PNAS
www.pnas.org

Supplementary Information for

Decreasing measles burden by optimizing campaign timing

Niket Thakkar, Syed Saqlain Ahmad Gilani, Quamrul Hasan and Kevin A. McCarthy

Correspondence should be addressed to nthakkar@idmod.org

This PDF file includes:

Supplementary text
Figs. S1 to S18
Tables S1 to S3
References for SI reference citations

Supporting Information Text

1. Data sources and processing

Data used throughout the paper comes from three major sources: WorldPop population and birth rasters (1–3), Pakistan's most recent Demographic and Health Survey (4), and lab reports from Pakistan's measles surveillance system. In this section, we provide some details on where this data was obtained and how it was processed.

Lab confirmed and rejected cases. Lab reports come from Pakistan's WHO affiliated lab in compliance with WHO surveillance standards (5). Briefly, all individuals presenting with fever, rash, and one of the 3 C's (cough, coryza, or conjunctivitis) are suspected of having measles and should in principle have blood drawn for lab confirmation. In practice however, not all suspected measles cases are lab confirmed, and we model selection as a time-independent, binomial process. Blood samples are then sent to a national lab where IgM assays are used to confirm or reject measles infection. The lab reports used in this study record lab activity over time (27,551 records from 2012-2018) with entries for each serum sample giving a location, birth-date, caregiver-reported dose history, and classification (measles positive or negative).

We can verify Pakistan's lab testing system is operational throughout the modelled period by plotting lab rejected cases over time. This is done in Fig. S1, demonstrating that lab rejection is consistently nonzero from 2012-2018. Notably, the lab rejection time series has significant seasonality; however, as others have demonstrated (6), this is likely due to seasonal variation in other infections with measles-like symptoms.

Population. Populations rasters for Pakistan in 2010 and 2015 were obtained from the WorldPop website (worldpop.org.uk) (1–3), and pixel level population estimates were aggregated by district to produce the map in Fig. S2. Data was aggregated further to province or national level by summing the population by district.

Population estimates were linearly interpolated in time. This is shown for Sindh in Fig. S3. Linear interpolation between WorldPop estimates creates a ramp from 2010 to 2015. The time series is then extrapolated by assuming a constant population on either end.

Birth rate. Province-level birth rate estimates were also obtained from WorldPop by processing 2010, 2012, 2015, and 2020 live-birth rasters in the same method as above. These rasters covered all the provinces modeled except for Gilgit Baltistan.

To obtain an estimate for Gilgit Baltistan, we used the 2012-13 DHS. The method we used varies from DHS specifications (7) by not using age-based fertility as an intermediate step and by using a fixed three-year window to count births. All estimates were computed in consideration of the complex survey design and corresponding sample weights using the `survey` package in `R` (8, 9), and we verified that the birthrates based on WorldPop estimates (where available) and these province-level DHS estimates are comparable. Birth rates are summarized in Table S1.

Vaccination coverage. Estimates of measles containing vaccine (MCV) coverage were obtained directly from the 2012-2013 DHS (4) and assumed constant throughout the modeled time period. These estimates, plotted in Fig. S4, were taken as estimates of single-dose routine immunization coverage since the survey asks this question of 12-23 month olds.

The WHO also recommends a second MCV dose (MCV2) at 15 months of age. Coverage for MCV2 was estimated by analyzing the percentage of rejected measles cases with 2 vaccines self reported in their dose histories. This method is very coarse, and estimating uncertainty is difficult. For our modeling purposes, we assume coverage uncertainty is small compared to transmission and reporting uncertainty, and it's therefore neglected. Still, more accurate estimation of MCV2 coverage is a topic of current research. Here, MCV2 rates were also assumed constant in the model time period. Coverage values are summarized in Table S2.

In the model, coverage numbers are used to calculate the number of births missed by RI. This calculation requires an assumption about MCV take (the probability that vaccination results in immunity), which is known to vary with age and depend on environmental factors like MCV cold-chain. This is a complicated issue, and in line with others (10), we greatly simplify the calculation by assuming MCV1 take is 90% and MCV2 take is 99% throughout the entire country during the entire model period. In that case, adjusted births, B_t, can be calculated from total births $\tilde{B}_t$, via

$$B_t = \tilde{B}_t \left[1 - 0.9V_{1,t}(1 - V_{2,t}) - 0.99V_{1,t}V_{2,t}\right], \quad [1]$$

where $V_{1,t}$ and $V_{2,t}$ are MCV1 and MCV2 coverages respectively. This calculation incorporates all coverage information into the model, and we have found further that model projections are robust to reasonable changes in MCV take (from 80% to 95%). It is worth noting however, that considerable uncertainty is being ignored when we approximate the semi-monthly number of children added to the susceptible pool by this deterministic time series. In principle, heterogeneity in the waning of maternal antibodies (11), the timing if MCV doses, and the birth-rate from semi-month to semi-month all contribute to uncertainty in the number of births over time. However, for our purposes of projecting total infections from 2018-2021, we anticipate that transmission stochasticity is a larger driver of overall uncertainty, and noise in the number of children missed by RI can be ignored.

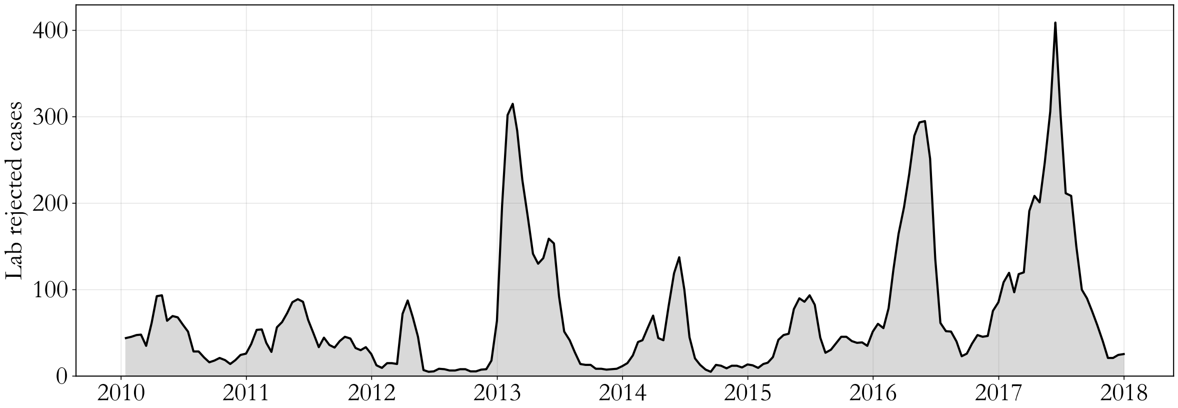

Fig. S1. Lab rejected cases plotted over time demonstrate that lab confirmation persists throughout the modelled time period.

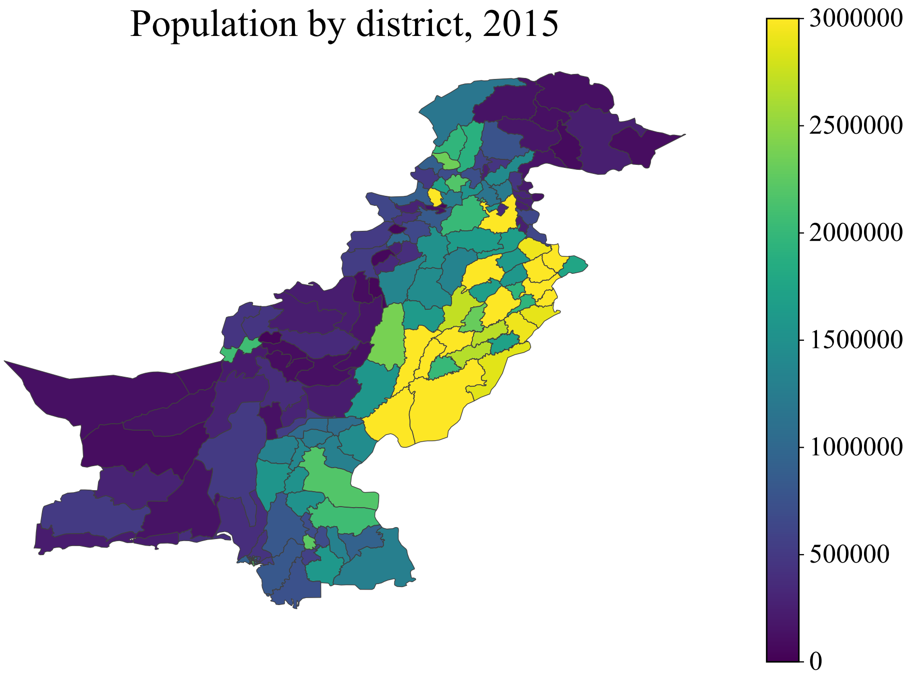

Fig. S2. The population estimates used in this study were generated by aggregating WorldPop rasters to district level, shown here in 2015.

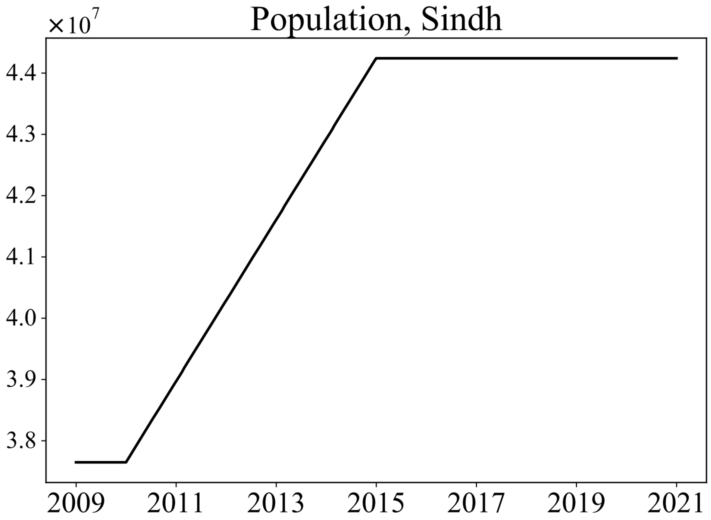

Fig. S3. Both population and birth-rate are linearly interpolated in time with constant extrapolation. This produces a ramp between rasters with plateaus on either end.

Province	Birth rate (WorldPop)	Birth rate (2012 DHS)
Balochistan	1.004	1.193
Gilgit Baltistan	Not available	1.0373
Islamabad	1.038	0.832
Khyber Pakhtoon	1.103	1.122
Punjab	1.059	1.142
Sindh	1.0823	1.091

Table S1. Time averaged birthrates (per 100k, per semi-month) computed via WorldPop rasters and the DHS. Since the WorldPop rasters have better time resolution, we default to WorldPop estimates where they're available, using the DHS only in Gilgit Baltistan.

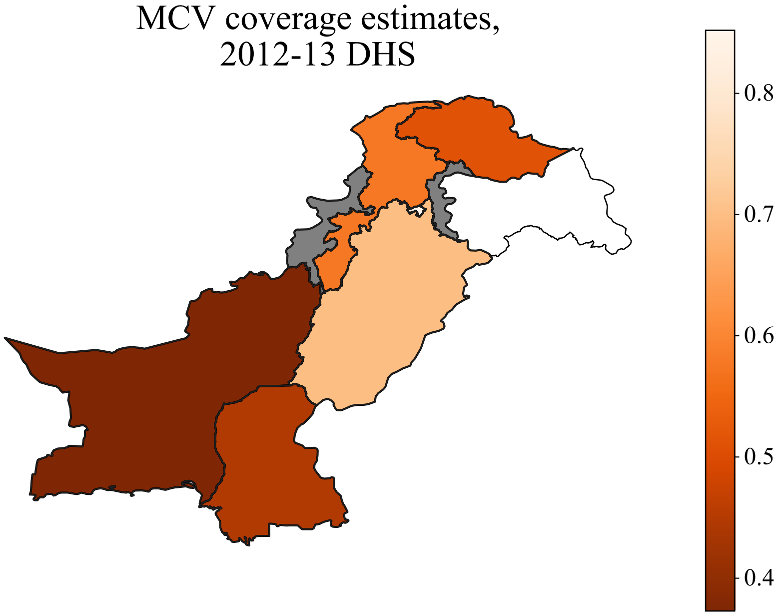

Fig. S4. Coverage of the first dose of routine vaccination was taken from the 2012-13 DHS measles vaccination coverage in 12-23 month olds by province and assumed constant throughout the modeled time period.

Province	MCV1 coverage (DHS)	MCV2 coverage (rejected cases)
Balochistan	0.381	0.079
Gilgit Baltistan	0.510	0.314
Islamabad	0.852	0.214
Khyber Pakhtoon	0.577	0.222
Punjab	0.696	0.563
Sindh	0.451	0.273

Table S2. MCV1 and MCV2 coverage estimates via the DHS and rejected measles cases. For MCV2 coverage estimates, rejected measles cases are assumed to be a sample of children from a given region. MCV2 coverage is estimated by computing the fraction of samples with 2 or more self-reported MCV doses.

2. A TSIR model with vaccination campaigns

Our model of measles is an extension of the time series susceptible-infected-recovered (TSIR) model of Finkenstädt and Grenfell(12), incorporating the effects of vaccination campaigns into calibration. Inspired by the traditional SIR model (13), we assume the interacting susceptible population, S_t, and infectious population, I_t, obey discrete time, stochastic dynamical equations,

$$S_t = (1 - \mu_{t-1})\,(B_t + S_{t-1} - I_t) \qquad [2]$$

$$I_t = \beta_t I_{t-1}^{\alpha} S_{t-1} \varepsilon_t, \qquad [3]$$

where time is taken in semi-monthly increments corresponding to the average duration of a measles infection (11). This choice allows us to compare I_t, the prevalence at time t, with observed cases, C_t, a measure of incidence. In the above equations, B_t is the estimate of the number of births missed by RI per semi-month described previously, μ_t is the efficacy of any SIA at time t, β_t is the time-dependent transmission rate which we assume has annual periodicity, and α accounts for inhomogeneous mixing (12). Transmission uncertainty is modeled by ε_t, a log-normally distributed random variable with zero mean and constant variance, and reporting uncertainty is modeled by assuming

$$C_t \sim \text{Binom}\left\{ I_t, p = \frac{1}{\rho} \right\}, \qquad [4]$$

where p is the lab-reporting-rate, the probability of a measles infection to be selected for lab-confirmation. All other sources of uncertainty, in population estimates for example, are assumed to be negligibly small.

In this section, we describe the method used to fit this model to lab-confirmed measles cases from Pakistan. Given SIA coverages μ_t, this requires two regressions, one to reconstruct the underlying susceptible population and another to determine the transmission seasonality.

Reconstructing S_t. Susceptible reconstruction requires us to account for uncertainty due to under-reporting and to relate observed cases to underlying prevalence. Under the binomial reporting model, we have

$$[C_t | I_t, p] = \binom{I_t}{C_t} p^{C_t} (1-p)^{I_t - C_t}, \qquad [5]$$

where $[\cdot]$ denotes a probability distribution or density. Via Bayes' theorem with a uniform prior enforcing $I_t \geq C_t$, this implies

$$[I_t | C_t, p] = \binom{I_t}{C_t} p^{C_t+1} (1-p)^{I_t - C_t}, \qquad [6]$$

where we continue to assume p is known. Using a moment generating function approach, this yields

$$\mathrm{E}\left[I_t | C_t, p\right] = \frac{1}{p}(C_t + 1) - 1, \qquad [7]$$

$$\mathrm{Var}\left[I_t | C_t, p\right] = \frac{1-p}{p^2}(C_t + 1). \qquad [8]$$

Setting these results aside for a moment, we return to Eq. 2. Following Finkenstädt and Grenfell (12), we assume $S_t = \bar{S} + Z_t$, where $\bar{S}$ is an average susceptible population and Z_t are time-dependent fluctuations around this average. Inserting this decomposition into Eq. 2,

$$\begin{aligned} \bar{S} + Z_t &= (1-\mu_{t-1})(\bar{S} + Z_{t-1} + B_t - I_t) \\ \implies Z_t &= -\mu_{t-1}\bar{S} + (1-\mu_{t-1})(Z_{t-1} + B_t - I_t). \end{aligned} \qquad [9]$$

This relation can be usefully rewritten in linear-algebra notation. Treating time dependent quantities as column vectors with ordered entries for each semi-month in the modeling period,

$$\mathbf{Z}_t = \mathbf{A}\mathbf{B}_t - \mathbf{A}\mathbf{I}_t - \bar{S}\mathbf{D}\mathbf{m}_t, \qquad [10]$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots \\ (1-\mu_0) & (1-\mu_0) & 0 & 0 & 0 & \dots \\ (1-\mu_0)(1-\mu_1) & (1-\mu_0)(1-\mu_1) & (1-\mu_1) & 0 & 0 & \dots \\ \vdots & \vdots & (1-\mu_1)(1-\mu_2) & (1-\mu_2) & 0 & \dots \\ \vdots & \vdots & \vdots & (1-\mu_2)(1-\mu_3) & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \dots & \dots \end{bmatrix} \qquad [11]$$

and

$$\mathbf{D} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & (1-\mu_0) & 0 & 0 & 0 & \dots \\ 0 & (1-\mu_0)(1-\mu_1) & (1-\mu_1) & 0 & 0 & \dots \\ \vdots & \vdots & (1-\mu_1)(1-\mu_2) & (1-\mu_2) & 0 & \dots \\ \vdots & \vdots & \vdots & (1-\mu_2)(1-\mu_3) & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \dots & \dots \end{bmatrix} \qquad [12]$$

In other words, **A** and **D** are linear transformations (both dimension $(t+1) \times (t+1)$) based on the cumulative product of the μ_t time series. Finally,

$$\mathbf{m}_t = \begin{bmatrix} 0 \\ \frac{\mu_0}{1-\mu_0} \\ \frac{\mu_1}{1-\mu_1} \\ \frac{\mu_2}{1-\mu_2} \\ \vdots \\ \frac{\mu_{t-2}}{1-\mu_{t-2}} \\ \frac{\mu_{t-1}}{1-\mu_{t-1}} \end{bmatrix}, \qquad [13]$$

a time-series based on past SIA efficacy. Under the assumption that $\mathbf{B}_t$ and $\mathbf{m}_t$ are known with negligible uncertainty, Eqs. 7 and 8 imply

$$\mathrm{E}[\mathbf{Z_t}|\mathbf{C}_t, p] = \mathbf{A}(\mathbf{B}_t + 1) - \frac{1}{p}\mathbf{A}(\mathbf{C}_t + 1) - \bar{S}\mathbf{D}\mathbf{m}_t, \qquad [14]$$

$$\mathrm{Var}[\mathbf{Z_t}|\mathbf{C}_t, p] = \mathbf{A}\left(\frac{1-p}{p^2}(\mathbf{C}_t + 1)\right)\mathbf{A}^T \propto \mathbf{C}_t + 1 \qquad [15]$$

where scalar addition to column vectors is applied element-by-element. Under the assumption that fluctuations in S_t will be small, Eq. 14 motivates a weighted least-squares regression of $\mathbf{A}(\mathbf{B}_t + 1)$ against $\mathbf{A}(\mathbf{C}_t + 1)$ and $\mathbf{D}\mathbf{m}_t$ to approximate $1/p$. Weights are assumed equal to $1/\sqrt{\mathbf{C}_t + 1}$ and regression residuals then approximate $\mathrm{E}[\mathbf{Z}_t]$. Thus, this procedure gives an estimate of p and the fluctuations in underlying susceptible population, $\mathbf{Z}_t$.

As a brief aside, we note that the reporting rates, p, are an interesting model output on their own, and we are continuing to study the spatial heterogeneity in p throughout Pakistan. One interesting comparison is plotted in Fig. S5. Here, the rate per 100k per year of rejected (i.e. non-measles) lab reports is averaged for each province from 2012-2017, offering an independent metric for the activity of the surveillance system in each province. We find that our inferred reporting rates, which are based on confirmed cases, are highly correlated, lending confidence to our modeling procedure.

Estimating β_t. Analysis of Eq. 3 offers a route to estimate the remaining parameters. Taking the log, we have

$$\log I_t = \log \beta_t + \alpha \log I_{t-1} + \log S_{t-1} + \log \varepsilon_t, \qquad [16]$$

where $\mathrm{E}[\log \varepsilon_t] = 0$ and $\mathrm{Var}[\log \varepsilon_t] = \sigma$. Here we follow Finkenstädt and Grenfell (12) exactly, and we assume

$$\log S_t = \log\left(\bar{S} + Z_t\right) \approx \log \bar{S} + \frac{Z_t}{\bar{S}}, \qquad [17]$$

where we have used the first order Taylor expansion for the log under the assumption that Z_t is small. Inserting this result into Eq. 16, we find

$$\log I_t = \log\left(\bar{S}\beta_t\right) + \alpha \log I_{t-1} + \frac{Z_{t-1}}{\bar{S}} + \log \varepsilon_t. \qquad [18]$$

This motivates the second regression. The first term is treated as an intercept for each semi-month of a given year and α and $1/\bar{S}$ are slopes. Z_t is replaced with $\mathrm{E}[Z_t]$ from the previous regression, and I_t is taken to be $[(C_t + 1)/p] - 1$ with p also from the previous regression. Finally, σ, the variance associated with transmission uncertainty, is estimated using the regression's residual variance.

This approach estimates β_t by maximizing the model's auto-regressive prediction accuracy. On the surface, this procedure is disconnected from underlying, mechanistic explanations of the annual variation in measles transmission; however, we can still gain insight into measles epidemiology by analyzing inferred seasonality's structure.

Evidence from other settings suggests that measles transmission fluctuates with changes in person-to-person contact due to weather (14) or population migration (15, 16). This motivates similar analysis in Pakistan. As we note in the main text (Fig. 1), our inferred low transmission season corresponds to Pakistan's hot, rainy summer, suggesting that weather may play a role in decreasing contact rates and suppressing measles transmission. This correspondence simultaneously increases our confidence in our inference.

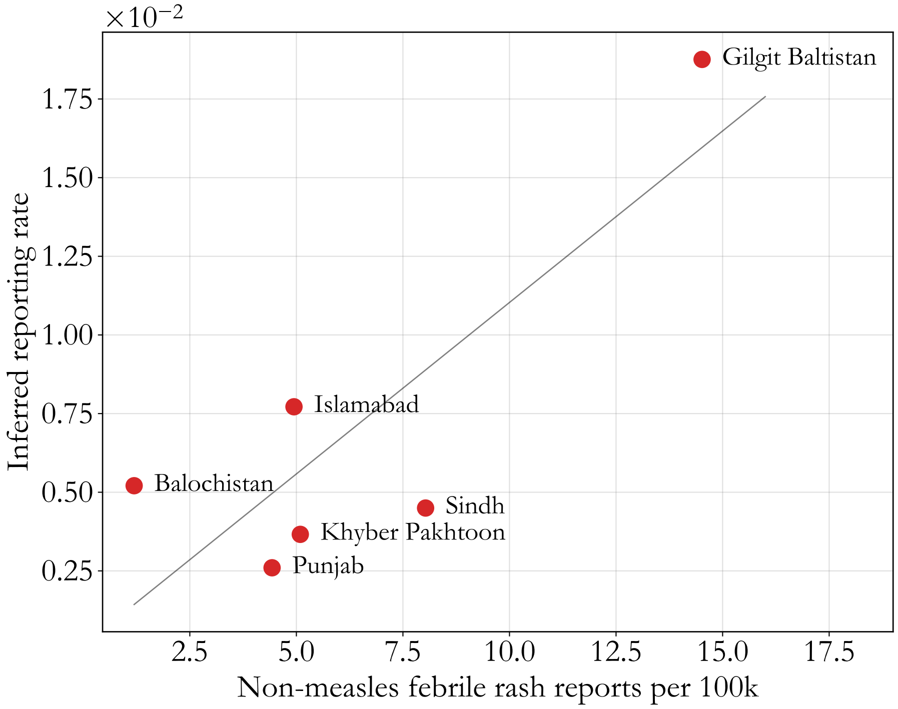

Fig. S5. Rejected cases, so-called non-measles febrile rash reports, per 100k per year averaged from 2012-2017 is a potential indicator of the activity of a province-wide surveillance system. We find that our inferred reporting rates (red dots) and computed rejected case rates are highly correlated (grey line is a least squares fit).

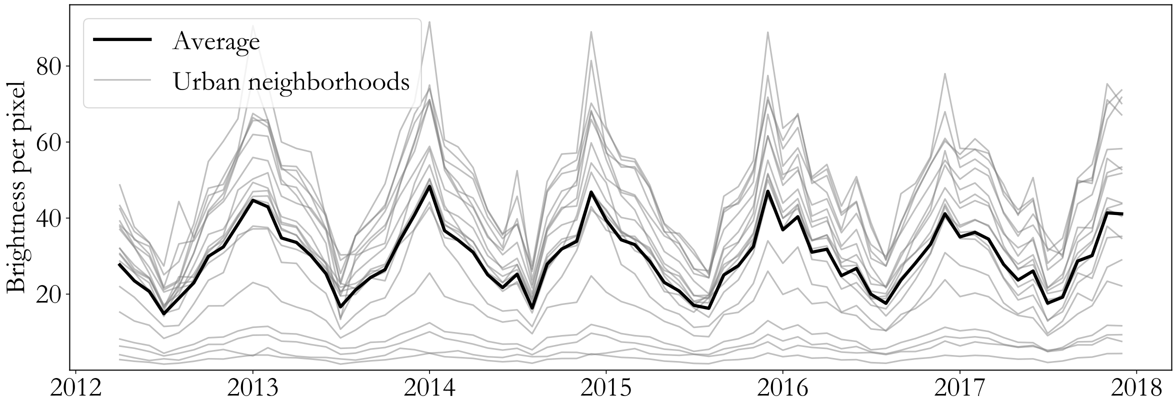

Fig. S6. Nighttime light satellite imagery (17) can be used to estimate rural-to-urban migration at high time resolution (15). Here, we plot brightness per pixel near Karachi and Lahore from 2012 to 2018 for each district in the cities (grey) and then averaged globally over districts (black). All time series are based on 69 composite satellite images, showing strong annual seasonality potentially correlated with annual rural-to-urban migration.

Furthermore, inspired by Ref (15), we can use nighttime light satellite imagery (17) to understand Pakistan's population migration from rural to urban areas. More specifically, we obtain monthly composite images using nighttime data from the Visible Infrared Imaging Radiometer Suite (17, 18), and we calculate average brightness per pixel in the bounding box corresponding to the borders of Pakistan's largest cities, Karachi and Lahore. These time series are plotted in Fig. S6, showing strong annual seasonality across districts. Averaging that data by month from 2012 to 2018 gives a time series likely correlated to Pakistan's annual urban population fluctuations (15).

These average fluctuations are plotted in the inset of Fig. S7, with districts in Karachi and Lahore in grey and a global average in black. Across districts, we see that nighttime light brightness decreases in the summer and begins to increase in October. Black dots in Fig. S7 compare our monthly β_t estimates (Fig. 1 in the main text) to average nighttime light brightness (standard deviation error bars), showing a high degree of correlation (0.725) and a strong linear relationship (black line, 95% CI in grey). This suggests, in agreement with work by others (15), that measles seasonality in Pakistan is driven in part by annual rural-to-urban migration and a corresponding increase in population density and person-to-person contact rates.

Estimating μ_t. The method outlined in the previous subsections can be used to estimate all model unknowns except SIA efficacies. As described in the main text, approximating μ_t can be reduced to estimating a single parameter since Pakistan has had a relatively small number of measles SIAs.

We assume that $\mu_t = \mu P_t$, where μ is an SIA efficacy parameter common to all campaigns from 2012-2017 and P_t is an estimate of the fraction of the population targeted in a given campaign. P_t is approximated using population estimates at district level in combination with WHO descriptions of regions targeted by each SIA (19). Age ranges of the campaigns (6 month to 9 years in 2012 and 2013, 6 month to 10 years in 2014 and 2015, and 6 months to 5 years in 2017) represent a small additional correction and are incorporated by computing the fraction of lab-confirmed cases within each age range relative to cases within the 2014-15 age range (see Fig. S8). Furthermore, SIA dates are rounded to the nearest semi-month. For example, in 2017, two districts in Sindh were targeted, representing $\sim 18\%$ of Sindh's total population while the age-range covers approximately 78% of cases, implying that $P_{t=8/15/2017} \approx 0.18 \times (0.78/0.94) \approx 0.15$. For reference, nonzero P_t are tabulated in Table S3.

Specifying μ_t thus requires us to estimate a single parameter. We do this by iterating the regression procedure above for a set of possible μ and comparing goodness-of-fit. More specifically, given a hypothetical μ, we estimate the remaining free parameters, compute the average long-term extrapolation (Fig. 2 in the main text), and calculate R^2 for the 2015-17 time-period. Restriction to the later years is made to emphasize fit quality closer to when competing SIA policies are to be tested. This method is closely related to the profile likelihood optimization procedures employed by others (20, 21).

The results of this approach applied to the Pakistan data are plotted in Fig. S9. R^2 values (grey, dashed) are smoothed with a 30 point, centered rolling average (black) and the highest R^2 value is selected (red circle). This yields $\mu \approx 0.40$ with an $R^2 \approx 0.6$ for the 2015-2017 model period.

3. Model testing

As shown in Fig. 2 in the manuscript, the model is tested by assessing extrapolation ability at semi-monthly and long-term time-scales. Those tests show a high goodness of fit ($R^2 \approx 0.89$ for short-term prediction and $R^2 \approx 0.35$ for long-term prediction), and demonstrate that the model captures all major outbreak events across 5 years of extrapolation.

The model can also be tested by withholding data. This is shown in Fig. S10, where data after Q1 2017 (orange squares) is reserved to test model performance. The model, fit to data from 2012 onward (black dots) accurately captures the 2013 and 2016 outbreaks (blue time series) and simultaneously predicts the severity of the 2017 outbreak (green time series). Moreover, we expect model performance to improve when fit to the full dataset. These tests taken together indicate that the model infers underlying susceptibility and predicts future burden with reasonable accuracy.

4. Comparing future timing-efficacy combinations

To evaluate the impact of future SIAs, we sample the calibrated model and sum the infections calculated with an extrapolation starting at end of the lab reports in 2017. In the main text, for simplicity, we restrict this comparison to a future efficacy of 0.4, equal to the historically precedented value according to our inference procedure above. In principle however, the efficacy in the future could change due to annual fluctuations in campaign accessibility (see Fig. S7), future changes in program performance, and a number of other ancillary factors. Here we discuss results as a function of both efficacy and timing.

Fig. S11 summarizes the expected infections as a function of SIA policy. Each dot is the average of 5000 samples of infections from 2018-2021 with SIAs at a particular time (x-axis) and efficacy (colorbar, which indicates the absolute value and the value relative to the historical average). Timing behaves just as discussed in the main text, with a precipitous decrease in cases over the course of the 2018 low transmission season leading to an optimum for a given efficacy at October-November.

Contours at future SIA efficacy equal to 0.3, (red), 0.4 (orange), and 0.5 (green) show that increasing SIA efficacy universally decreases total infections, just as we would expect. This type of analysis is useful when trying to understand the trade-off between delaying (potentially increasing burden) and increasing time to plan (potentially increasing efficacy). Generally speaking, we find that substantial changes in efficacy ($\sim 10\%$) are needed to justify adjustments away from the optimal October-November timing.

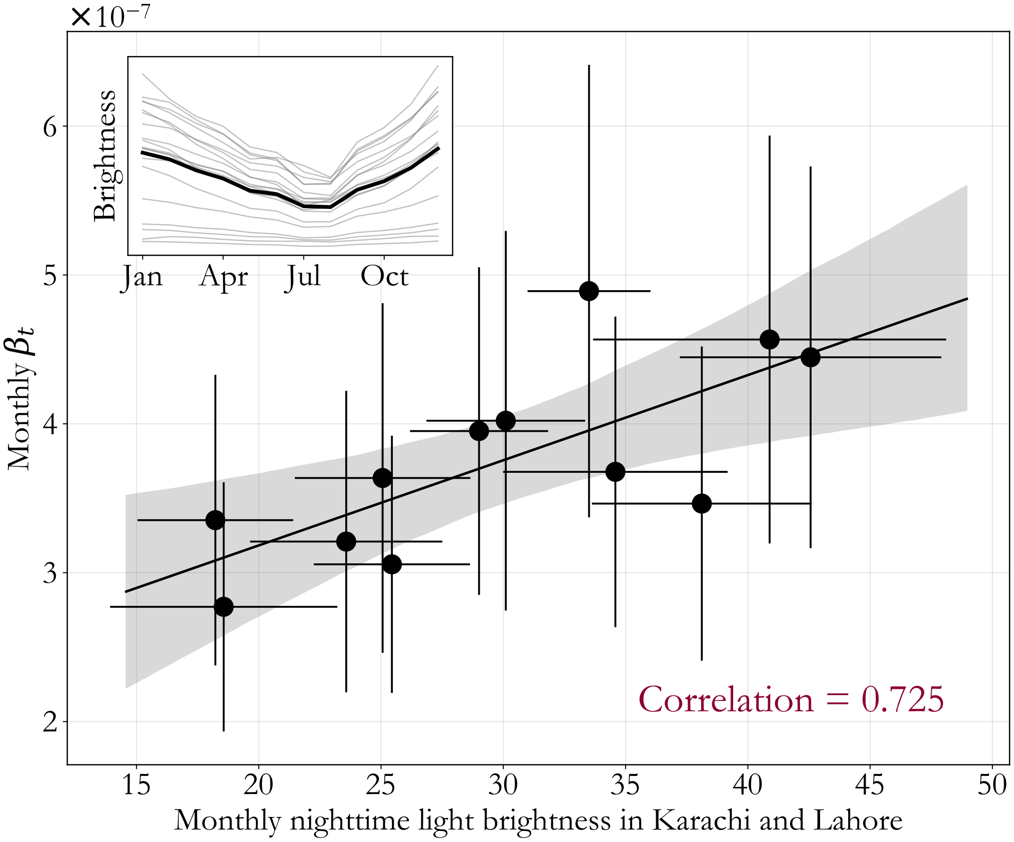

Fig. S7. Correlating rural-to-urban migration with measles transmission. In the inset, we plot average brightness per pixel near Karachi and Lahore, averaged by month from 2012 to 2018 for each district in the cities (grey) and then averaged globally over districts (black). Plotting the global average against inferred measles transmission seasonality, β_t, (black dots, standard deviation error bars) shows that measles seasonality and nighttime light brightness are highly correlated with a strong, linear relationship (black line, 95% CI in grey).

Province	Date	Target population fraction (P_t)
AJK	2014-12-15	1.00
Balochistan	2012-12-31	0.81
	2015-04-30	1.00
FATA	2015-08-31	1.00
Gilgit Baltistan	2015-05-31	1.00
Islamabad	2015-02-28	1.00
Khyber Pakhtoon	2014-05-31	1.00
Punjab	2013-05-15	0.66
	2015-02-15	1.00
Sindh	2013-05-15	0.66
	2014-05-31	1.00
	2017-08-15	0.15

Table S3. Tabulated target population fractions from the WHO SIA calendar and SIA reports (19). For all dates not included, $P_t = 0$. Note that, while we have this information for AJK and FATA, lack of other data (RI coverage for example) prevents modeling, and we ignore those campaigns in the analysis.

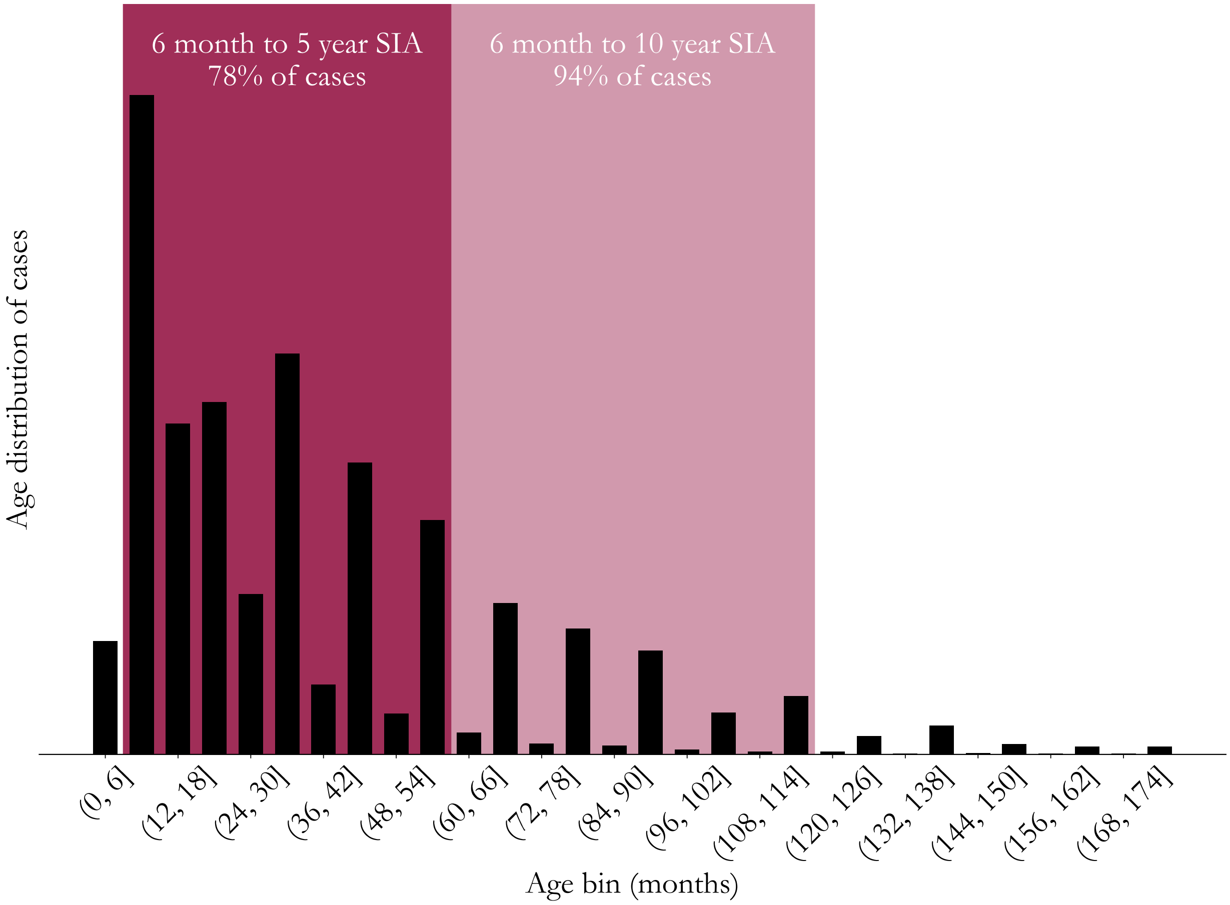

Fig. S8. Campaign age ranges can be compared by calculating the number of lab confirmed cases covered by the age range. Black bars are ages associated with lab-confirmed cases binned into 6 month age bins, showing a peak at low ages as expected for a high burden setting (6). Red shaded regions indicate 2 age ranges, 6 months to 5 years and 6 months to 10 years, corresponding to the 2017 and 2014-15 campaigns respectively. These regions cover 78% and 94% of the age distribution. Since the coverage for the 2014-15 campaign is near total, we use that campaign as a base case from which to set P_t for the others.

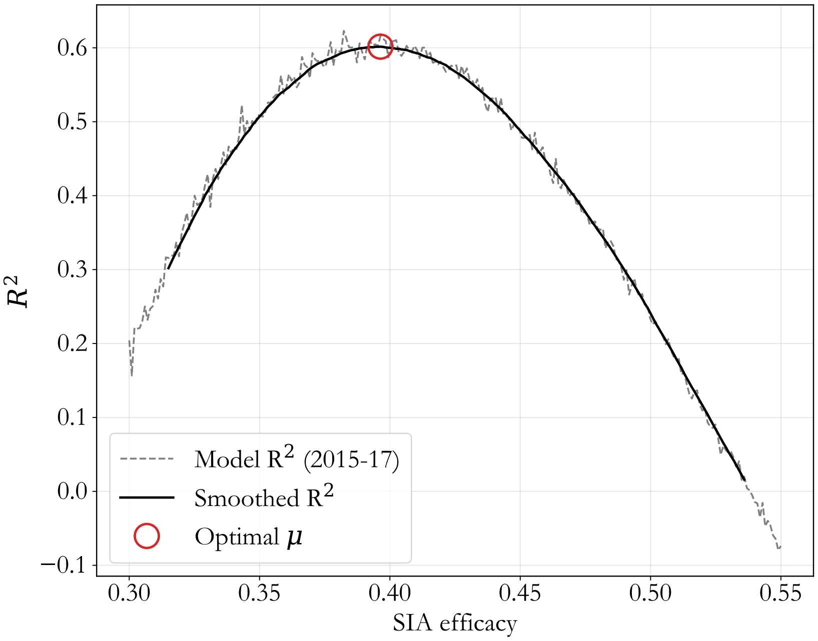

Fig. S9. Optimization for μ is carried out by repeated fitting and model testing. Comparing goodness-of-fit (measured via R^2 in the 2015-2017 model period) yields $\mu \approx 0.4$.

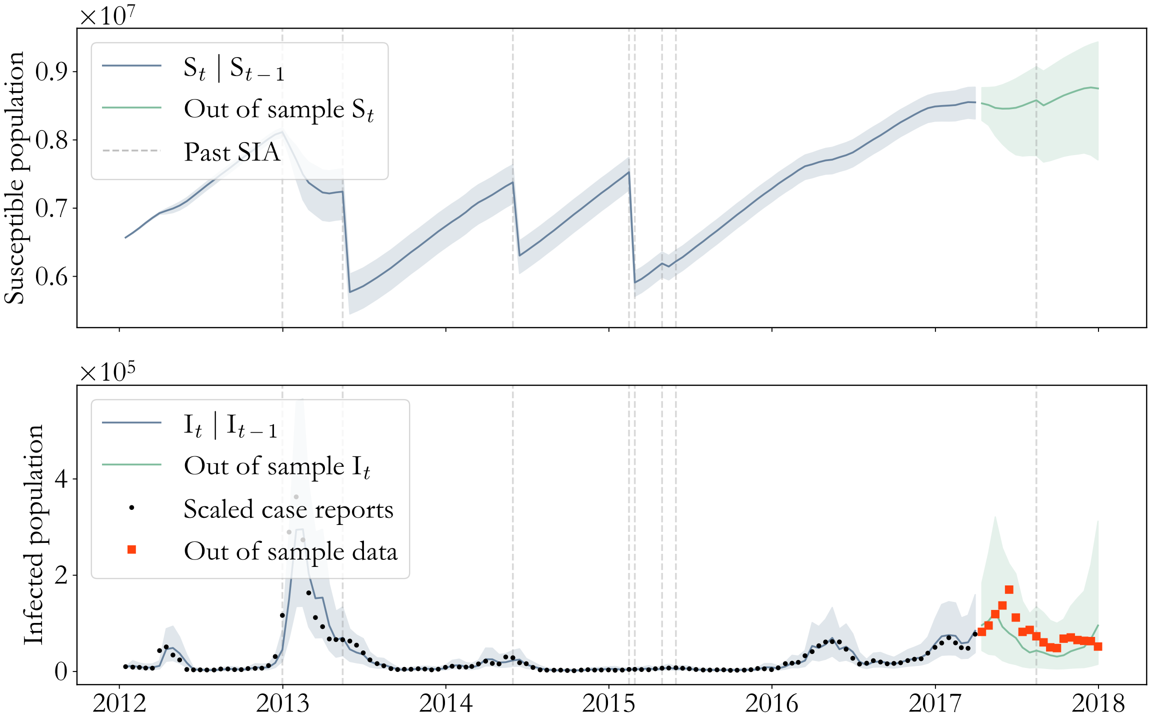

Fig. S10. To test extrapolation performance, we withhold incidence data after Q1 2017 (orange squares) and fit the model to data from 2012 to Q1 2017 (black dots). The model still fits the data accurately (blue time series) while simultaneously predicting the 2017 outbreak (green).

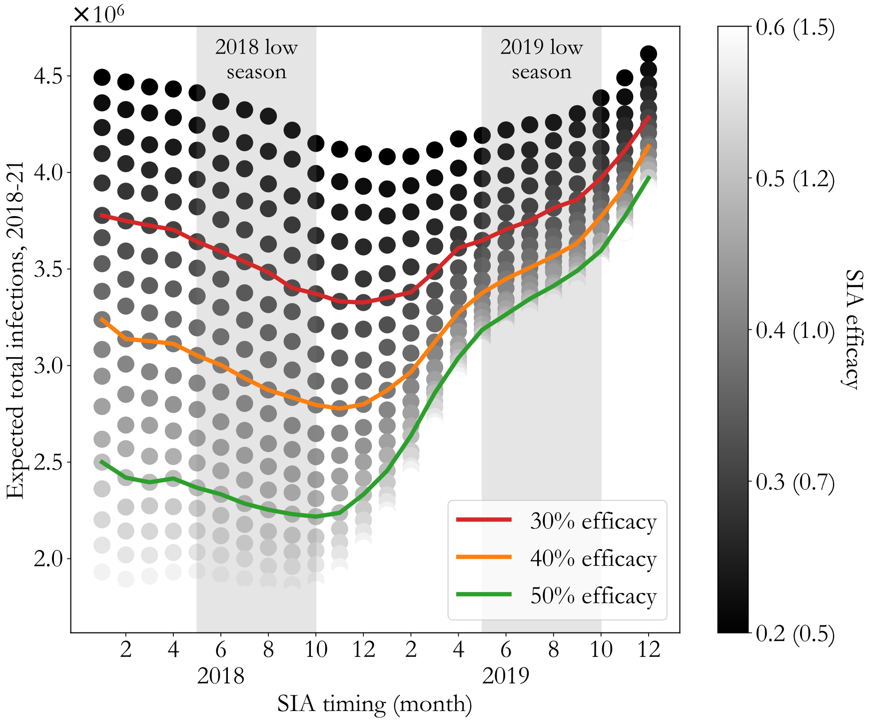

Fig. S11. Once calibrated, the model can be used to test future SIAs with varying efficacy and timing. Dots represent expected infections from 2018-21 as a function of timing with SIA efficacy proportional to color (colorbar, with value relative to inferred μ indicated). Efficacy contours at 0.3 (red), 0.4 (orange), and 0.5 (green) qualitatively demonstrate that infections fall regardless of timing as efficacy increases.

5. Estimating herd immunity's contribution

As indicated by Eqs. 2 and 3, SIAs decrease burden directly (by immunizing susceptible individuals who would have otherwise been infected) and indirectly (via herd immunity, the decrease in probability of subsequent infections at lower S_t). Given a parameterized model, we can estimate each contributions relative effect on total SIA impact.

By comparing a particular SIA policy's total burden from 2018-2021 with the estimated burden over the same period in the absence of a campaign, we can approximate the number of infections averted by the campaign. Simultaneously, analysis of the corresponding S_t time series allows us to calculate the number of susceptible individuals vaccinated in a campaign, placing an upper limit on the SIA's direct effects. Taken together, these two estimates can be compared to calculate the infections averted indirectly, i.e. by herd immunity.

This analysis is summarized in Fig. S12. The average, relative contributions of direct effects (blue) and herd immunity (orange) to the total number of infections averted is plotted for campaigns conducted in each month of 2018. We see that herd immunity effects are maximized at roughly 20% in the inferred low transmission season, indicating that susceptible accumulation offers the opportunity to immunize a larger number of individuals yielding larger herd protection.

6. Testing province-level models

The model can be generalized to province level by assuming transmission related parameters (i.e. those in Eq. 3) remain qualitatively the same and refitting the remaining parameters using data relevant to the province in question. More specifically, we assume β_t from the national level can be scaled by the fraction of the total population in the province of interest to model contact rate's dependence on total population. Then, S_t is reconstructed using the method above with birth, coverage, and case data specific to the province being modeled.

Sparsity of data at the province level necessitates these fairly severe modeling assumptions even though there is no general reason to believe transmission seasonality is qualitatively similar across provinces. On top of this, we ignore transmission and importation between provinces, an effect which has likely occurred in past outbreaks (22). Incorporation of inter-province correlation is a topic of current research.

Even under these assumptions, we can verify that our province level models forecast effectively, indicating that we have captured dominant effects at this spatial scale. In Figs S13-S18, calibration tests like those of Fig. 2 in the main text are shown for each province. We see universally that long-term extrapolation is possible and accurately predicts future outbreaks, even in provinces with relatively small population. This gives us confidence in the models' 2018-2021 projections, which are shown in these figures in the absence of a 2018 SIA.

References

1. Tatem AJ (2017) Worldpop, open data for spatial demography. *Scientific data* 4:170004–170004.
2. Tatem AJ, et al. (2014) Mapping for maternal and newborn health: the distributions of women of childbearing age, pregnancies and births. *International journal of health geographics* 13(1):2.
3. Gaughan AE, Stevens FR, Linard C, Jia P, Tatem AJ (2013) High resolution population distribution maps for Southeast Asia in 2010 and 2015. *PloS one* 8(2):e55882.
4. (2013). National Institute of Population Studies - NIPS/Pakistan and ICF International, "Pakistan Demographic and Health Survey 2012-13".
5. (2003). World Health Organization, "WHO-recommended standards for surveillance of selected vaccine preventable diseases".
6. Prada JM, Metcalf CJE, Ferrari MJ (2018) Improving measles incidence inference using age-structured serological data. *Epidemiology & Infection* 146(13):1699–1706.
7. Rutstein SO (2006) Guide to DHS statistics.
8. Lumley T (2004) Analysis of complex survey samples. *Journal of Statistical Software* 9(1):1–19. R package verson 2.2.
9. Lumley T (2016) survey: analysis of complex survey samples. R package version 3.32.
10. Simons E, et al. (2012) Assessment of the 2010 global measles mortality reduction goal: results from a model of surveillance data. *The Lancet* 379(9832):2173–2178.
11. Plotkin S, Orenstein W, Offit P, Edwards KM (2017) *Plotkin's Vaccines.* (Elsevier), 7 edition.
12. Finkenstädt BF, Grenfell BT (2000) Time series modelling of childhood diseases: a dynamical systems approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 49(2):187–205.
13. Keeling MJ, Rohani P (2011) *Modeling infectious diseases in humans and animals.* (Princeton University Press).
14. Ferrari MJ, et al. (2008) The dynamics of measles in sub-Saharan Africa. *Nature* 451(7179):679.
15. Bharti N, et al. (2011) Explaining seasonal fluctuations of measles in Niger using nighttime lights imagery. *Science* 334(6061):1424–1427.
16. Bharti N, Djibo A, Tatem AJ, Grenfell BT, Ferrari MJ (2016) Measuring populations to improve vaccination coverage. *Scientific reports* 6:34541.
17. (2018). National Oceanic and Atmospheric Administration, "Version 1 VIIRS Day/Night Band Nighttime Lights", (https://www.ngdc.noaa.gov/eog/viirs/).
18. Mills S, Weiss S, Liang C (2013) Viirs day/night band (dnb) stray light characterization and correction in *Earth Observing Systems XVIII.* (International Society for Optics and Photonics), Vol. 8866, p. 88661P.

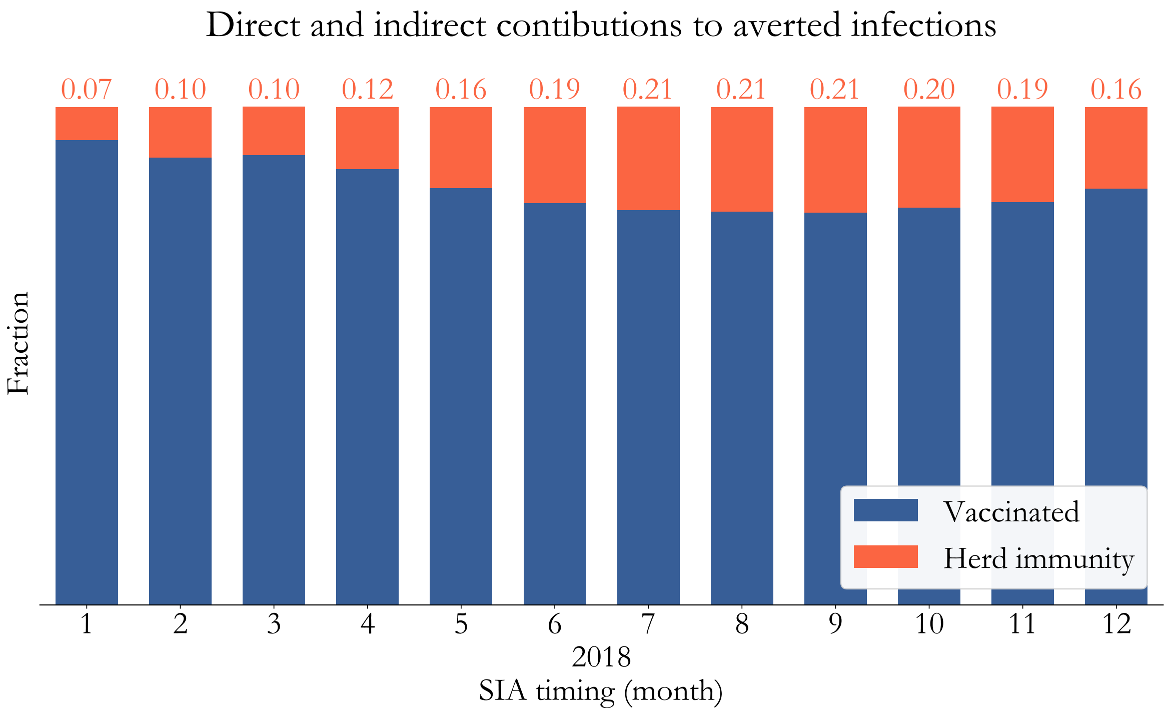

Fig. S12. Relative contributions of vaccination (blue) and herd immunity (orange) to total averted infections are compared for SIAs conducted in each month of 2018. We find that herd immunity effects are maximized at approximately 20% during the low transmission season.

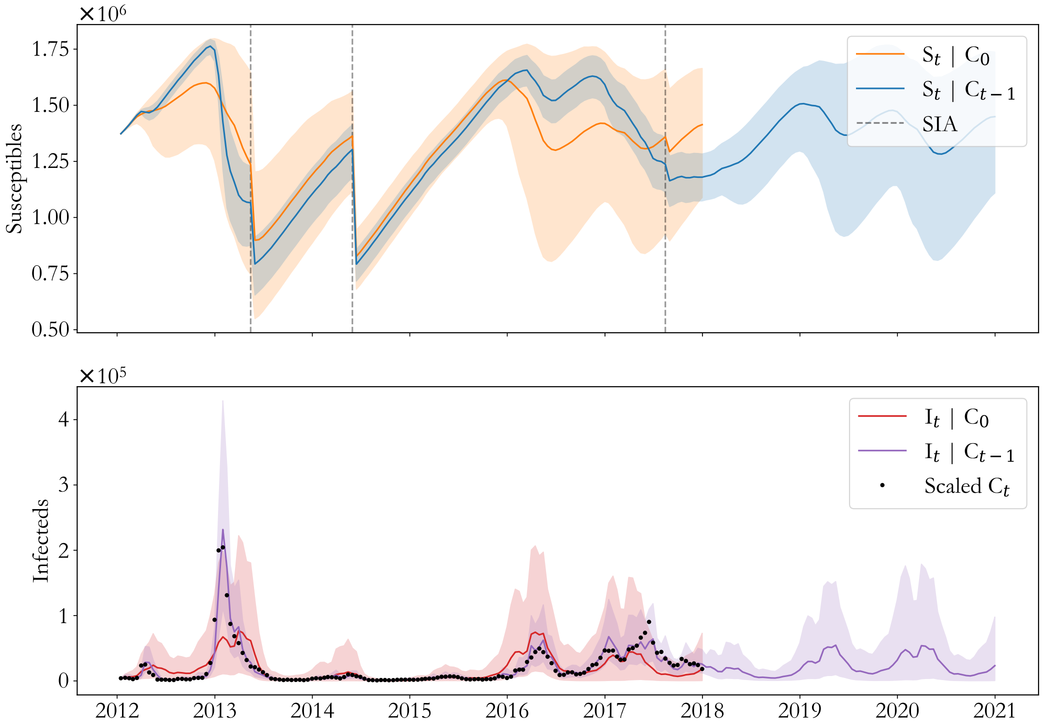

Fig. S13. Model testing and performance for Sindh.

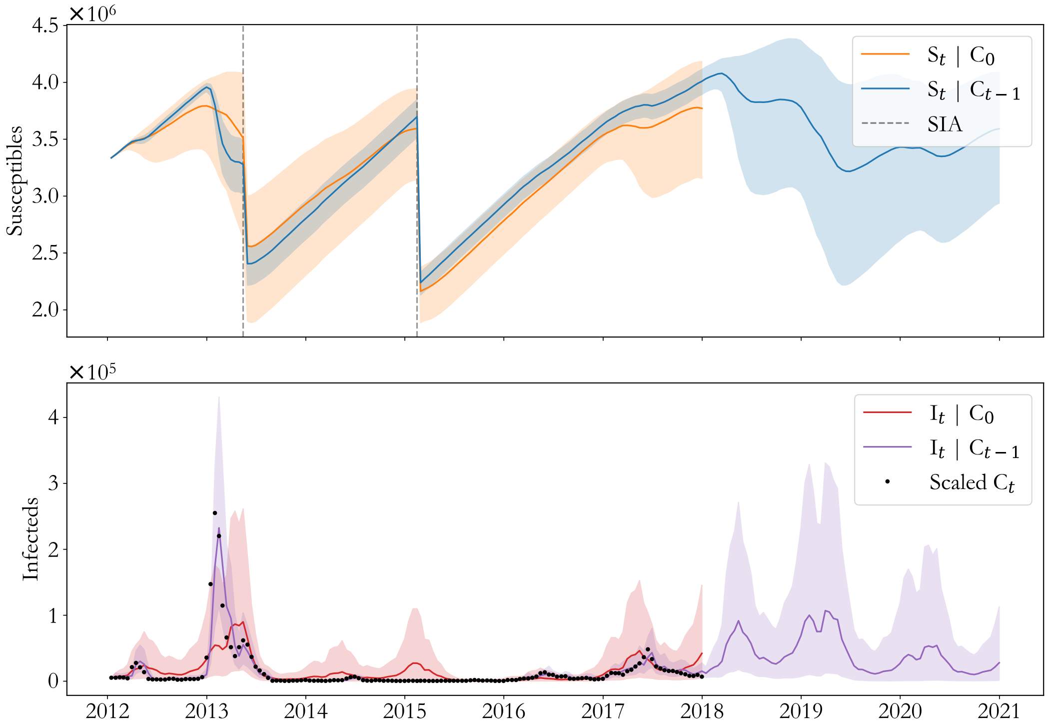

Fig. S14. Model testing and performance for Punjab.

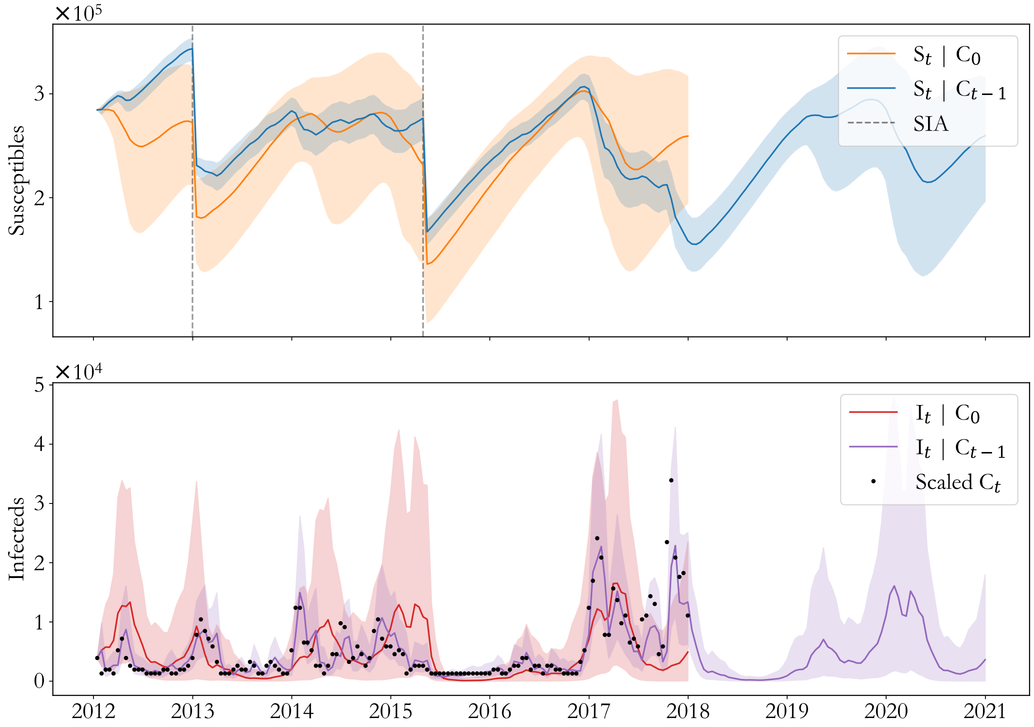

Fig. S15. Model testing and performance for Balochistan.

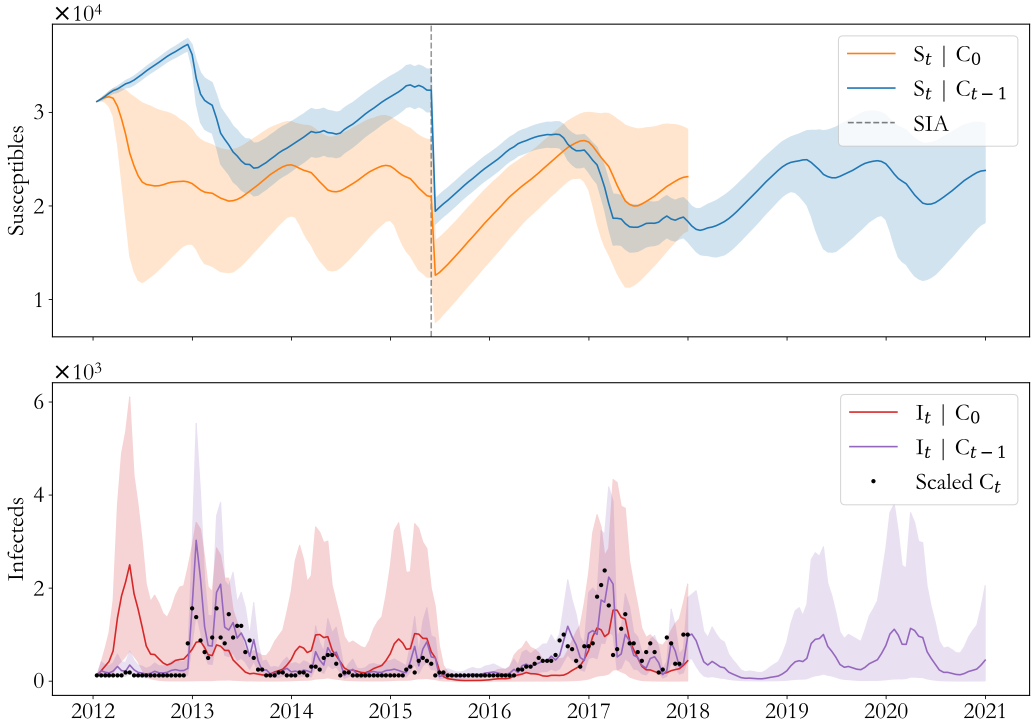

Fig. S16. Model testing and performance for Gilgit Baltistan.

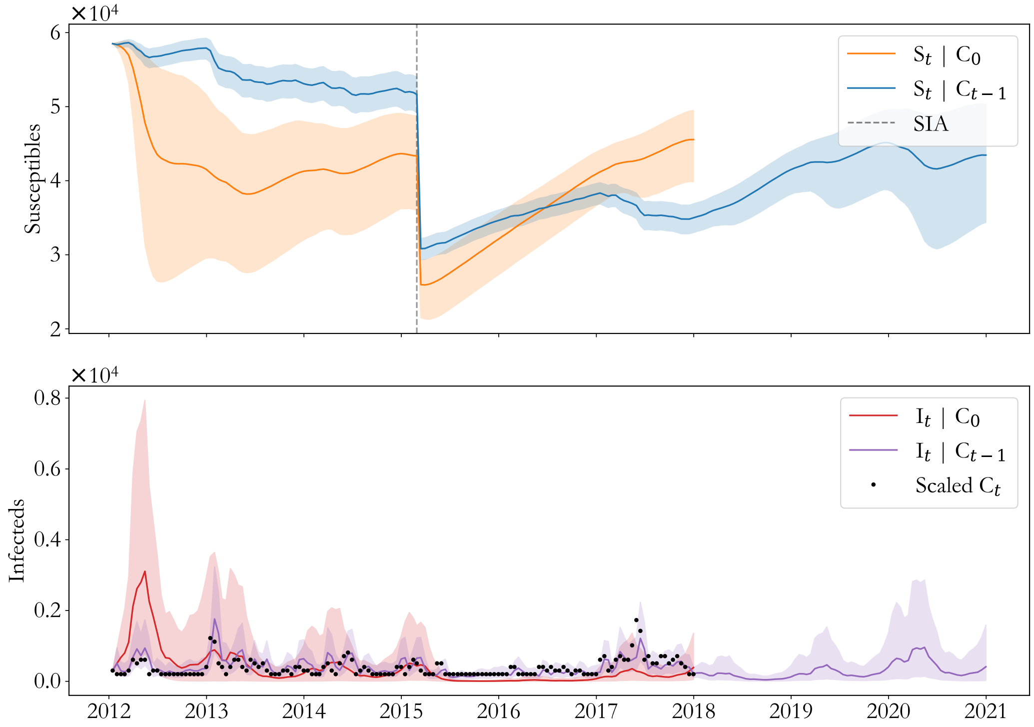

Fig. S17. Model testing and performance for Islamabad.

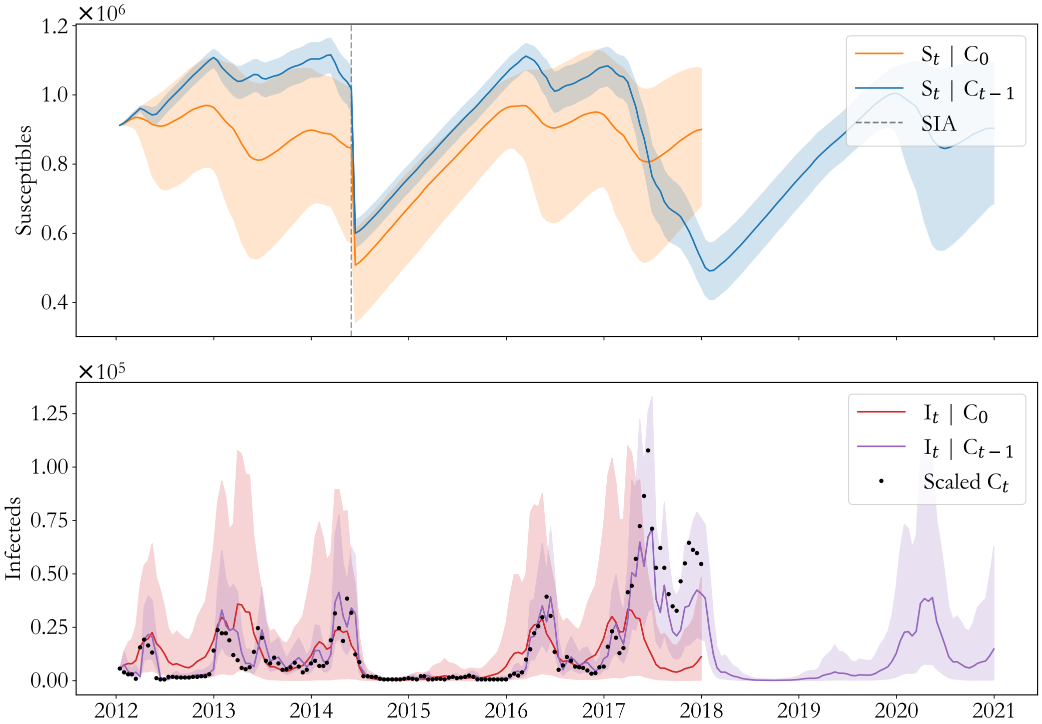

Fig. S18. Model testing and performance for Khyber Pakhtoon.

19. (2017). World Health Organization. “Summary of measles-rubella supplementary immunization activities”, (http://www.who.int/immunization/monitoring_surveillance/data/en/).
20. Ionides EL, Bretó C, King AA (2006) Inference for nonlinear dynamical systems. *Proceedings of the National Academy of Sciences* 103(49):18438–18443.
21. Ionides EL, Nguyen D, Atchadé Y, Stoev S, King AA (2015) Inference for dynamic and latent variable models via iterated, perturbed Bayes maps. *Proceedings of the National Academy of Sciences* 112(3):719–724.
22. Khan T, Qazi J (2014) Measles outbreaks in pakistan: causes of the tragedy and future implications. *Epidemiology Reports* 2(1):1.