Epidemiological Models: Heterogeneity and Endogeneity

Glenn Ellison

Massachusetts Institute of Technology

June 29, 2021

Classic SIR Model

The SIR model (Kermack and McKendrick 1927) remains an elegant way to bring out many intuitions.

Many economics papers use it or slight variants as their pandemic model.

- Unit mass of agents are Susceptible, Infectious, or Recovered
- Assume fractions $S(t)$, $I(t)$, $R(t)$ evolve as:

$$\begin{aligned} \dot{I}(t) &= S(t)I(t)R_0\gamma - \gamma I(t) \\ \dot{R}(t) &= \gamma I(t) \\ \dot{S}(t) &= -S(t)I(t)R_0\gamma \end{aligned}$$

Motivation:

1. Infectious agents recover at Poisson rate γ. They are infectious for $1/\gamma$ periods on average.
2. Infectious agents interact in a manner that would transmit with prob. $R_0\gamma dt$ in each dt. Totals R_0 on average while infectious.

Properties of the SIR Model

1. R_0 is a critical parameter.

   When $R_0 < 1$ the infection-free state $(S, I, R) = (1, 0, 0)$ is a locally stable steady state. A small infection introduced will die out.

   When $R_0 > 1$ the $(1, 0, 0)$ steady state is unstable and a small infection will spread.

2. Growth rates

   The growth rate, $g(t) = \frac{d}{dt} \log(I(t))$, of the infectious is $g(t) = \gamma R_0 S(t) - \gamma$.

   Practically, people estimate R_0 early in a pandemic with a multistage approach: Estimate γR_0 from case/hospitalization growth rates, estimate γ from contact tracing data; divide to get R_0.

   Estimating the R_0 is of a new variant is harder because you need to separately estimate $S(0)$.

Properties of the SIR Model

3. Herd immunity thresholds

   Let $\overline{S}$ be such that $(S, 0, 0)$ is a stable steady state if and only if $S < \overline{S}$.

   The threshold is incredibly important. The effective R_0 can be temporarily reduced by limiting activity, wearing masks, quarantining, etc., but the infection will resume spreading once restrictions are removed if $S > \overline{S}$ (unless it has been completely eradicated).

   In the classic SIR model $\overline{S} = 1/R_0$.

   - For the original strain if $R_0 \approx 2.3$ this will be about 0.44.
   - The Boris Johnson (α) variant may have $R_0 \approx 3.5 \implies \overline{S} \approx 0.29$.
   - The Modi (δ) variant may have $R_0 \approx 5.5 \implies \overline{S} \approx 0.18$.

   One can think of vaccinations as reducing $S(t)$.

   SIR suggests we'll need to vaccinate (or infect) 80% with a 90% effective vaccine to reach herd immunity against the α variant.

Properties of the SIR Model

4. Overshooting and Eventual Infection

   An SIR model calibrated to COVID-19 predicts that an uncontrolled epidemic reaches herd immunity quickly.

   Many people (10-15%) are infectious when herd immunity is reached. This leads to substantial "overshooting".

   Write $S(\infty)$ for the fraction who escape infection. Eventually there will be $(1 - S(\infty))R^0$ interactions of an Infectious with another. The number of times an individual is hit is Poisson with mean $(1 - S(\infty))R^0$. This implies $S(\infty) = e^{-(1-S(\infty))R^0}$.

   The difference can be large. With R_0=2.3 we have $\overline{S} = 0.44$ and $S(\infty) \approx 0.14$.

   In several economic analyses policies are motivated by overshooting (and hospital capacity).

Limitations of the SIR Model

Epidemiologists modeling short-run dynamics typically augment the SIR model in many ways: adding a state for exposed but not yet infectious, adding a quarantine state, making recovery time distribution more flexible, etc. These are particularly important for the feasibility of controlling the epidemic via testing, contact tracing, quarantines, etc.

Two other concerns very important to the overall trajectory are:

1. Endogeneous Behavior

   The SIR model treats R_0 as a fixed primitive. In practice, people will alter their behavior in multiple ways when infection rates are high.

2. Heterogeneity

   In practice, contact rates vary dramatically depending on whether one rides public transportation, frequents bars, works in a jobs that involves close contact, lives in a crowded home, etc.

   Also, when thinking about new variants, it is as if contact rates as vary across vaccination statuses.

Heterogeneous SIR Model with Uniform Matching

Consider an SIR variant with N groups.

- Activity levels $R_{01}, R_{02}, \ldots, R_{0N}$ differ.
- Members of group i have contacts with prob. $R_{0i}\gamma dt$.
- i's contacts are in group j with probability $R_{0j}/\sum_k R_{0k}$.

This motivates a system of differential equations:

$$\begin{aligned}
\dot{I}_i(t) &= S_i(t)\sum_j \beta_{ij} I_j(t) - \gamma I_i(t) \\
\dot{S}_i(t) &= -S_i(t)\sum_j \beta_{ij} I_j(t) \\
\dot{R}_i(t) &= \gamma I_i(t)
\end{aligned}$$

with $\beta_{ij} \equiv \gamma R_{0i}\frac{R_{0j}}{\sum_k R_{0k}}$.

Properties of the Heterogeneous Uniform SIR Model

To explore whether small infections spread, linearize changes in infection rates in a neighborhood of $I = 0$ as $\dot{I}(t) \approx A^S I(t)$.

The largest eigenvalue of A^S is easy to find. It implies

1. Case Distribution

   The principal eigenvector is $v_1 = (S_1^0 R_{01}, \ldots, S_N^0 R_{0N})$. Soon after an epidemic starts cases will be distributed in these proportions.

2. Growth Rates

   Once cases align as above a small epidemic will grow at rate $\gamma\left(\frac{\sum_i S_i^0 R_{0i}^2}{\sum_i R_{0i}} - 1\right)$.

   Early growth will resemble that of a classic SIR model with parameter $\overline{R}_0 \equiv \frac{\sum_i R_{0i}^2}{\sum_i R_{0i}} = \sum_i \frac{R_{0i}}{\sum_k R_{0k}} R_{0i} = E(R_{0i}) + \frac{\text{Var}(R_{0i})}{E(R_{0i})}$.

   Early estimates of growth rates will have estimated this weighted average, not the arithmetic mean of the R_{0i}.

Heterogeneous SIR Model with Homophilic Matching

Many agents with high R_0 will naturally interact with others with high R_0: public transportation, bars, neighborhoods with crowded housing, COVID skeptics, etc.

Suppose the probability a contact of someone in group i is in group j is

$$p_{ij} = \begin{cases} h + (1-h)\frac{R_{0j}}{\sum_k R_{0k}} & \text{if } j = i \\ (1-h)\frac{R_{0j}}{\sum_k R_{0k}} & \text{if } j \neq i. \end{cases}$$

With total contacts from group i as above we get infection dynamics

$$\dot{I}_i(t) = S_i(t) \sum_j \beta_{ij}^h I_j(t) - \gamma I_i(t), \text{ with}$$

$$\beta_{ij}^h = \begin{cases} \gamma R_{0i}(h + (1-h)\frac{R_{0j}}{\sum_k R_{0k}}) & \text{if } j = i \\ \gamma R_{0i}(1-h)\frac{R_{0j}}{\sum_k R_{0k}} & \text{if } j \neq i. \end{cases}$$

The parameter $h \in [0,1]$ captures the degreee of homophily in matching.

Properties of the Heterogeneous Model with Homophily

The model is not as tractable as the uniform model, but one can derive several resuts about herd immunity:

1. Each population must be internally stable.
   If $S_i^0 h R_{0i} > 1$ for any i, then $(S^0, 0)$ is unstable.
2. Herd immunity again depends on something like a weighted average
   Otherwise, $(S^0, 0)$ is unstable if $\sum_i \frac{R_{0i}}{\sum_k R_{0k}} \frac{1}{1-hS_i^0 R_{0i}}(S_i^0 R_{0i} - 1) > 0$ and stable if $\sum_i \frac{R_{0i}}{\sum_k R_{0k}} \frac{1}{1-hS_i^0 R_{0i}}(S_i^0 R_{0i} - 1) < 0$.
3. Homophily shrinks the herd immunity region
   If a disease-free equilibrium $(S^0, 0)$ is unstable for some h it is unstable for all larger h. (Homophily matters when there is heterogeneity in contact rates.)

Implications: Herd Immunity

Many discuss COVID as homogenous SIR with parameter $\overline{R}_0$.

This leads one to expect a herd immunity threshold of $\hat{S} = 1/\overline{R}_0$.

For example, in the uniform model with $R_0 = (3.5, 1.5, 1, 0.5, 0.5)$, $\overline{R}_0 = 2.3$ and $\hat{S}_0 = 0.44$. This suggests we need 56% vaccinated.

1. Potentially large overestimation of required infections/vaccinations

   The uniform model can reach herd immunity with 14% vaccinated.

   With homophilic matching and $h \approx 1$ we need 21% vaccinated.

2. Vaccine targeting

   In the uniform model, herd immunity is always reached most quickly by vaccinating those with the highest R_{0i}.

   Vaccine allocation involving hassle costs is unfortunate.

   Also affects calculations when effectiveness is a choice variable.

   In the homophilic model, optimal targeting is less extreme.

Implications: Difficulties in Calibration and Forecasting

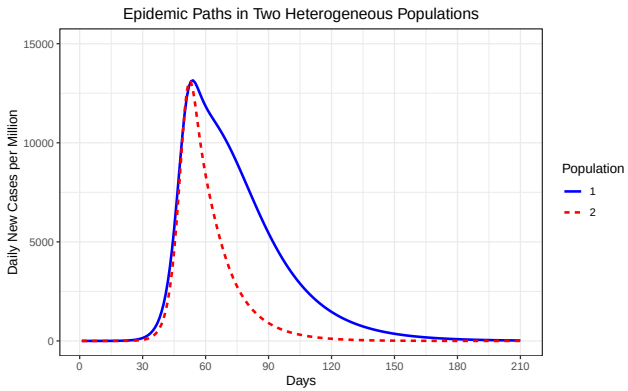
Predicting the course of an epidemic is inherently difficult.

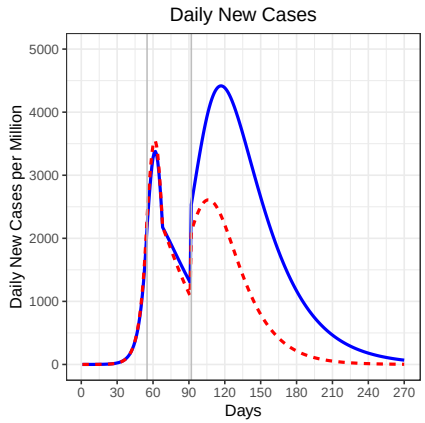
1. Estimating the parameters of a heterogeneous model is hard.

   Early growth rates let one estimate $\overline{R}_0$. Some have augmented with contact survey data, but these may not reflect COVID transmission. Parameters of low-activity groups will be particularly difficult.

2. Epidemic paths can depend on activity levels in low-activity groups

3. Effects of policy relaxations and new variants can also depend greatly on details of lower-activity groups.





Implications: Overshooting

The overshooting in the classic SIR model may not be practically relevant: people would isolate on their own if anything like 10% were currently infectious, so epidemics proceed more slowly.

Overshooting becomes more salient in heterogeneous models.

1. Extra dimension of overshooting

   Overshooting is a group-by-group phenomenon. Infections of the less active contribute little to herd immunity.

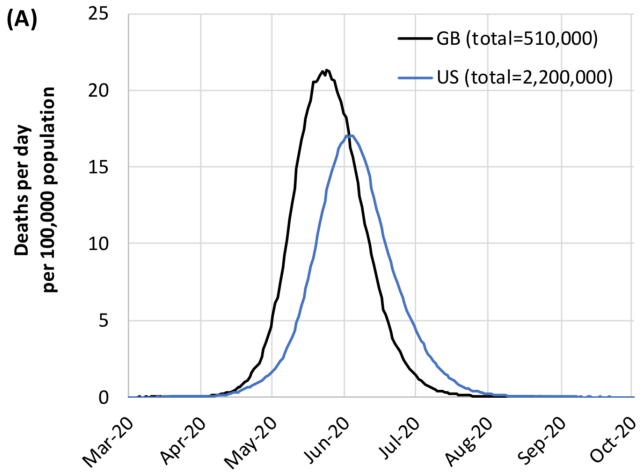
2. Temporary lockdowns

   Temporary lockdowns become more appealing when vaccination campaigns will have an effect more quickly.

   Uncontrolled epidemics initially grow in proportion to activity. The optimal path to herd immunity has even more concentrated cases. Targeted lockdowns can get us closer to this path.

Endogenous Interaction

The time path of the COVID pandemic has looked very little like the early forecasts of an uncontrolled epidemic.



It has been widely recognized that such dynamics are implausible when people have the option to self-isolate.

Economists and epidemiologists have approached this in two ways:

1. Reduced form responses
2. Rational activity levels

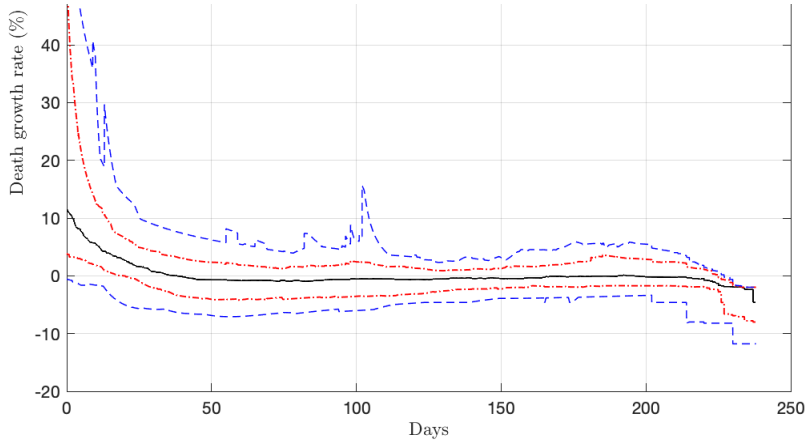
Endogenous Interaction: Reduced Form Responses

A classic approach is assume to to assume directly that distancing in response to concurrent death (or infection) rates leads to dynamics like

$$\dot{I}(t) = \frac{1}{(1 + a\dot{D}(t))^k} R_0 S_i(t) I(t) - \gamma I(t)$$

Calibrations by Atkeson et al. (2021) and Weitz et al. (2020) support endogenous distancing.

Infection/death growth rates drop much more quickly than the decline in $S(t)$ could account for. They then stabilize instead of dropping further.



Endogenous Interaction: Reduced Form Responses

A reduced-form model can easily account for main features of the epidemic: the low peak and very slow (almost plateau-like) declines: we just need activity at the steady state death rate to keep R_t near one.

Other features don't fit as well: mobility data suggests activity increased even before the peak; multiple peaks

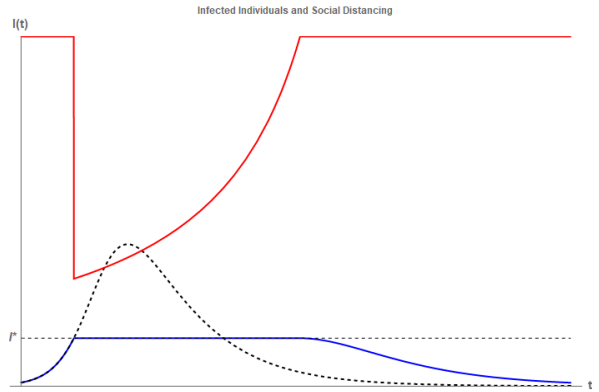
One augmentation the departures suggest is pandemic fatigue.

Endogenous activity models can have cycles driven by the lag between infections and deaths. Seasonal and variant-driven changes in transmissability are probably more important in practice.

Endogenous Interaction: Rational Responses

In a myopic model no agents will distance when I is low, all will distance when I is high, and agents will mix for intermediate I. Mixing will keep the risk where agents are indifferent.

This effect remains present in models with foward-looking agents and can lead to long-lasting infection plateaus.

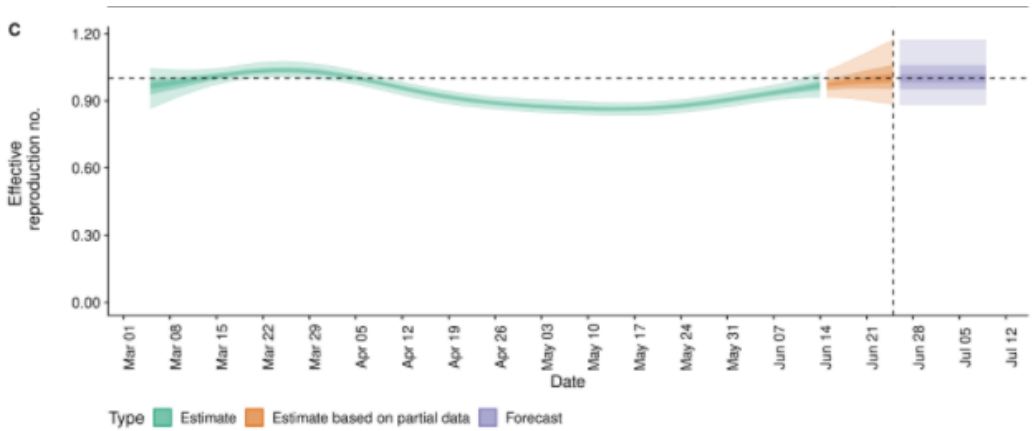


Endogenous social distancing improves social welfare relative to the SIR model. Social distancing is not socially optimal because individuals exert a negative externality when infected. A social planner would start distancing earlier and lower the plateau.

Accounting for effects on endogenous distancing is potentially important to many policy analyses.

Endogenous Interaction: $R_t \approx 1$

Estimates of the effective transmission rate (which would be $S(t)R_0$ in SIR) show that it has stayed remarkably close to one for a very long time.



In the heterogneous SIR model this requires that $\sum_i R_{0i}\left(S_i(t)R_{0i} - 1\right) \approx 0$ where I am now using R_{0i} for endogenous activity rates.

Groups with R_{0i} well below one can't contribute much to the sum.

Hence, estimates suggest that (pre-vaccines) the endogenous models must have worked pretty well even for those who stayed active.

References

1. Overview
   - Avery, Bossert, Clark, Ellison, Ellison, *Jrnl of Econ Perspectives*, 2020
2. Heterogeneous SIR Models
   - Diekmann, Hesterbeck, Metz, *Journal of Mathematical Biology*, 1990
   - Dushoff, Levin, *Mathematical Biosciences*, 1995
   - Ellison, *Covid Economics*, 2020.
3. Analyses of COVID Mitigation Policies
   - Acemoglu, Chernozhukov, Werning, Whinston, *AER Insights*, 2021
   - Baqaee, Farhi, Mina, Stock, *Brookings Papers*, 2020
   - Eichenbaum, Rebelo, Trabandt, *NBER* wp 26882, 2020
   - Favero, Ichino, Rustichini, 2021

References

4. Endogenous Distancing: Reduced Form
   - Atkeson, *NBER* wp 28760, 2021
   - Atkeson, Kopecky, Zha, Minnesota Fed Staff Report 618, 2021
   - Funk, Salathe, Jansen, *Journal of the Royal Society Interface*, 2010
   - Weitz, Park, Eksin, Dushoff, *PNAS* 2020
5. Endogenous Distancing: Rational Distancing
   - Farboodi, Jarosch, Shimer, *NBER* wp 27059, 2020
   - Kremer, *QJE*, 1996
   - McAdams, *Covid Economics*, May 2020
   - Reluga, *PLOS Computational Biology*, 2010
   - Toxvaerd, *Covid Economics*, 2020
6. Empirics on Endogenous vs. Policy-Driven Distancing
   - Chernozhukov, Kasaha, Schrimpf, *arXiv:2005.14168*, 2020
   - Goolsbee, Syverson, Becker-Friedman wp 2020-80, 2020