Estimating and Simulating a SIRD Model of COVID-19

Jesús Fernández-Villaverde and Chad Jones

April 13, 2020

(Very preliminary and incomplete)

Outline

- Basic model
  - Social distancing via a time-varying β
- Estimation and simulation
  - Different countries and states
  - Robustness to parameters
  - “Forecasts” from each of the last 7 days
  - Standard errors coming soon
- How much can we relax social distancing?

Basic Model

Notation

- Number of people who are (stocks):

$$\begin{aligned} S &= \text{Susceptible} \\ I &= \text{Infective} \\ R &= \text{Recovered} \\ D &= \text{Dead} \end{aligned}$$

- Constant population size is N

$$S_t + I_t + R_t + D_t = N$$

SIRD Model: Overview

- Susceptible get infected at rate $\beta I_t/N$

  New infections = $\beta I_t/N \cdot S_t$

- Infections resolve at Poisson rate γ, so the average number of days until resolution is $1/\gamma$ so $\gamma = .2 \Rightarrow 5$ days.

- Resolution happens in one of two ways:
  - Death: fraction δ
  - Recovery: fraction $1 - \delta$

SIRD Model: Laws of Motion

$$\Delta S_{t+1} = \underbrace{-\beta S_t I_t / N}_{\text{new infections}}$$

$$\Delta I_{t+1} = \underbrace{\beta S_t I_t / N}_{\text{new infections}} - \underbrace{\gamma I_t}_{\text{resolving infections}}$$

$$\Delta R_{t+1} = \underbrace{(1-\delta)\gamma I_t}_{\text{recover}}$$

$$\Delta D_{t+1} = \underbrace{\delta \gamma I_t}_{\text{die}}$$

$$R_0 = D_0 = 0$$

New notation (terrible) R_0: Initial infection rate

- Let $s_0 \equiv S_0/N$ = fraction susceptible
- Terrible recycled notation: “Reproduction number” $R_0 \equiv \beta/\gamma$
    - β = rate at which you get the virus from one infective person
    - $1/\gamma$ = average time infected
    - So R_0 = expected number of infections generated by one sick person when no herd immunity ($s_0 \approx 1$)

Basic Properties of Differential System (Hethcote 2000)

- If $R_0 s_0 > 1$, the disease spreads, otherwise declines immediately
- Initial exponential growth rate is $\beta - \gamma$
- As $t \to \infty$, the total fraction of people ever infected, e^*, solves (assuming $s_0 \approx 1$)

$$e^* = -\frac{1}{R_0} \log(1 - e^*)$$

Long run is pinned down by R_0 (and death rate),
γ affects timing

Social Distancing

- What about a time-varying infection rate β_t?
  - Disease characteristics – fixed, homogeneous
  - Regional characteristics (NYC vs Montana) – fixed, heterogeneous
  - Social distancing – varies over time and space
- Model: assume two key parameters β_0 and β^*
- Economy decays exponentially from β_0 to β^* at rate λ:

$$\beta_t = \beta_0 e^{-\lambda t} + \beta^*(1 - e^{-\lambda t})$$

$\Rightarrow$ *can think about initial* $R_0 = \beta_0/\gamma$

and final $R_0^* = \beta^*/\gamma$

Estimates and Simulations

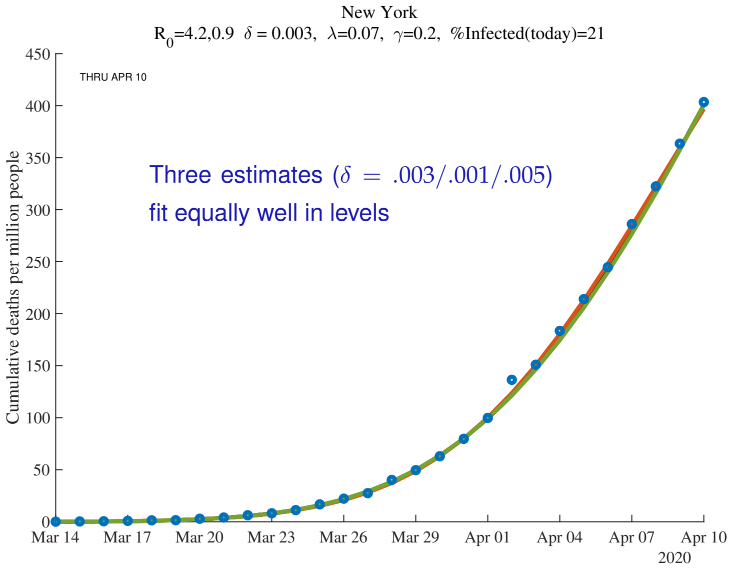
Estimation: Countries and States

- Parameters that are fixed and homogeneous
  - $\gamma = 0.2$: average duration is 5 days (or $\gamma = 0.1$)
  - $\delta = 0.003$ (Heinsberg, Germany random sampling)
  - $\lambda = 7\%$: β_t falls halfway to new value each 10 days
- Parameters that vary across countries/states
  - β_0 and β^*
  - I_0: initial number of infections (gets timing right)
- Objective function:
  - Equally weighted SSR for Cumulative deaths (logs) and Daily deaths (logs)

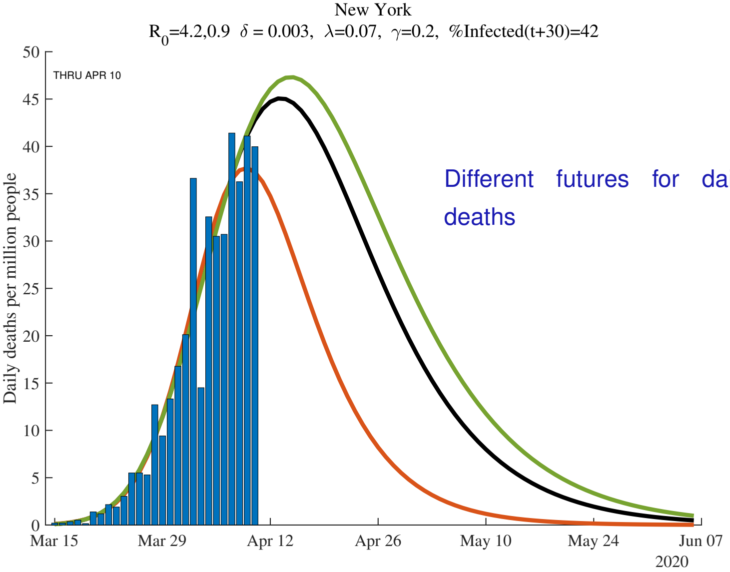
Guide to Graphs

- 7 days of forecasts: ROY-G-BIV (old to new, low to high)
  - Black=current
  - Red = oldest, Orange = second oldest, Yellow =third oldest...
  - Violet (purple) = one day earlier
- For robustness graphs, same idea
  - Black = baseline (e.g. $\delta = .003$)
  - Red = lowest parameter value (e.g. $\delta = .001$)
  - Green = highest parameter value (e.g. $\delta = .005$)

New York: Cumulative Deaths per Million People ($\delta = .003/.001/.005$)

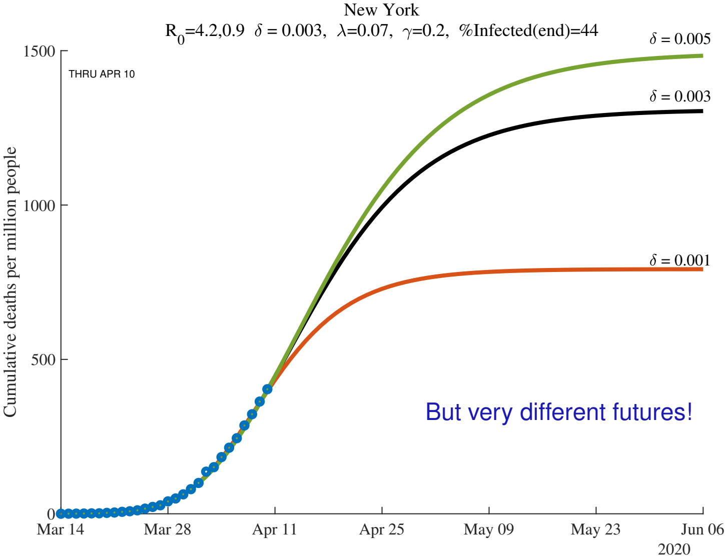


New York: Daily Deaths per Million People ($\delta = .003/.001/.005$)



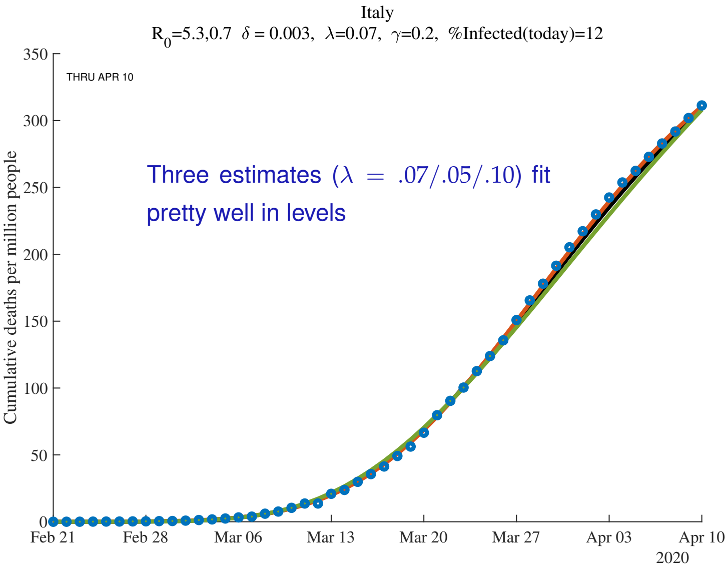
Different futures for daily deaths

New York: Cumulative Deaths per Million (Future, $\delta = .003/.001/.005$)



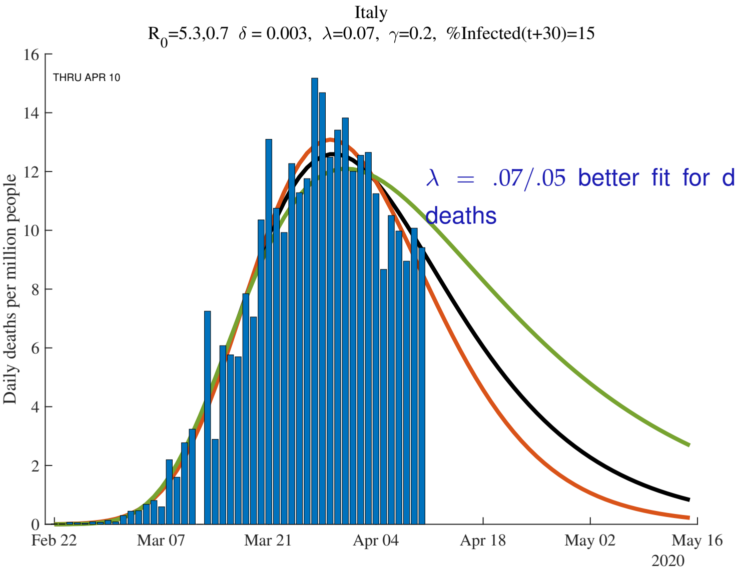
But very different futures!

Italy: Cumulative Deaths per Million People ($\lambda = .07/.05/.10$)



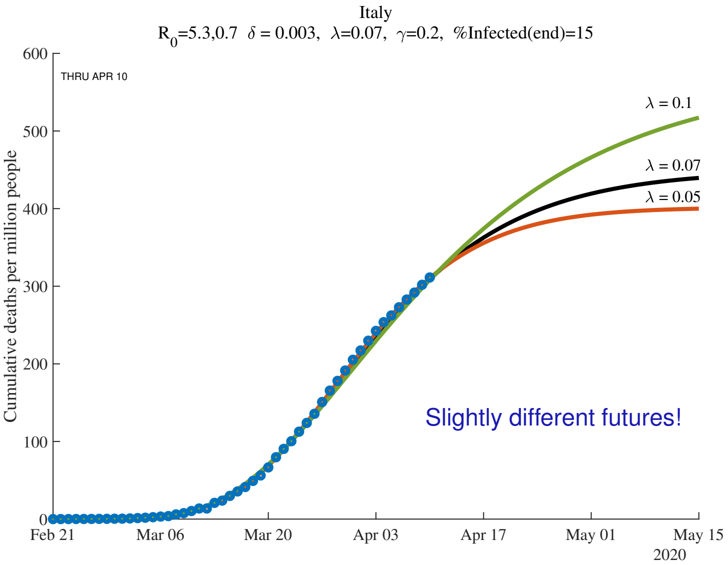
Three estimates ($\lambda = .07/.05/.10$) fit pretty well in levels

Italy: Daily Deaths per Million People ($\lambda = .07/.05/.10$)

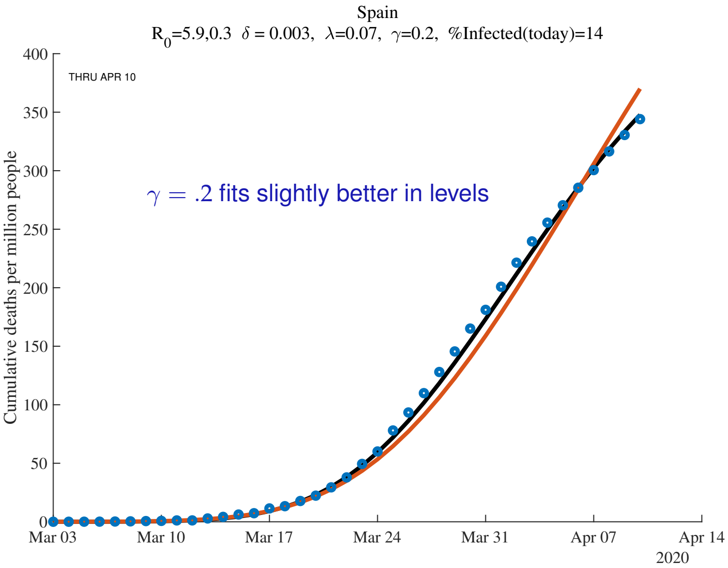


$\lambda = .07/.05$ better fit for daily deaths

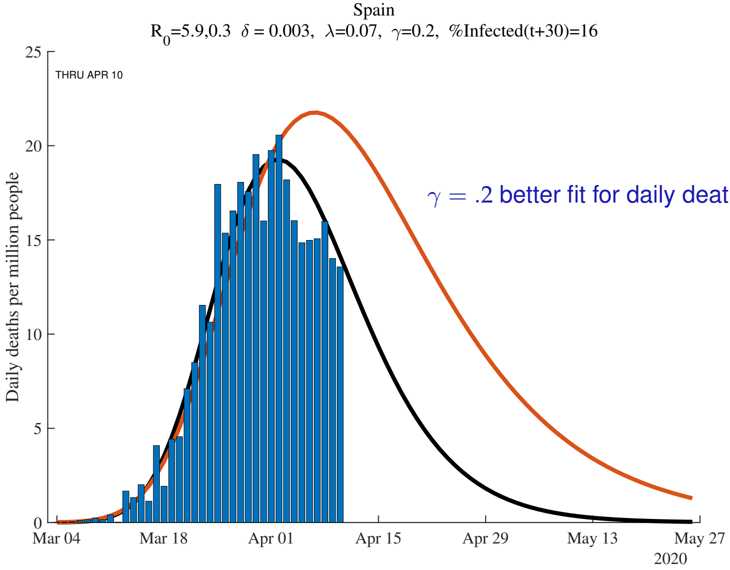
Italy: Cumulative Deaths per Million (Future, $\lambda = .07/.05/.10$)



Spain: Cumulative Deaths per Million People ($\gamma = .2/.1$)

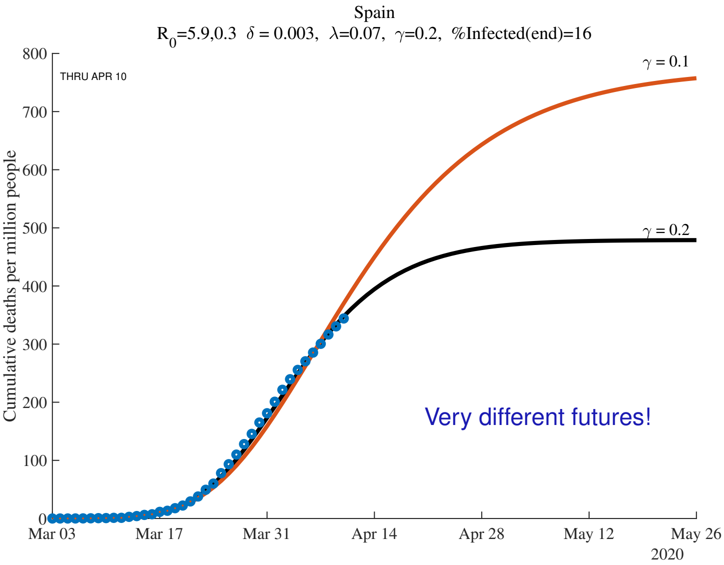


Spain: Daily Deaths per Million People ($\gamma = .2/.1$)



$\gamma = .2$ better fit for daily deaths

Spain: Cumulative Deaths per Million (Future, $\gamma = .2/.1$)

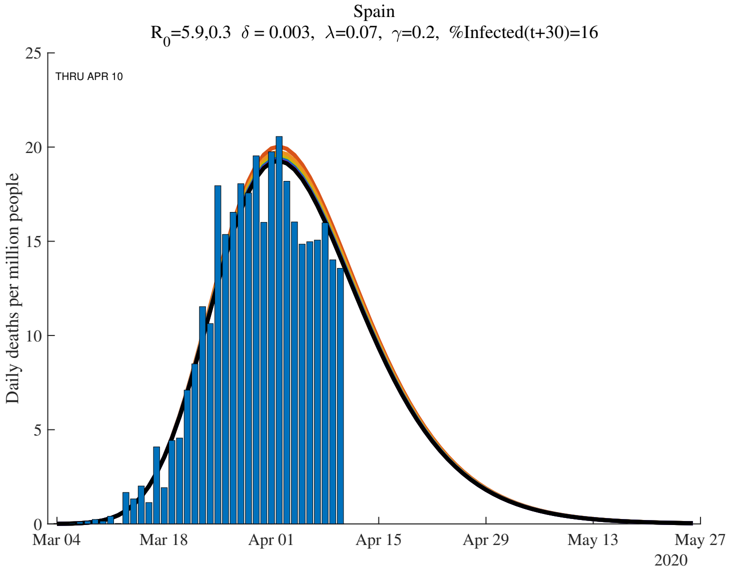


Repeated “Forecasts” from the past 7 days of data

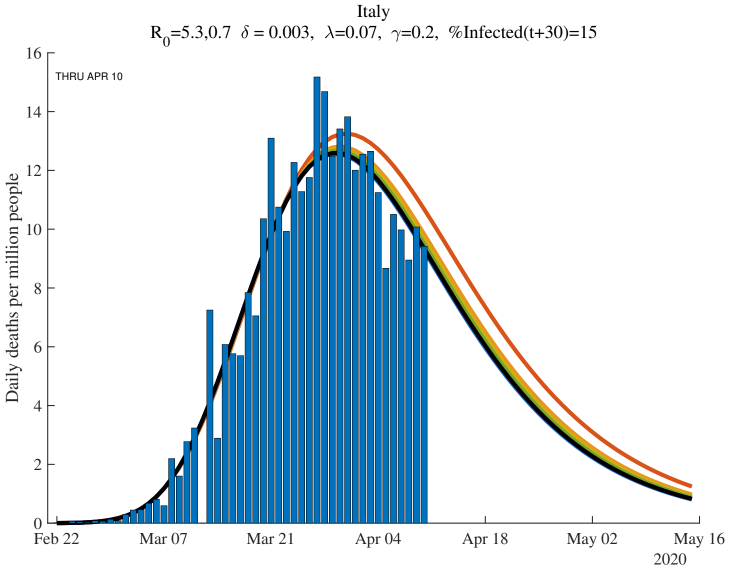
– After peak, forecasts settle down.

– Before that, very noisy!

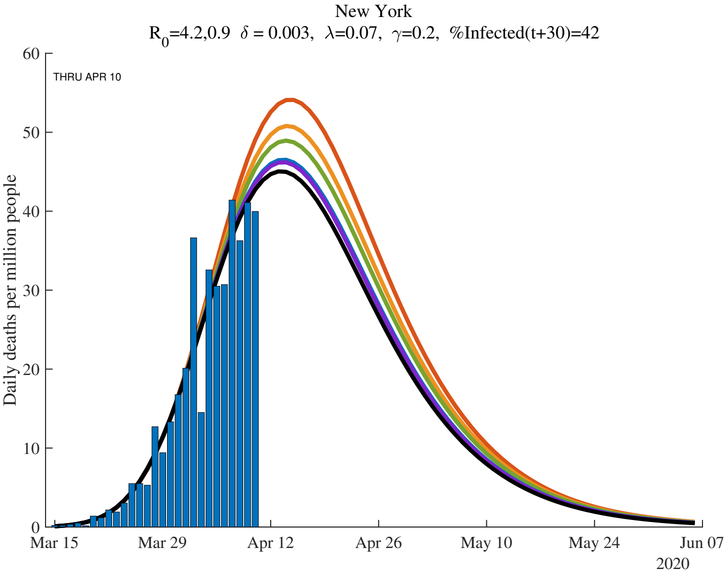
Spain (7 days): Daily Deaths per Million People



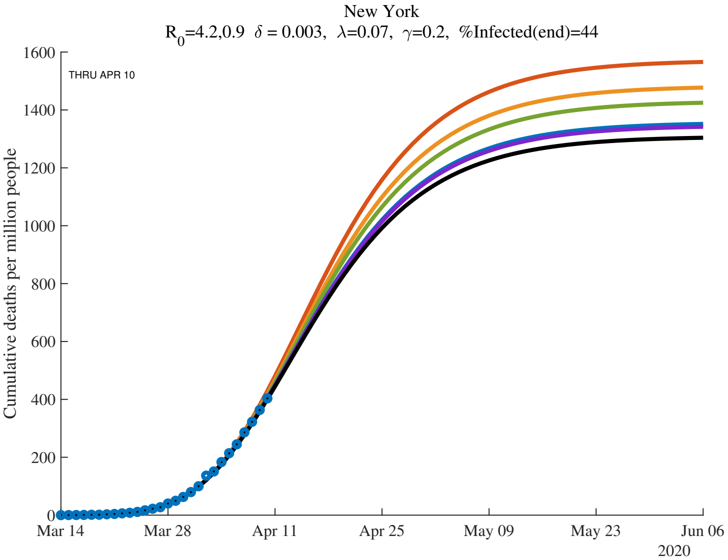
Italy (7 days): Daily Deaths per Million People



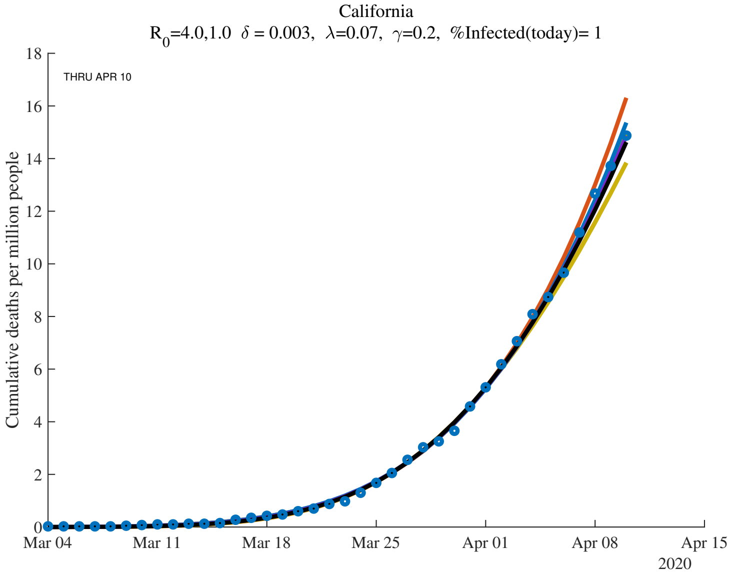
New York (7 days): Daily Deaths per Million People



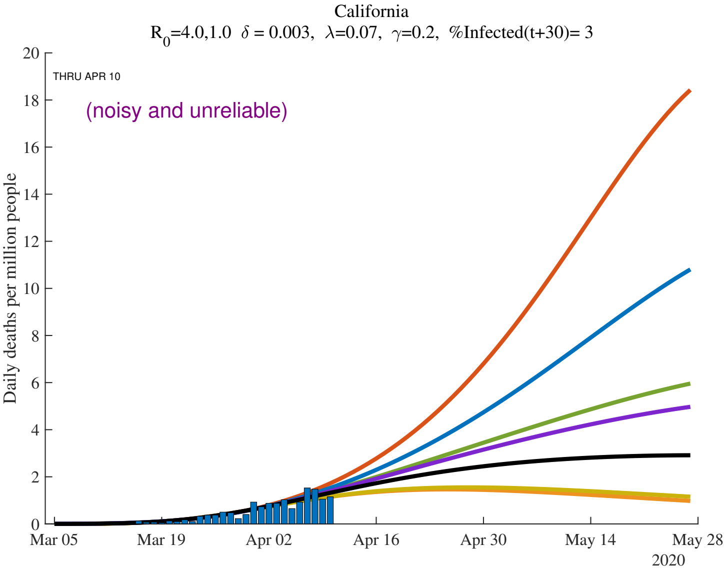
New York (7 days): Cumulative Deaths per Million (Future)



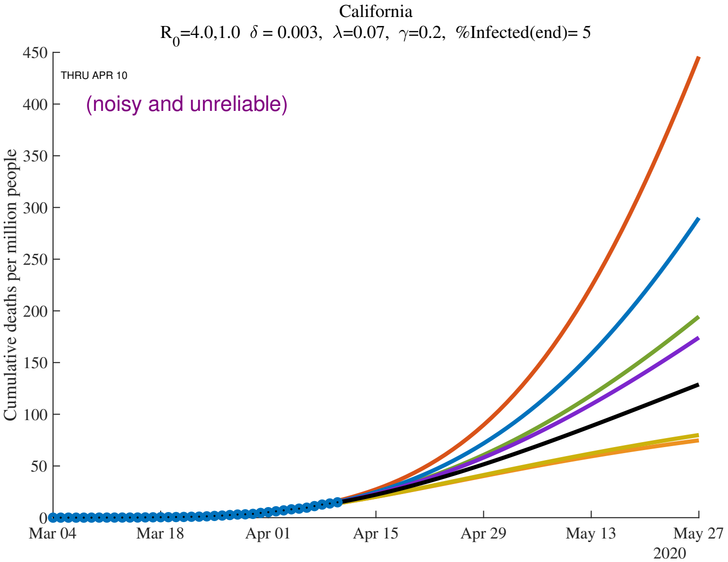
California (7 days): Cumulative Deaths per Million



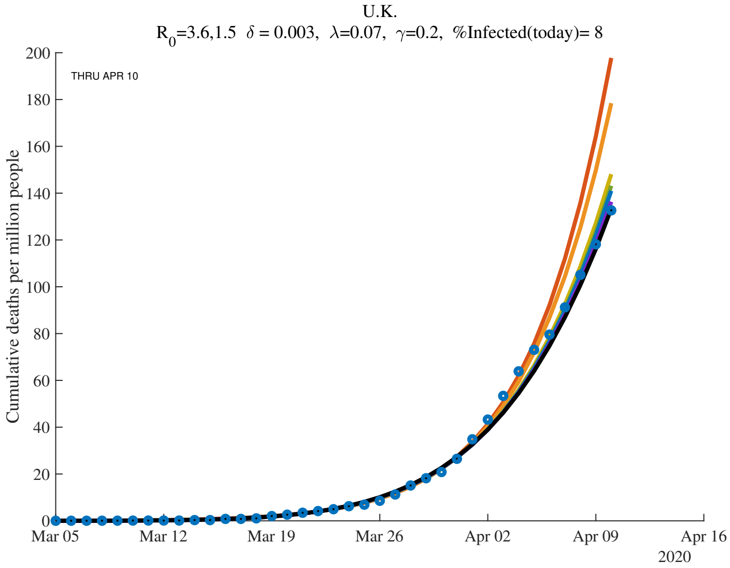
California (7 days): Daily Deaths per Million People



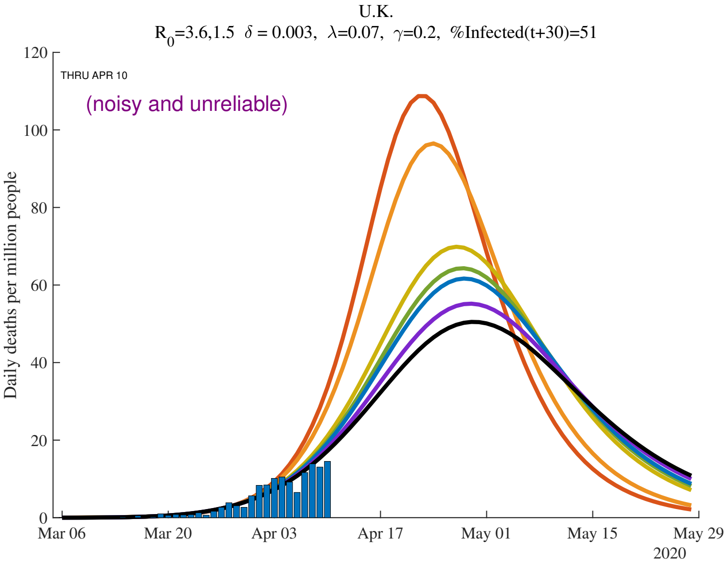
California (7 days): Cumulative Deaths per Million (Future)



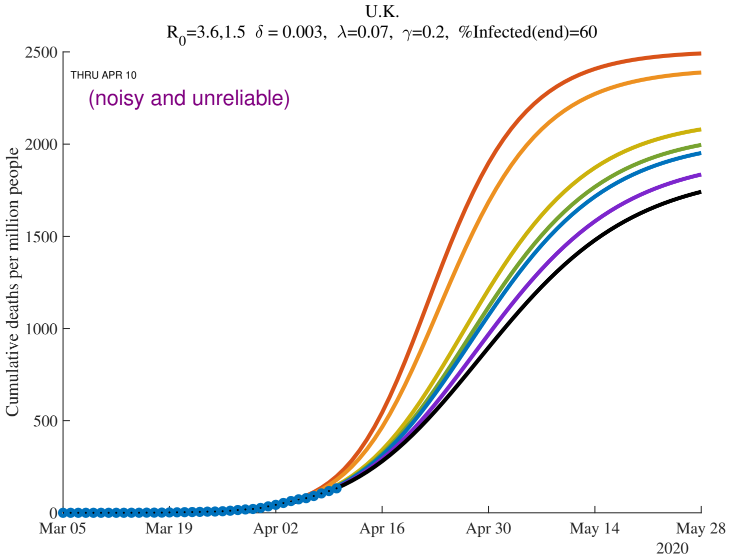
U.K. (7 days): Cumulative Deaths per Million



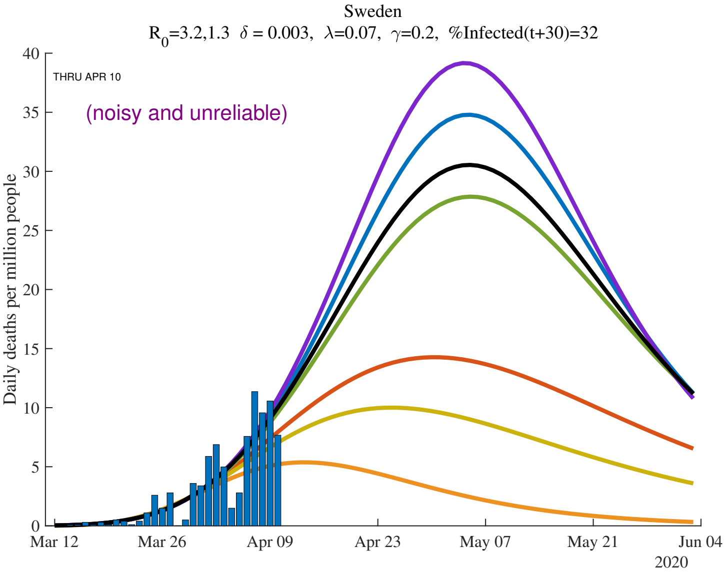
U.K. (7 days): Daily Deaths per Million People



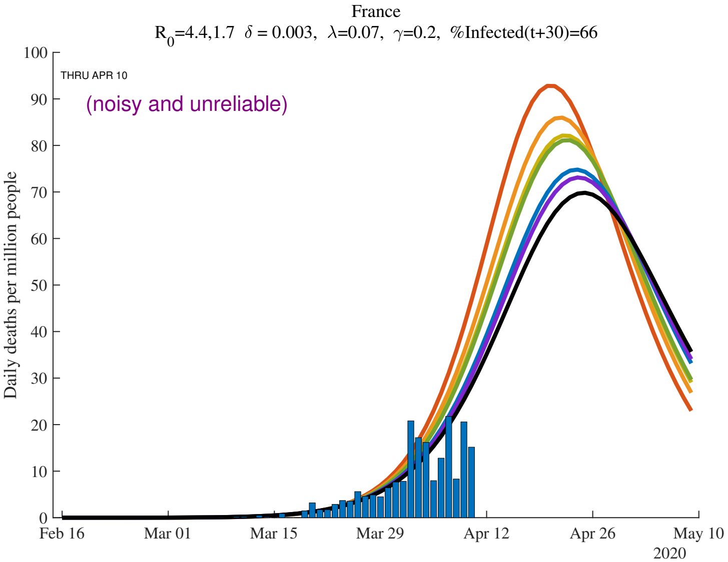
U.K. (7 days): Cumulative Deaths per Million (Future)



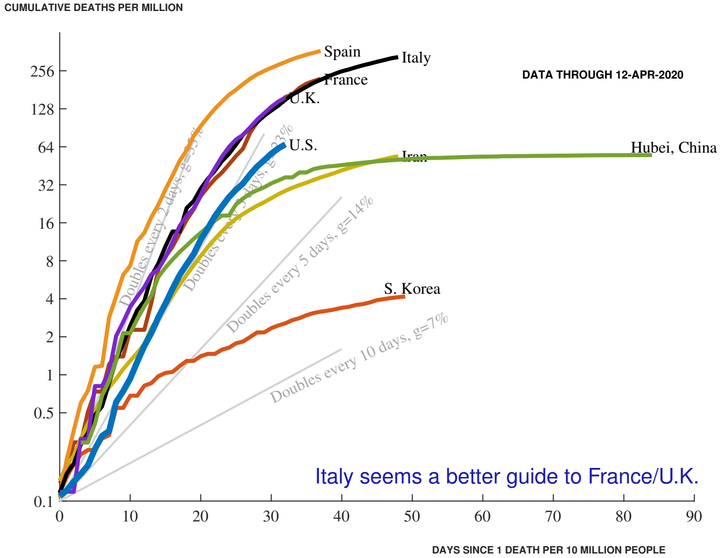
Sweden (7 days): Daily Deaths per Million People



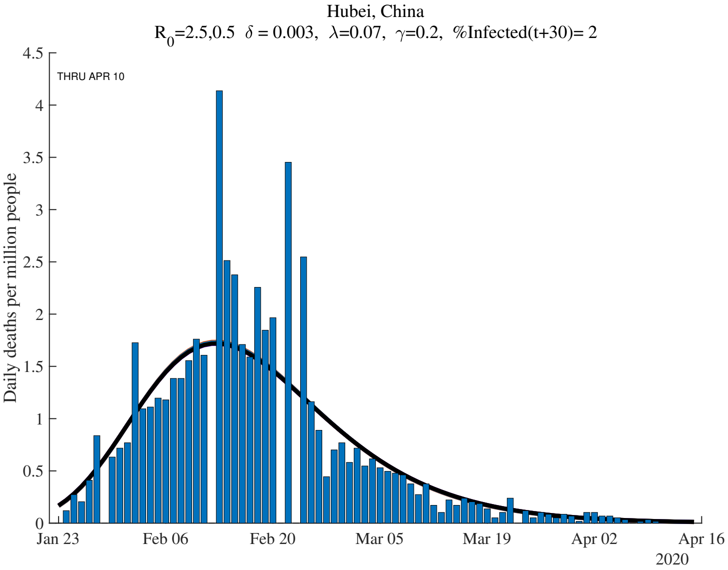
France (7 days): Daily Deaths per Million People



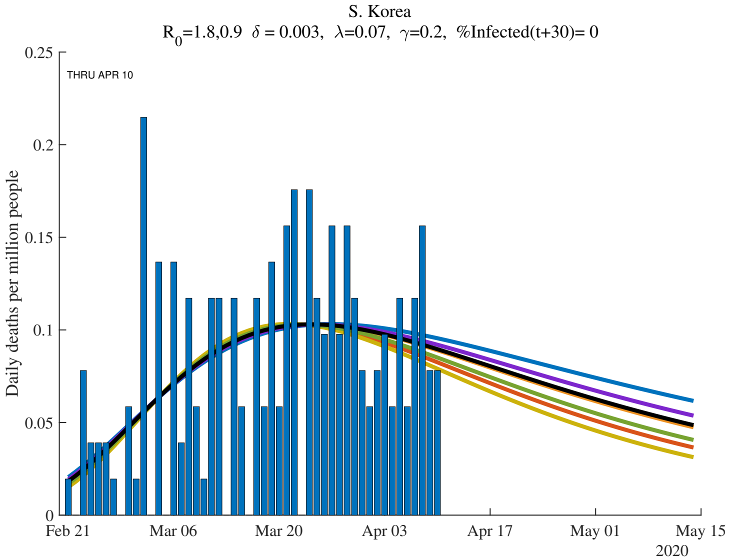
Cumulative Deaths per Million, Log Scale



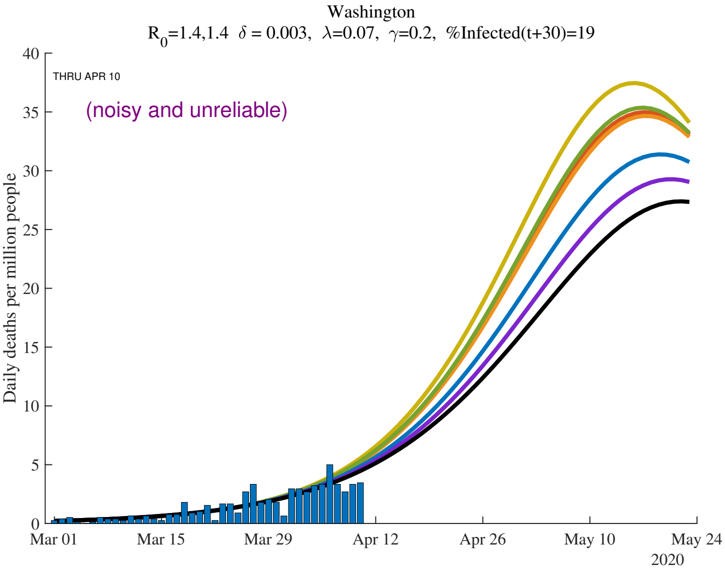
Hubei, China (7 days): Daily Deaths per Million People



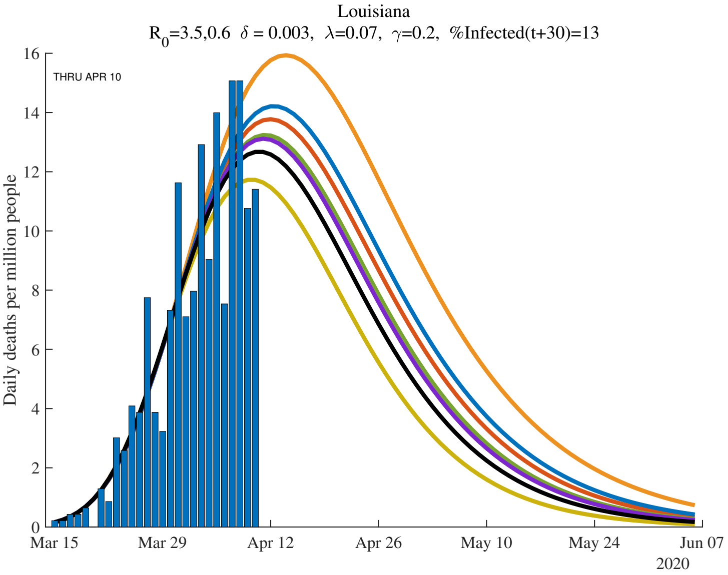
S. Korea (7 days): Daily Deaths per Million People



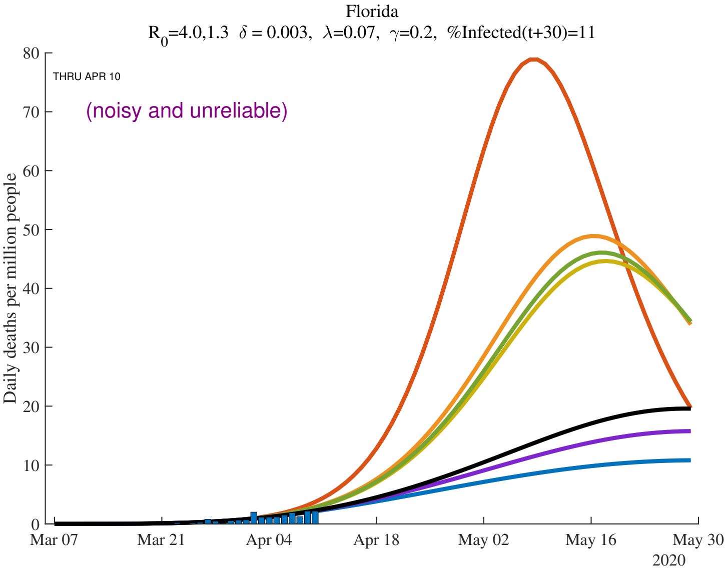
Washington (7 days): Daily Deaths per Million People



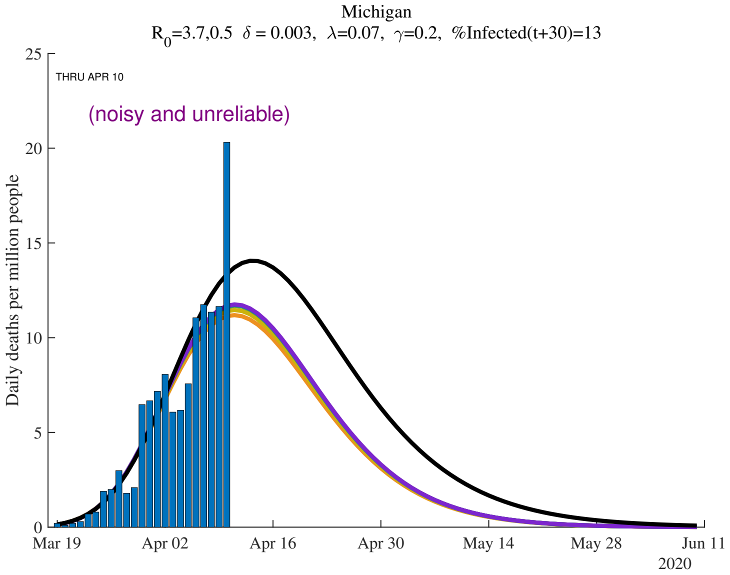
Louisiana (7 days): Daily Deaths per Million People



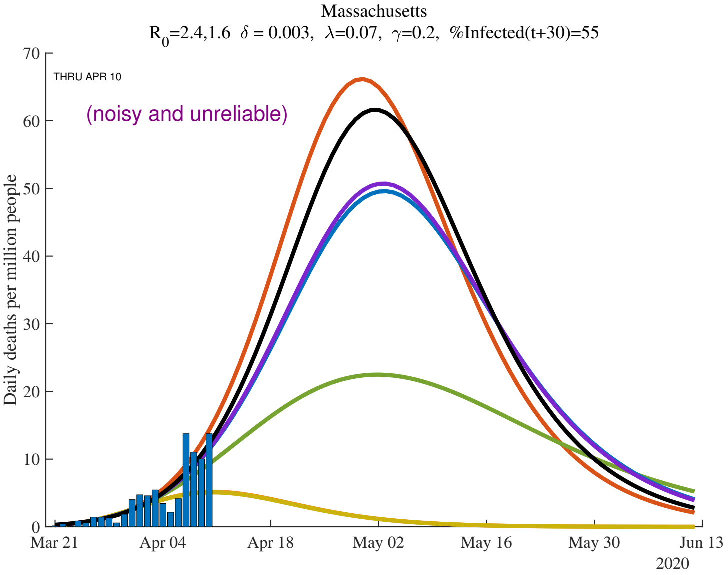
Florida (7 days): Daily Deaths per Million People



Michigan (7 days): Daily Deaths per Million People



Massachusetts (7 days): Daily Deaths per Million People



Assessing Uncertainty

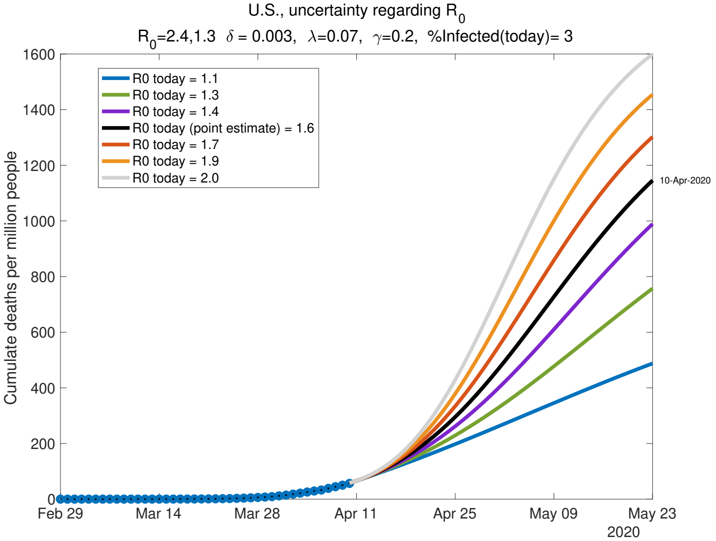
How Do We Evaluate Uncertainty?

- Basic SIRD model is deterministic
    - There is an implicit appeal to a law of large numbers holding in populations.
    - How do we map it into data?
    - Measurement error (huge undercount of deaths in Italy and Spain, massive under-reporting of infections). However, unlikely to be classical measurement error.
    - Un-modeled shocks (we have not specified them explicitely).
- Thus, assessing uncertainty is not straightforward.

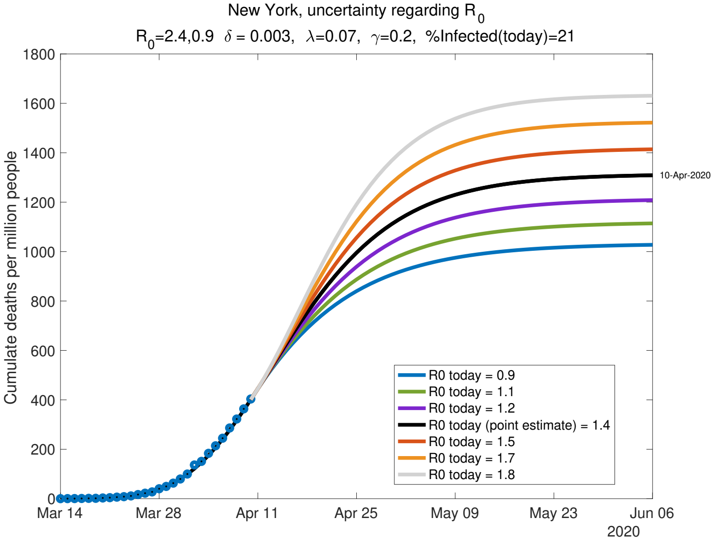
Our Approach

- A simple empirical Bayesian method
  - Specify a prior on β, γ, and δ centered around our point estimates and having a variance based on plausibility/our reading of the medical literature.
  - You sample from the prior and forecast the future behavior of the model based on the sampled parameter values.
  - Informative about the properties of the model and range of likely outcomes
- We tried to compute standard errors.
  - We got numerically unstable values.
  - However, our provisional values are within the range of our priors (if anything, our computed standard errors feel "too small").

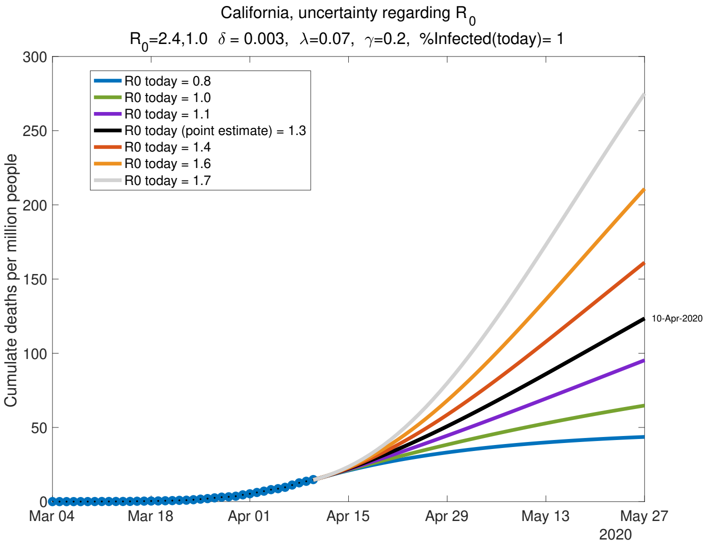
R_0 is key for dynamics...



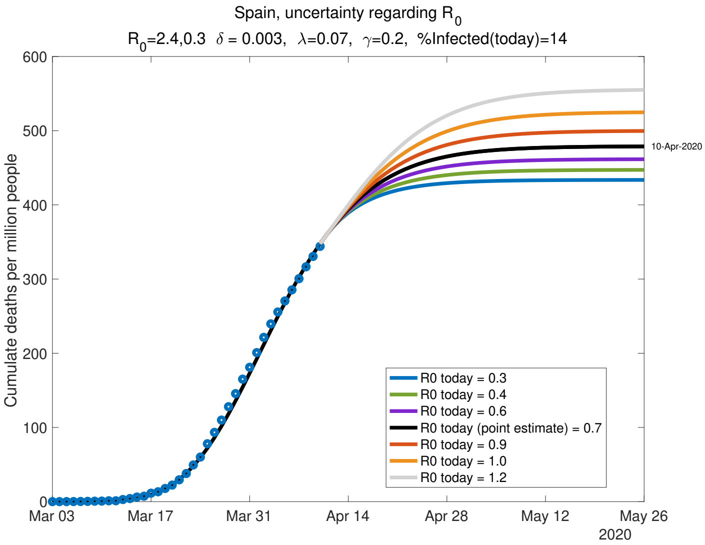
R_0 is key for dynamics...



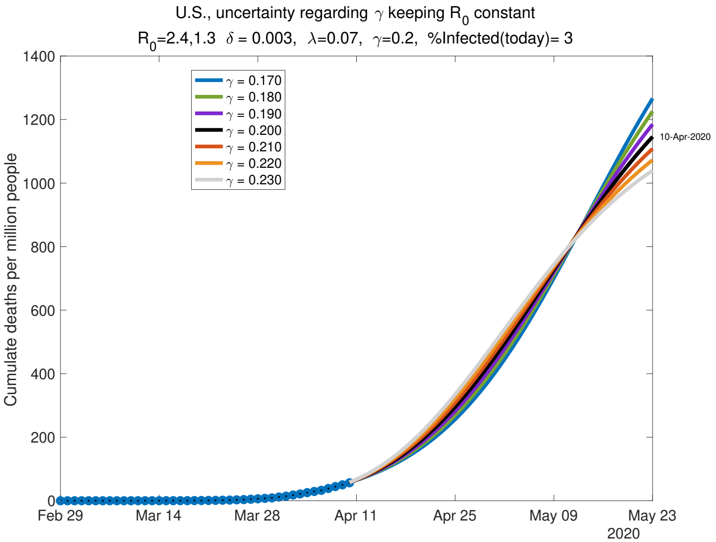
R_0 is key for dynamics...



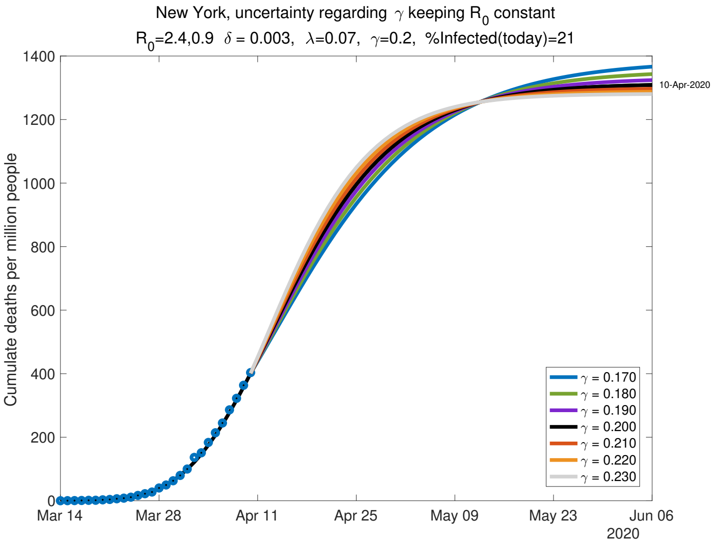
R_0 is key for dynamics...



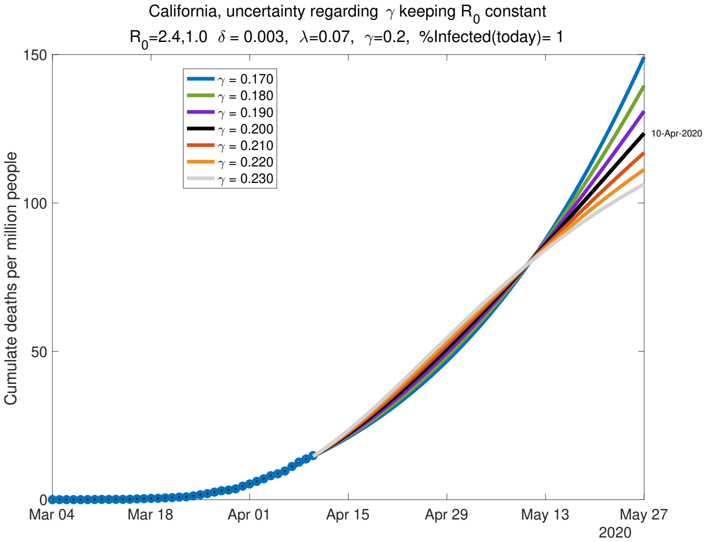
γ not so much



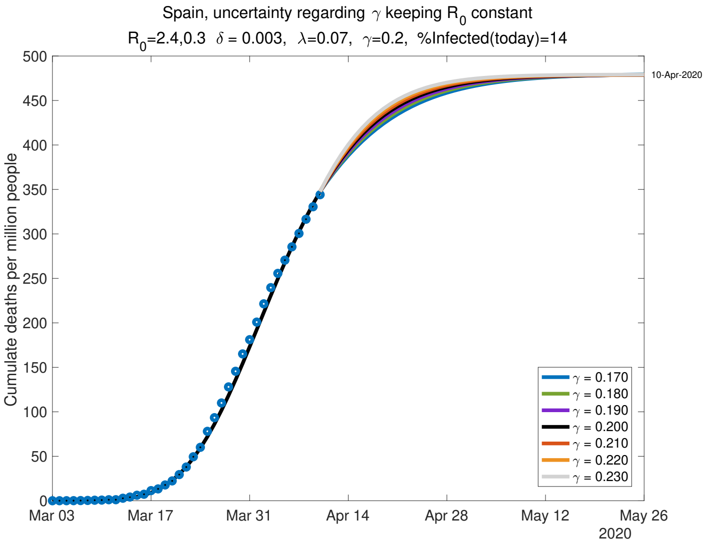
γ not so much



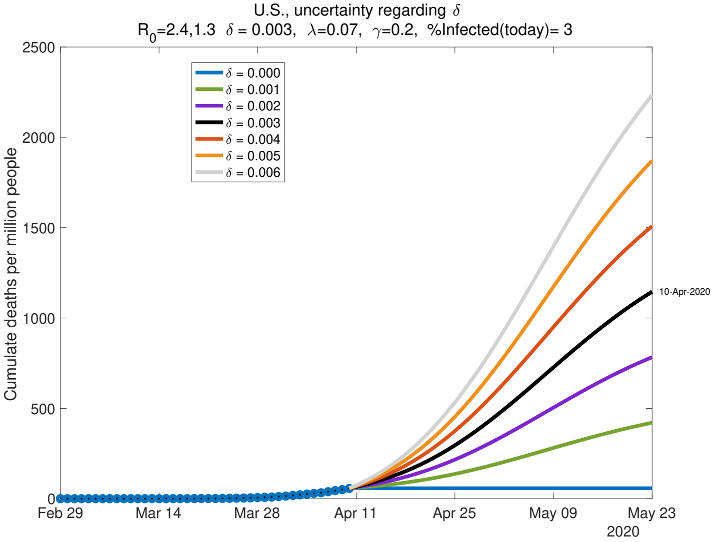
γ not so much



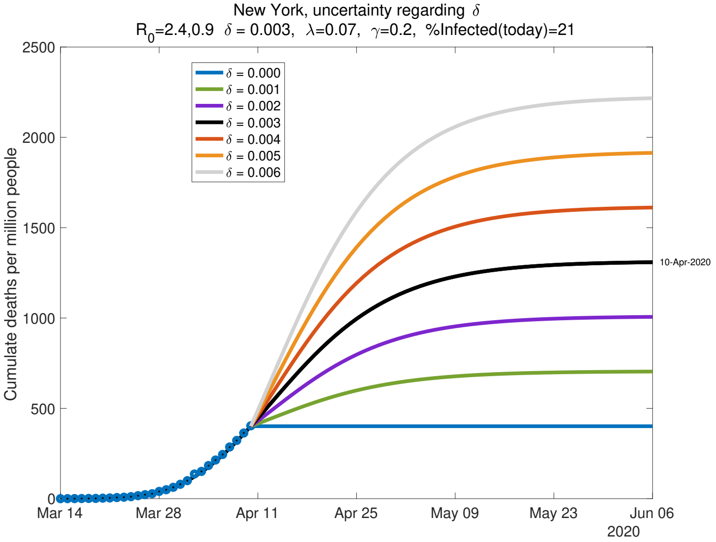
γ not so much



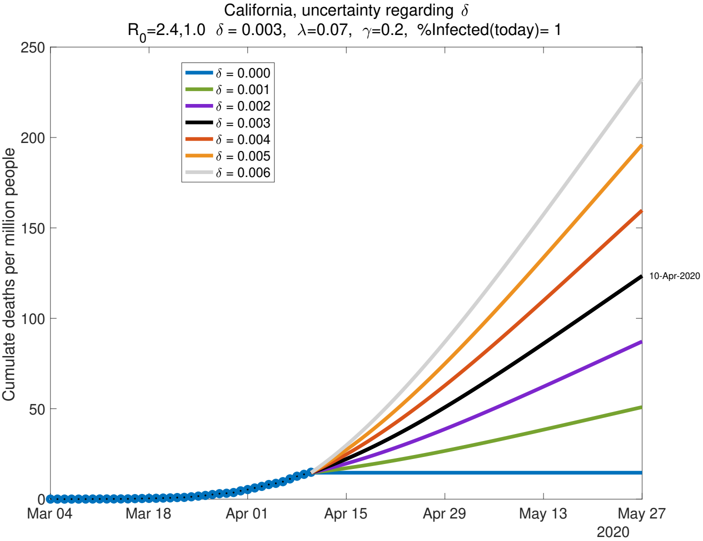
δ matters for deaths, not for dynamics



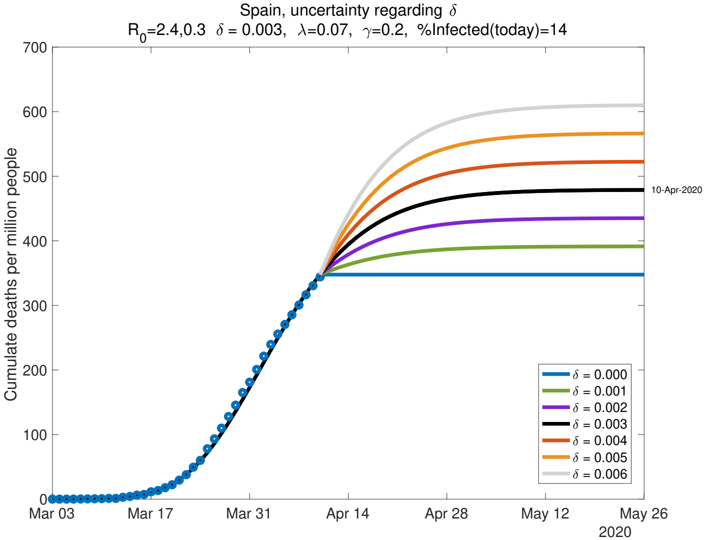
δ matters for deaths, not for dynamics



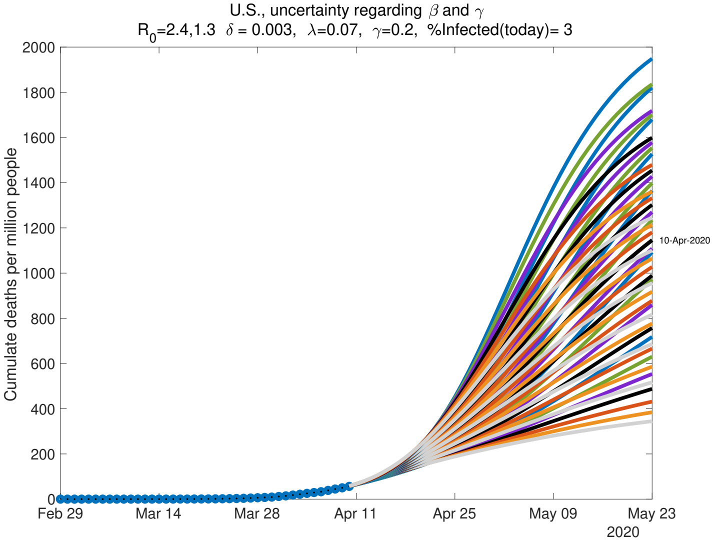
δ matters for deaths, not for dynamics



δ matters for deaths, not for dynamics

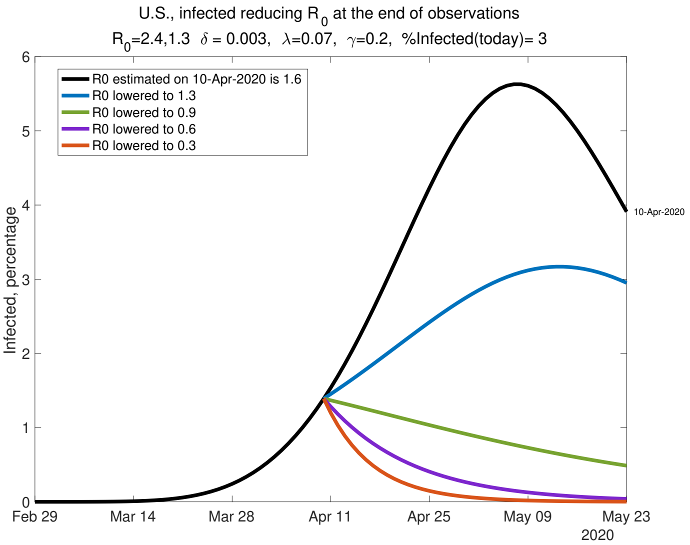


We can plot joint draws

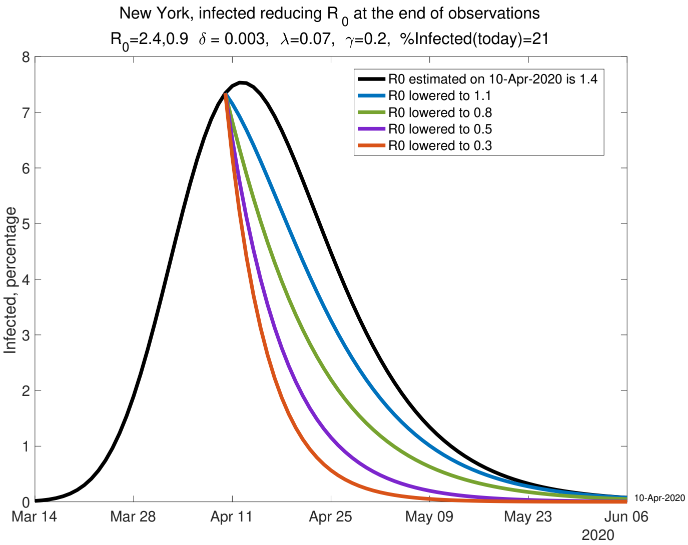


Policy Counterfactuals

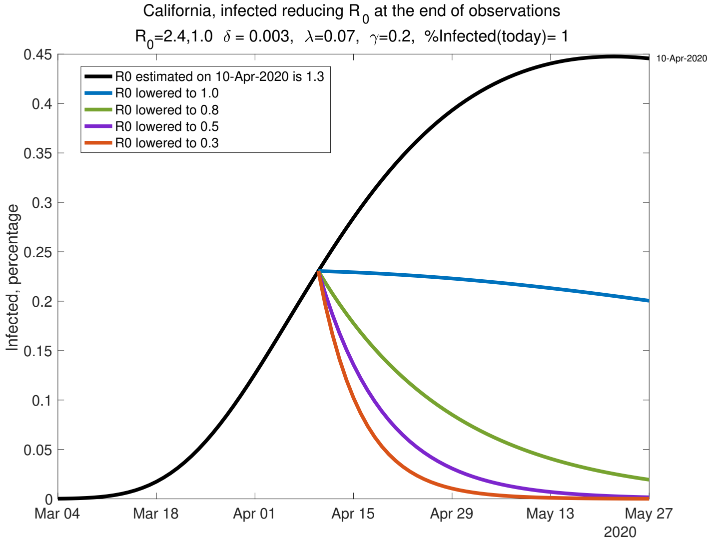
Reducing R_0 has a huge impact...



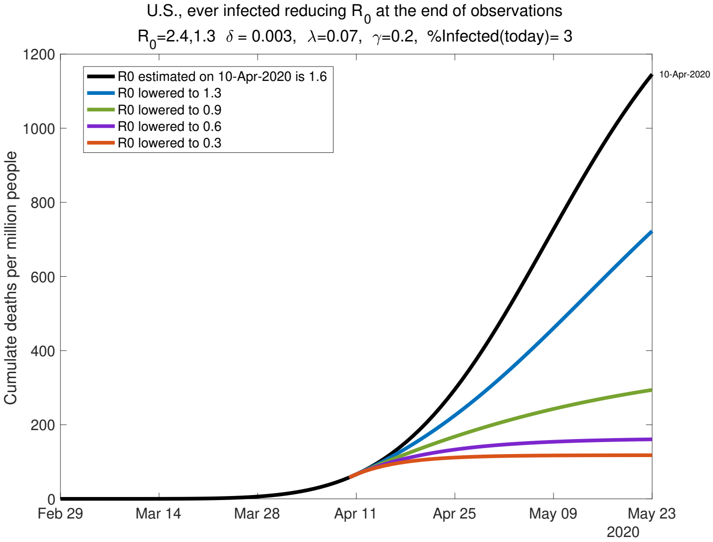
Reducing R_0 has a huge impact...



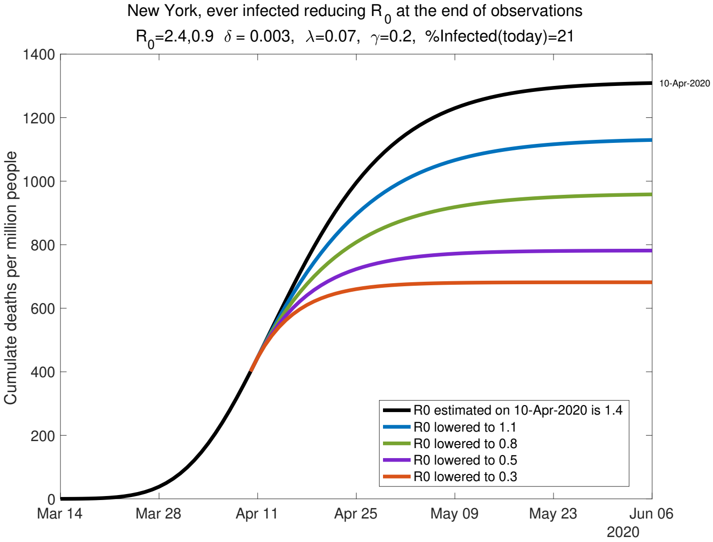
Reducing R_0 has a huge impact...



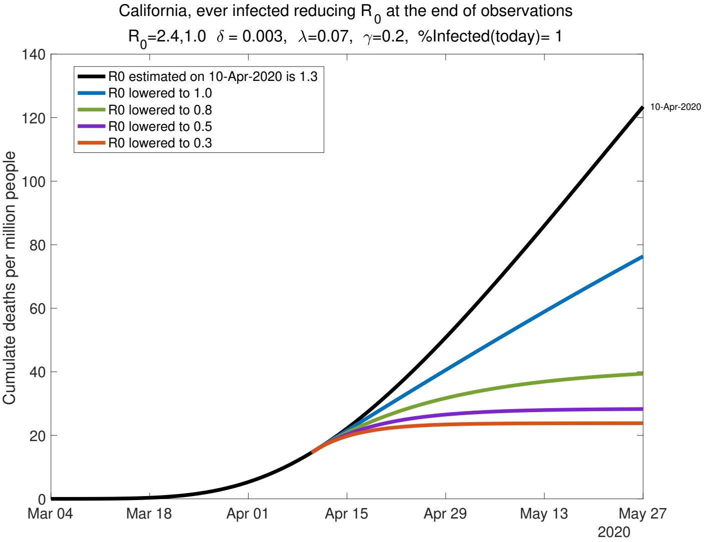
...but it has a decreasing marginal effects on cumulate deaths



...but it has a decreasing marginal effects on cumulate deaths



...but it has a decreasing marginal effects on cumulate deaths



Reopening and Herd Immunity

Percent Ever Infected would be very informative

	— Percent Ever Infected (today) —		
	$\delta = .001$	$\delta = .003$	$\delta = .005$
New York	58	21	13
California	2	1	0
Italy	35	12	7
Spain	40	14	8
U.K.	22	8	5
France	34	13	8
Hubei, China	5	2	1
S. Korea	0	0	0
Sweden	14	5	3
Louisiana	22	7	4
Florida	3	1	1
Washington	8	3	2
Michigan	19	6	4
Massachusetts	15	5	3

Herd Immunity

- How far can we relax social distancing?
- Let $s(t) = S(t)/N$ = the fraction still susceptible
  - The disease will die out as long as

$$R_0(t)s(t) < 1$$

  - That is, if the "new" R_0 is smaller than $1/s(t)$
  - Today's infected people infect fewer than 1 person on average
- We can relax social distancing to raise $R_0(t)$ to $1/s(t)$

Herd Immunity and Opening the Economy?

	R_0	R_0^*	Percent Susceptible t+30	R_0(t+30) with no outbreak	Percent way back to normal
New York	4.2	0.9	57.7	1.7	26.0
California	4.0	1.0	96.9	1.0	-0.3
Hubei, China	2.5	0.5	98.2	1.0	25.0
Italy	5.3	0.7	85.3	1.2	10.0
Spain	5.9	0.3	84.1	1.2	15.3
U.K.	3.6	1.5	49.1	2.0	25.8
France	4.4	1.7	34.3	2.9	45.0
S. Korea	1.8	0.9	99.8	1.0	11.0
Sweden	3.2	1.3	68.5	1.5	8.8
Louisiana	3.5	0.6	87.0	1.1	18.7
Florida	4.0	1.3	89.1	1.1	-5.3
Washington	1.4	1.4	81.3	1.2	-456.2
Michigan	3.7	0.5	86.7	1.2	19.8
Massachusetts	2.4	1.6	44.7	2.2	76.7

Thanks!