

Appendix S1 – Additional Figures

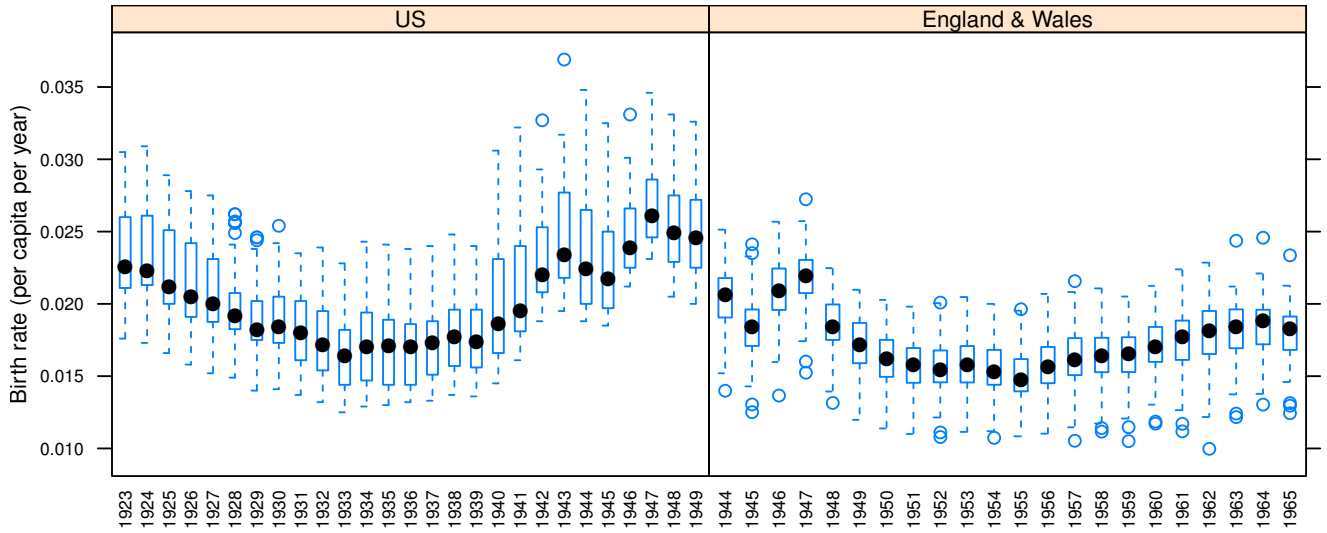


Figure S1. Birth rates adjusted for infant mortality. All rates are yearly per capita. The U.S. figure shows state birth rates, while the England & Wales figure shows city rates.

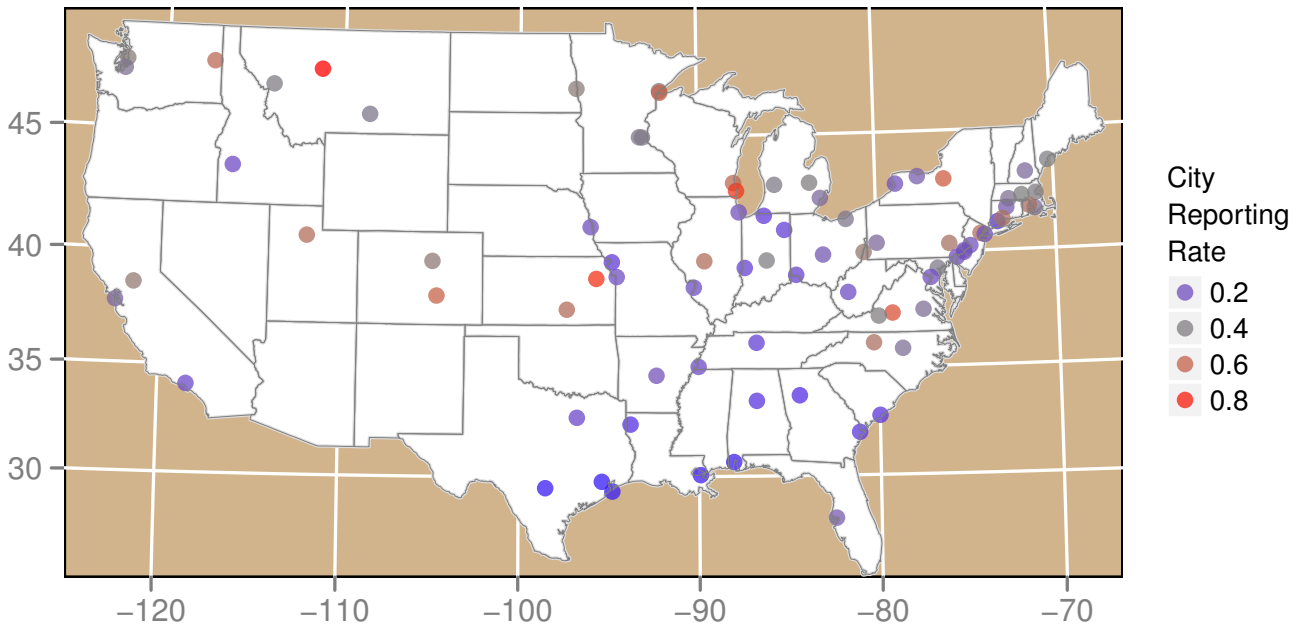


Figure S2. Sampled U.S. cities, showing estimated reporting rates.

Country	Inferred Parameter	Model Term	Estimate	Std. Error	t value	Pr(> t)
US	Mean	(Intercept)	-6.392	0.179	-35.710	< 10 ⁻⁹
US	Mean	logN	0.339	0.034	9.963	< 10 ⁻⁹
US	SD	(Intercept)	2.217	0.151	14.712	< 10 ⁻⁹
US	SD	logN	-0.205	0.029	-7.143	< 10 ⁻⁹
England & Wales	Mean	(Intercept)	-6.457	0.263	-24.579	< 10 ⁻⁹
England & Wales	Mean	logN	0.432	0.051	8.532	< 10 ⁻⁹
England & Wales	SD	(Intercept)	2.494	0.188	13.276	< 10 ⁻⁹
England & Wales	SD	logN	-0.295	0.036	-8.125	< 10 ⁻⁹

Table S1. Descriptive linear models (separated in table by horizontal lines) for each country and inferred parameter against log N. This gives a closed-form expression for the CDF of log incidence as a function of N, which was used to compute the *extinction boundary probability B* for a range of population sizes.

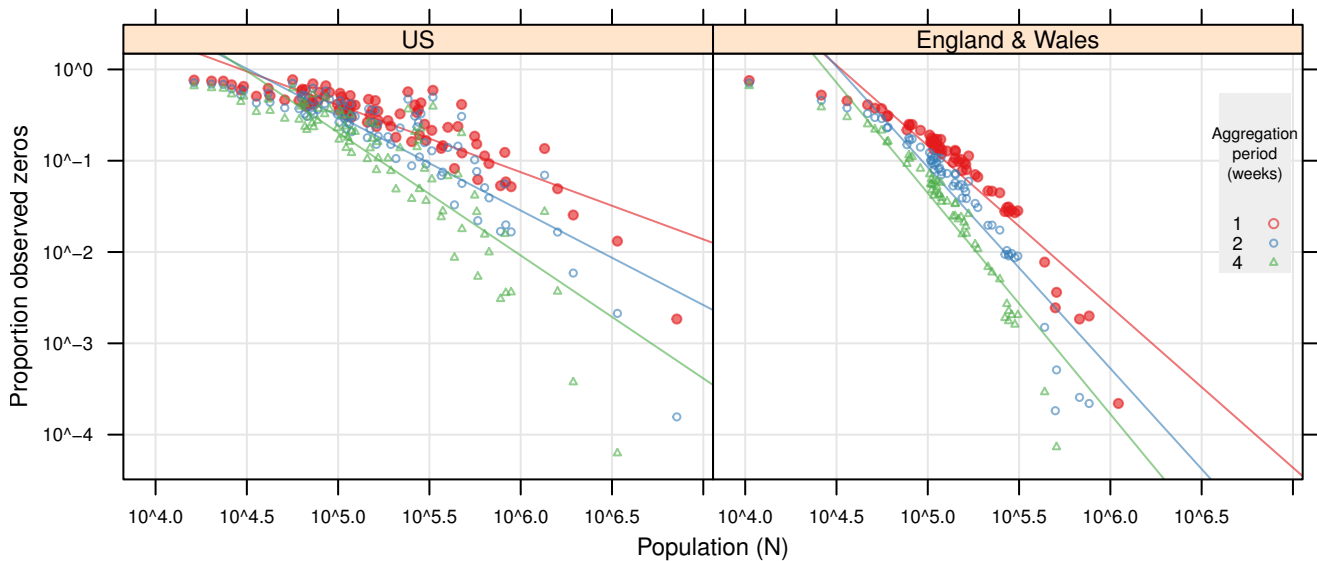
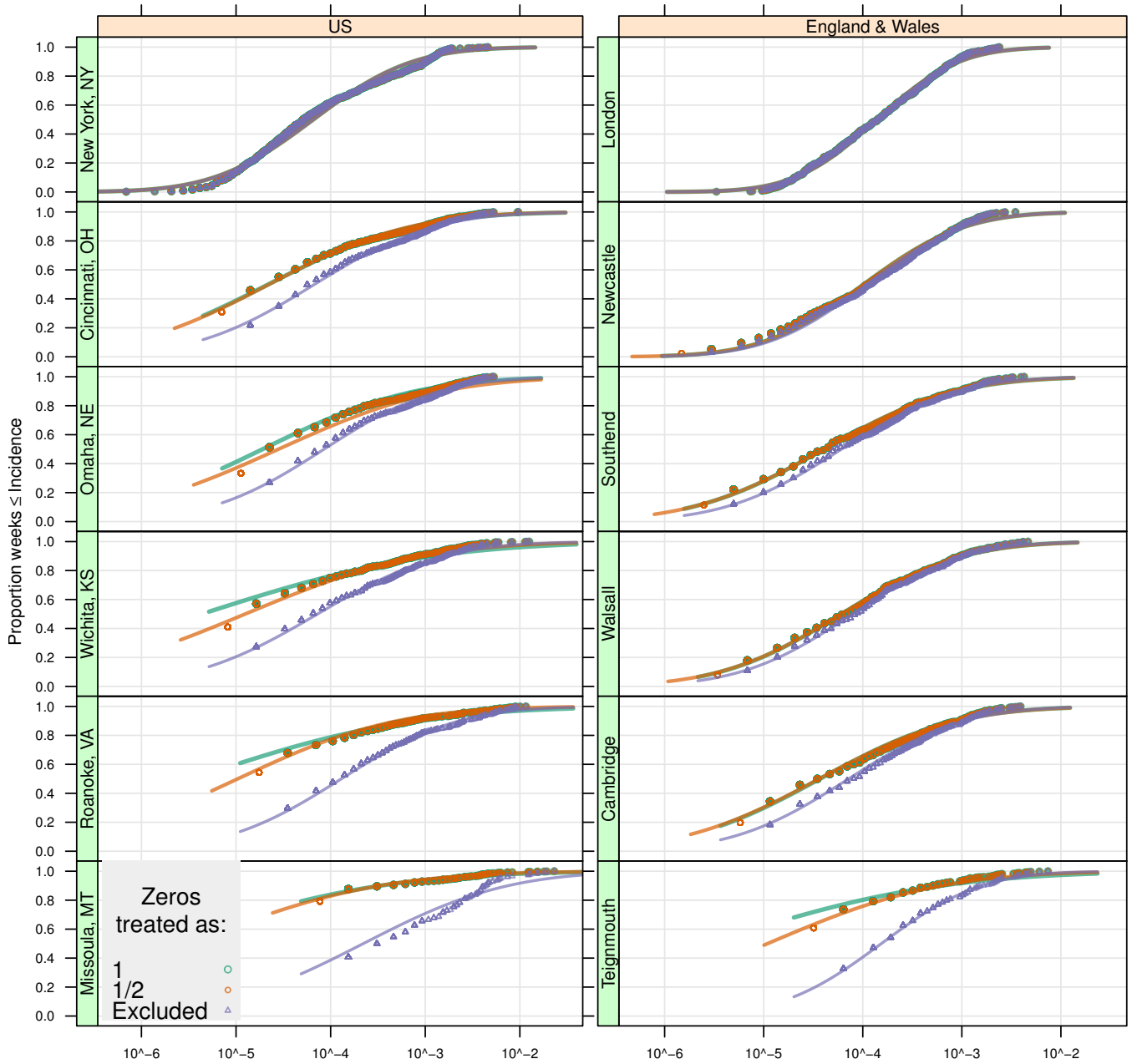


Figure S3. Classic CCS curve for varying levels of temporal aggregation. Observed weekly cases from simulations were summed over varying number of weeks. The mean proportion of observed zeros per sampling period was computed for each city. Linear regressions for each sample period are overlaid.



Incidence (inferred cases per capita per week)

Figure S4. The ECDF (points) and fitted normal distribution (lines) of log incidence (ξ) for a sample of (population-ordered) cities. Here zeros are treated either as 1 or 1/2 actual case, or are excluded. Excluding zeros significantly shifts the ECDF and resulting fit, particularly in the U.S. Discretization of the ECDF is clearly evident at low incidence.

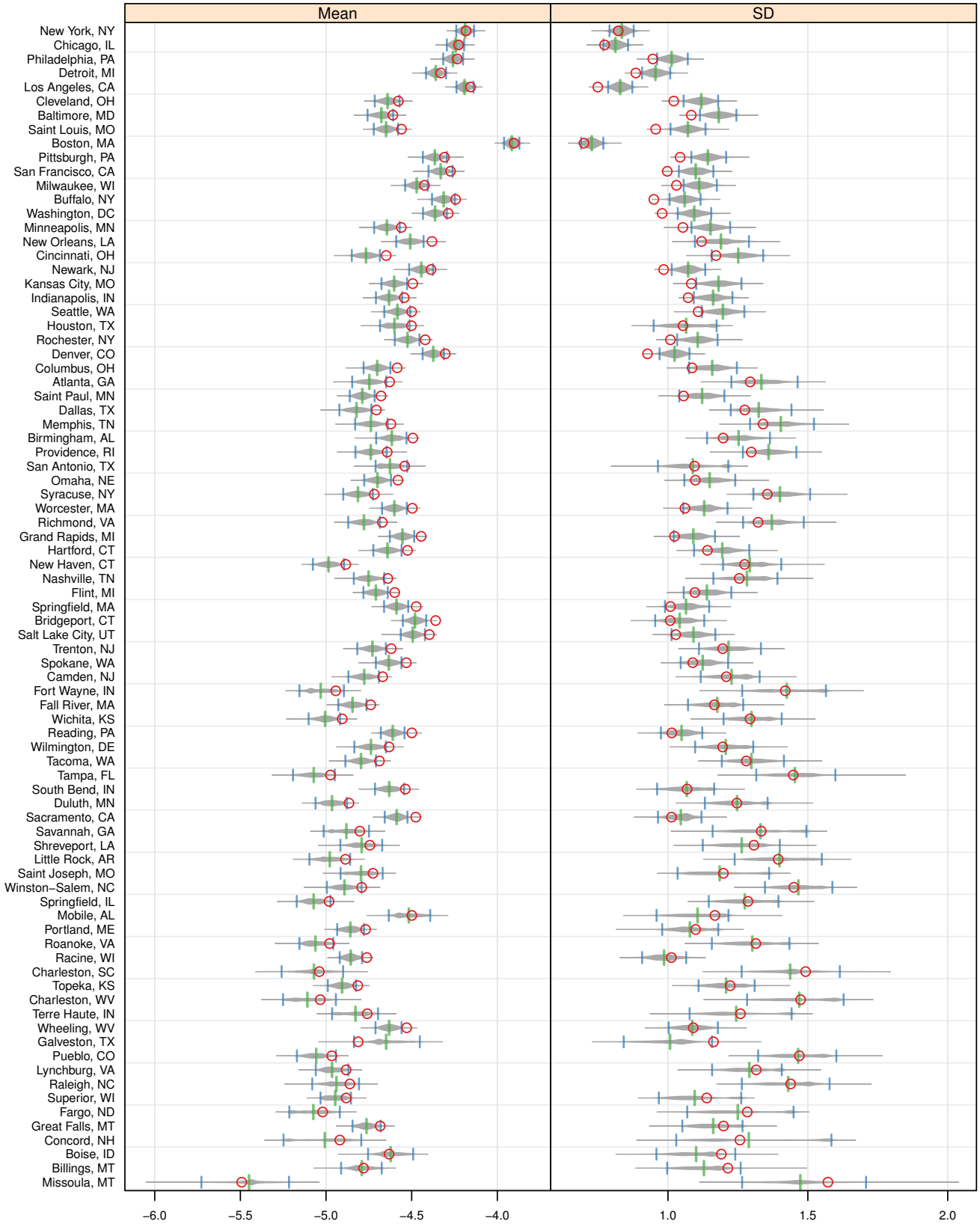


Figure S5. U.S. distribution of parameters inferred from parametric bootstrapped case reports. Red circle shows value inferred from data. Central green line shows bootstrap mean. Blue lines show 95% confidence interval. See Figure S7 for details.

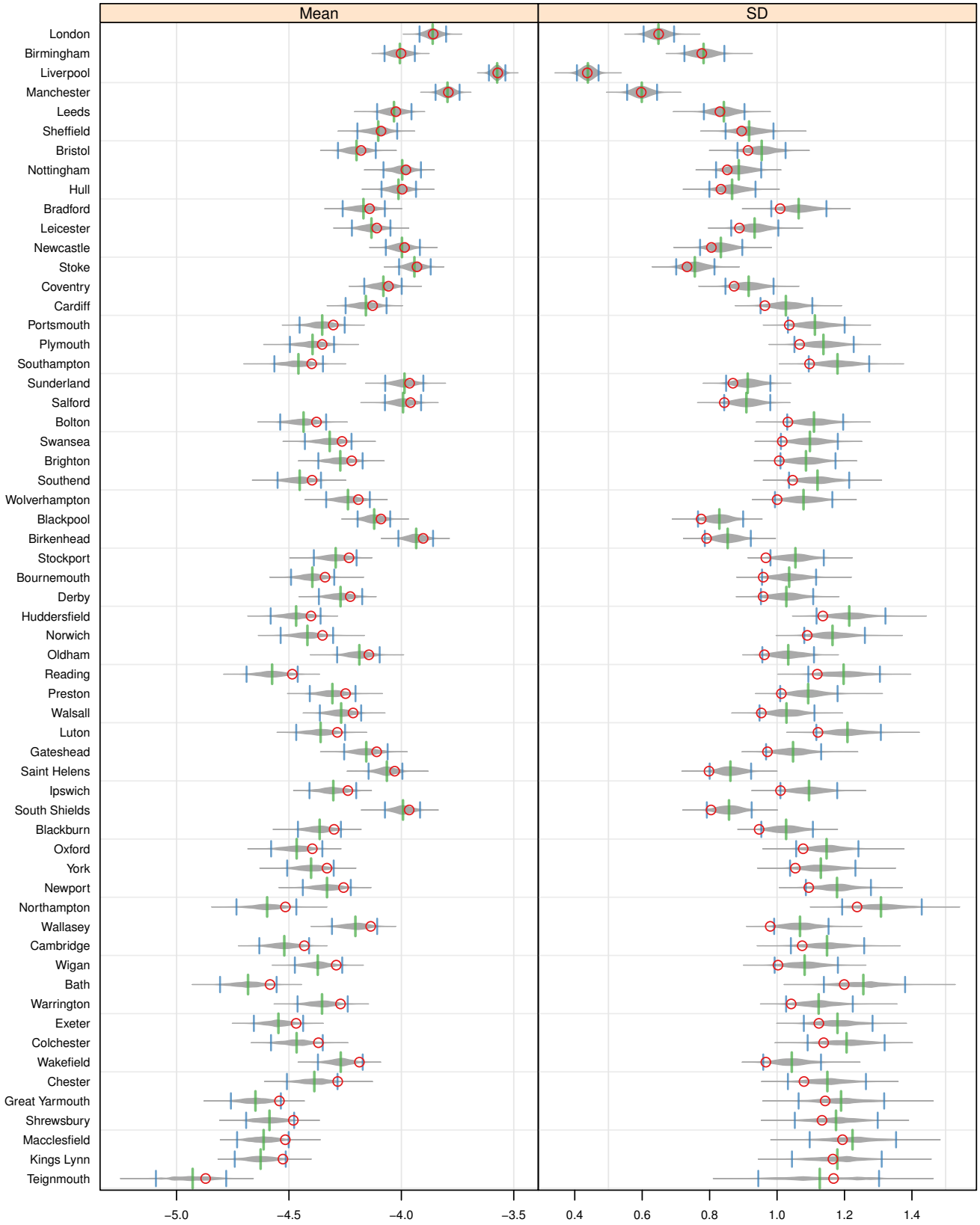


Figure S6. England & Wales distribution of parameters inferred from parametric bootstrapped case reports. Red circle shows value inferred from data. Central green line shows bootstrap mean. Blue lines show 95% confidence interval. See Figure S7 for details.

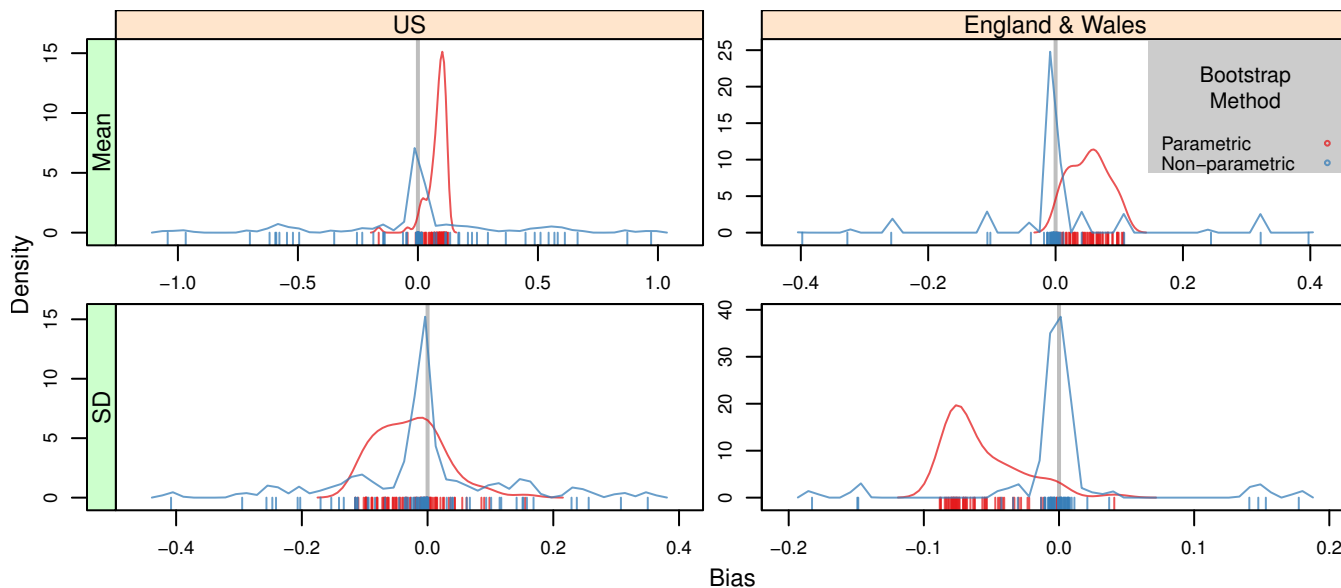


Figure S7. Bias of inferred parameters, as estimated from synthetic (bootstrap) datasets of case reports. For non-parametric bootstraps, case reports were sampled with replacement from original timeseries, excluding NAs. For parametric bootstraps, inferred cases were constructed by sampling ξ_i from $N(\hat{\mu}_i, \hat{\sigma}_i)$. Each ξ_i was anti-logged and multiplied by N_i to yield \hat{C}_i . C_i were obtained from \hat{C}_i via binomial sampling, with a per-trial success rate of r_i . One thousand replicates were evaluated for each process. Bias is the distance between the bootstrap mean and the value inferred from data. Parametric bootstraps show consistent bias with lower variation of bias. The mean bias of non-parametric bootstraps is approximately zero, with very high bias for a small number of populations. See Figures S5 and S6 for parametric results.

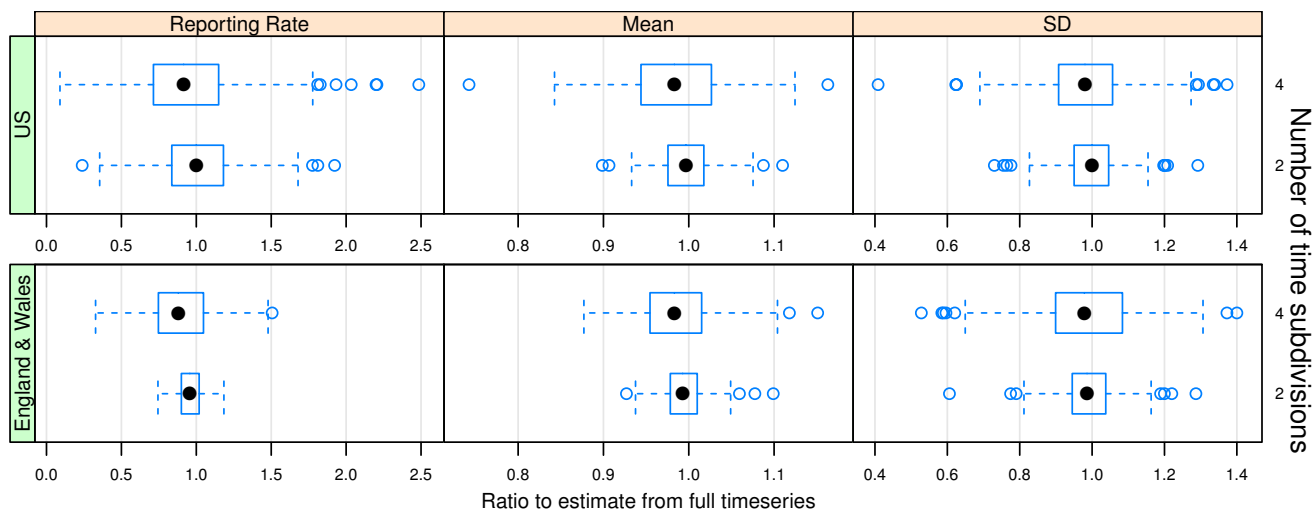


Figure S8. Ratio of inferred estimates from N_t equal-length subdivisions of timeseries to that of full timeseries (~ 20 years for each country), for $N_t = 2$ or 4 . Mean = $\hat{\mu}$ and SD = $\hat{\sigma}$. R_0 is held constant at 20. Reporting rate is the estimate most sensitive to time length, particularly in the U.S. For $N_t = 2$, the inter-quartile range (IQR) of estimated reporting rates is within $\pm 20\%$ of the full length value for the U.S., and within $\pm 10\%$ in England & Wales. This suggests that reporting rate is relatively conserved over time for many populations. For $N_t > 4$ (timeseries less than 5 years), estimates diverge significantly.

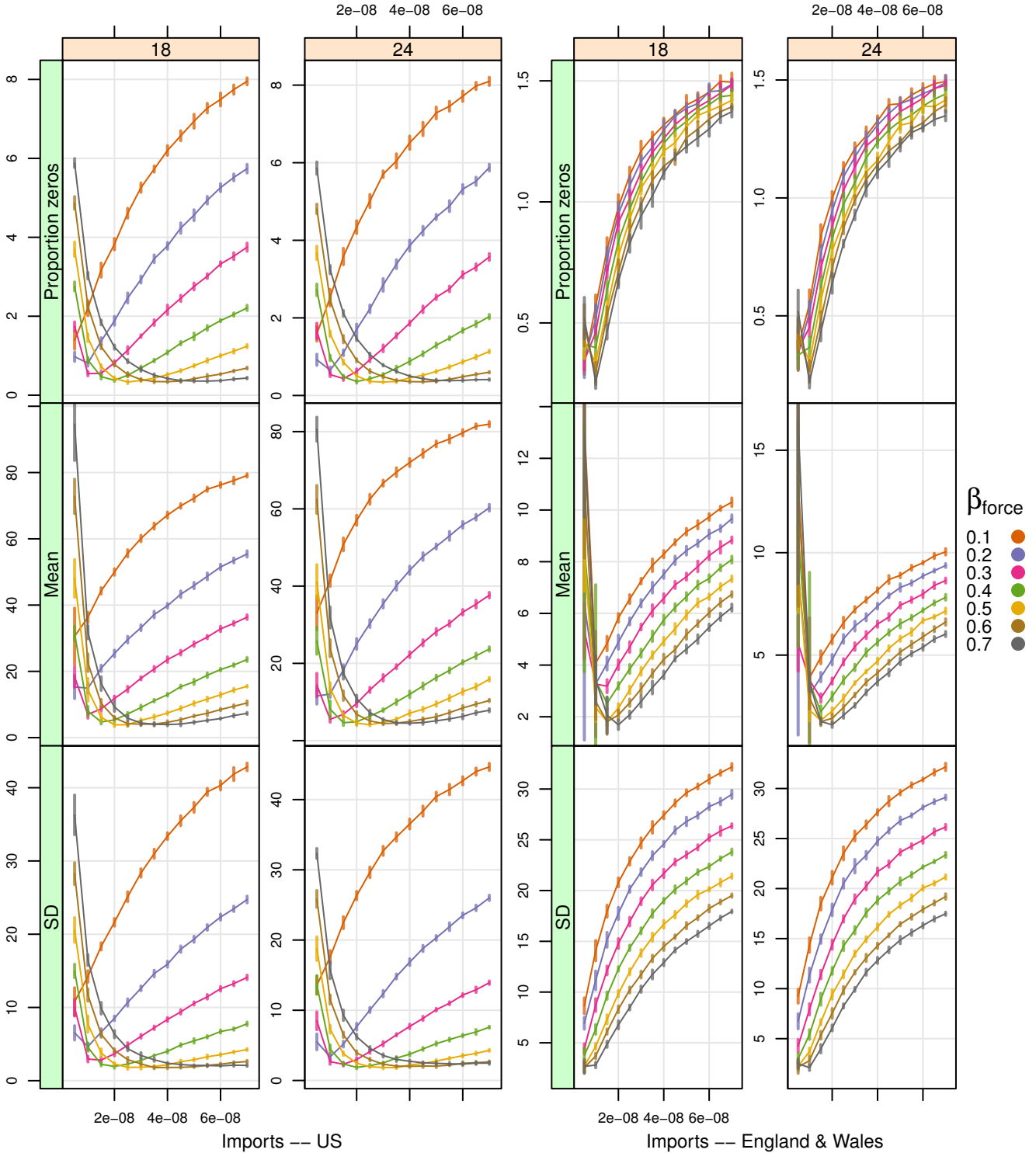


Figure S9. An ensemble of 10 realizations was constructed for each parameter set (columns show R_0 ; see Table S3 for tested parameter ranges). The residual sum of squares (RSS_δ) between model and data was computed for each realization and measure $\delta \in \{P_0, \hat{\mu}, \hat{\sigma}\}$. The ensemble mean RSS_δ is shown above, with vertical bars showing \pm one standard deviation. For each δ , the condition with minimum ensemble mean RSS was identified. A t-test was conducted between the minimum set of RSS and each other condition's RSS set. Any parameter set that was greater than the minimum set with $p < 0.05$ was considered inferior to the best set for that measure, with the remaining sets considered equivalent. A final parameter set was chosen to maximize the number of measures in which it was equivalent to the best condition.

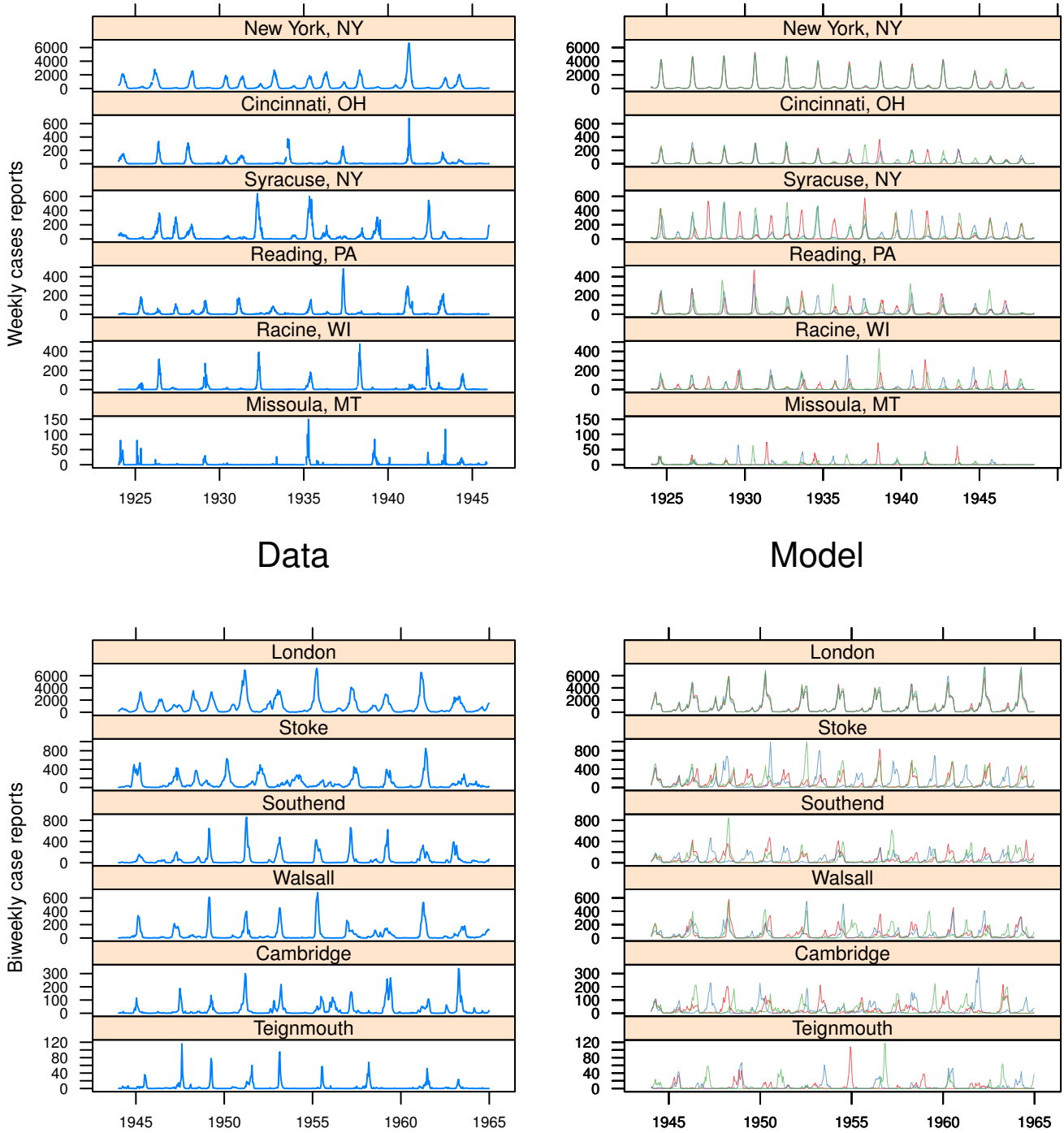


Figure S10. Simulation/data comparison of timeseries for select cities. Cities are ordered by descending population, with evenly-spaced population ranks. Data is shown on the left, and three randomly-selected simulations are shown on the right. To simplify comparison, matching cities share a common Y axis. For model parameters, see Table S4.

Appendix S2 – Epidemiological Model Details

Event	Change in State	Transition Rate
Birth (discounting infant mortality)	$(S, E, I, R) \rightarrow (S + 1, E, I, R)$	$\nu(t)N$
Immigration of susceptible	$(S, E, I, R) \rightarrow (S + 1, E, I, R)$	$\mu_+(t)(\frac{1}{R_0})N$
Emigration of susceptible	$(S, E, I, R) \rightarrow (S - 1, E, I, R)$	$\mu_-(t)(\frac{1}{R_0})N$
Immigration of recovered	$(S, E, I, R) \rightarrow (S, E, I, R + 1)$	$\mu_+(t)(1 - \frac{1}{R_0})N$
Death or emigration of recovered	$(S, E, I, R) \rightarrow (S, E, I, R - 1)$	$\mu_-(t)(1 - \frac{1}{R_0})N$
Exposure due to imports	$(S, E, I, R) \rightarrow (S - 1, E + 1, I, R)$	$\eta \hat{\beta} S \frac{\sum_{j \neq i} I_j}{\sum_j N_j}$
Exposure due to internal dynamics	$(S, E, I, R) \rightarrow (S - 1, E + 1, I, R)$	$\hat{\beta} S \frac{I}{N}$
Infection	$(S, E, I, R) \rightarrow (S, E - 1, I + 1, R)$	σE
Recovery	$(S, E, I, R) \rightarrow (S, E, I - 1, R + 1)$	γI

Table S2. Events and corresponding transition rates in the stochastic *SEIR* model for population i .

Parameter	Range
$\nu(t)$	Specified by demographic data
$\mu(t)$	Specified by demographic data
η	$[10^{-9}, 10^{-6}]$
$R_0 = \frac{\beta_0}{\gamma}$	$[15, 25]$
β_{force}	$[0.1, 0.5]$
$\hat{\beta}$	$\beta_0(1 + \beta_1 \sin(\frac{2\pi t}{365}))$ per day
σ	1/8 per day
γ	1/5 per day

Table S3. Parameter values.

Country	R_0	β_{force}	η	School Forcing Function
US	18	0.50	2.50e-08	Sin
England & Wales	24	0.70	1.00e-08	Term Time

Table S4. Table of final epidemiological model parameters. School term in England & Wales as per Keeling *et al.* (2001).

Each model realization was initialized at the equilibrium values of the equivalent non-seasonal deterministic model and run over years where demographics were available; the U.S. model was run from 1910 through 1949, and the England & Wales model was run from 1944 through 1964. Model results outside the observed time series range were discarded.

We tested a number of model innovations including term-time forcing, seasonal forcing of η , and scaling η with population size. We tested each model innovation over a range of parameter values. Specifically, we varied the transmission rate, strength of seasonal forcing, size of the import term, and R_0 in turn while holding all other parameters constant. For further analysis, we selected the simplest possible model structure as it yielded results comparable to more complex models.