The cohort effect in childhood disease dynamics
ELECTRONIC SUPPLEMENTARY MATERIAL

Daihai He[1,2] **and David J.D. Earn**[2]

[1]Department of Applied Mathematics,
Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong;

[2]Department of Mathematics and Statistics,
McMaster University, Hamilton, Ontario, Canada, L8S 4K1

June 18, 2016 @ 13:09

Contents

1 Data

All the data studied in this paper are available either from the journal website as Electronic Supplementary Material or from the International Infectious Disease Data Archive (IIDDA) at

http://IIDDA.McMaster.ca.

The specific data files are listed below.

1.1 Measles incidence

Measles incidence in the UK were published in the Registrar General's Weekly Returns for England and Wales. Monthly measles notifications in New York City, 1928–1973, were tabulated by London and Yorke [1, 2] (a longer measles time series for NYC was recently digitized and made available by Hempel and Earn [3]).

- Weekly reported measles in England and Wales, 1948–1966:
  meas_uk_ew_1948-66_wk.csv.
- Weekly reported measles in London, England, 1944–1994:
  meas_uk_london_1944-94_wk.csv
- Weekly reported measles in Liverpool, England, 1944–1994:
  meas_uk_liverpool_1944-94_wk.csv
- Monthly reported measles in New York City, USA, 1928–1963:
  meas_us_nyc_1928-63_mn.csv

1.2 Annual population sizes and birth rates

Published annual population size and numbers of live births in England and Wales were obtained from "Population estimates for England and Wales by total persons, males and females – Mid-1838 to Mid-2013" and "Live birth 1938–2013", UK Office for National Statistics (http://www.ons.gov.uk).

- Yearly births in London, UK, 1944–1994:
  bth_uk_lon_1944-94_yr.csv
- Yearly births in Liverpool, UK, 1944–1994:
  bth_uk_lpl_1944-94_yr.csv
- Yearly Birth rate in New York City, USA, 1925–1963:
  bthrt_us_ny_1925-63_yr.csv

1.3 Average population sizes and birth rates

Table S1 gives the average population sizes and average birth rates for the four locations studied in this paper. The source for E&W as a whole and the cities of London and Liverpool is

http://www.visionofbritain.org.uk/.

The source for New York City is the NYC Department of Health's *Summary of Vital Statistics* for 1961 for the city of New York, which can be obtained from:

http://www.nyc.gov/html/doh/downloads/pdf/vs/1961sum.pdf.

Table S1: Population sizes and birth rates during the periods of biennial cycles of measles epidemics examined in this paper.

Place	Population	Birth rate
E&W (UK)	45,000,000	0.02 / yr
London (UK)	3,300,000	0.017 / yr
Liverpool (UK)	760,000	0.021 / yr
NYC (US)	7,800,000	0.021 / yr

2 Realistic Age-Structured (RAS) model

Schenzle's [4] realistic age structured (RAS) model contains 21 age classes, each of which is subdivided into four compartments: susceptible (S_i), exposed (E_i), infectious (I_i), and recovered (R_i), for $i = 0, 1, \ldots, 20$. The population size in each age class is $N_i = S_i + E_i + I_i + R_i$. Births occur into the lowest age class continuously. The simple transmission rate $\beta(t)$ of the SEIR model (**equation** (1)) is replaced by a transmission matrix $\beta_{ij}(t)$, which includes term-time forcing in primary school (ages 6–10). Initial school entry and grade-wise movement of cohorts occur on the first day of school each year.

2.1 RAS model equations

For convenience, we write the force of infection in age class i as

$$\lambda_i = \sum_{j=0}^{20} \beta_{ij} I_j \,, \tag{S1}$$

where β_{ij} is given in **Table S4**. Following Schenzle [4], we assume that deaths occur only in the oldest age class (20+). Newborns enter S_0. Thus, for age-class-0

$$\dot{S}_0 = \nu N_{20} - \lambda_0 S_0 \tag{S2a}$$
$$\dot{E}_0 = \lambda_0 S_0 - \sigma E_0 \tag{S2b}$$
$$\dot{I}_0 = \sigma E_0 - \gamma I_0 \tag{S2c}$$
$$\dot{R}_0 = \gamma I_0 \tag{S2d}$$

For age-class-1 to age-class-19

$$\begin{aligned} \dot{S}_i &= -\lambda_i S_i && \text{(S3a)} \\ \dot{E}_i &= \lambda_i S_i - \sigma E_i && \text{(S3b)} \\ \dot{I}_i &= \sigma E_i - \gamma I_i && \text{(S3c)} \\ \dot{R}_i &= \gamma I_i && \text{(S3d)} \end{aligned}$$

Finally, in age-class-20+ there are deaths as well:

$$\begin{aligned} \dot{S_{20}} &= -\lambda_{20} S_{20} - \mu S_{20} && \text{(S4a)} \\ \dot{E_{20}} &= \lambda_{20} S_{20} - \sigma E_{20} - \mu E_{20} && \text{(S4b)} \\ \dot{I_{20}} &= \sigma E_{20} - \gamma I_{20} - \mu I_{20} && \text{(S4c)} \\ \dot{R_{20}} &= \gamma I_{20} - \mu R_{20} && \text{(S4d)} \end{aligned}$$

On the first day of school each year, we have

$$\begin{aligned} S_0 &= 0 && \text{(S5a)} \\ E_0 &= 0 && \text{(S5b)} \\ I_0 &= 0 && \text{(S5c)} \\ R_0 &= 0 && \text{(S5d)} \end{aligned}$$

and for $i = 1, \ldots, 19$,

$$\begin{aligned} S_i &= S_{i-1} && \text{(S6a)} \\ E_i &= E_{i-1} && \text{(S6b)} \\ I_i &= I_{i-1} && \text{(S6c)} \\ R_i &= R_{i-1} && \text{(S6d)} \end{aligned}$$

Finally,

$$\begin{aligned} S_{20} &= S_{20} + S_{19} && \text{(S7a)} \\ E_{20} &= E_{20} + E_{19} && \text{(S7b)} \\ I_{20} &= I_{20} + I_{19} && \text{(S7c)} \\ R_{20} &= R_{20} + R_{19} && \text{(S7d)} \end{aligned}$$

2.2 Initial conditions

The initial proportions of the population in each age class are given in **Table S2**. This initial age distribution is based on the published age structure of the UK in 1971:

> http://www.ons.gov.uk/ons/interactive/uk-population-pyramid---dvc1/index.html

This initial setting turns out to be irrelevant, since Schenzle's [4] model approaches a different equilibrium age structure: All age classes (except for 20+) converge to an equilibrium population size x, where $x = \nu(N - 20x)$ and hence [5]

$$x = \frac{N}{20 + 1/\nu} = 642857\,. \qquad \text{(S8)}$$

The initial conditions for the state variables are

$$S_i = 0.001N\,, \tag{S9a}$$
$$E_i = I_i = 0.000001N\,, \tag{S9b}$$
$$R_i = N_i - S_i - E_i - I_i\,, \tag{S9c}$$

where N_i is the initial proportion of the population in age-class i (**Table S2**).

Table S2: Initial proportions of the population in each age class.

Age Classes	Symbols	Percentage
0–4	$N_0, \ldots, N_3$	1.6
4–9	$N_4, \ldots, N_8$	1.7
9–11	N_9, N_{10}	1.6
11–13	N_{11}, N_{12}	1.5
13–20	$N_{13}, \ldots, N_{19}$	1.4
21+	N_{20}	69

2.3 Parameter values

Several fixed parameter values are listed in **Table S3**.

Table S3: Fixed parameters for the RAS model.

Parameter	Symbol	Value
Population size	N	45 million
Birth rate	ν	$0.02\,\mathrm{yr}^{-1}$
Death rate	μ	$0.02\,\mathrm{yr}^{-1}$
Mean latent period	σ^{-1}	8 days
Mean infectious period	γ^{-1}	5 days

Since the number of parameters in the RAS model is large (441 matrix entries β_{ij}), simplifying assumptions are typically made to reduce the number of free parameters. We follow previous work ([4], [6], [7, p. 85]) and assume that β_{ij} is a symmetric matrix with identical entries for sequences of age classes (**Table S4**, multiplied by an overall factor b_0). The function $b(t)$ in the second diagonal entry of **Table S4** is

$$b(t) = \begin{cases} b_2 & \text{during school terms,} \\ 0 & \text{otherwise.} \end{cases} \tag{S10}$$

Our best-fit parameter values are given in **Table S5**. Note that the best-fit RAS model yields $E_V = 0.486$, whereas the best-fit term-time cohort-entry model yields $E_V = 0.229$.

Table S4: Transmission matrix structure for the RAS model. Each entry is actually multiplied by the factor b_0 given in **Table S5**.

	age 0–5	age 6–9	age 10–19	age 20+
age 0–5	2.175	2.175	0.975	0.6
age 6–9	2.175	$2.175 + b(t)$	0.975	0.6
age 10–19	0.975	0.975	0.975	0.6
age 20+	0.6	0.6	0.6	0.6

Table S5: Fitted parameters for the RAS model.

Parameter	Symbol	Value
Overall transmission factor	b_0	398.6
School transmission enhancement	b_2	18.23
Reporting ratio	η	0.61

3 Alternative formulation of cohort entry model

The total population size, $N = S + E + I + R$, does not appear explicitly in **equation** (1). If the population size is constant (so $B = \mu N$) then it has no dynamical effect. If $B = \nu N$ (where ν is the *per capita* birth rate) and $\beta = \beta'/N$ (where β' is constant) then the equations for the proportions of individuals in each compartment (S/N, E/N, *etc.*) do not contain N so, again, population size has no dynamical effect [8]. If B is constant but $B \neq \mu N$ then population size *does* affect the dynamics. Which formulation is appropriate depends on the biological context and the questions that are being addressed. For childhood infectious diseases in modern cities, **equation** (1) with constant B ($B \neq \mu N$) is the most successful in terms of predicting qualitative dynamical changes in observed incidence time series [9–11].

Note that if we re-express the SEIR model in terms of proportions of the population in each compartment then B appears instead of μ in the mortality terms of **equation** (1) [8], in which case there are several places where we need to substitute $B \to \widetilde{B}(t)$. He [12] investigated this alternative formulation of the cohort entry model and found results that are qualitatively identical and quantitatively nearly identical as those presented in this paper. In particular, the topology of the bifurcation tree (**Figure 5**(c)) is identical and the positions ($\mathcal{R}_0$) of the bifurcations differ only slightly. See **Figure S5**(a).

As mentioned in the introduction, the cohort-entry model is similar to a two-age-class limit of the RAS model. He [12, Figure 3.19, p. 101] compared the cohort-entry model with a two-age-class model and found that they had similar, but not identical, dynamics.

4 Transmission rate reconstruction

Figure S1 and **Figure S2** show the equivalents of **Figure 2** and **Figure 3** for the SEIR model with term-time forcing but without the cohort effect (**Figure S1**) and with the cohort effect but without any seasonal forcing (**Figure S2**).

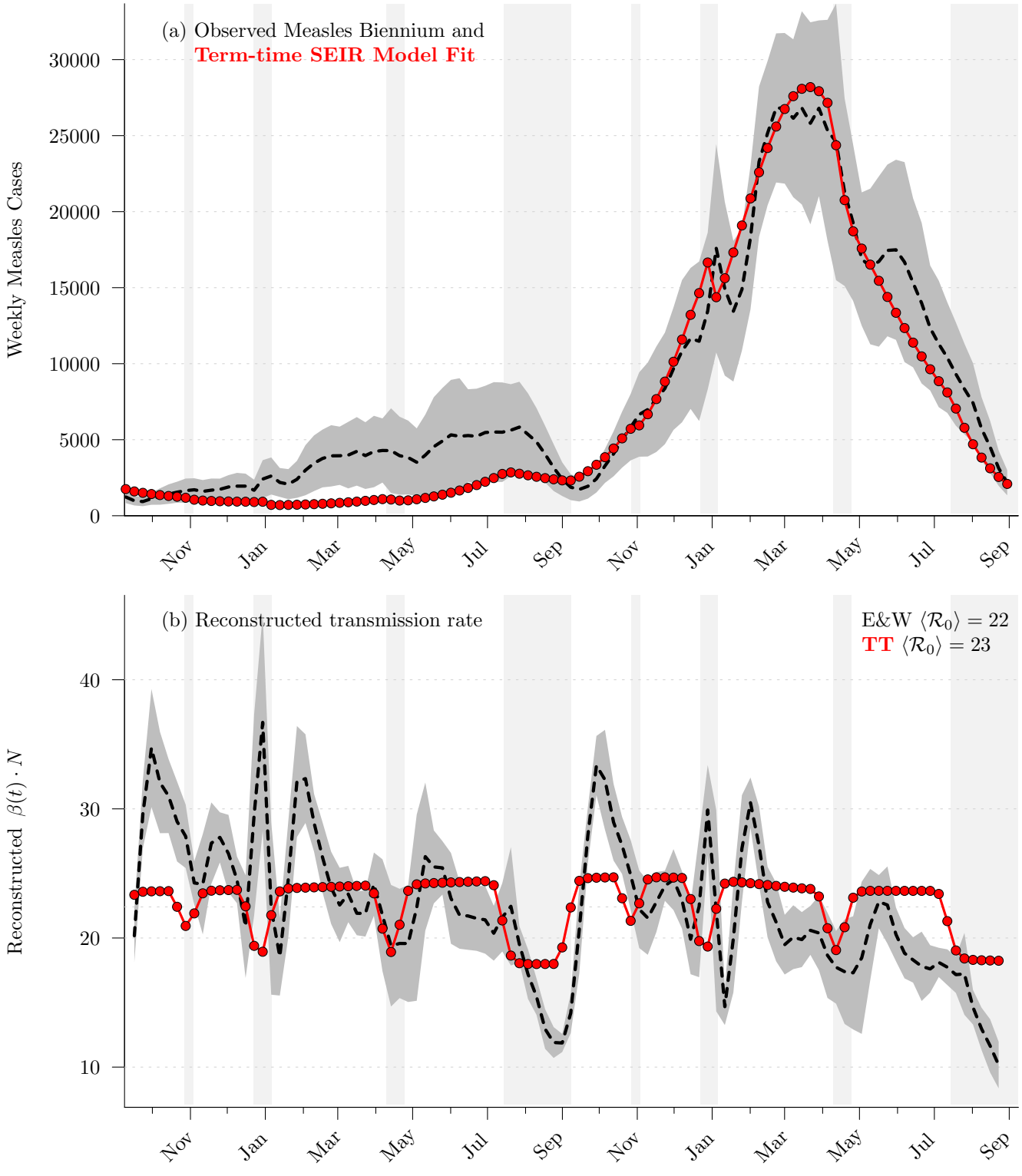

Figure S1: Equivalent of **Figure 2** for the SEIR model with term-time forcing.

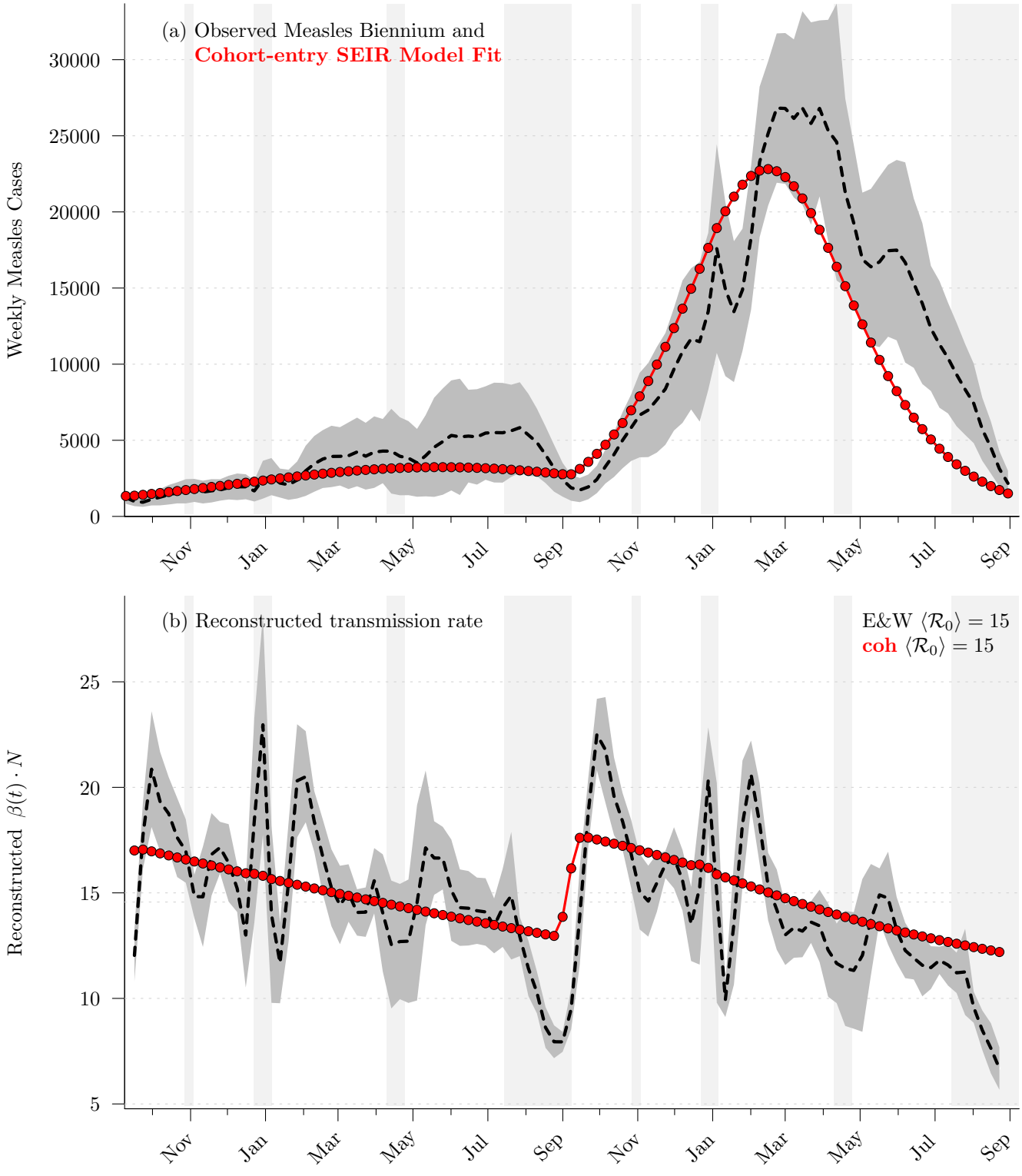

Figure S2: Equivalent of **Figure 2** for the cohort entry SEIR model.

5 City-level measles and demographic data

Figure 1 shows the aggregate E&W weekly measles data studied by Fine and Clarkson [13], together with annual rates of birth and susceptible recruitment. **Figure S3**(a) shows the E&W aggregate data again, while **Figure S3**(b–d) show the corresponding data for three individual cities (London and Liverpool in the the UK and New York City).

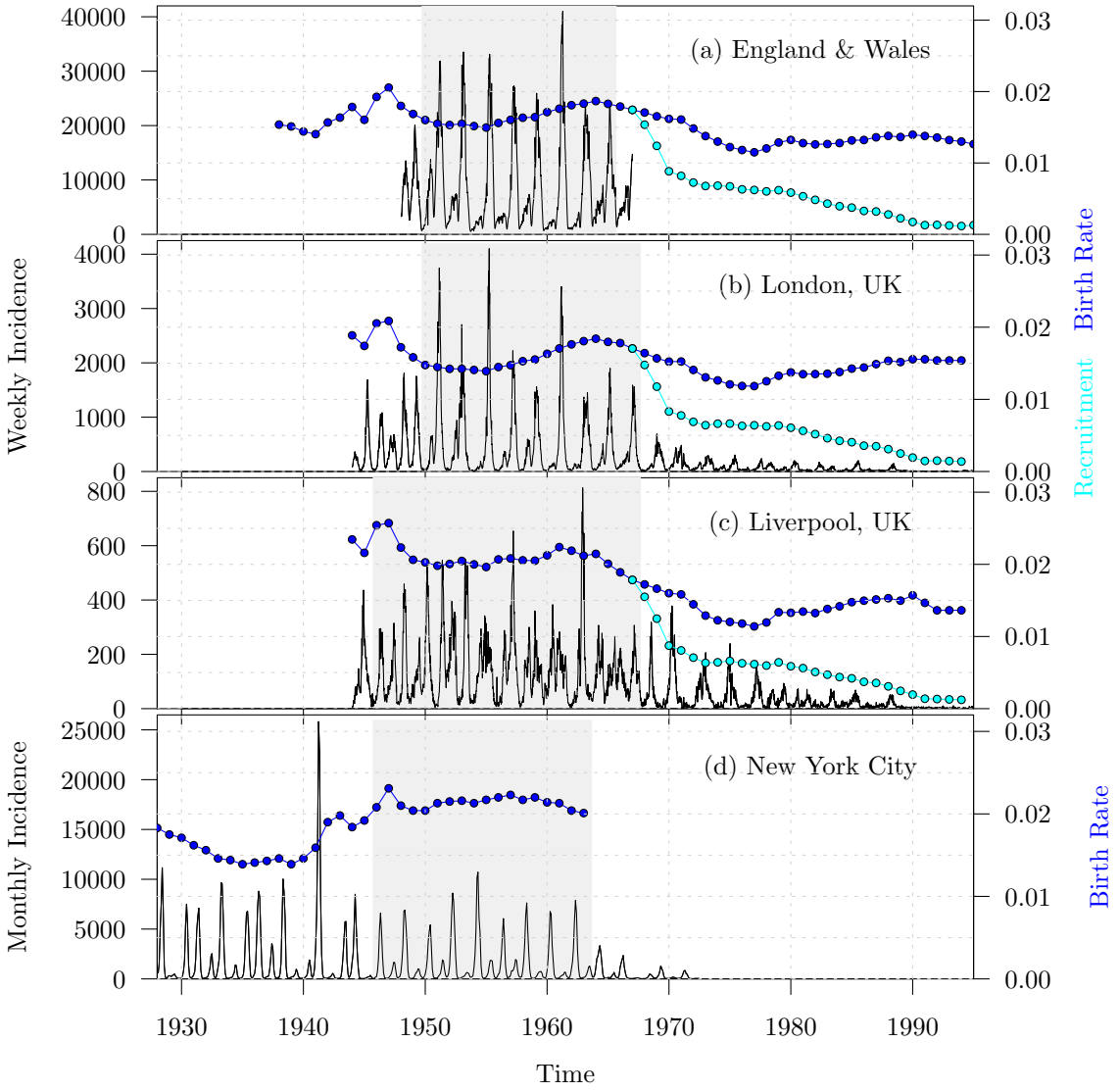

Figure S3: Measles incidence and concurrent demographic data in England and Wales and three individual cities. The four panels show measles incidence (black solid curve, weekly for the UK and monthly for the US), annual birth rate (blue) and susceptible recruitment (cyan) in England & Wales, London (UK), Liverpool (UK) and New York City (US), respectively. In each panel, the time interval that we used to compare with the biennial attractor of various models is highlighted with grey shading.

6 City-level measles biennium fits

Figure S4 shows the equivalent of **Figure 4** for the three city-level measles time series shown in **Figure S3**.

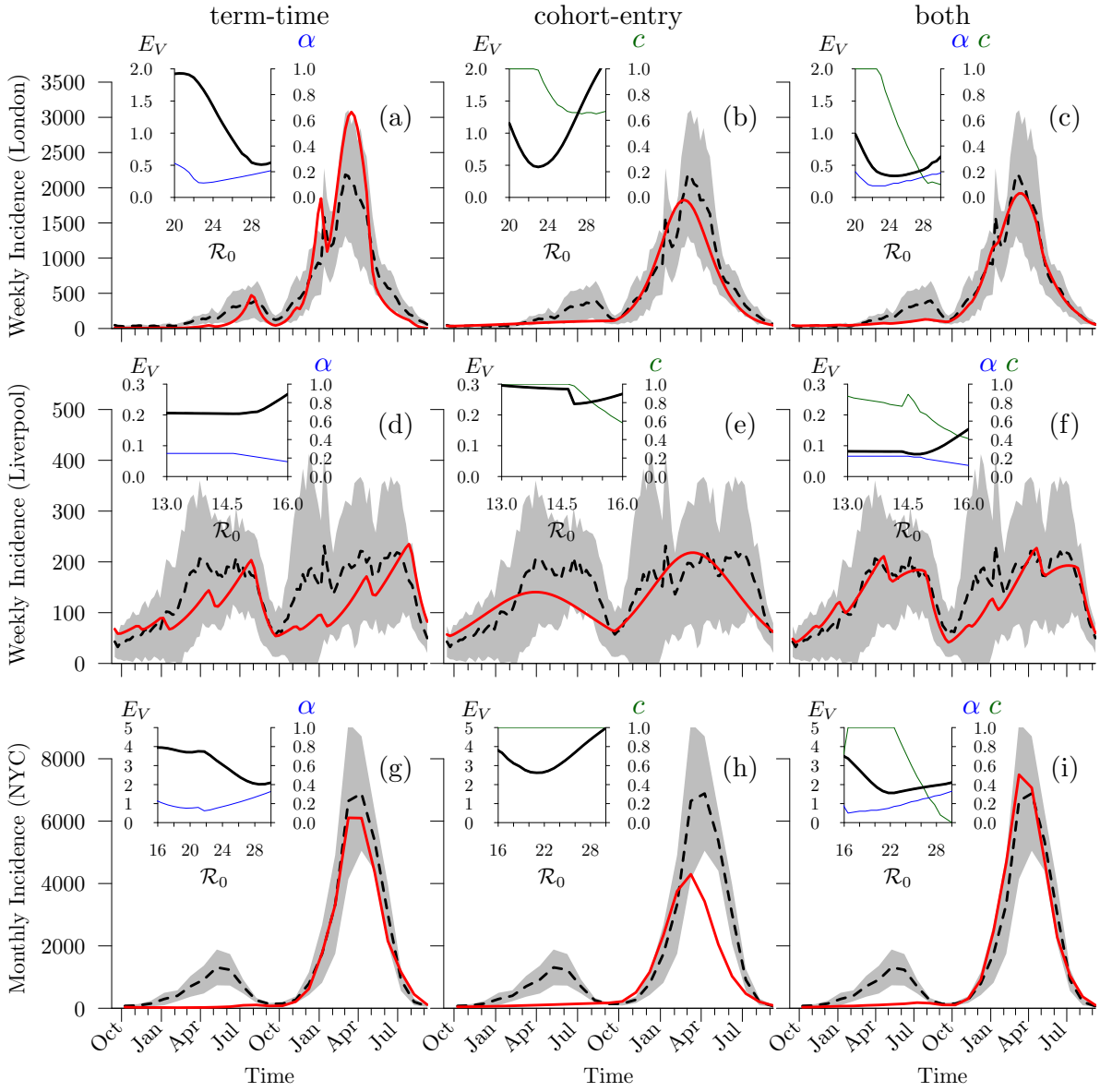

Figure S4: Fitting models by trajectory matching to measles incidence data for three cities (*cf.* **Figure 4** for the equivalent for the aggregated E&W measles data). From top to bottom: London (UK), Liverpool (UK), and New York City (US). From left to right: term-time forcing only, cohort-entry only, and both effects. In all of nine main panels, the black dashed curve shows the observed average measles biennium during the period highlighted with grey shading in **Figure S3**, while the red solid curve shows the fitted model simulation. The estimated parameter values are given in **Table S6**. The inset in each panel shows the fitting error E_V (**equation** (9)) as a function of $\mathcal{R}_0$ (black solid curve) and the estimated term-time forcing amplitude (α, blue) or cohort-entry proportion (c, green). School days in E&W, London, and Liverpool were [7,100] (*i.e.,* from 7th to 100th day of the year), [115,199], [252,300], and [308,356]; while in New York City, school days were [3,48], [58,102], [114,179], and [251,357] [14].

7 Parameter estimates and fitting errors

Table S6 shows our best-fit parameter estimates and associated fitting errors for the fits of the various models to the measles biennium in the locations we studied. In all places, including the cohort effect substantially reduces the fitting error (E_V, §**2.3.4**). In addition, the differences in AIC allow us to infer that these improvements are significant.

Table S6: Parameter estimates obtained by fitting epidemic models to measles notification time series in four locations. Parameters that are not fitted are indicated with '–'.

Place	Model	c	α	$\mathcal{R}_0$	η	E_V	ΔAIC
E&W (UK)	term-time (fixed $\mathcal{R}_0 = 17$)	–	0.25	–	0.25	2.867	270.4
	term-time	–	0.15	22.6	0.44	0.813	58.7
	cohort-entry	1.00	–	17.2	0.43	0.732	50.3
	both	1.00	0.16	17.1	0.49	0.229	0.0
	RAS	–	–	13.4*	0.61	0.486	24.7
London (UK)	term-time	–	0.19	29.0	0.41	0.510	16.3
	cohort-entry	1.00	–	22.9	0.40	0.473	12.5
	both	0.65	0.11	24.5	0.42	0.334	0.0
Liverpool (UK)	term-time	–	0.24	14.9	0.40	0.201	11.3
	cohort-entry	0.98	–	14.8	0.45	0.236	15.0
	both	0.81	0.21	14.6	0.47	0.073	0.0
NYC (US)	term-time	–	0.28	28.5	0.10	2.017	45.3
	cohort-entry	1.00	–	21.2	0.07	2.611	107.1
	both	1.00	0.17	22.5	0.11	1.562	0.0

* For the RAS model, $\mathcal{R}_0$ was estimated as the mean of the reconstruction in **Figure 2**. The fitted RAS parameters are b_0, b_2 and η; see **Table S5**.

For England and Wales as a whole and for New York City, the best-fit cohort proportion is $c = 1$. Consequently, a further model selection step could be performed, fitting reduced models in which c and α are each equal to 0 or 1. However, this step would not change which model is selected.

For all the models we considered, finding the best-fit parameters listed in **Table S6** uniquely determined a solution, because there was always a unique biennial attractor for the best-fit parameter values. (Coexisting stable cycles do exist for the cohort model [12, Figures 3.8 and 3.9, p 73], but we have found no evidence of coexisting, distinct biennial attractors.) The proportion susceptible, $S(0)$, at the start of the fitted biennial cycle, is listed for each model in **Table S7**.

Table S7: Initial susceptible proportion $S(0)$ from trajectories fitted to E&W measles. If the system were in equilibrium then the proportion susceptible would be $1/\mathcal{R}_0$.

Model	$\mathcal{R}_0$	$1/\mathcal{R}_0$	$S(0)$
Term-time SEIR	17 (fixed)	0.059	0.053
Term-time SEIR	22.6	0.044	0.037
Cohort-entry SEIR	17.2	0.058	0.063
Both	17.1	0.058	0.062
RAS	13.4	0.075	0.067

8 Dynamical structure in the (α, c) plane with different $\mathcal{R}_0$

Figure 5(d) summarizes the key features of the dynamical structure of the cohort entry SEIR model with term-time forcing for the specific basic reproduction number $\mathcal{R}_0 = 17$. **Figure S5** shows the same information for other values of $\mathcal{R}_0$.

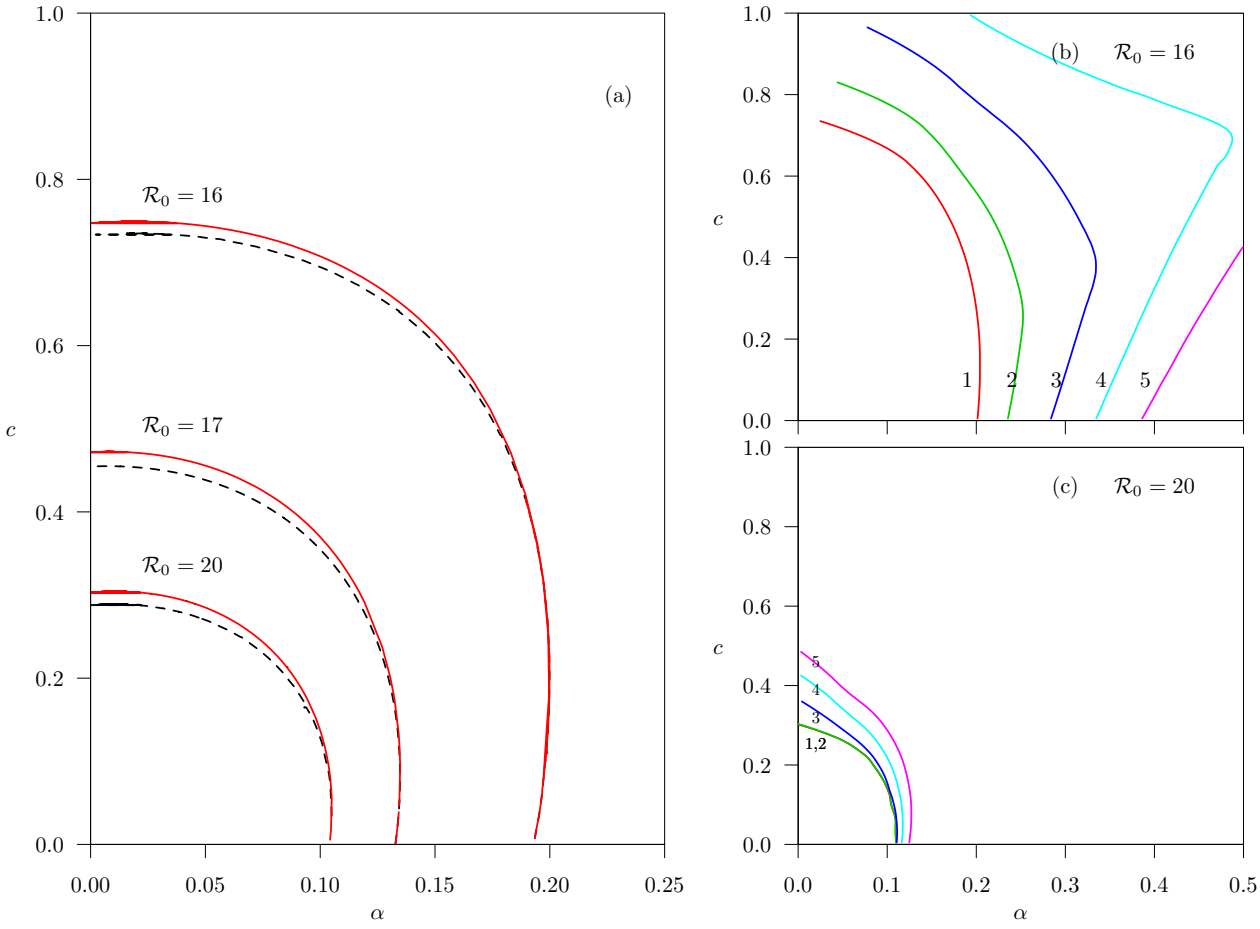

Figure S5: The period-doubling (from annual to biennial) bifurcation curve in the two-dimensional parameter plane (term-time forcing amplitude α *versus* cohort proportion c) with different $\mathcal{R}_0$. (a) The red solid curves show the case of cohort forcing on both birth and death terms, while the black dash curves show the case of cohort forcing on birth term only (*cf.* §3). Panels (b) and (c) show the peak height ratio contour curve associated with biennial cycles for (b) $\mathcal{R}_0 = 16$ and (c) $\mathcal{R}_0 = 20$. The five contour curves in each panel correspond to peak height ratio (major year over minor year) of 1 to 5 as in **Figure 5**(d).

9 Transient dynamics of the cohort entry SEIR model

As mentioned in §3.3, in addition to the topological invariance of the principal bifurcation tree (which summarizes the asymptotic dynamics of the model), the transient dynamics of the cohort-entry model are also qualitatively identical to those of the term-time forced model. **Figure S6** shows the period of the transient oscillations that damp out on a stable annual or biennial cycle in the term-time SEIR model (a) and the cohort entry SEIR model (b).

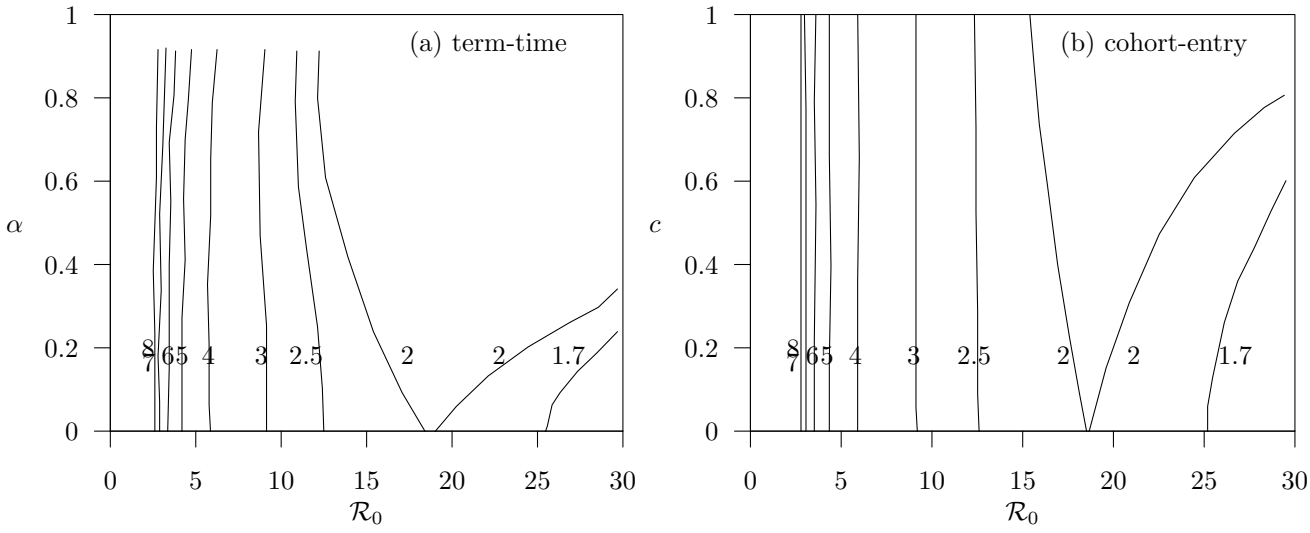

Figure S6: A comparison of contour plots of constant natural damping frequency $T_{f,1}$ in the parameter planes of transmission rate seasonality α versus $\mathcal{R}_0$ for the time-time SEIR model (a) and cohort entry proportion c versus $\mathcal{R}_0$ for the cohort-entry SEIR model (b). The natural damping period contour curves $T_{f,1} = 1.7$, 2, 2.5, 3, 4, 5, 6, 7, and 8 years are shown in both panels; between the two $T_{f,1} = 2$ curves there is a $T_{f,1} = 2$ phase-locked region. Panel (a) reproduces Figure 1a of [10].

References

[1] London, W, Yorke, JA (1973) Recurrent outbreaks of measles, chickenpox and mumps. I. Seasonal variation in contact rates. *American Journal of Epidemiology* 98:453–468.

[2] Yorke, JA, London, W (1973) Recurrent outbreaks of measles, chickenpox and mumps. ii. systematic differences in contact rates and stochastic effects. *American Journal of Epidemiology* 98:468–482.

[3] Hempel, K, Earn, DJD (2015) A century of transitions in New York City's measles dynamics. *Journal of the Royal Society Interface* 12:20150024.

[4] Schenzle, D (1984) An age-structured model of pre- and post-vaccination measles transmission. *IMA Journal of Mathematics Applied in Medicine and Biology* 1:169–191.

[5] Bolker, B (1993) Chaos and complexity in measles models: A comparative numerical study. *IMA Journal of Mathematics Applied in Medicine and Biology* 10:83–95.

[6] Bolker, BM, Grenfell, BT (1993) Chaos and biological complexity in measles dynamics. *Proceedings of the Royal Society of London, Series B, Biological Sciences* 251:75–81.

[7] Keeling, MJ, Rohani, P (2008) *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, Princeton, New Jersey).

[8] He, D, Earn, DJD (2007) Epidemiological effects of seasonal oscillations in birth rates. *Theoretical Population Biology* 72:274–291.

[9] Earn, DJD, Rohani, P, Bolker, BM, Grenfell, BT (2000) A simple model for complex dynamical transitions in epidemics. *Science* 287:667–670.

[10] Bauch, CT, Earn, DJD (2003) in *Dynamical Systems and Their Applications in Biology*, Fields Institute Communications, eds Ruan, S, Wolkowicz, G, Wu, J (American Mathematical Society, Toronto) Vol. 36, pp 33–44.

[11] Earn, DJD (2009) in *Mathematical Biology*, IAS/Park City Mathematics Series, eds Lewis, MA, Chaplain, MAJ, Keener, JP, Maini, PK (American Mathematical Society) Vol. 14, pp 151–186.

[12] He, D (2006) Phd (McMaster University, Canada).

[13] Fine, PEM, Clarkson, JA (1982) Measles in England and Wales — I: an analysis of factors underlying seasonal patterns. *International Journal of Epidemiology* 11:5–14.

[14] Bauch, CT, Earn, DJD (2003) Transients and attractors in epidemics. *Proceedings of the Royal Society of London, Series* B 270:1573–1578.