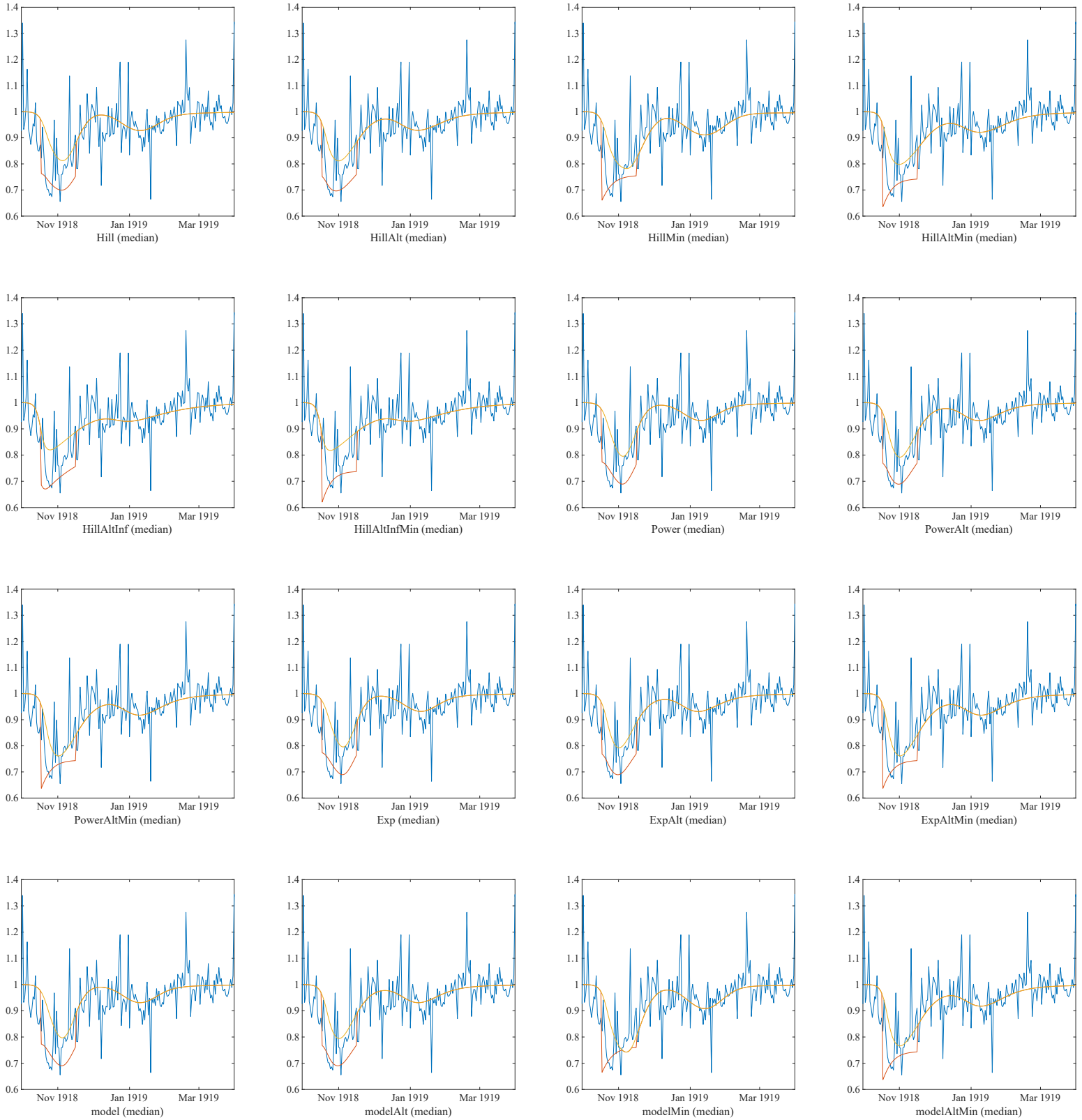

Figure 13: Economic activity (as proxied by ridership), actual and predicted.