Supplementary Materials for "Comparing drivers and dynamics of tuberculosis (TB) in California, Florida, New York and Texas."

Sourya Shrestha[1,*], Andrew N. Hill[2], Suzanne M. Marks[2], David W. Dowdy[1]

[1]Department of Epidemiology, Johns Hopkins School of Public Health, Baltimore, MD 21205, USA
[2]Center for Disease Control and Prevention, Atlanta, GA 30329,USA

[*]To whom correspondence should be addressed; E-mail: sourya@jhu.edu.

Model Details

Demography

The population in each of the state was modeled to have a constant *per capita* birth rate and an age-specific mortality rate, based on a Siler mortality model [1], where the age-specific mortality rate of an individual of age a was $\mu(a) = \mu_a \exp(-\mu_b a) + \mu_c + \mu_d \exp(\mu_f a)$. Parameters μ_a, μ_b, μ_c, μ_d, and μ_f were fit separately for each of the four states to match the age-specific population distributions of foreign-born and US-born populations. Data and model fits are compared in Fig. 2 in the main text, and the parameter estimates are presented in Table E-1.

Immigration

We modeled the immigration of foreign-born individuals explicitly as arrivals of new individuals into the population. The rate of immigration was taken to be constant over time and fit separately for each of the four states such that the resulting size of foreign-born populations were consistent with the size of foreign-born as observed in the American Community Survey of 2014 [2]. The fitted immigration rates are presented in Table E-1, and proportions of foreign-born populations in data and model-simulations are compared in Fig. 2 in the main text. Additionally, nativity and age of the immigrants were also fit to match the age and nativity distributions in the data. To fit the nativity distribution, we set the probability of a new immigrant being from a given region to be equal to the proportion of the immigrants from the same region. To fit the age distribution of the immigrations in the population, we assumed the age distribution of the arrivals to be given by the following function: age $\sim 100 \times \text{Beta}(a_\alpha, a_\beta)$. The parameters corresponding to the fitted distribution are presented in Table E-1, and age-distribution of foreign born individuals in data and model simulations are compared in Fig. 2 in the main text.

Importation of TB

We model importation of TB infection as a function of the region of the origin of the immigrant, and the age of the immigrant. Consider an immigrant of age a from a region with TB prevalence of A^*. To estimate the likelihood that this immigrant has TB infection, we first estimate the annual rates of TB infection in the region of origin that is consistent with the prevalence in that region. We do so using a simple compartmental TB transmission model as described by the set of differential equations below [3].

$$\begin{aligned}
\frac{\mathrm{d}U}{\mathrm{d}t} &= \mu\, L + \mu_A\, A - \lambda\, U \\
\frac{\mathrm{d}L}{\mathrm{d}t} &= [1-p]\,\lambda\, U - [\phi + \mu + p\,\xi\,\lambda]\, L + \omega\,\tau\, A \\
\frac{\mathrm{d}A}{\mathrm{d}t} &= p\,\lambda\,[U + \xi\, L] + \phi\, L - [\omega\,\tau + \mu_A]\, A
\end{aligned}$$

Here, we model the transitions between the compartments U (consisting of uninfected individuals), L (consisting of individuals of Latent TB infections), and A (consisting of individuals with active TB disease). The default model parameters are: background mortality rate, $\mu = 0.02$ per year, TB-related mortality rate, $\mu_A = 0.1$ per year, probability of rapid progression, $p = 0.14$, reactivation rate, $\phi = 0.0015$ per year, average duration of TB infection, $\omega = 1$ years, TB treatment success probability, $\tau = 0.95$, degree of partial immunity imparted by previous infection, $1 - \xi = 0.67$, in a standard population of $N = 100000$.

The force of infection (i.e., the risk of TB faced by an uninfected individual) in this model is $\lambda = \beta A$. The steady-state analysis of this model yields a solution for transmission rate β for a given level of TB prevalence A^* at the steady state, which is as follows:

$$\beta(A^*) = \frac{-q_B + \sqrt{q_B^2 - 4\,q_A\,q_C}}{2\,q_A},$$

where,

$$\begin{aligned}
q_A &= p\,\xi\,(\frac{N}{A^*} - 1) \\
q_B &= \frac{N\,(\phi + \mu\,p + \mu)}{A^{*2}} - \frac{(p\,\xi\,\mu_A + (1-p)\,\mu_A + \omega\,\tau + \phi + \mu\,p)}{A^*} \\
q_C &= -\frac{(\mu_A\,\phi + \mu\,(\mu_A + \omega\,\tau))}{A^{*\,2}}
\end{aligned}$$

Furthermore, the force of infection consistent with a steady-state prevalence of A^*, $\lambda(A^*) = \beta(A^*)\,A^*$. Now, assuming that an individual of age a is exposed to TB infection at the intensity of $\lambda(A^*)$, the probability that the individual is infected by age a is hence $1 - \exp(-a\,\lambda(A^*))$.

Relevant TB prevalence for all 8 regions were estimated for three time points, 1990, 2000 and 2010 (See Table E-2). Country level TB prevalence for Mexico, China and India were based on WHO estimates. Prevalence for Europe and Afica are based on very rough estimate of TB prevalence in the region. Prevalence for Asia (excluding China and India) and Latin America (excluding Mexico and Canada) were based on weighted average of TB prevalence in some of the prominent countries in those regions, where the weights reflect the size of the immigrants from the country, hence accounting for state-level differences in the composition. For estimates in Asia (excluding China and India) we included TB prevalence estimates in South Korea, Philippines and Vietnam, for estimates in Latin America (excluding Mexico and Canada) we included TB prevalence estimates in Cuba, Dominican Republic, Haiti, Jamaica, El Salvador, Colombia, Ecuador, Peru and Guyana. For the estimate for rest of the world, we used the TB prevalence in Canada.

We used the prevalence estimates in 1990 for the pre-1993 "burn-in" phase of the model, prevalence estimates in 2000 for model estimates between 1993-2013, and prevalence estimates in 2010 for model projections between 2013-2025. The estimated age- and origin-specific probabilities of being infected with TB at arrival, for each of the four states in the three time periods are shown in Fig. E-1.

Natural history of TB

The natural history of tuberculosis in an individual was captured via transitions of the individuals between four compartments: Uninfected, LTBI, Active TB, and Post-treatment. We assumed the model population to be homogeneously mixing regardless of age or nativity. Hence the force of infection of TB infection (TB hazard rate) on an individual is the same regardless of age or nativity. We modeled the transmission of of TB infection to have secular trends. In particular, the transmission rate of TB $\beta = \beta_1 \exp(-\beta_2\,t)$, where β_2 is the annual rate of decline of transmission.

We modeled the reactivation rate to be age-dependent, and to decrease over time since the time of infection, similar to the model by Vynnycky et al. [4]. Mathematically, the reactivation rate of latent TB acquired τ years ago, in an individual of age a, $\phi(\tau, a) = \phi_a\,(\phi_3 + \exp(-\phi_4\,\tau))$, where, ϕ_a is:

$$\phi_a = \begin{cases} \phi_1, & \text{if age, } a < 10 \text{ years} \\ \phi_2, & \text{if age, } a \geq 20 \text{ years} \\ \phi_1 + \frac{(\phi_2 - \phi_1)\,(a-10)}{10}, & \text{otherwise} \end{cases}$$

We defined long-term reactivation rate as $\lim_{\tau\to\infty}\phi(\tau,a)=\phi_a\,\phi_3$, and recent transmissions as infections that occurred at most two years in the past. Individuals with prior history of TB infection (those that are currently infected with latent TB, and those that have been successfully treated for TB) were modeled to be susceptible to reinfection: we assumed that prior TB infection imparted partial immunity, the range for strength of the protection as shown in Table. E-1.

Parameter	**California**	**Florida**	**New York**	**Texas**
Demography				
Annual *per capita* birth rate	0.019	0.018	0.018	0.019
Age-specific background mortality rates, given by: $\mu(a) = \mu_a \exp(-\mu_b a) + \mu_c + \mu_d \exp(\mu_f a)$				
μ_a	0.195	0.195	0.195	0.195
μ_b	4.5	4.5	4.5	4.5
μ_c	1×10^{-5}	0.8×10^{-5}	0.8×10^{-5}	0.8×10^{-5}
μ_d	2.5×10^{-4}	2.3×10^{-4}	2.3×10^{-4}	2.3×10^{-4}
μ_f	0.065	0.06	0.0625	0.069
TB-specific mortality rate (per year)	0.1	0.1	0.1	0.1
Immigration				
Immigration rate (net, per year)	0.0096	0.0052	0.0061	0.0044
Age distribution of immigrants on arrival: age $\sim 100 \times \text{Beta}(a_\alpha, a_\beta)$				
a_α	6.1	4.9	5	4.7
a_β	16	15.5	16.5	20
Estimated mean age at arrival (in years)	27.6	24.0	23.3	19.0
TB-specific†				
Pre-1993 annual rates of decline in transmission rate	0-0.02			
Post-1993 annual rates of decline in transmission rate	0-0.4			
Average duration of active TB (in years) until diagnosis and treatment initiation	0.67-1.33			
Protection imparted by previous TB infection (%)	40-70			
TB-treatment success proportion	0.9-0.97			
Reactivation rate of latent TB acquired τ years ago, in an individual of age a, $\phi(\tau, a) = \phi_a\,(\phi_3 + \exp(-\phi_4\,\tau))$, where, ϕ_a is: $\phi_a = \begin{cases} \phi_1, & \text{if age, } a < 10 \text{ years} \\ \phi_2, & \text{if age, } a \geq 20 \text{ years} \\ \phi_1 + \frac{(\phi_2 - \phi_1)\,(a-10)}{10}, & \text{otherwise} \end{cases}$				
ϕ_1	0.01–0.05			
ϕ_2	0.05–0.2			
ϕ_3	0.001–0.01			
ϕ_4	0.5–2			

Table E-1: Model parameters. †: TB-specific parameters were fit by maximizing the likelihood function, where the parameters were sampled from the given range.

Table E-2: TB prevalences (per 100,000) in 1990, 2000 and 2010 by regions of origin of foreign-born for the four states.

Region	California	Florida	New York	Texas
1990				
Mexico	122	122	122	122
Latin America (excl. Mexico and Canada)	105	134	208	92
China	215	215	215	215
India	465	465	465	465
Asia (excl. China and India)	715	709	600	637
Africa	350	350	350	350
Europe	30	30	30	30
Rest	11	11	11	11
2000				
Mexico	42	42	42	42
Latin America (excl. Mexico and Canada)	69	103	138	59
China	170	170	170	170
India	438	438	438	438
Asia (excl. China and India)	515	507	425	438
Africa	300	300	300	300
Europe	20	20	20	20
Rest	8	8	8	8
2010				
Mexico	28	28	28	28
Latin America (excl. Mexico and Canada)	58	77	97	49
China	108	108	108	108
India	269	269	269	269
Asia (excl. China and India)	348	342	305	299
Africa	300	300	300	300
Europe	15	15	15	15
Rest	6.3	6.3	6.3	6.3

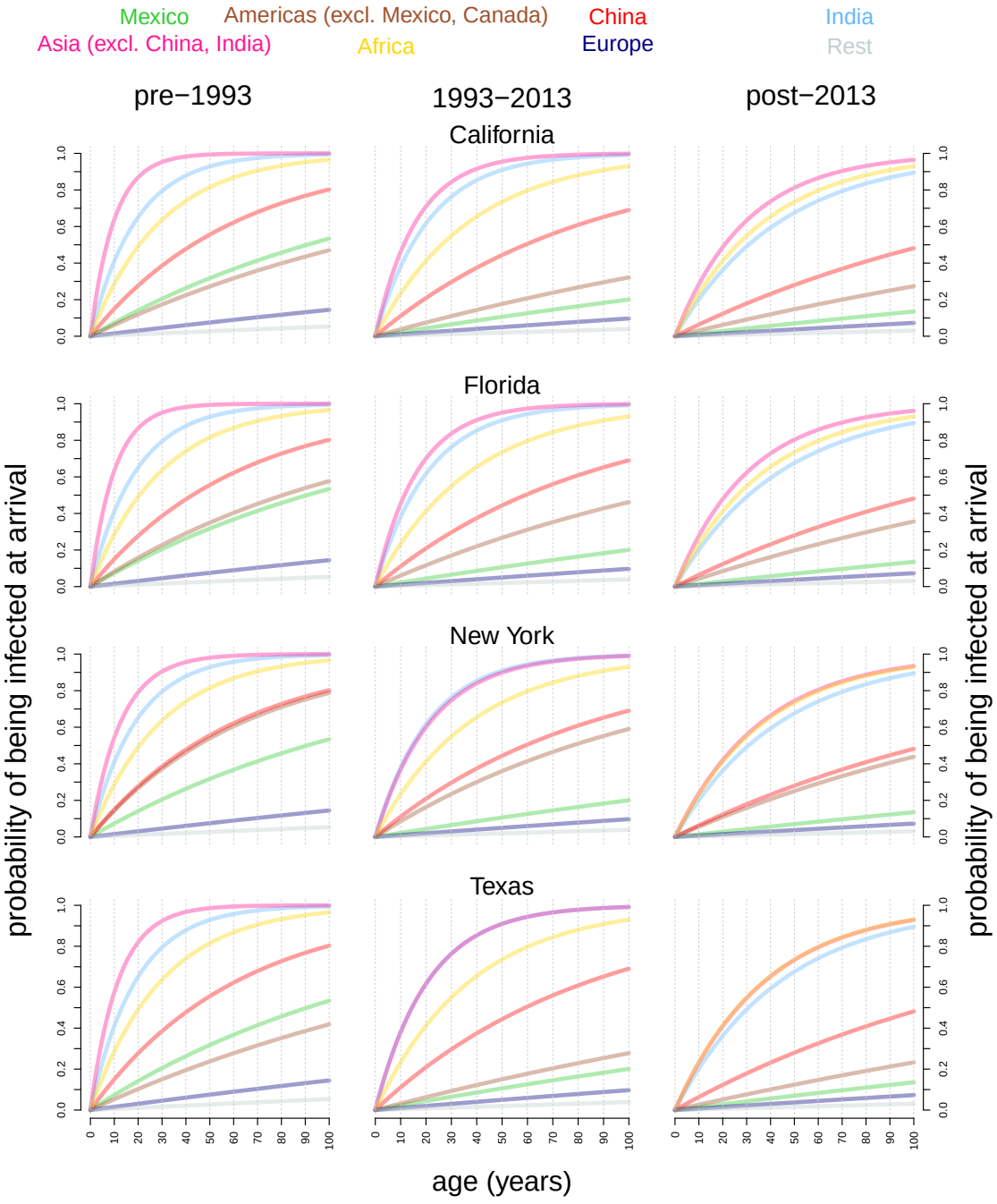

Figure E-1: Predicted probability of LTBI among foreign-borns at arrival. Shown in each panel is the probability that a foreign-born individual of given age (on the x-axis) from a given region (by color) being infected of TB at arrival. Sorted column-wise are probabilities for time periods, before 1993 (left), between 1993-2012 (middle) and after 2013 (right); and sorted row-wise are probabilities for each of the four states – (from top to bottom) California, Florida, New York, and Texas.

Model fitting

Likelihood function

We used a likelihood-based approach to fit our model to the data. In this framework, the likelihood of a model with a set of parameters θ, given data x is $\mathcal{L}(\theta|x) = f(x|\theta)$, where $f(x|\theta)$ is the probability of generating data x with parameter set θ. For each of the four states, the data we aimed to fit were the number of TB cases that were stratified by age and origin during different time periods, $x_{\text{time, age, origin}}$, where, time $\in$ {1993–1997, 2001–2005, 2009–2013}, age $\in$ {0–2, 25–45, 45–64, 65+}, and origin $\in$ {US-born, foreign-born}. The likelihood (and the log-likelihood) functions were defined as products (and sums) of conditional likelihoods, $f(x_{\text{time, age, origin}}|\theta)$:

$$\mathcal{L}(\theta) = \prod_{\text{time}} \prod_{\text{age}} \prod_{\text{origin}} f(x_{\text{time, age, origin}}|\theta)$$

$$\log \mathcal{L}(\theta) = \sum_{\text{time}} \sum_{\text{age}} \sum_{\text{origin}} \log f(x_{\text{time, age, origin}}|\theta).$$

We assumed the likelihood function to take a binomial functional form:

$$f(x_{\text{time, age, origin}}|\theta) = \text{Binomial}(x_{\text{time, age, origin}}; n = n_{\text{time, age, origin}}, p = \frac{c_{\text{time, age, origin}}}{\text{pop}_{\text{time, age, origin}}}).$$

Here, $n_{\text{time, age, origin}}$ is the age- and origin-specific population size during the specified time period, $c_{\text{time, age, origin}}$ is the age- and origin-specific model-based projections of TB cases during the specified time period, and $\text{pop}_{\text{time, age, origin}}$ is the age- and origin-specific simulated population sizes.

Likelihood maximization and profile likelihood

Likelihood functions were maximized separately in each of the four state. To maximize the likelihood function in a given state, we used Latin Hypercube Sampling to generate 200,000 parameter combinations, in which each parameter was drawn from uniform distributions with ranges specified in Table E-1. We simulated the model with each of parameter combinations, and calculated likelihood using the likelihood function described above. We averaged the likelihoods over five repeated simulations. Maximum likelihood estimate (MLE) was taken to be the set of parameters with the maximum likelihood estimate.

To construct likelihood profiles, we held the parameter being profiled at the point being profiled and maximized likelihoods over all other parameters. We carried out this maximization across a series of profile points, and then fit a line through them to construct likelihood profiles. The profile describes the strength of evidence pertaining to the focal parameter over the profiled range. The value corresponding to the largest likelihood estimate was considered a maximum likelihood estimate (MLE), and values corresponding to the likelihood estimates 1.98 units below maximum were taken to be 95% confidence interval. The 95% confidence interval was taken to be $\chi^2_1(.95)/2 \approx 1.98$ log-likelihood units below the maximum — univariate confidence limits using the χ^2 distribution.

Table E-3: Model-based estimates LTBI prevalences (in %) among US- and foreign-borns in the four states (in 2013) compared with reported LTBI prevalence in the US (based on Interferon Gamma Release Assay (IGRA) test in the NHANES study for 2011-2012 [5].

Age Categories	California	Florida	New York	Texas	US
US-born					
0-24	0.3	0.3	0.3	0.3	6-14: 0.7 15-25: 2.2
25-44	2.6	2.4	2.4	1.8	1.9
45-64	5.5	5.2	5.3	3.7	3.6
65 +	10	10.6	10	6.5	5.2
Foreign-born					
0-24	24.3	12.7	18.6	13.6	6-14: 2.6 15-25: 7.1
25-44	31.4	17.9	25.7	18.9	12
45-64	38.2	23.0	32.4	25.1	23.5
65 +	42.7	26.5	35.9	27.3	32.1

Supplementary results

Data and model simulations

We used the MLE model as the baseline model for estimations and projections. The simulations from the MLE model and the data are compared in Fig. E-2.

LTBI prevalences

Based on the MLE model, we estimated/projected age-specific LTBI prevalence among US- and foreign-born in each of the four states in 2013. The model-based estimation of LTBI prevalence in 2013 among US-born were very similar across the four states (except for generally lower prevalences in Texas). The LTBI prevalence among foreign-born, though, varied considerably across the four states: they were between 20 to 45% in California, 15 to 40% in New York, 10 to 30% in Texas and Florida. These are reflective of the demographic composition of foreign-born in each of the state. For instance, significant fraction of foreign-borns comprised of individuals from Asian countries in California and New York, where the TB prevalences tended to higher. Asian immigrants were comparatively lower in Florida and Texas. Large fraction of immigrants in Florida were comprised of individuals from Latin America, particularly Cuba and Mexico, where the TB prevalences have been considerably lower. In Table E-3, we compare the estimated prevalences with the reported prevalences in the NHANES survey [5].

Recent transmissions

We calculated the proportion of incident TB cases that were recent transmission (which was defined as cases resulting from infections that occurred at most two years ago) in the four states between 1993 and

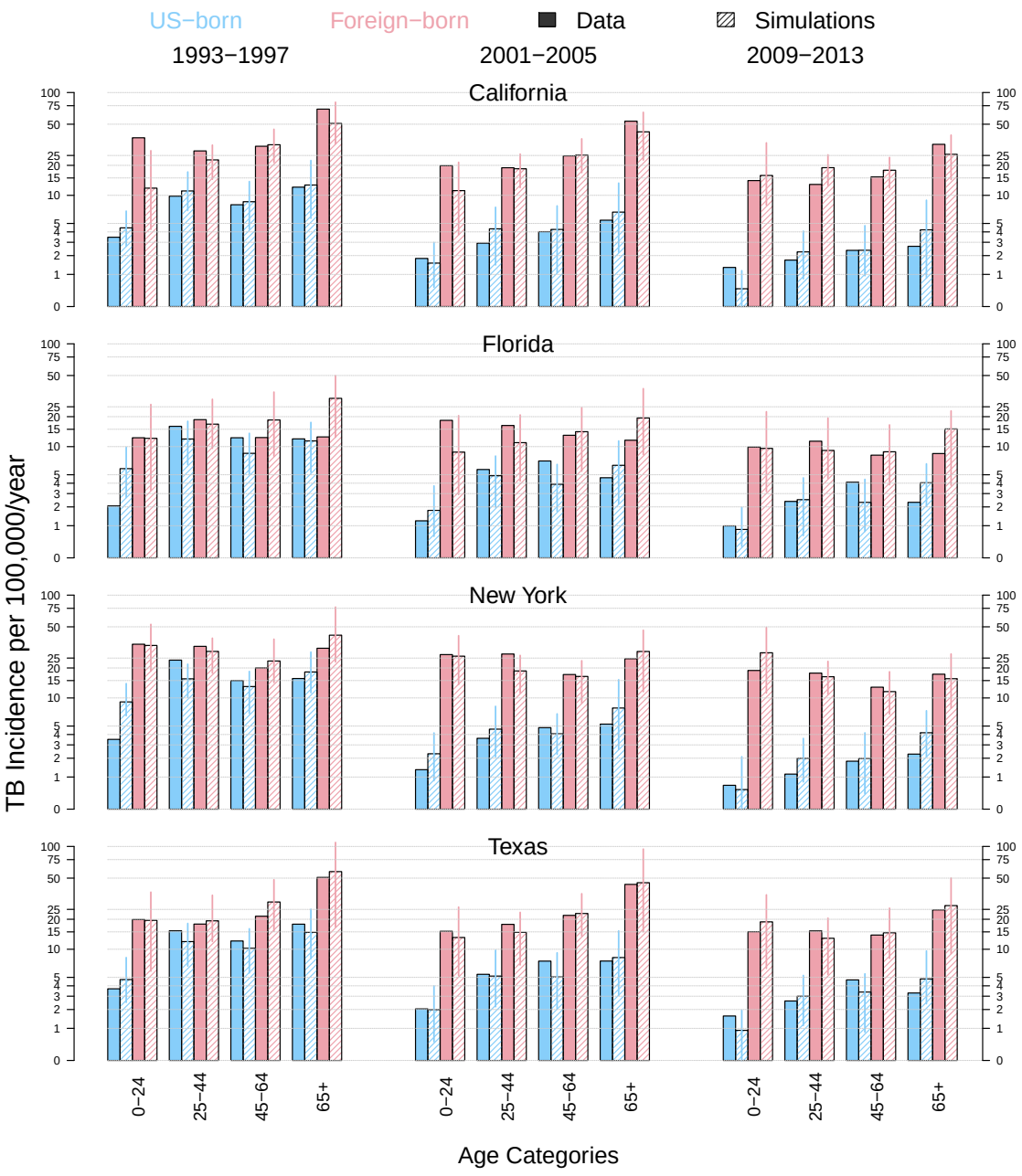

Figure E-2: Comparisons of TB incidences, categoried by age and origin (US- vs. foreign-born) in data and model simulations for each of the four states. The bar charts show TB incidences in the four states, California, Florida, New York and Texas, row-wise from top to bottom, respectively, for 5-year time-periods spanning 1993-1997 (left column), 2001-2005 (middle column), and 2009-2013 (bottom column). The TB incidences are categorized by age (4 age-categories are labeled at the bottom) and by origin (US-born in light blue and foreign-born in pink). The solid bars show estimates based on data: TB case report data were obtained from the CDC's Online Tuberculosis Information System (OTIS) data repository, and data on demographics were obtained from the American Community Survey. The hatched bars show model-based estimates: shown are medians (and 95% range) in 100 replicated simulations of the model with maximum-likelihood estimates for each of the four states.

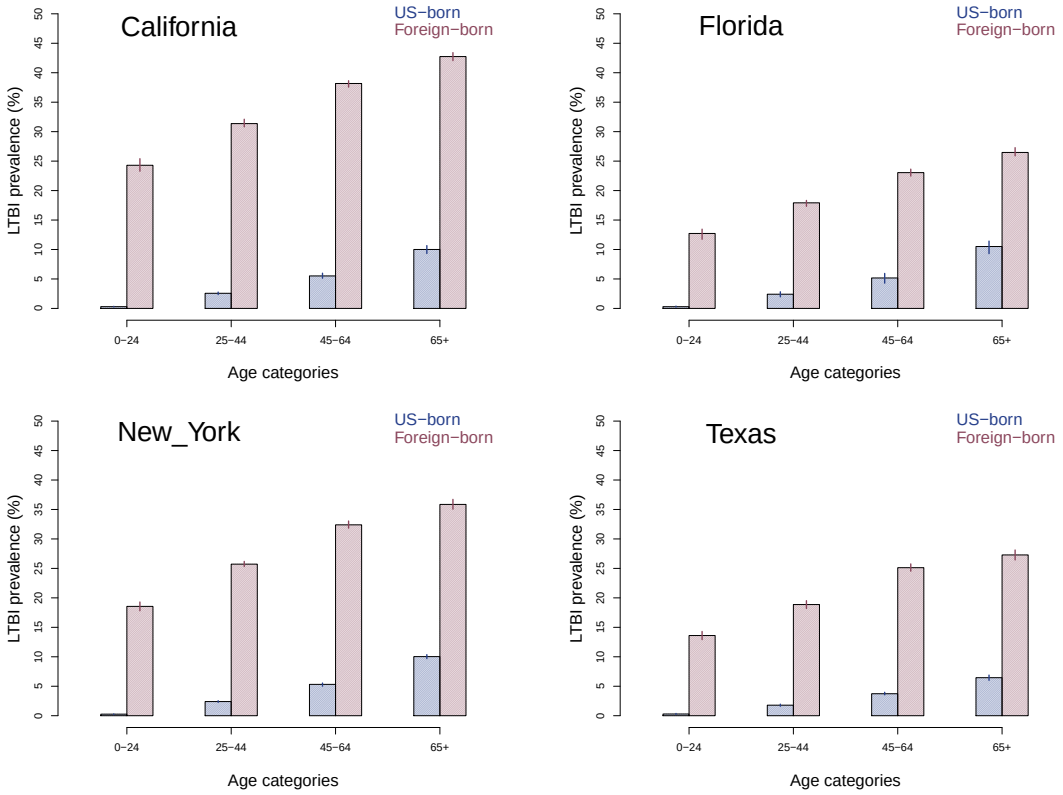

Figure E-3: Model-based estimates of LTBI prevalences in four states in 2013. Shown are model-based estimates for LTBI prevalence in 2013 in the four states; [top-left] California, [top-right] Florida, [bottom-left] New York, and [bottom-right] Texas. Shown in red are estimates for foreign-born individuals, and shown in blue are estimates for US-born individuals.

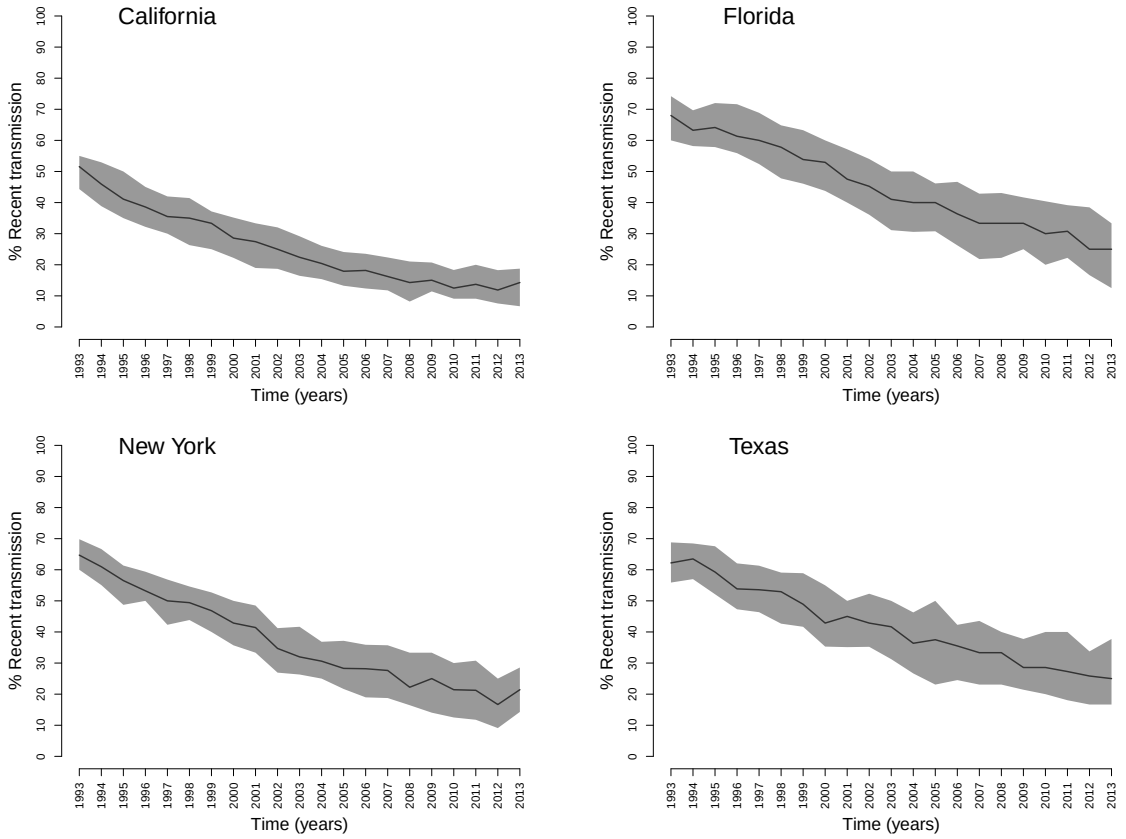

Figure E-4: Model-based estimates of percentage of TB cases that are recent transmission in four states between 1993 to 2013. Shown are proportions of new TB cases that were recent transmissions: Recent transmissions were defined as infections that occurred at most two years in the past. Shown are median (in solid lines) and interquartile range (grey shaded area).

2013. As can be seen in Fig. E-4, the proportion of TB cases that were recent transmission were estimated to decline in all four states. Declines were higher in New York and California compared to Florida and Texas, consistent with higher declines in transmission rates in New York and California compared to Florida and Texas. The estimated percentage of TB cases that were recent transmission, were between 10% and 20% in New York and California, and between 20% and 30% in Florida and Texas.

Ensemble approach

The model simulations and projections presented in the main text were generated using state-specific MLE models—a set of parameters that maximized the likelihood function. To account for variability in the parameters in the simulations and projections, we considered an alternative "ensemble" approach. In the approach, we considered all parameter combinations within the estimated 95% confidence interval to generate the simulations and projections. For each state, 95% confidence interval was taken to be within 8.46 log-likelihood units of the MLE. This estimate was based on 95% confidence interval of a χ^2 distribution with 9 degrees of freedom, 9 being the number of parameter fitted. We then generated 100 replicate simulations from each of these selected models, and constructed weighted medians and percentiles, with weights taken to be proportional to the likelihoods of the parameter estimates.

These simulations and projections, shown in Fig. E-5, were very similar to the simulations and projections generated from the MLE model, suggesting that the MLE-based simulations and projections are fairly robust to variation in parameters within the estimated 95% confidence interval.

Sensitivity to heterogeneous mixing

In the baseline model, we assumed homogeneous mixing between individuals regardless of nativity and age. Here, we explore the sensitivity of the results to heterogeneity in mixing with nativity. In particular, we explore a scenario where the mixing within groups of foreign-born and US-born are twice as more frequent to the mixing between the groups. We followed the same procedure of calibrating the model as with the baseline model.

The inferred annual rates of decline in transmission rates were 9.8%(95% range:4–11) in California, 1.3%(0–8) in Florida, 6.9%(7–12) in New York, and 4.7(0–9) in Texas (see Fig. E-6). These estimates were slightly lower (but generally within the 95% range) compared to the estimates from the baseline model with homogeneous mixing: 11%(5–13) in California, 6.8%(0–9) in Florida, 10%(9–15) in New York, and 5.5%(3–13) in Texas.

The inferred reactivation rates (per 100 person-year) were 0.039 (0.04–0.05) in California, 0.032(0.02–0.05) in Florida, 0.03(0.03–0.05) in New York, and 0.055(0.04–0.08) in Texas (see Fig. E-7). These estimates were similar to the estimates from the baseline model with homogeneous mixing: 0.046(0.04–0.06) in California, 0.037(0.02–0.05) in Florida, 0.036(0.03–0.05) in New York, and 0.058(0.04–0.08) in Texas.

The model-based simulations are projections are shown in Fig. E-8. Compared to projections based on the homogeneous model, these projections were very similar for California and New York, and slightly higher projections for Florida and Texas (e.g., 2025 projection of TB incidence in the entire population in Florida was 3.4 per 100,000/year, compared to 2.5; and in Texas it was 3.6 per 100,000/year, compared to 2.9).

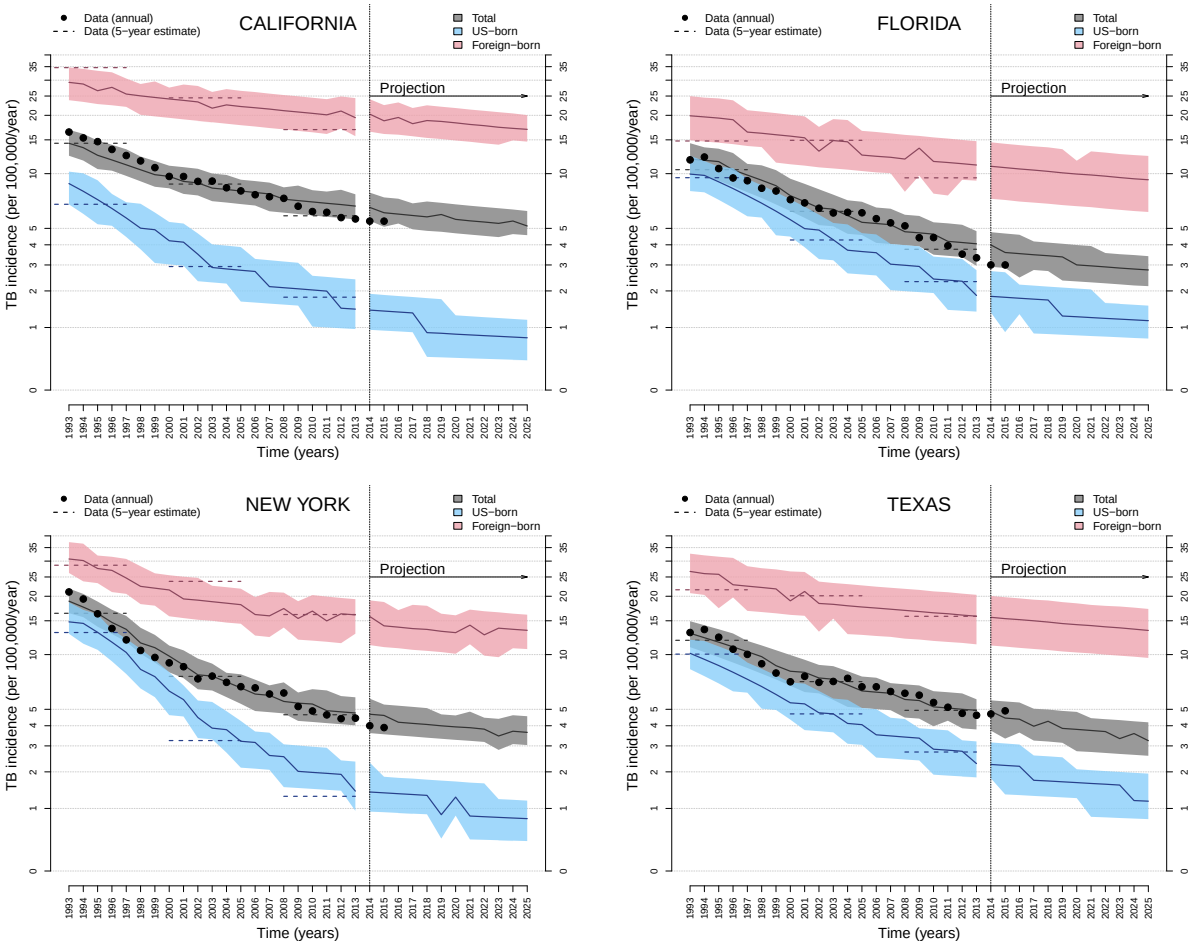

Figure E-5: Model-based simulations and projections of trends in TB incidence in four states, using an ensemble approach. Shown are model-based simulations in trends of tuberculosis (TB) incidence between 1993-2013 followed by projections up to 2025 for each of four states, namely, California [top left], Florida [top right], New York [bottom left] and Texas [bottom right]. The simulations are based on the state-specific ensemble models, and projections are the continuation of model simulations (with continued decline in the transmission rates at the post-1993 estimate). For each panel, shown are weighted medians (in solid lines) and weighted interquartile range (shaded area) ensemble models. Shown in black dots are data for annual TB incidence between 1993 and 2015. Shown in dashed lines are the estimated TB incidence (based on data) in each state across three 5-year periods of 1993-1997, 2001-2005, and 2009-2013. The foreign-born population is represented in pink, US-born in light blue, and the total population in grey.

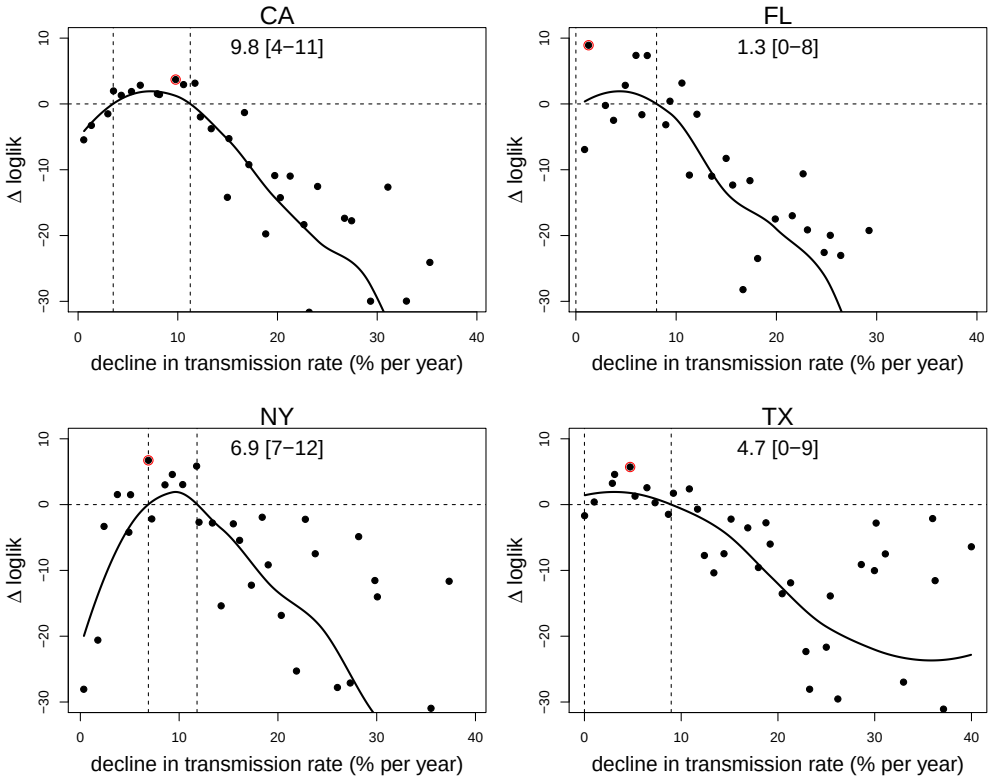

Figure E-6: Inferred annual declines in tuberculosis transmission rates under a model with differential mixing. Likelihood profiles (shown in a log scale as a difference with Δloglik=0 representing the estimated 95% threshold) for the annual rates of decline in the transmission rate (in % per year) in the four states—California [top-left], Florida [top-right], New York [bottom-left], and Texas [bottom-right]. Each point on the profile represents the log-likelihood of each estimate (on the y-axis) maximized over all parameters with the decline in transmission held at the level on the x-axis. The point circled red shows the maximum likelihood estimate (MLE), and the two dashed vertical lines show the estimated 95% confidence interval (smoothed estimate of log-likelihood no lower than 1.92 less than the MLE). Compared to Fig. 4 in the main text here the mixing within groups of foreign-born and US-born are twice as more frequent compared to mixing between the two groups.

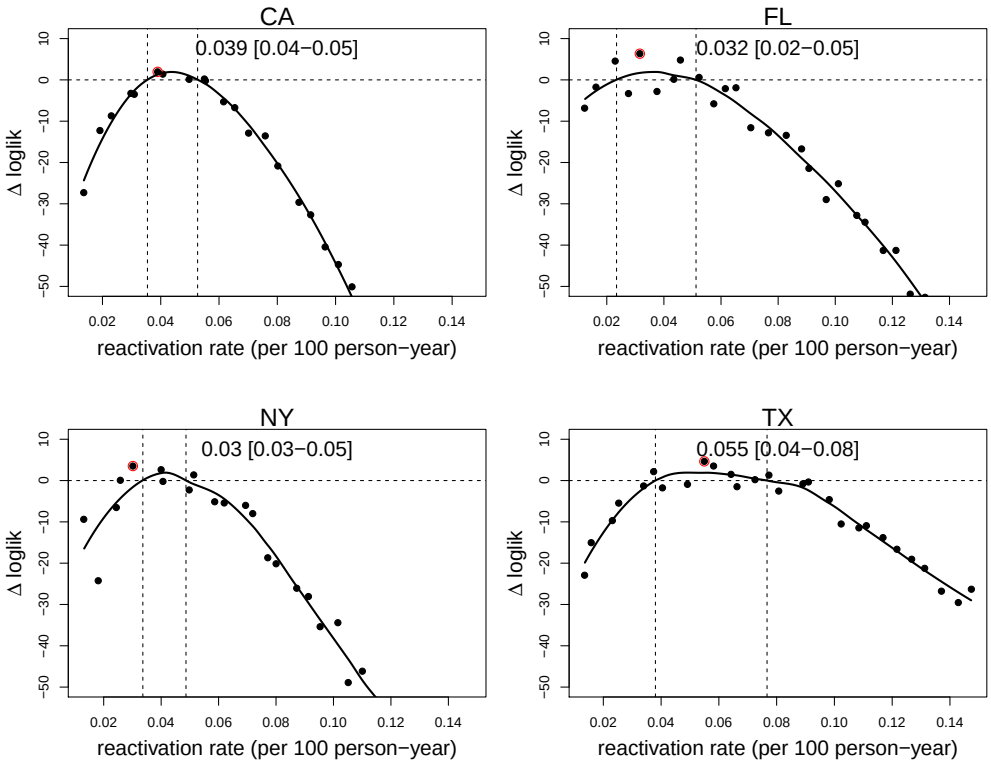

Figure E-7: Inferred tuberculosis reactivation rates. Likelihood profiles (shown in a log scale as a difference with Δloglik=0 representing the estimated 95% threshold) for reactivation rate in the four states— California [top-left], Florida [top-right], New York [bottom-left], and Texas [bottom- right]. The point circled red shows the maximum likelihood estimate (MLE), and the two dashed vertical lines show the estimated 95% confidence interval. Compared to Fig. 5 in the main text here the mixing within groups of foreign-born and US-born are twice as more frequent compared to mixing between the two groups.

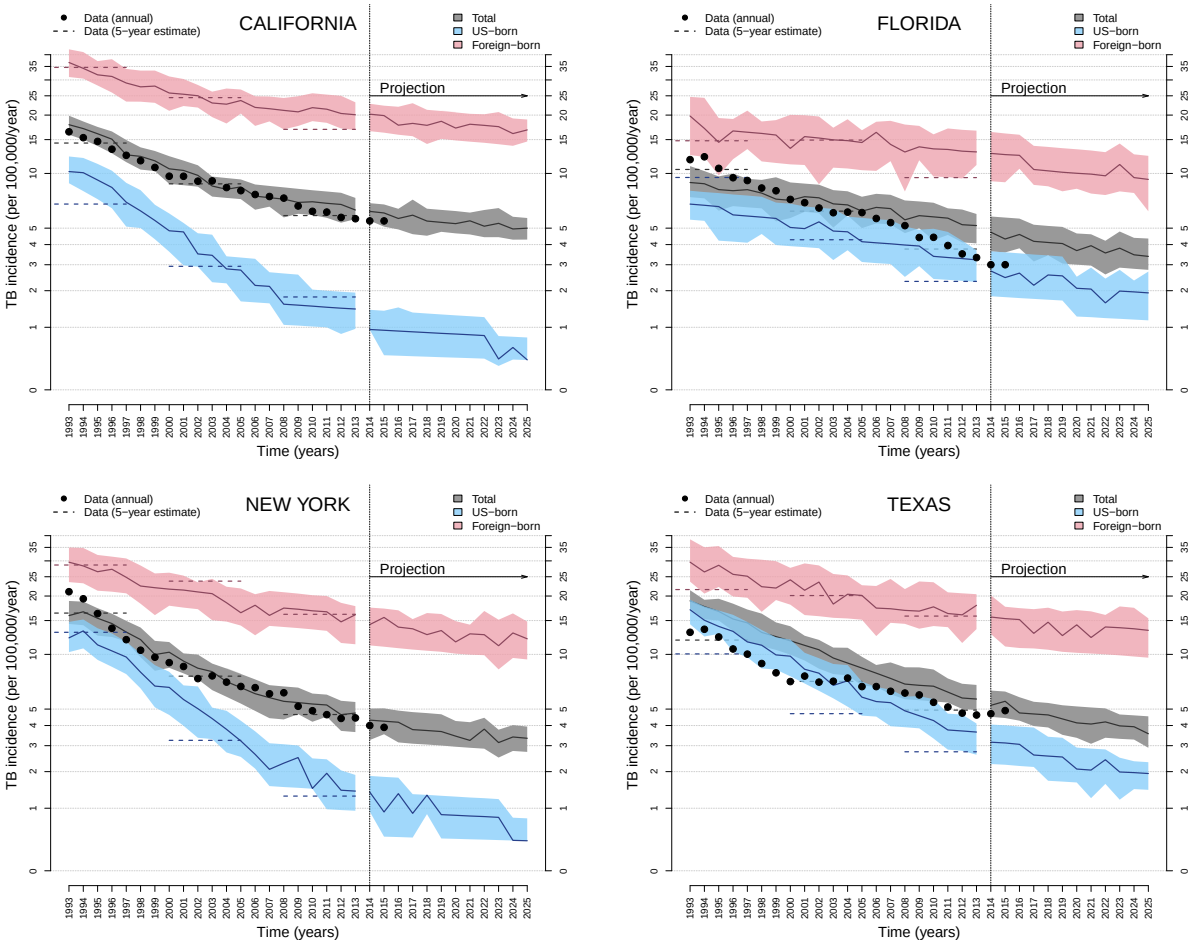

Figure E-8: Model-based simulations and projections of trends in TB incidence in four states, under heterogenous mixing. Shown are model-based simulations in trends of tuberculosis (TB) incidence between 1993-2013 followed by projections up to 2025 for each of four states, namely, California [top left], Florida [top right], New York [bottom left] and Texas [bottom right]. The simulations are based on the state-specific MLE models, and projections are the continuation of model simulations (with continued decline in the transmission rates at the post-1993 estimate). For each panel, shown are weighted medians (in solid lines) and weighted interquartile range (shaded area) ensemble models. Shown in black dots are data for annual TB incidence between 1993 and 2015. Shown in dashed lines are the estimated TB incidence (based on data) in each state across three 5-year periods of 1993-1997, 2001-2005, and 2009-2013. The foreign-born population is represented in pink, US-born in light blue, and the total population in grey. Compared to Fig. 6 in the main text here the mixing within groups of foreign-born and US-born are twice as more frequent compared to mixing between the two groups.

References

[1] Siler, W., 1979 A competing-risk model for animal mortality. *Ecology* **60**, pp. 750–757. ISSN 00129658.

[2] US Census Bureau; American Community Survey, 2014. 2014 America Community Survey, 5-year estimate.

[3] Shrestha, S., Knight, G. M., Fofana, M., Cohen, T., White, R. G., Cobelens, F. & Dowdy, D. W., 2014 Drivers and trajectories of resistance to new first-line drug regimens for tuberculosis. *Open Forum Infectious Diseases* **1**. (doi:10.1093/ofid/ofu073).

[4] Vynnycky, E. & Fine, P., 1997 The annual risk of infection with *Mycobacterium tuberculosis* in England and Wales since 1901. *The International Journal of Tuberculosis and Lung Disease* **1**, 389–396.

[5] Miramontes, R., Hill, A. N., Yelk Woodruff, R. S., Lambert, L. A., Navin, T. R., Castro, K. G. & LoBue, P. A., 2015 Tuberculosis infection in the united states: Prevalence estimates from the national health and nutrition examination survey, 2011-2012. *PLoS ONE* **10**, e0140881. (doi: 10.1371/journal.pone.0140881).