Economically-Motivated Interactions and Disease Spread

Wilbur John Coleman II*

2020-08-10

Abstract

Human interaction is largely centered on interactions at home, school, work, and shopping. This economically-motivated partial spatial compartmentalization of our lives, in particular the persistent and recurring interaction with a very limited number of people, is surely of first-order importance for the modeling of disease spread that depends on human contact. This paper models the movement of people throughout the day to match data on the distribution of households by size and age of occupants, the distribution of K-12 schools and colleges by size and staff-student ratios, and the distribution of office and retail establishments by number of employees. The spread of disease calibrated to COVID-19 characteristics is significantly less in this model than in a corresponding SIR model in which people interact in one large location. Simulations of this model also reveal that large-scale compliance to a mitigation policy only directing symptomatic people to stay home is sufficient to control the widespread of a disease such as COVID-19, without any noticeable employment effects. A partial lockdown policy is also shown to control the widespread of such a disease, but at a cost of considerable employment effects.

*Duke University (coleman@duke.edu).

1 Introduction

This paper uses data on the distribution of households, businesses, and schools by number and age of people to capture how our economically-motivated movement between home, work, school, and shopping determines the spatial aspects of disease spread and mitigation policies. I estimate that, without any mitigation policy, about 30 percent of infections occur at home, 12 percent occur in an office, 11 percent occur at school and 47 percent occur in retail locations. Although the initial spread of the disease is faster in this type of model relative to an SIR model, due to the economically-motivated partial spatial compartmentalization of this model, the new-infection curve is significantly flattened relative to a standard SIR model. Remarkably, as long as less than about 60 percent of infected people are asymptomatic, I estimate that a mitigation policy in which all symptomatic people stay home is sufficient to prevent the widespread of a disease calibrated to COVID-19 characteristics, even if other family members without symptoms are free to go about their business. In contrast, a partial lockdown mitigation policy that is sufficiently broad to prevent the widespread of such a disease has disastrous employment consequences.

The introduction of multiple layers into the SIR model initially developed by Kermack and McKendrick (1927) is not new. See, for example, Dickison, at. al. (2012), Yagan, et. al. (2013), Zuzek, et. al., (2015), and Gevertz, et. al., (2020). What's new here is the adaption of such a model to environments in which families spend about a third of each day with each other at home, about a third at separate school or work locations, and the remainder in retail establishments of various sorts, and calibrating such a model of human interaction to match data on the distribution of households, school environments, work environments, and retail establishments. This approach can be used to evaluate a mitigation policy,[1] either self-imposed or "mandated," in which people quarantine themselves at home with family members who may or may not be infected, and as well can evaluate various lockdown policies in which some schools and office and retail establishments are directed to close. An attractive data-driven alternative is to use historical individual tracking data to capture actual physical contact, as done recently by Kucharski, et. al., (2020) and Ruiyn, et., al. (2020). This approach is limited, though, by the lack of available data, the likely incompleteness of tracking the entire population even in a confined geographic area, and the reliance on one historical episode to make inferences about future outcomes.

[1]See also the study of non-pharmaceutical interventions by Bayham, et. al. (2015), Chowdhury (2020), Ferguson, et. al., 2020), and Sofonea, et. al. (2020)

Various papers in economics have also contributed to this literature, such as Acemoglu, et. al., (2020), Alvarez, et. al., (2020), Atkeson (2020a, 2020b), Eichenbaum, et. al., (2020), Farboodi, et. al., (2020), Garriga, et. al., (2020), Jones, et. al., (2020), Jones and Fernandez-Villaverde (2020), Rampini (2020), Rowthorn, et. al., (2020), and Stock (2020). Many of these papers identified and studied important economic consequences and trade-offs involving various mitigation policies. An essential element to any of these papers is the physical disease-spreading environment within which people interact. This paper argues that economics may have an important contribution to this dimension of the problem as well.

2 eSIR Model

I will refer to the model as an eSIR model,[2] where "e" designates "economically-motivated interactions" in a way described here. In the model, people visit various locations, labeled as home, school, work, and shopping. The distribution of households[3] by size of household and age of occupants is chosen to match data from the 2010 American Community Survey.[4] All children aged 6-17 attend K-12 and some young adults aged 18-22 attend college, with the distribution of school size and college attendance chosen to match data from the National Center for Education Statistics (NCES)[5] in 2018. Young adults not in college are assumed to be in the labor force. Labor force participation rates of adults aged 23-64 and 65+ are chosen to match data from the Bureau of Labor Statistics (BLS) in 2018, with adults 23-64 not in the labor force labeled as non-workers and those 65+ labeled as retired. Employment is either in an office environment in which co-workers only interact with each other, a K-12 school or college campus in which co-workers interact with each other and students, or a retail establishment in which co-workers interact with each other and shoppers. The distribution of office and retail establishment size is chosen to match data from the Statistics of U.S. Businesses (SUSB)[6], which reports the distribution of establishments by number of workers. The staff/student ratio for K-12 and college are chosen to match data from NCES. Retail establishments are further divided into Retail-A and Retail-B, where some workers are employed at Retail-A during the day and some workers are employed at Retail-B during the evening. Everybody is thus allocated to one of these mutually exclusive types: office

[2]The model is more fully presented in Appendix A.
[3]All data are more fully presented in Appendix B.
[4]https://www.census.gov/programs-surveys/acs
[5]https://nces.ed.gov/
[6]https://www.census.gov/programs-surveys/susb.html

worker, K-12 worker, college worker, retail-A worker, retail-B worker, K-12 student, college student, non-worker 18-65, pre-school 1-5, and retired.

A day is divided into three periods[7] and each person is given a daily planned location profile that depends upon their type as listed in Table 1. A person's home, school, or work location is the same for each day, but a person's shopping location is different for each day of the week. A person's weekly planned location profile repeats itself from one week to the next. The allocation of establishments between Retail-A and Retail-B is chosen to equate the shopper/worker ratio between these two.

Table 1: Planned Location Profile

Person Type	**period 1 (Night)**	**period 2 (Day)**	**period 3 (Evening)**
office/school/retail-A worker	home	work	shopping
retail-B worker	home	shopping	work
student K-12/College	home	school	home
non-worker 18-65	home	home	shopping
pre-school 1-5/retired	home	home	home

As in a standard variant of the SIR model, people are either *susceptible, exposed, infectious*, or *recovered*,[8] with a random transition time from one compartment to another. Infection of a susceptible person from the virus at a particular location depends on the fraction of infectious (contagious) people at that location. Let I/N denote the fraction of infectious people at a location, and denote the probability of becoming exposed during a period at that location by $(\beta/3)(I/N)$. Denote the probability of exiting an exposed/latent phase in each period by $\alpha/3$ and the probability of exiting an infectious phase in each period by $\gamma/3$. There is of course tremendous uncertainty regarding parameter values, and as an economist it is far beyond my expertise to add to this debate, so I will simply use the parameter estimates in a recent paper of Li, et. al. (2020). Li, et. al., differentiate between a transmission rate for documented and undocumented infections, with an average transmission rate during the initial phase of infection spread of 0.6866, which is the value I will use. The parameter values are listed in Table 2.

One other parameter that is especially important regarding any mitigation policy

[7]The parameter $p = 3$ in the notation in Appendix A.

[8]The model described in Appendix A distinguishes between a recovered and death state, but to focus only on the spread of the disease, the simulations include only a recovered state.

Table 2: Parameter Values

Parameter	Value
β (determines prob. exposed)	0.6866
α (prob. exit exposed phase)	0.2710
γ (prob. exit infectious phase)	0.2882

is, given infection, the likelihood of being asymptomatic. It seems rather difficult to know what fraction of infected people are asymptomatic, with estimates of ranging from 20 to 80 percent. In the simulations I will consider a broad range of values, but I will think of a benchmark value of 50 percent.

Define a minimal, no-intervention policy response to the spread of infection as one in which everybody follows their planned schedule. A mitigation policy is a “directive” that people stay home depending on their health status and whether or not the number infected exceeds a threshold. A compliance parameter determines the fraction of people that follow the directive.

3 Simulation Results

Following an index case, the simulation tracks how people move through various locations and change their health status. To begin, I simulate an SIR model for a large city in which everybody is located in one large location. I then simulate an eSIR model with a mitigation policy that a symptomatic person should stay home until recovered. Also considered is the dependence of the eSIR simulation on assumptions regarding the compliance to the mitigation policy, a new-case threshold for the implementation of the mitigation policy, and the fraction of asymptomatic people that are infected with the virus.

3.1 Compliance

Figure 1, Panel A, shows the spread of new infections for the standard SIR model as well as the eSIR model with a mitigation policy that all symptomatic people stay home, but with four different levels of compliance. The new-infection curve for the standard SIR model exhibits a rather sharp peak, in part because I use the parameter values that Li, et. al., estimate for the initial spread of the virus. Under the minimal, no-intervention policy (also 0 compliance), the new-infection curve is significantly flattened and pushed out to the future.

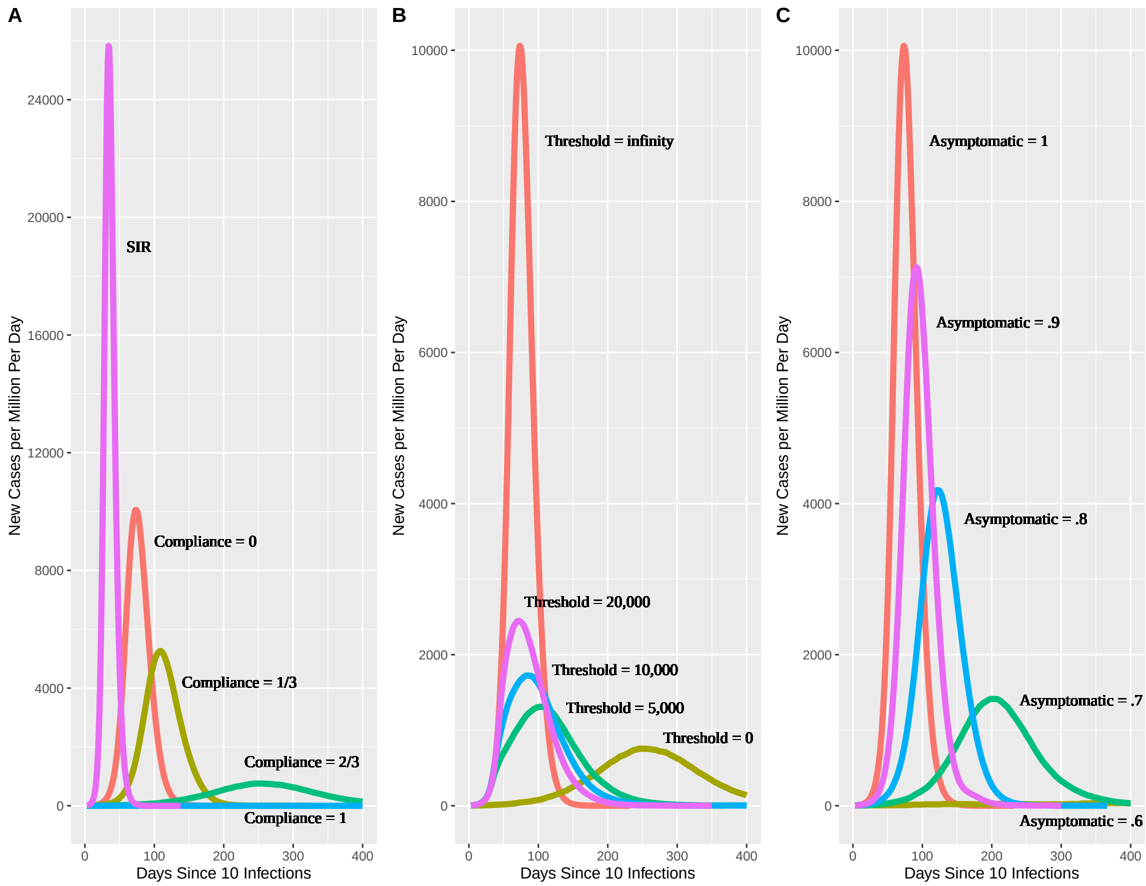

Figure 1: eSIR Simulation of New Cases. Mitigation policy: stay home if symptomatic. Compliance = fraction in compliance with mitigation policy. Threshold = new-case threshold for mitigation policy implementation. Asymptomatic = fraction asymptomatic. Panel **A**: Threshold = 0, Asymptomatic = .5, SIR = one large location. Panel **B**: Compliance = 2/3, Asymptomatic = .5. Panel **C**: Compliance = 1, Threshold = 0. All data are averages of 100 simulations, each starting with 10 infections chosen at random, with the resulting time series presented as a 5-day centered moving average.

The curve with 1/3 compliance is flattened even more so, and with 2/3 compliance yet even more, and the peak of 2/3 compliance is shifted to the future by about 200 days relative to the minimal policy. With 100 percent compliance to the mitigation policy, the disease does not become widespread so the new infection curve is essentially flat.

3.2 Mitigation Policy Threshold

Figure 1, Panel B, displays simulations of the model with a mitigation policy that does not go into effect until a certain number of cases are revealed. As shown in this figure, for 2/3 adherence to the mitigation policy, if the mitigation policy is implemented below a threshold of 5,000 cases, the peak of the new-infection curve does not rise by much, although the peak is shifted earlier by about 150 days. For thresholds higher than 5,000 the peak begins to rise sharply, but the shifting forward in time is much less. However, even at a threshold of 20,000, the peak is significantly flattened relative to the minimal, no-intervention policy.

3.3 Fraction Asymptomatic

Figure 1, Panel C, displays the model with 100 percent compliance, but where the fraction of asymptomatic infections ranges from .6 to 1 of total infections. As shown in this figure, even if up to 60 percent of infectious people are asymptomatic and thus are not able to follow the directive of staying home, the disease does not become widespread.

Table 3: R_0 for SIR and eSIR for various levels of compliance to a mitigation policy directing symptomatic people to stay home, assuming 50 percent asymptomatic, and the mitigation policy implemented immediately. Results are averages of 10,000 simulations, each starting with one index case chosen at random.

	Fraction		eSIR: Compliance			
Person Type	**of Pop.**	**SIR**	**0**	**1/3**	**2/3**	**1**
office worker	0.240	2.38	3.35	2.50	1.72	1.32
retail-A worker	0.065	2.38	3.04	2.76	1.94	1.59
retail-B worker	0.167	2.38	3.51	2.65	1.90	1.42
K-12 worker	0.011	2.38	1.61	1.38	1.52	0.79
college worker	0.007	2.38	1.40	1.35	1.20	1.00
K-12 student	0.167	2.38	2.38	1.83	1.61	1.37
college student	0.034	2.38	1.76	2.08	1.21	1.34
non-worker 18-65	0.124	2.38	2.01	1.61	1.24	0.93
pre-school 1-5	0.079	2.38	1.27	1.23	1.08	0.88
retired	0.107	2.38	0.60	0.45	0.44	0.42
weighted average		2.38	2.49	1.97	1.48	1.18

Table 4: Fraction of new infections by location for various levels of compliance to a mitigation policy directing symptomatic people to stay home, assuming 50 percent asymptomatic, and the mitigation policy implemented immediately. Results are averages of 100 simulations, each starting with 10 infections chosen at random.

		eSIR: Compliance			
Location Type	**SIR**	**0**	**1/3**	**2/3**	**1**
home		0.30	0.36	0.44	0.52
office		0.12	0.12	0.09	0.07
school		0.11	0.10	0.11	0.09
retail		0.47	0.43	0.36	0.31
all (frac of pop)	0.84	0.46	0.31	0.06	0.00006

3.4 R_0 and Infection Location

For the compliance levels reported in Fig. 1, Panel A, Table 3 reports R_0 for various categories of people, as well as for the SIR model. R_0 is computed from simulations by assigning as *infectious* one person chosen from the population at random, and tracking the number of infections caused by this person. If a person became exposed at a location with the index case and $n-1$ other infected people present, then the index case is recorded as infecting $1/n$ people. As seen from this table, R_0 is not generally lower for the no-compliance eSIR model than for the SIR model with the same parameter values, and on average is higher. Evidently, the disease initially spreads faster in the no-compliance eSIR model, but is more contained due to the spatial compartmentalization of that model. For the eSIR model, R_0 differs significantly across types of people, with office and retail workers having the highest R_0, and pre-school children and retired people having the lowest. As expected, R_0 generally falls as compliance rises. For 100 percent compliance, R_0 is on average slightly above 1, which is sufficient in the eSIR model to prevent the widespread of the disease as exhibited in Fig. 1.

Also for the compliance levels reported in Fig. 1, Panel A, Table 4 reports the location of new infections. Without a mitigation policy, most infections occur in a retail location, with the second most at home. With an increasing compliance to a mitigation policy that symptomatic people stay home, as expected an increasing number of infections occur at home. With 2/3 compliance to the mitigation policy, most infections occur at home. This table also reports the total fraction of infections for the course of the disease. For the SIR model, about 84 percent of the population eventually become infected. The total fraction infected is only about half that at 46 percent for the eSIR model with no mitigation policy. With 100 percent adherence to the mitigation policy, the total number of cases becomes very small. For the U.S. with a population of roughly 330 million, the total number of cases would be about 20,000.

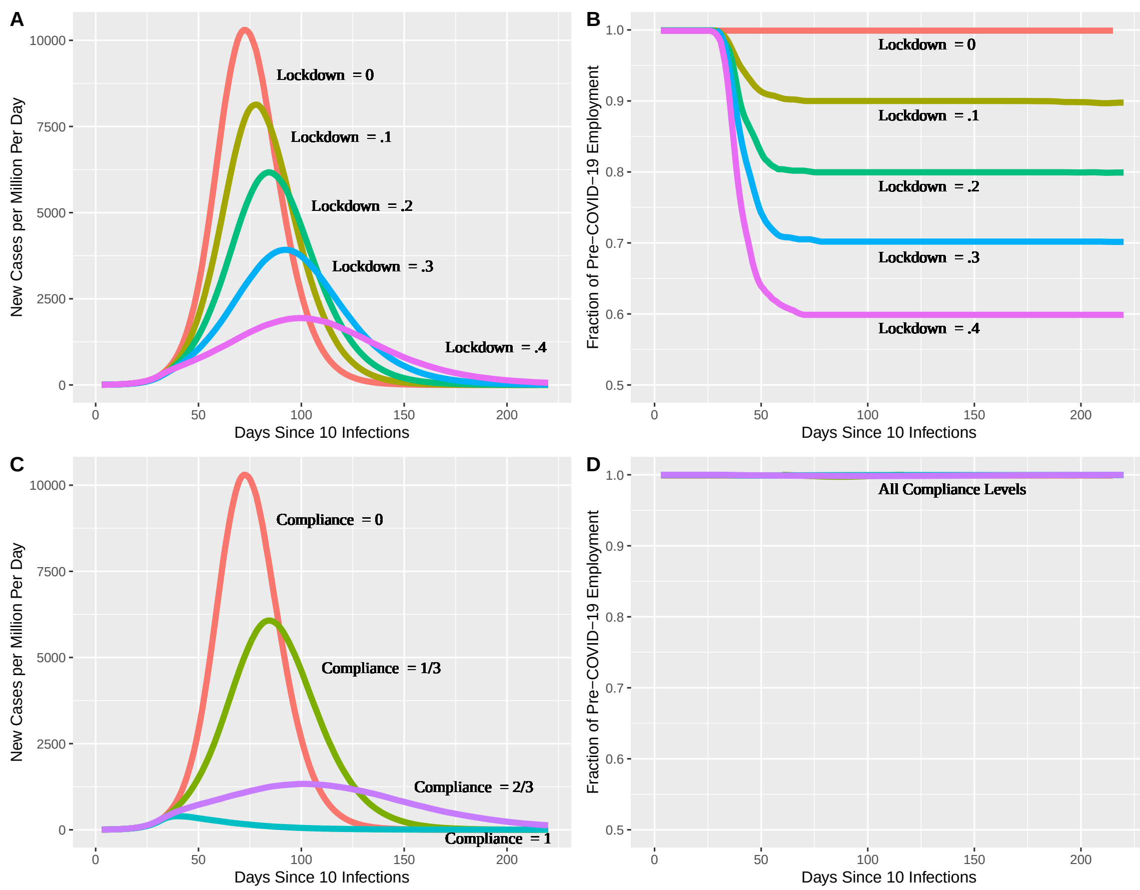

Figure 2: eSIR Simulation with Lockdown. Mitigation policy: stay home if symptomatic versus indiscriminate lockdown. All panels: 5,000 new-case threshold for mitigation policy implementation and .5 of cases asymptomatic. Lockdown = fraction of all establishments closed and people who would otherwise visit those establishments staying home during the planned time of visit. Compliance = fraction in compliance with stay-home-if-symptomatic mitigation policy. For the indicated mitigation policy, Panels **A** and **C** report daily new cases and Panels **B** and **C** report fraction of pre-COVID-19 employment. All data are averages of 100 simulations, each starting with 10 infections chosen at random, with the resulting time series presented as a 5-day centered moving average.

3.5 Comparison to Lockdown Policies

Fig. 2 compares the mitigation policy of directing symptomatic people to stay home versus directing a fraction of all non-home locations to shutdown and people who would otherwise visit those locations to stay home. The lockdown policy is indiscriminate and randomly shuts down a fraction of all non-home locations, including office and retail establishments. For various lockdown fractions, Panel **A** reports daily new cases and Panel **B** reports fraction of pre-COVID-19 employment, and similarly for various compliance fractions, Panel **C** reports daily new cases and Panel **D** reports fraction of pre-COVID-19 employment. As can be seen, a partial lockdown policy is able to obtain a significant reduction in new cases, but with a severe effect on reducing employment, whereas directing only symptomatic people to stay home can also obtain a significant reduction in new cases, but with a negligible effect on employment.

3.6 Reducing the Probability of Infection

The eSIR model can also be used to simulate the effects of other policies, such as the wearing of masks or social distancing, to reduce the probability of infection. Fig. 3 displays the reduction in new cases due to a reduction in the probability of infection. A 30 percent reduction in the probability of infection, if achievable through the wearing of masks or other measures, seems able to prevent the widespread of COVID-19. In this regard, note that Li, et. al. (2020) estimate an average reduction in the probability of transmission from an early infection phase to a later infection phase of 59 percent (from .6866 to .3900). It seems difficult, though, to breakdown this reduction into a part due to mask-wearing and other social distance measures and a part due to only symptomatic people staying at home.

4 Summary

For the spread of a disease such as COVID-19, the particular way in which we move about between home, work, school, and shopping matters. Especially important is the home location, not only because it captures persistent interaction with a limited number of people for a significant fraction of the day, but because it plays a crucial role in a quarantine-based mitigation policy. Simulations reveal that a mitigation policy in which all symptomatic people stay home is sufficient to control the spread of COVID-19, which is a result that holds even if up to 60 percent of the population is asymptomatic, and as well even if other family

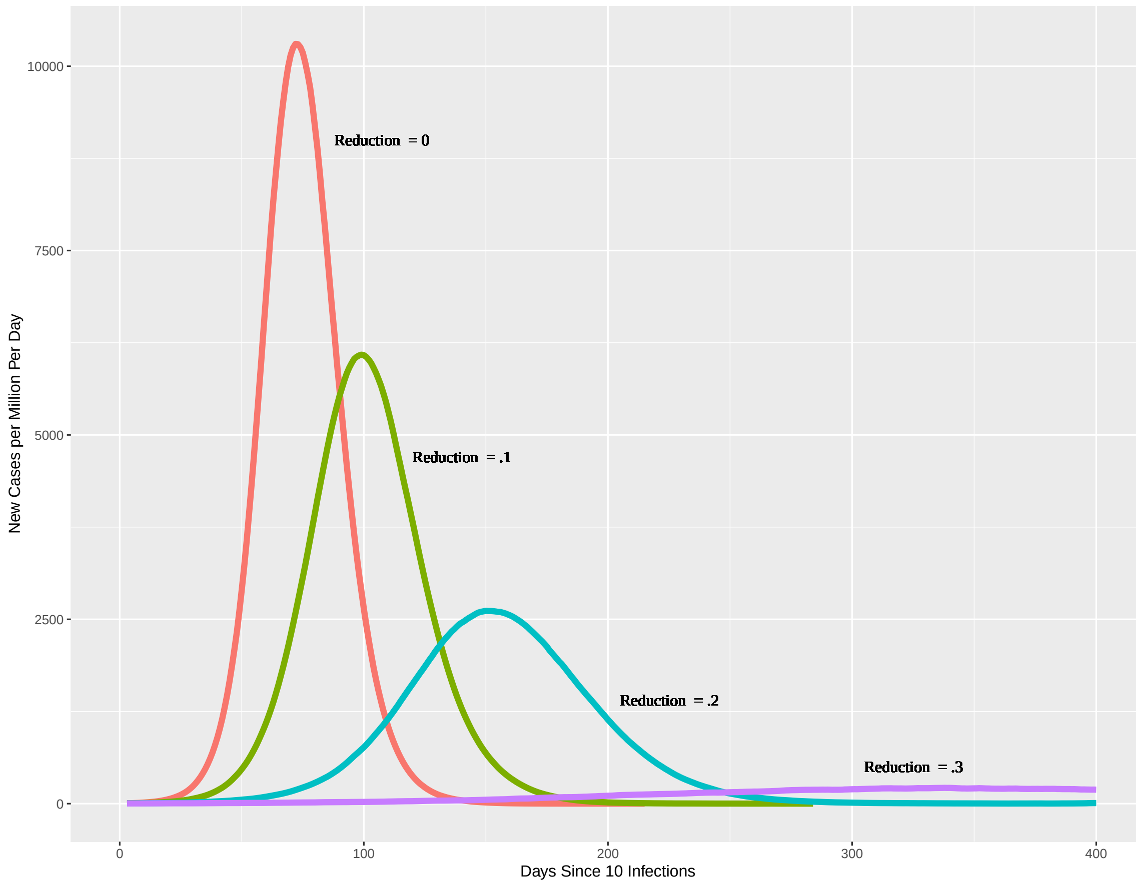

Figure 3: eSIR Simulation and Infection Probability. Mitigation policy: mask wearing and social distancing. 5,000 new-case threshold for mitigation policy implementation. Reduction = fraction reduction of infection probability (new probability = (1 - Reduction)*β). All data are averages of 100 simulations, each starting with 10 infections chosen at random, with the resulting time series presented as a 5-day centered moving average.

members are free to go about their business. This seems important for three reasons. One, it may explain the better success of some countries over others in managing the spread of COVID-19. Two, such a mitigation policy avoids the devastating economic consequences of a more severe mitigation policy. And three, it highlights the immense value to any policy, such as the dissemination of information on symptoms and/or widespread testing, that encourages infected people to stay home, or otherwise reduce the probability of infection.

A Appendix: eSIR Model Details

Divide time into equal length intervals, indexed by $t = 0, 1, 2, ...,$ and suppose a day is comprised of p intervals. At the beginning of time t there are N_t people, indexed by $i = 1, ..., N_t$, who have every lived up to time t. Person i has characteristic a_i from a set of possible characteristics A, such as gender or period of birth (age).[9] The economy is comprised of $L+1$ locations, indexed by $\ell = 0, ..., L$. During each time interval, a person is located at one and only one location, and over time a person potentially visits different locations. Denote the location of person i during time t by ℓ_{it}. For any ℓ such that $\ell \in \{0, 1, ..., L\}$, define

$$\chi_\ell(\ell_{it}) = 1 \text{ if } \ell_{it} = \ell,$$

and zero otherwise, so the number of people at location ℓ during time t is

$$N_{\ell t} = \sum_{i=1}^{N_t} \chi_\ell(\ell_{it}).$$

Let the zero location $\ell = 0$ be an absorbing state of death. Events during a period potentially lead to a person dying at the end of a period, hence entering a location of death at the beginning of the next period.

Newborns can add to the population that potentially increases N_t over time. Denote the number of newborns during time t by B_t and associate with each newborn an index i and a location ℓ_{it}, so the population at the beginning of any time $t > 1$ equals

$$N_t = N_{t-1} + B_{t-1} - \sum_{i=1}^{N_{t-1}+B_{t-1}} \chi_0(\ell_{it}).$$

Including newborns in the death location $\ell = 0$ admits the possibility that some newborns may die from the disease.

As in a standard variant of the SIR model, people are either *susceptible, exposed, infectious, recovered*, or *dead*, denoted by health status h. A susceptible person is healthy with a positive probability of becoming exposed if coming into contact with an infectious person. An exposed person is in a latent phase, which means the person is not contagious and does not exhibit symptoms, and has a constant probability of exiting a latent phase

[9] Denote the period of birth by b_i, so age at time t equals $t - b_i$. For the initial population N_0, assume the period of birth is a negative number that represents initial age.

to an infectious phase during each time interval, say $\alpha(t, a_i)/p$.[10] An infectious person is contagious and may or may not exhibit symptoms and has a constant probability of recovery during each time interval, say $\gamma(t, a_i)/p$, and if not recovered, a constant probability of dying during each time interval, say $\mu(t, a_i)/p$.[11] A recovered person is assumed to be immune from re-infection. The dependence on t and a_i can capture heterogeneity such as a dependence on age, or that only some people suffer severe illness that potentially leads to death. Aside from capturing one cost of the disease, including the state of death in the analysis allows for a mitigation policy to depend on the more readily measurable variable, especially early in an epidemic, that counts the number of fatalities.

Infection of a susceptible person from the virus at a particular location depends on the fraction of infectious people at that location. Let $I_{\ell t}$ denote the number of infectious people at location $\ell \in \{1, 2, ..., L\}$ during time t, given by

$$I_{\ell t} = \sum_{i=1}^{N_t} \chi_\ell(\ell_{it}) \chi_{infectious}(h_{it}),$$

where $\chi_{infectious}(h_{it}) = 1$ if $h_{it} = infectious$ and zero otherwise. The probability of infection at location ℓ during time t for a susceptible person is presumed to be, as in the SIR model,

$$\frac{\beta(t, a_i)}{p} \frac{I_{\ell t}}{N_{\ell t}}.$$

A person can be at a location either as a worker or not, from which can be derived three different types of locations. A location with only workers can be thought of as a business that does not interact directly with customers (e.g., an office), a location with both workers and non-workers can be thought of a business that interacts directly with customers (e.g., a retail establishment or school), and a location with only non-workers can be thought of as a place in which people meet but no work is performed (e.g., home). Denote work status as d, where $d = 0$ means not a worker and $d = 1$ means a worker.

A realization of an economy is thus a list of locations $\{0, ..., L\}$, a sequence of populations $\{N_t\}$ with characteristics $\{a_i\}$, and a sequence of locations $\{\ell_{it}\}$, health status $\{h_{it}\}$, and work status $\{d_{it}\}$.

[10]For comparison purposes, I will scale all probabilities by p, which is motivated by an approximation such as $dS/dt \approx p(S_{t+\frac{1}{p}} - S_t)$, where t is in units of days.

[11]Simulations of the model do not include a death state, so $\mu(t, a_i) = 0$.

B Appendix: People and Locations

I will think of an economy as a large city in the U.S., with the distribution of people and locations estimated to match statistics for the entire country. People begin the day from home, with the distribution of homes by number and age of people estimated to match that of the entire U.S. The important issue regarding a home, and other locations as well, is that at various intervals of time the same people get together in a home. During the day, people go to work, school, or visit retail establishments, and at the end of the day return home. At work people interact with the same co-workers each day and at school students interact with the same fellow students. In addition, workers at school interact with students, and workers in retail establishments interact with shoppers. The mixing of family members, co-workers, students and shoppers potentially exposes people to each other throughout the city.

B.1 Households and Homes

According to the Current Population Survey (CPS),[12] in 2016 there are about 125,819,000 households in the U.S. Table 5 list the number of households broken down by household size. Based on this table, I will only consider households up to 7 members, as households up to 6 members already includes 98.7 percent of all households.

Table 5: U.S Households by Size in 2016 (000's)

Size	**Total**	**Percent**
One member	35,388	28.1
Two members	42,785	34.0
Three members	19,423	15.4
Four member	16,267	12.9
Five members	7,548	6.0
Six members	2,813	2.2
Seven or more members	1,596	1.3
All households	125,819	100

The American Community Survey (ACS)[13] 2010 dataset provides a sample that contains the age of each household member, which can be used to estimate the number of households by the specific age composition of its members. Suppose there are $G = 5$ age

[12] https://www.census.gov/programs-surveys/cps.html

[13] https://www.census.gov/programs-surveys/acs

categories: under 6, 6 - 17, 18 - 22, 23 - 64, 65+. Suppose each household must have a member greater than age category g_{min}, where for ACS $g_{min} = 2$ (i.e., a head of household must be at least 18 years old). Let a type of household be a composition of members differentiated by age category. The number of distinct types of households that can be formed that have up to n_{max} members is thus

$$\sum_{n=1}^{n_{max}} \left[\binom{n+G-1}{n} - \binom{n+g_{min}-1}{n} \right].$$

With $n_{max} = 7$, there are 756 distinct types of households in the U.S. From ACS, I computed the frequency of each, and I assume that a large city has a distribution of households similar to that of the entire U.S. The top 15 types of households are listed in Table 6, which represents 79 percent of all types of households. Given the total population of a city, this determines the number of homes, the distribution of homes by the number and age of residents, and the demographic composition of the city population by age.

Table 6: Top 15 Households in U.S. in 2010 by Age of Members (000's)

age category of household member, oldest to youngest							**Percent of**
first	**second**	**third**	**fourth**	**fifth**	**sixth**	**seventh**	**all households**
23-64	23-64						18%
23-64							17%
65+							10%
65+	65+						7%
65+	23-64						4%
23-64	23-64	6-17	6-17				4%
23-64	23-64	6-17					4%
23-64	23-64	under 6					3%
23-64	23-64	23-64					2%
23-64	23-64	6-17	under 6				2%
23-64	6-17						2%
23-64	23-64	18-22					2%
23-64	23-64	under 6	under 6				2%
23-64	18-22						1%
23-64	23-64	6-17	6-17	6-17			1%

B.2 Employment and Schooling

According to the Bureau of Labor Statistics (BLS), in 2018 the labor force participation rate amongst adults 25-54 was 82.1 percent and adults 55-64 was 65.0 percent.[14] I will thus assume that 78 percent of all adults 23-64, the age category used in this paper, are employed (for this exercise, I assume full employment).[15] Also, BLS estimates that 27 percent of adults 65-74 and 8.7 percent of adults 75+ are in the labor force, so I will assume that 19 percent of adults 65+ are employed.[16]

I assume all people in age category 6-17 are in school. According to the National Center for Education Statistics (NCES),[17] in 2018 the fraction of high school graduates that attended some form of college was 69 percent and the fraction of high school students that graduated high school (with their cohort) was 85 percent, so I assume 55 percent of people in the age category 18-22 category are in college, while the remainder are in the labor force (assume to be employed). Given the city population size and age distribution, this determines the total number of employees, as well as the number students in K-12 and college.

B.3 K-12 and College

All schooling data are obtained from NCES for the 2016-2017 school year. The average public school size for K-12 was 528 students.[18] There were about 15 students per staff at a pubic school,[19] which translates into about 35 staff per school. There were about 4,726 post-secondary degree-granting institutions in the U.S., with about 20 million students, which translates into about 4,232 students per school.[20] There were about 3,920,836 staff at these schools, which translates into 830 staff per school.[21] These data are summarized in Table 7. Given the number of students in K-12 and college, this determines the number of K-12 schools along with the number of K-12 workers, as well as the number of colleges and the number of college workers.

[14] `https://www.bls.gov/emp/tables/civilian-labor-force-participation-rate.htm`

[15] I used a weighted average using population estimates in 2018 for 23-54 and 55-64, assuming labor force participation for 23-24 that are not in school is the same as for 25-54.

[16] A weighted average based on 2018 population estimates.

[17] `https://nces.ed.gov/`

[18] `https://nces.ed.gov/programs/digest/d18/tables/dt18_214.40.asp`

[19] `https://nces.ed.gov/pubs2007/ruraled/tables/table3_12.asp`

[20] `https://nces.ed.gov/programs/digest/d17/tables/dt17_317.10.asp?current=yes` and `https://nces.ed.gov/programs/digest/d17/tables/dt17_303.20.asp`

[21] `https://nces.ed.gov/pubs2012/2012156rev.pdf`

Table 7: U.S. School Size in 2016-2017

Type	Students per School	Staff per School
primary/secondary	528	35
post-secondary	4,232	830

B.4 Work and Establishments

The Statistics of U.S. Businesses (SUSB)[22] by the U.S. Census reports annual data on establishments by number of employees. As described by the U.S. Census, "An establishment is a single physical location at which business is conducted or services or industrial operations are performed."[23] The data are reported in Table 8.[24] Given the number of employees in a city net of those that work in a K-12 school or college, this determines the number of business of various sizes.

Table 8: Establishment Employment in 2017Q1

Establish. Empl. Class	Number of Establish.	Total Employment	Employment per Estab.	Freq. Establish.
< 5	6,171,807	8,642,604	2*	.6290
5 - 9	1,453,674	9,664,903	7	.1482
10 - 19	1,031,800	13,994,592	14	.1052
20 - 49	725,722	21,859,357	30	.0740
50 - 99	240,881	16,527,937	69	.0246
100 - 249	135,496	20,285,252	150	.0138
250 - 499	34,114	11,568,064	339	.0035
500 - 999	11,756	8,018,994	682	.0012
1000+	6,328	13,907,553	2198	.00006

*Decimal value = 1.4, but 2 is used to capture co-worker interactions.

B.5 Employment, Occupations, and Customers

The Occupational Employment Statistics[25] program of BLS reports the number of employees by major occupation group. These data for 2018 are summarized in Table 9. I categorized

[22] https://www.census.gov/programs-surveys/susb.html
[23] https://www.census.gov/programs-surveys/susb/about/glossary.html
[24] source: https://data.bls.gov/cew/apps/data_views/data_views.htm#tab=Tables
[25] https://www.bls.gov/oes/

each group as being either an office worker in which there is no or minimal interaction with customers and a retail worker in which there is likely to be significant interaction with customers. Based on this breakdown, about 49 percent of employment are in the office category, with the remaining 51 percent in the retail category. Given the overall number of employees that work in a business, this determines the breakdown of business type into an office or retail establishment.

Table 9: Occupation by Employment and Customers

Occupation	Employment (000's)	Customers
Management	10,193.3	no
Business and financial operations	8,589.5	no
Computer and mathematical	4,674.4	no
Architecture and engineering	2,699.1	no
Life, physical, and social science	1,322.7	no
Healthcare support	4,315.6	no
Building and grounds cleaning and maintenance	5,756.9	no
Office and administrative support	22,973.6	no
Farming, fishing, and forestry	1,096.9	no
Construction and extraction	7,212.5	no
Production	9,452.4	no
Community and social service	2,730.1	yes
Legal	1,346.0	yes
Education, training, and library	9,647.4	yes
Arts, design, entertainment, sports, and media	2,899.8	yes
Healthcare practitioners and technical	9,107.8	yes
Protective service	3,560.9	yes
Food preparation and serving related	13,664.0	yes
Personal care and service	7,117.2	yes
Sales and related	15,728.6	yes
Installation, maintenance, and repair	6,108.1	yes
Transportation and material moving	10,841.0	yes
fraction no customer	0.49	
fraction customer	0.51	

B.6 Random Allocation

Households are created in proportion to the observed frequency of households by number and age of occupants. The total number of households are created to match a given population, equal to one million in the simulations. Each person in age group 18-22 is randomly allocated to being either a college student or worker, with probability of being a student equal to the observed fraction of this age group that are students. Each person in age group 23-64 is randomly allocated to being either in the labor force or a non-worker, with probability of being in the labor force equal to the observed labor force participation rate of this age group. Each person in age group 65+ is randomly allocated to being either in the labor force or retired, with probability of being in the labor force equal to the observed labor force participation rate of this age group. Each person allocated to being a worker is randomly allocated to being an office worker or non-office worker, with the probability of being an office worker equal to the observed fraction of office workers. Each person in age group 6-17 is randomly allocated to a K-12 school, and each person labeled as a college student is randomly allocated to a college, where the number of K-12 schools and colleges are chosen to match data on student size. Office workers are randomly allocated to offices of various sizes to match the observed distributed of establishment sizes. Non-office workers are randomly allocated to K-12 schools, colleges, retail-A establishments and retail-B establishments of various sizes to match staff-student ratios at the schools and to match the observed distributed of establishment sizes. Shopping in retail-A locations are chosen randomly with probability equal to the number of workers as a fraction of total retail-A workers, and similarly for shopping in retail-B locations.

References

[1] Acemoglu, Daron, Victor Chernozhukov, Ivan Werning, Michael D. Whinston, "Optimal Targeted Lockdowns in a Multi-Group SIR Model," NBER Working Paper 27102, 2020.

[2] Alvarez, Fernando, David Argente, and Francesco Lippi, "A Simple Planning Problem for COVID-19 Lockdown," NBER Working Paper 26981, 2020.

[3] Atkeson, Andrew "How Deadly Is COVID-19? Understanding The Difficulties With Estimation Of Its Fatality Rate," Working Paper 26965, National Bureau of Economic Research, 2020.

[4] ——, "What will be the Economic Impact of COVID-19 in the US? Rough Estiamtes of Disease Scenarios," NBER Working Paper 26867, 2020.

[5] Bayham, Jude, Nicolai V. Kuminoff, Quentin Gunn, and Eli P. Fenichel, "Measured voluntary avoidance behaviour during the 2009 A/H1N1 epidemic," *Proceedings of the Royal Society B: Biological Sciences*, 282 (1818), 20150814, 2015

[6] Chowdhury, Rajiv Chowdhury, Kevin Heng, Md Shajedur Rahman Shawon, Gabriel Goh, Daisy Okonofua, Carolina Ochoa?Rosales, Valentina Gonzalez-Jaramillo, Abbas Bhuiya, Daniel Reidpath, Shamini Prathapan11, Sara Shahzad, Christian L. Althaus, Nathalia Gonzalez?Jaramillo, Oscar H. Franco, "Dynamic interventions to control COVID-19 pandemic: a multivariate prediction modelling study comparing 16 worldwide countries," *European Journal of Epidemiology*, 35:389-399, 2020. https://doi.org/10.1007/s10654-020-00649-w.

[7] Dickison, M., S. Havlin, and H. E. Stanley, "Epidemics on interconnected networks." *Phys. Rev. E.* 85, 66109, 2012.

[8] Eichenbaum, Martin, Sergio Rebelo, and Mathias Trabandt, "The Macroeconomics of Epidemics," Working Paper 26882, National Bureau of Economic Research, 2020.

[9] Farboodi, Maryam, Gregor Jarosch, and Robert Shimer, "Internal and External Effects of Social Distancing in a Pandemic," Working Paper 27059, National Bureau of Economic Research, 2020.

[10] Ferguson, Neil M, Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Zulma Cucunub, Gina Cuomo-Dannenburg, Amy Dighe, Ilaria Dorigatti, Han Fu, Katy Gaythorpe, Will Green, Arran Hamlet, Wes Hinsley, Lucy C Okell, Sabine van Elsland, Hayley Thompson, Robert Verity, Erik Volz, Haowei Wang, Yuanrong Wang, Patrick GT Walker, Caroline Walters, Peter Winskill, Charles Whittaker, Christl A Donnelly, Steven Riley, Azra C Ghani, "Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand." Imperial College COVID-19 Response Team, 2020.

[11] Fernandez-Recio, Juan, "Modelling the evolution of COVID-19 in high-incidence European countries and regions: estimated number of infections and impact of past and future intervention measures," *J. Clin. Med.*, 9(6), 1825, 2020. https://doi.org/10.3390/jcm9061825.

[12] Garriga, Carlos, Rody Manuelli, and Siddhartha Sanghi, "Optimal Management of an Epidemic: An Application to COVID-19. A Progress Report," April 2020, mimeo.

[13] Gevertz, Jana, James Greene, Cynthia Hixahuary Sanchez Tapia, Eduardo D Sontag, "A novel COVID-19 epidemiological model with explicit susceptible and asymptomatic isolation compartments reveals unexpected consequences of timing social distancing," *medRxiv*, 2020.05.11.20098335, 2020. doi: https://doi.org/10.1101/2020.05.11.20098335.

[14] Jones, Callum J, Thomas Philippon, and Venky Venkateswaran, "Optimal Mitigation Policies in a Pandemic: Social Distancing and Working from Home," Working Paper 26984, National Bureau of Economic Research April 2020.

[15] Jones, Charles, I. And Jesus Fernandez-Villaverde, "Estimating and Simulating a SIRD Model of COVID-19 for Many Countries, States, and Cities," working paper, Stanford University, 2020.

[16] Kermack, William Ogilvy and A. G. McKendrick A. G., "A contribution to the mathematical theory of epidemics," *Proc. R. Soc. Lond.* A 115: 700-721, 1927. http://doi.org/10.1098/rspa.1927.0118.

[17] Kucharski, Adam J., Petra Klepac, Andrew J. K. Conlan, Stephen M Kissler, Maria L Tang, Hannah Fry, Julia R Gog, and W. John Edmunds, "Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study," www.thelancet.com/infection, published online June 16, 2020. https://doi.org/10.1016/S1473-3099(20)30457-6.

[18] Rampini, Adriano A, "Sequential Lifting of COVID-19 Interventions with Population Heterogeneity," Working Paper 27063, National Bureau of Economic Research, 2020.

[19] Rowthorn, Robert and Flavio Toxvaerd, "The Optimal Control of Infectious Diseases via Prevention and Treatment," Technical Report 2013, Cambridge-INET Working Paper, 2020.

[20] Ruiyun Li, Sen Pei, Bin Chen, Yimeng Song, Tao Zhang, Wan Yang, Jeffrey Shaman, "Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2)," *Science*, 489-493, 2020.

[21] Sofonea, Mircea T., Bastien Reyn, Baptiste Elie, Ramss Djidjou-Demasse, Christian Selinger, Yannis Michalakis, Samuel Alizon, "Epidemiological monitoring and control perspectives: application of a parsimonious modelling framework to

the COVID-19 dynamics in France," *medRxiv*, 2020.05.22.20110593, 2020. doi: https://doi.org/10.1101/2020.05.22.20110593.

[22] Stock, James H, "Data Gaps and the Policy Response to the Novel Coronavirus,? Working Paper 26902, National Bureau of Economic Research, 2020.

[23] Yagan, O., D. Qian, J. Zhang, J. and D. Cochran. "Conjoining Speeds up Information Diffusion in Overlaying Social-Physical Networks." *IEEE JSAC Special Issue on Network Science* 31, 1038, 2013.

[24] Zuzek, L., H. Stanley, and L. Braunstein, "Epidemic Model with Isolation in Multilayer Networks." *Sci Rep*, 5, 12151, 2015. https://doi.org/10.1038/srep12151