Submitted to the Annals of Applied Statistics

SUPPLEMENTARY MATERIAL FOR "SPACE-TIME SMOOTHING OF COMPLEX SURVEY DATA: SMALL AREA ESTIMATION FOR CHILD MORTALITY"

By Laina D Mercer[1], Jon Wakefield[1,2], Athena Pantazis[3], Angelina M Lutambi[4] Honorati Masanja[4] and Samuel Clark[3,5,6,7,8]

[1]*Departments of Statistics University of Washington, USA*
[2] *Department of Biostatistics, University of Washington, USA*
[3] *Department of Sociology, University of Washington, USA,*
[4] *Ifakara Health Institute, Dar es Salaam, TZA*
[5]*Institute of Behavioral Science (IBS), University of Colorado at Boulder, Boulder, CO USA*
[6] *MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa*
[7] *INDEPTH Network, Accra, Ghana*
[8] *ALPHA Network, London, UK*

1. Details of Discrete Survival Model. We wish to estimate the under 5 mortality rate (U5MR) denoted by ${}_5q_0$. The U5MR is calculated as

$$ {}_5q_0 = 1 - \prod_{j=1}^{J}(1 - {}_{n_j}q_{x_j}). \tag{1.1}$$

As described in the main text, the complement of surviving each month of the interval $[x_j, x_j + n_j)$ is used to calculate ${}_{n_j}\widehat{q}_{x_j} = 1 - \left(1 - \text{expit}(\widehat{B}_j)\right)^{n_j}$ where $B_j = \text{logit}({}_1q_x)$ for $x \in [x_j, x_j + n_j)$. Since

$$\begin{aligned} 1 - {}_{n_j}\widehat{q}_{x_j} &= \left(1 - \text{expit}(\widehat{B}_j)\right)^{n_j} = \left(\frac{e^{\widehat{B}_j}+1}{e^{\widehat{B}_j}+1} - \frac{e^{\widehat{B}_j}}{e^{\widehat{B}_j}+1}\right)^{n_j} \\ &= \left(\frac{1}{e^{\widehat{B}_j}+1}\right)^{n_j} = \left(e^{\widehat{B}_j}+1\right)^{-n_j} \end{aligned}$$

(1.1) can be written in terms of $\widehat{\mathbf{B}}$ as

$${}_5\widehat{q}_0 = 1 - \prod_{j=1}^{J}(1 - {}_{n_j}\,\widehat{q}_{x_j}) = 1 - \prod_{j=1}^{J}\left(e^{\widehat{B}_j}+1\right)^{-n_j}.$$

2. Derivation of Standard Error for U5M. In a finite population of size N we may fit the model $Y_i|\boldsymbol{\beta} \sim \text{Binomial}(1, p_i)$, with $p_i = \exp(\boldsymbol{x}_i^{\mathrm{T}}\widehat{\boldsymbol{\beta}}))/[1+\exp(\boldsymbol{x}_i^{\mathrm{T}}\widehat{\boldsymbol{\beta}})]$ and $\boldsymbol{\beta}$ a $J \times 1$ vector. The finite population parameter $\boldsymbol{B}$ is the solution to the score equations:

$$\sum_{i=1}^{N} x_{ij}\left[y_i - \frac{\exp(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{B})}{1+\exp(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{B})}\right] = 0 \text{ for } j = 1, \ldots, J.$$

When a survey is taken, following Binder (1983), a design-based estimate of $\boldsymbol{B}$ is given by the solution to

$$\sum_{i\in S} w_i x_{ij}\left[y_i - \frac{\exp(\boldsymbol{x}_i^{\mathrm{T}}\widehat{\boldsymbol{B}})}{1+\exp(\boldsymbol{x}_i^{\mathrm{T}}\widehat{\boldsymbol{B}})}\right] = 0 \text{ for } j = 1, \ldots, J$$

where S represents the units included in the sample. The design-based variance of $\boldsymbol{B}$ is more difficult and the following is based on Roberts *et al.* (1987), using the notation of that paper. Suppose we have a saturated logistic model (as in our example) with I groups (factor levels). Let N_i be the true number of individuals who fall in group i and $N = \sum_{i=1}^{I} N_i$ the total number of individuals in the population. Also let N_{i1} be the number of individuals responding in group i. The ratio estimator of the proportion responding is

$$p_i = \frac{\widehat{N}_{i1}}{\widehat{N}_i} = \frac{\sum_{k\in S_i} w_k y_k}{\sum_{k\in S_i} w_k}$$

where S_i is the random set of the sampled units that fall in group i and w_k is the design weight associated with surveyed response y_k, $k = 1, \ldots, n$ (so that n is the size of the survey). Also let

$$w_i = \frac{\widehat{N}_i}{\widehat{N}} = \frac{\sum_{k\in S_i} w_k}{\sum_{k\in S} w_k}$$

where S is the random set of all the samples. The design-based variance estimator of $\boldsymbol{p} = [p_1, \ldots, p_I]^{\mathrm{T}}$ will be denoted $\widehat{\boldsymbol{V}}$, and obviously depends on the design chosen. The pseudo-MLEs of the fractions responding in group i will be denoted $\widehat{f}_i$. Then the variance-covariance of the estimator of $\boldsymbol{B}$ is given by equation (2.4) of Roberts *et al.* (1987):

$$\widehat{\text{var}}(\widehat{\boldsymbol{B}}) = n^{-1}\widehat{\Delta}^{-1}\boldsymbol{D}(w)\widehat{\boldsymbol{V}}\boldsymbol{D}(w)\widehat{\Delta}^{-1} \tag{2.1}$$

where $\Delta = \text{diag}(w_1\widehat{f}_1(1-\widehat{f}_1), \ldots, w_I\widehat{f}_I(1-\widehat{f}_I))$ and $\boldsymbol{D}(w) = \text{diag}(w_1, \ldots, w_I)$.

We choose to model the parameter $\eta = \text{logit}({}_5q_0)$. Ultimately we will use a Gaussian distribution for the first stage of our hierarchical model, so we

would like a parameter that can take values along the whole real line. We have

$$\begin{aligned}\eta = \text{logit}({}_5q_0) &= \log\left(\frac{{}_5q_0}{1-{}_5q_0}\right)\\ &= \log\left(\frac{1-\prod_{j=1}^{J}\left(e^{B_j}+1\right)^{-n_j}}{\prod_{j=1}^{J}\left(e^{B_j}+1\right)^{-n_j}}\right)\\ &= \log\left(\prod_{j=1}^{J}\left(e^{B_j}+1\right)^{n_j}-1\right)\end{aligned} \tag{2.2}$$

The asymptotic distribution of the MLE is $\widehat{\boldsymbol{B}} \sim N(\boldsymbol{B}, \Sigma)$ and from the delta method we can find the asymptotic distribution of $\hat{\eta}$:

$$\widehat{\eta} \sim N\left(\text{logit}({}_5q_0), \widehat{V}_{\text{DES}}\right) \tag{2.3}$$

where

$$\widehat{V}_{\text{DES}} = \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{B}}^T \widehat{\Sigma} \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{B}}. \tag{2.4}$$

where $\widehat{\Sigma}$ is $\widehat{\text{var}}(\widehat{\boldsymbol{B}})$ from (2.1). The weights depend on the design and $\widehat{\Sigma}$ can be extracted from the `svyglm()` function within the `survey` package.

Then, if we define:

$$\gamma = \prod_{j=1}^{J}\left(e^{B_j}+1\right)^{n_j}$$

and set $\eta = \log(\gamma - 1)$ and, after some algebra,

$$\frac{\partial \eta}{\partial B_j} \quad = \quad \frac{\gamma}{\gamma - 1} \times \left[n_j \times \text{expit}(B_j)\right].$$

The value of $\frac{\partial \eta}{\partial \boldsymbol{B}}$ is used in (2.4) for the asymptotic distribution, (2.3) which is used to derive $(1-\alpha)\%$ confidence intervals for ${}_5q_0$ as

$$\left[\text{expit}\left(\hat{\eta} + \sqrt{\widehat{V}_{\text{DES}}} \times z_{\alpha/2}\right), \text{expit}\left(\hat{\eta} + \sqrt{\widehat{V}_{\text{DES}}} \times z_{1-\alpha/2}\right)\right]. \tag{2.5}$$

3. Simulation to test coverage performance of derived SE. Pedersen and Liu (2012) discuss the difficulties associated with deriving a variance estimate in the context of child mortality estimates. DHS typically uses a jackknife estimator, $V_{\text{JACK}}({}_5\widehat{q}_0)$ of ${}_5\widehat{q}_0$, which for cluster sampling is

$$\widehat{V}_{\text{JACK}} = \frac{N_c - 1}{N_c} \sum_{c=1}^{N_c} ({}_5\widehat{q}_{0(c)} - {}_5\widehat{q}_0)^2 \tag{3.1}$$

where N_c is the number of clusters and ${}_5\widehat{q}_{0(c)}$ is the estimate based on all of the data while holding out the c-th cluster (Lohr, 2009, ch. 9). A 95% confidence interval for ${}_5q_0$ is based on

$$\left[\quad {}_5\widehat{q}_0 - 1.96 \times \sqrt{\widehat{V}_{\text{JACK}}} \quad , \quad {}_5\widehat{q}_0 + 1.96 \times \sqrt{\widehat{V}_{\text{JACK}}} \quad \right]. \tag{3.2}$$

3.1. *Data Generation.* A simulated dataset was created to asses the coverage properties of interval based on V_{DES} and V_{JACK}. To simulate the estimation process within one region 100,000 women were assigned to 500 clusters. The number of births for each woman were generated from a Poisson distribution with rate of 3. Each birth was assigned a calendar month between 0 and 119 (a 10 year period).

We considered three scenarios to generate the deaths within the first 60 months of life. Deaths were assigned based on a multinomial distribution with the following discrete hazards p_0 for the first month, p_1 for months 2–12 and p_2 for months 13–60. As we are only interested in death within the first 5 years, the remaining probability was assigned to a 61st category for death after the age of 5 years.

In the first scenario we assume that monthly probabilities of death are constant between clusters and assumed

$$\begin{aligned} \text{logit}(p_0) &= \alpha_0 \\ \text{logit}(p_1) &= \alpha_0 + \alpha_1 \\ \text{logit}(p_2) &= \alpha_0 + \alpha_2 \end{aligned}$$

where α_0, α_1, and α_2 were chosen to correspond to probabilities of 0.04, 0.004, and 0.001, respectively as shown in shown in Figure 1. This distribution of probability represents the highest risk in the first month of life, an elevated risk for the remainder of the first year, and the lowest risk for the following 4 years. These probabilities result in an expected U5MR of 124.5 deaths per 1,000, which is similar to the late 1990s U5MR in Tanzania.

In the second and third scenario we generated data with cluster-specific probabilities

$$\begin{aligned} \text{logit}(p_{0,c}) &= \alpha_0 + \epsilon_c \\ \text{logit}(p_{1,c}) &= \alpha_0 + \alpha_1 + \epsilon_c \\ \text{logit}(p_{2,c}) &= \alpha_0 + \alpha_2 + \epsilon_c \end{aligned}$$

for clusters $c = 1, ..., 500$. Scenario 2 had low between cluster variability with $\epsilon_c \sim N(0, \sigma = 0.1)$ resulting in the U5MR varying from 81.7 to 156.2 deaths per 1,000 births. Scenario 3 had higher between cluster variability with $\epsilon_c \sim N(0, \sigma = 0.3)$ resulting in the U5MR varying from 54.0 to 265.9 deaths per 1,000 births. The distribution of cluster-specific U5MRs are shown in Figure 2. As we are only interested in death within the first 5 years, the remaining probability was assigned to a 61st category for death after the age of 5. Code for data generation and the simulation can be found at http://faculty.washington.edu/jonno/software.html.

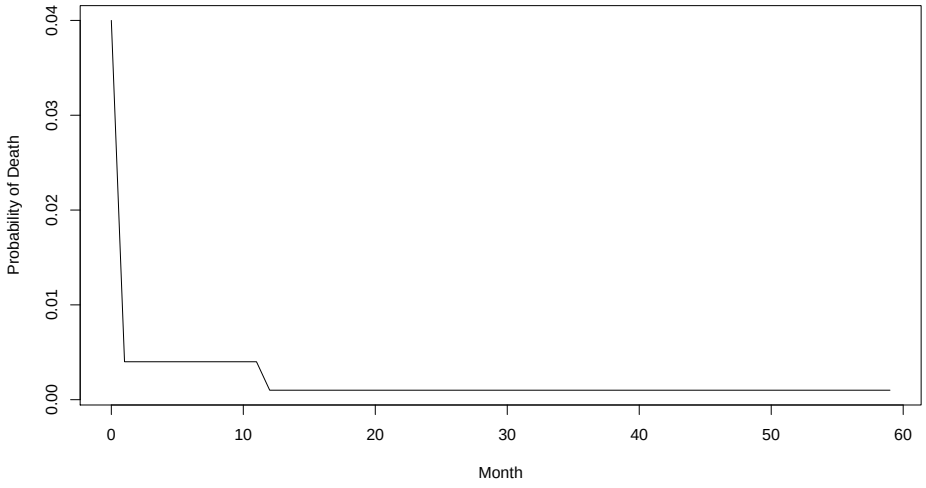

Fig 1: Monthly probability of death for the first 60 months.

When creating estimates for a particular 5 year calendar period from the household survey and HDSS data, births before or during this period contribute person-time from children who are born or die or are censored in other 5 year time periods. To mimic this aspect of the real data, births were simulated from a wider time period and then subsetted to include only

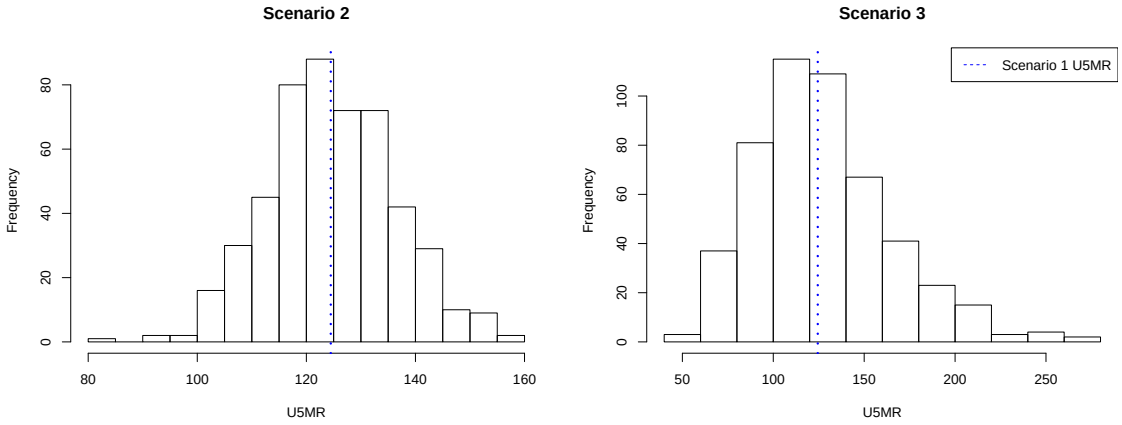

Fig 2: Simulated cluster-specific under 5 child mortality rates (U5MR).

observations within the relevant time period. Children who survive past 60 months only contribute person time for months 0–59. This subset included person-time from approximately 240,000 unique births from 91,000 mothers and the U5MR in these populations were 131.2, 133.3, and 136.4 per 1,000 for scenarios 1, 2, and 3, respectively.

3.2. *Simulation Procedure.* We employed a two stage cluster sampling design. At the first stage n_{c} clusters (analogous to enumeration areas in the TDHS designs) were randomly selected from the N_{c} available. At the second stage, suppose cluster c is selected, then n_{W} women were randomly selected from the $N_{\mathrm{W}c}$ total women within the selected cluster. The resulting sampling weights for a mother selected in cluster c is

$$w_{\mathrm{E}c} = \frac{N_{\mathrm{c}}}{n_{\mathrm{c}}} \times \frac{N_{\mathrm{W}c}}{n_{\mathrm{W}}}.$$

The number of clusters was set at $N_{\mathrm{c}} = 500$ for all simulations and the number selected at the first stage was one of $n_{\mathrm{c}} = 15, 25$. Sample sizes within clusters (n_{W}) varied by 5 within 10–30. For each combination of n_{c} and n_{W} we draw 1,000 samples and the corresponding delta method and jackknife confidence intervals were created based on $\widehat{V}_{\mathrm{DES}}$ and $\widehat{V}_{\mathrm{JACK}}$, respectively. The sample coverage of each interval type was calculated as the average number of intervals that contained the true population U5MR.

3.3. *Results.* The coverage of the delta method and jackknife intervals by number of clusters and within sample size cluster are shown in Figure

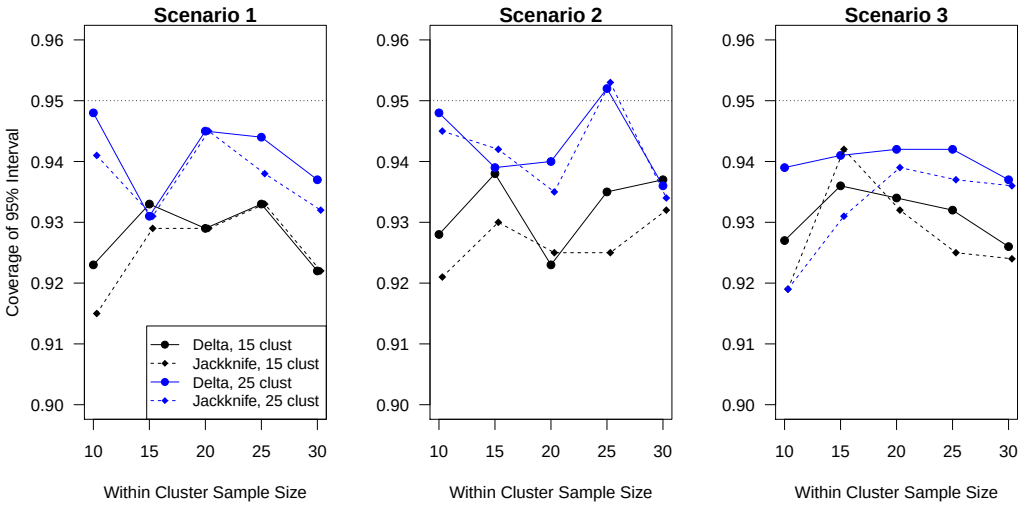

Fig 3: Coverage of jackknife and delta method 95% confidence intervals by number of clusters and within cluster sample size. Sample size values are offset horizontally to improve the visibility.

3. Results are much as one would expect from clustered sampling, coverage improves when there are more clusters and within a given number of clusters there is little gain in precision when increasing the sample size. Generally the performance of the delta method and jackknife intervals is very similar. Figure 4 displays the average 95% confidence interval width by number of clusters and sample size within cluster. As expected, the intervals narrow as either the number of first stage clusters or the number of samples within each cluster increases and scenario 3, which has the most between cluster variability, has the widest intervals.

The TDHS sampling scheme is generally around 25 clusters with a within sample size of approximately 20 women, which corresponds to the blue line at a sample size of 20. This suggest that TDHS intervals may be slightly anti-conservative. We prefer the delta method as it is generally applicable (i.e., to a variety of designs) and has a far smaller computational burden. We conclude that the asymptotic normal sampling distribution and the delta method variance result in sufficiently accurate confidence interval coverage for the cluster and sample sizes considered in our application. Consequently, we will use the asymptotic distribution with the delta method variance as a working likelihood.

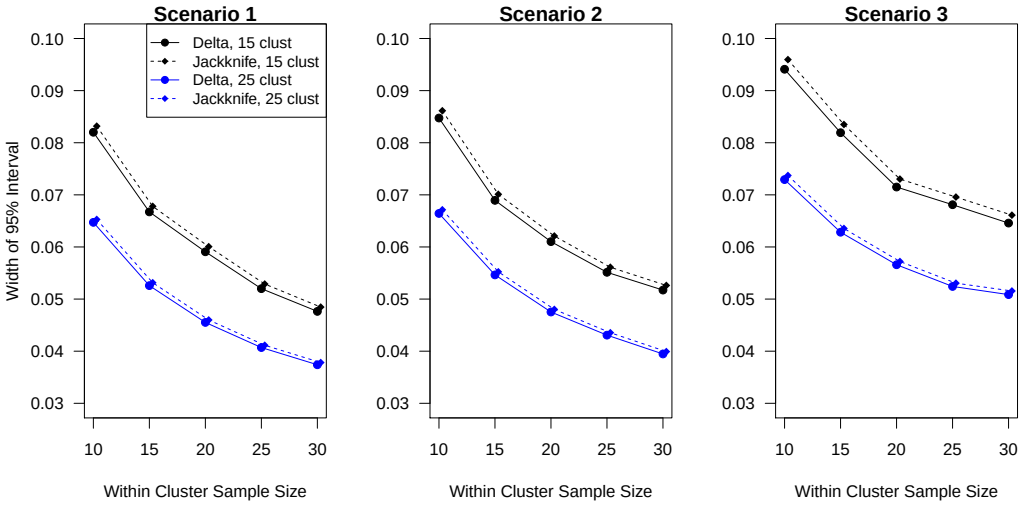

Fig 4: Mean width of jackknife and delta method 95% confidence intervals by number of clusters and within cluster sample size. Sample size values are offset horizontally to improve the visibility.

4. Hyperprior Specification. Figure 5 illustrates the sensitivity of point and interval estimates with three different prior distributions corresponding to 95% ranges for the residuals odds ratios of [0.5,2], [0.2,5], [0.1,10]. The first choice was used for the results presented in the paper. We see sensitivity for the spatial random effects, though the total spatial random effects contributions remain relatively constant since as the structured random effects increase, the unstructured random effects decrease. The structured temporal random effects are robust, which is reassuring since these provide the largest contribution to the overall variability; these are well-estimated, however, since the trend is strong. Similarly, the unstructured survey-area random effects are robust, but all of the standard deviations of the remaining independent random effects show modest increases as the prior moves further from zero. Examples of the code used to find the prior parameters for the ICAR, RW1 and RW2 models are included in Section 9.

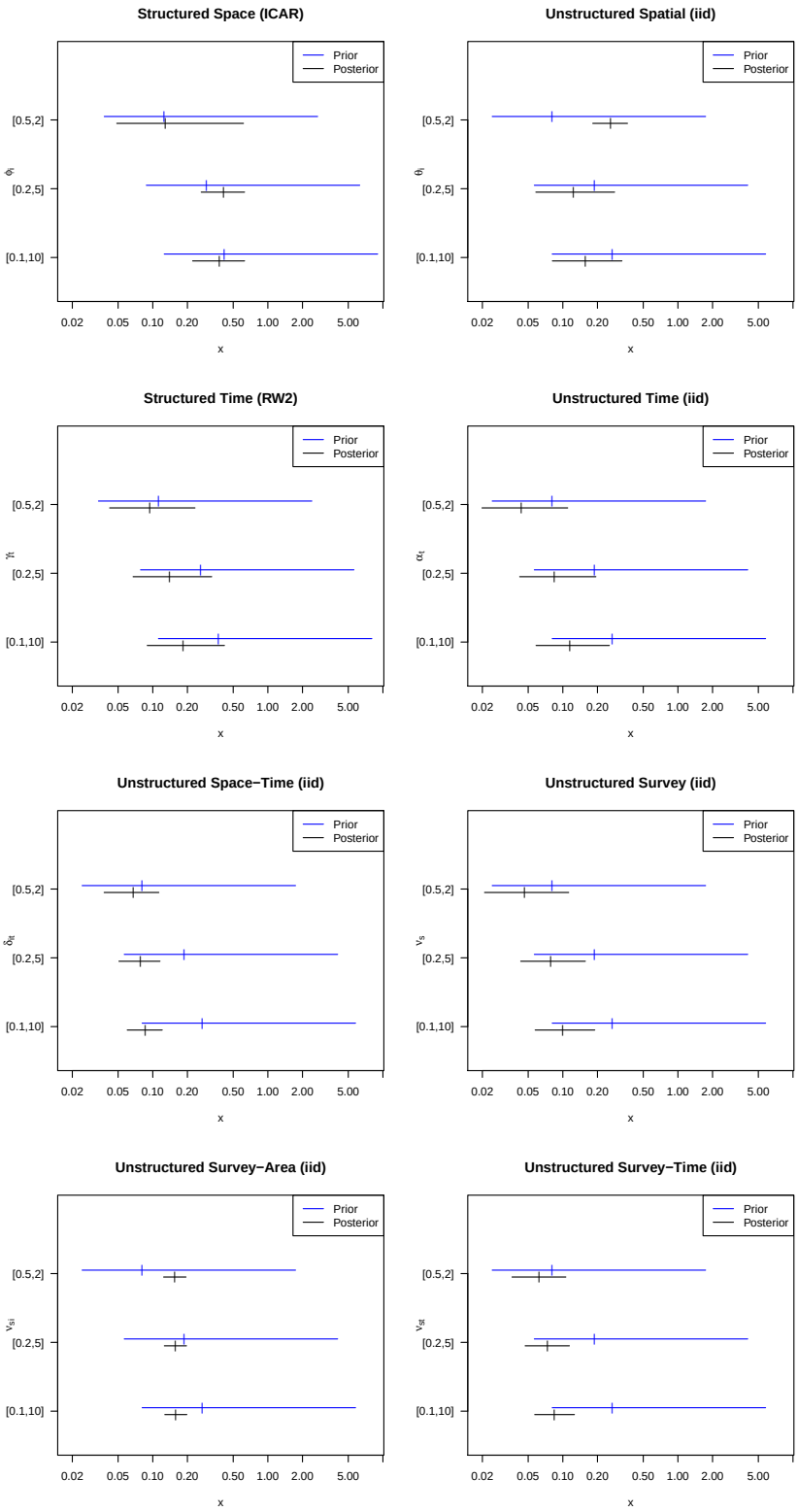

Fig 5: Prior sensitivity of the standard deviations of the eight random effects in the model. The three priors are based on 95% prior intervals on the residual odds ratios of [0.5,2], [0.2,5], [0.1,10].

5. Summary of Random Effects. In this section we present graphical summaries of the various random effects present in Model Vb of the Tanzania U5MR model. Figure 6 presents the posterior medians of the spatial (ICAR) and unstructured random effects (note that the scales on the different plots differ). All temporally structured random effects are from a RW2.

Figure 7 plots the unstructured temporal random effects versus period, along with the survey by period random effects. The unstructured random effects are relatively small in magnitude compared with the survey by period random effects.

Figure 8 displays the unstructured time random effects (α_t) compared with the structured time random effects (γ_t). The structured random effects have a much larger range than the unstructured effects.

Figure 9 displays structured time random effects (γ_t) by time. There is a noticeable negative trend in the random effects.

Figure 10 provides maps of the unstructured space-time random effects (δ_{it}). There is no clear spatial pattern to the high-valued and low-valued random effects by time period. So, based on this plot, there is little evidence of space-time interaction (which is consistent with the model comparison statistics given in the paper). Similarly, the plots in Figure 11 which display the survey-area random effects (ν_{is}) along with the magnitude of the survey-specific (ν_s) random effect, do not show similar spatial patterns over the different time periods.

6. Comparison of weighted and unweighted estimates. Figure 12 provides maps of the inverse-variance weighted Horvitz-Thompson regional estimates of child mortality. Unlike the smoothed maps provided in the main text these maps display some unlikely temporal trends. Figure 13 displays the differences in weighted and unweighted regional estimates of U5MR and associated variances for all surveys. We see that the weighting makes a significant impact on many of the region U5MR estimates and variances.

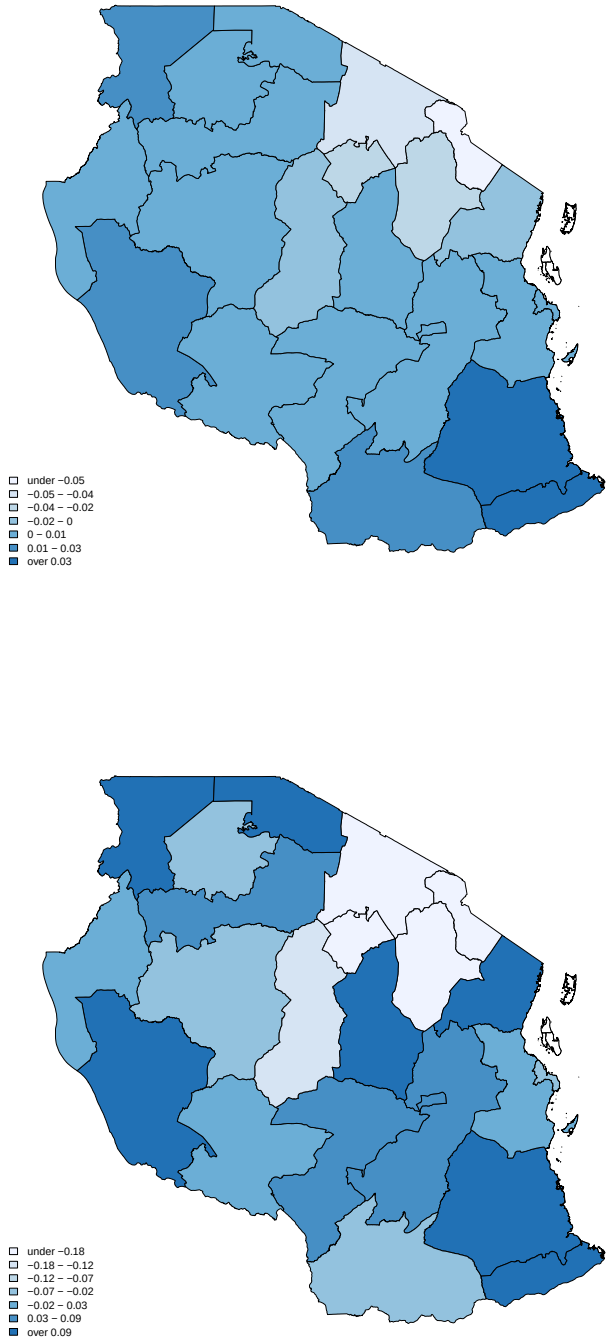

Fig 6: ICAR random effects, ϕ_i (top) and unstructured spatial random effects, θ_i (bottom).

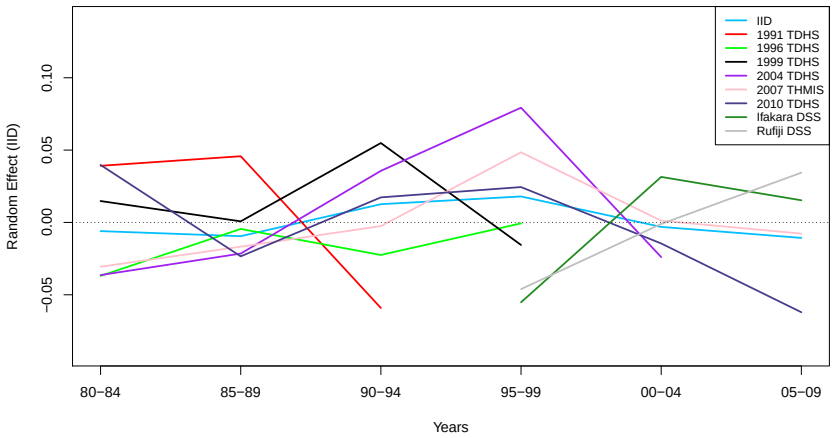

Fig 7: Unstructured time (α_t) and survey-time (ν_{st}) random effects.

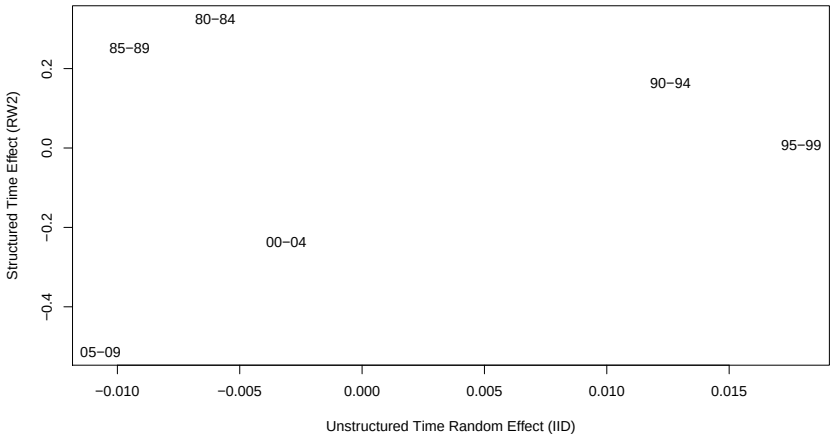

Fig 8: Unstructured time (α_t) and structured time (γ_t) random effects.

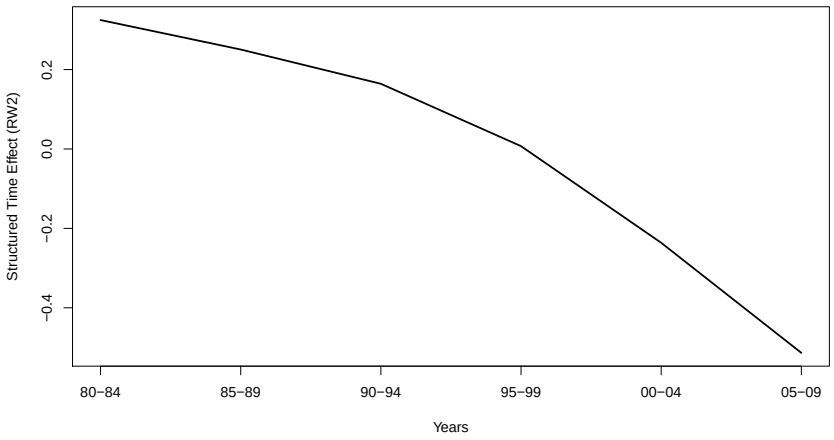

Fig 9: Structured time (γ_t) random effects.

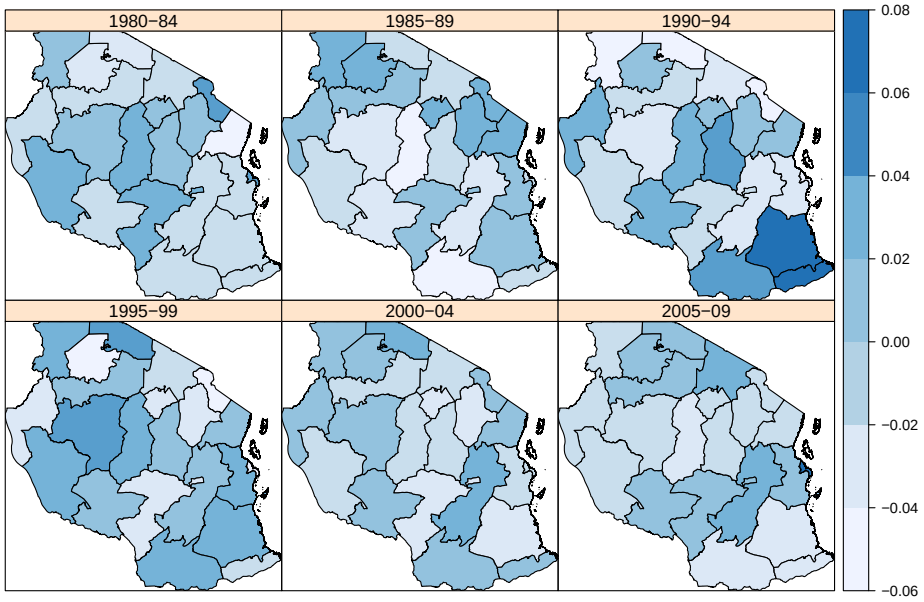

Fig 10: Unstructured space-time random effects (δ_{it}).

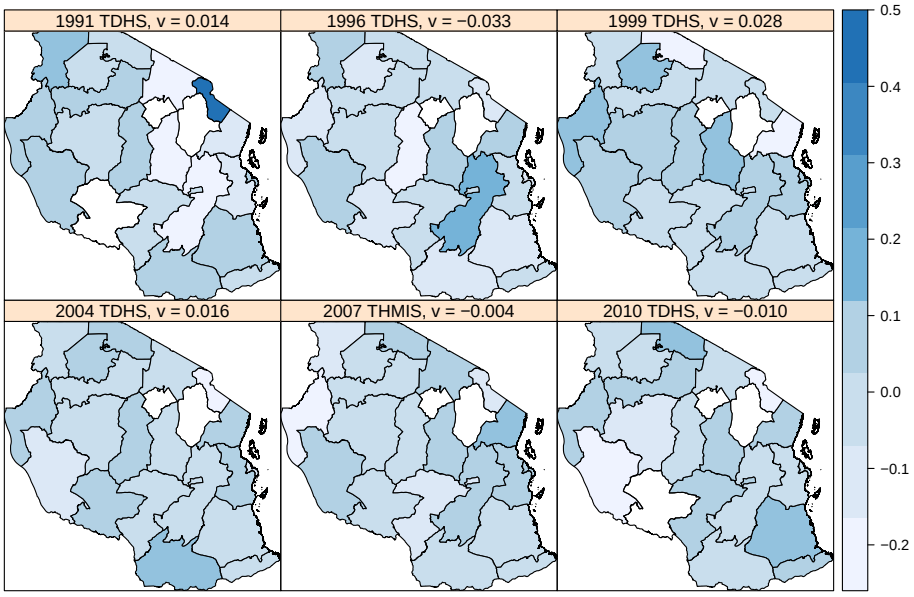

Fig 11: Survey (ν_s) and survey-area (ν_{si}) random effects. The median random effect (ν_s) is given in the heading of each plot. There are five Demographic and Health Surveys (DHS) and one Tanzania HIV and Malaria Indicator survey (THMIS).

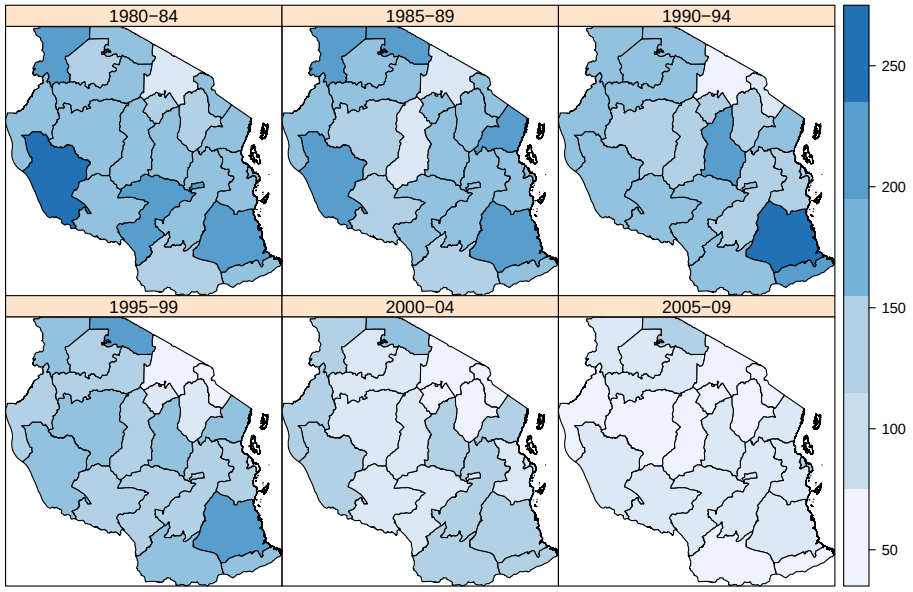

Fig 12: Inverse-variance weighted Horvitz-Thompson regional estimates of child mortality (per 1000 births).

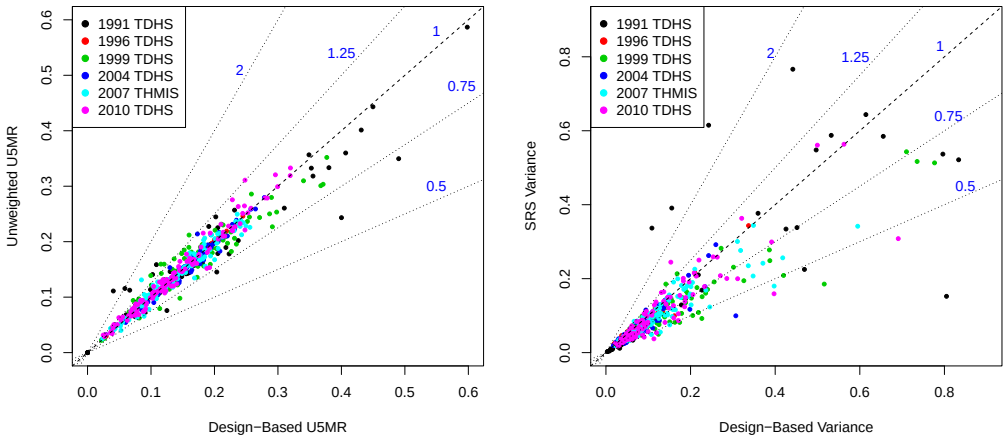

Fig 13: Comparison of design-based and unweighted estimates of U5MR and variances. Values in blue indicate the slope of the lines.

7. Decreasing Mortality. Figure 14 shows the smoothed values for each region, by each time point and a projection into the 2010–2014 period. Figure 15 shows the percent decrease in each region compared to the 1985–1989. The line at -66% corresponds to the the fourth millennium development goal of a two thirds reduction in child mortality by 2015.

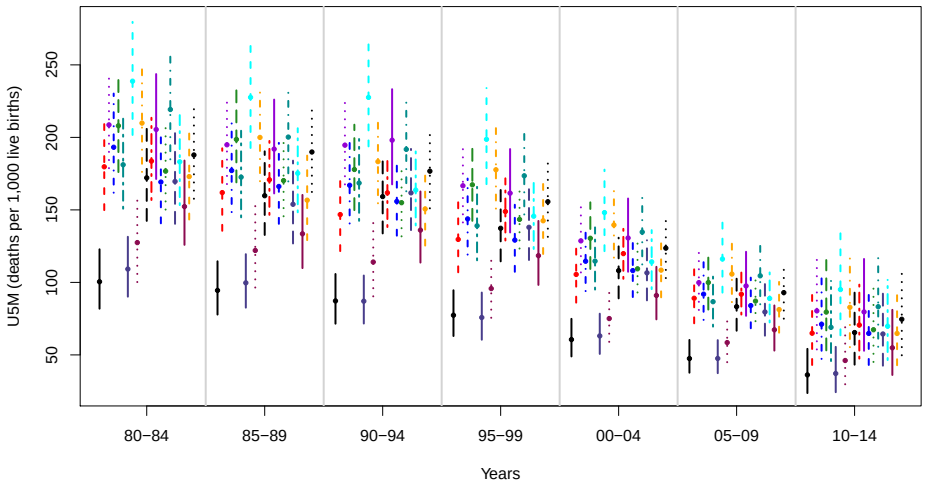

Fig 14: Posterior medians and 95% intervals for the 21 regions of Tanzania and a projection for 2010–2014.

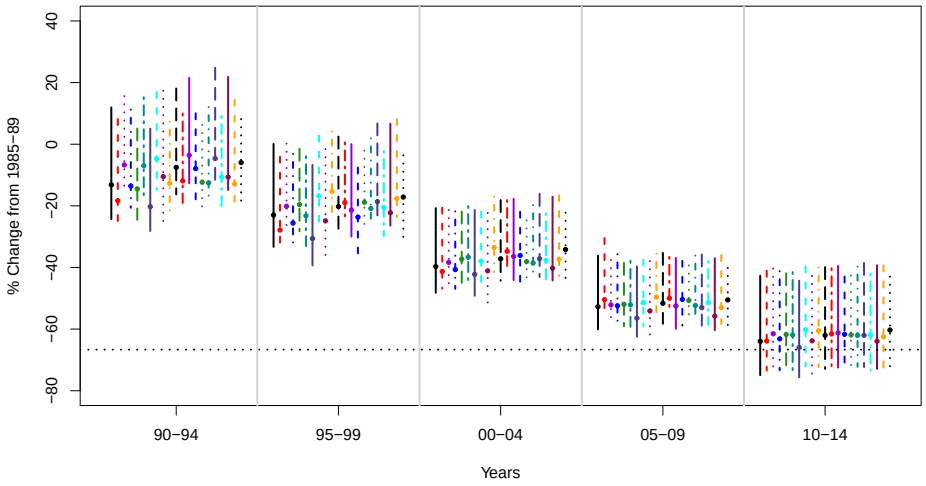

Fig 15: Percent reduction in region-specific child mortality since 1985–1989 with projections for 2010–2014.

8. Model Validation. Figure 16 displays the intervals created with the variance of the observed logit response around the posterior mean with the observed point estimates from each survey, by region for each time interval. Coverage proportions were 0.950, 0.917, 0.899, 0.905, 0.969, and 0.932 for the six time intervals and 0.932 over all observations.

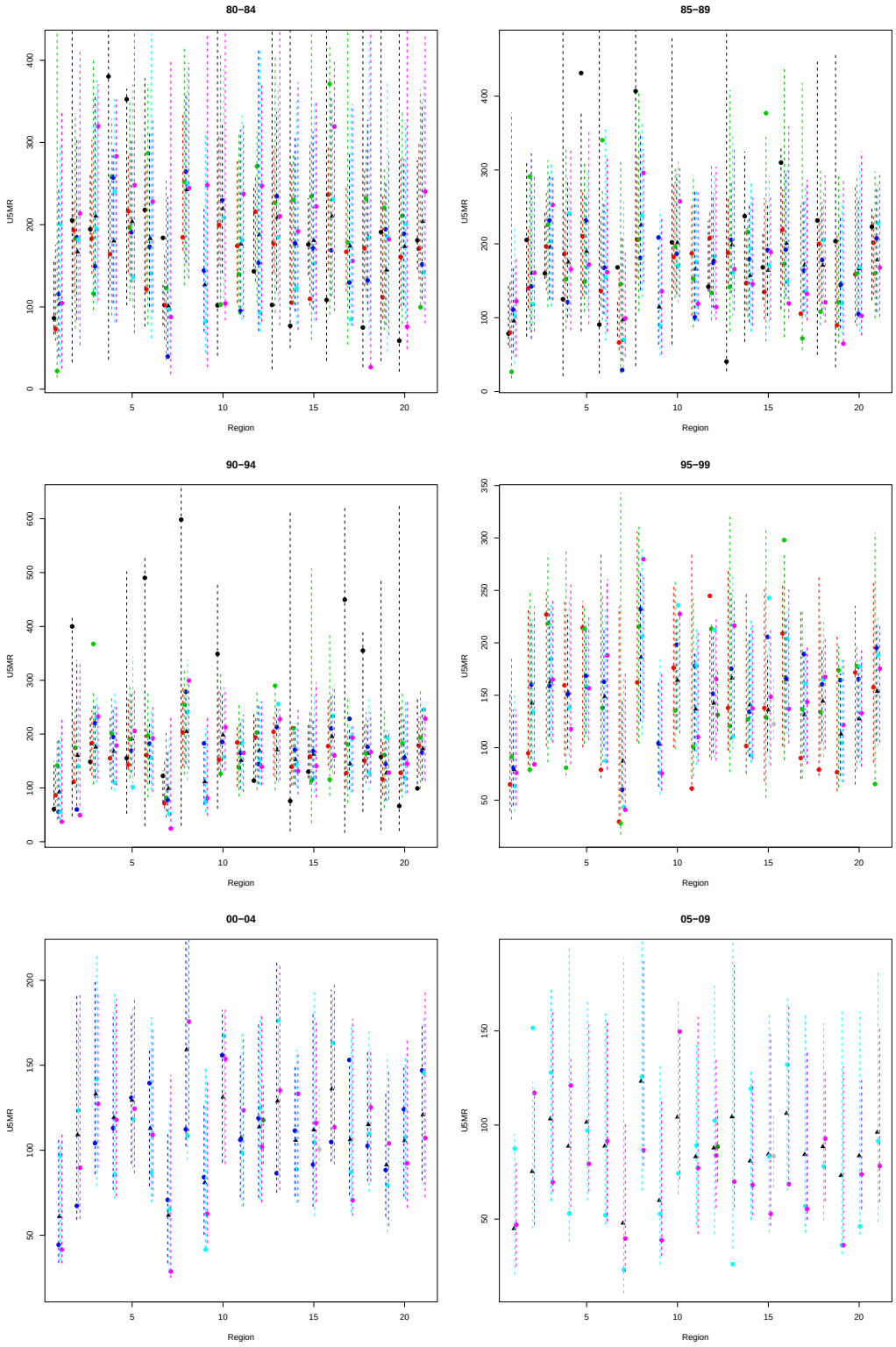

Fig 16: Intervals based on the variance of the observed logit response and region and time-specific direct estimates.

9. R Code. Section 9.1 provides examples of using the `survey` package in `R` to calculate the U5MR as well as a function to provide the intervals described in Section 2. Sections 9.2, 9.3, and 9.4 provide code for selecting the hyperprior for the Random Walk 1, Random Walk 2, and ICAR random effects, respectively. Section 9.5 fits the space-time model selected (in the main text) using the `inla()` function in `R`.

9.1. *Using the* ***survey*** *package.*

```
# read in child month data #
births<-read.dta("dhs2010ChildMonths.dta")

# --- setting up the design object --- #
options(survey.lonely.psu="adjust")
births$wt<-births$v005/1000000
my.svydesign <- svydesign(id= ~v021,
                          strata=~v022,nest=T,
                          weights= ~wt, data=births)


# subsetting design object #
design80<-subset(my.svydesign,per5=="80-84")

# fitting the logistic model #
glm80<-svyglm(died~factor(ageGrpD),design=design80,family=binomial)

# a function to calculate logit(U5M) and Variance #
get.est<-function(glm1){
       V<-vcov(glm.ob)

      ## under 5 child mortality for males in 00-02 ##
      betas<-summary(glm.ob)$coef[,1]
      ns<-c(1,11,12,12,12,12)
      probs<-expit(betas)

      u5m.est<-(1-prod((1-probs)^ns,na.rm=T))#*1000

      ## partial derivatives ##
      gamma<-prod((1+exp(betas))^ns)
      derivatives<-(gamma)/(gamma-1)*ns*expit(betas)
```

```
      ## Items to return ##
      var.est<-t(derivatives)%*%V%*%derivatives
      lims<-logit(u5m.est)+qnorm(c(0.025,0.975))*sqrt(var.est)
      return(c(u5m.est,expit(lims),logit(u5m.est),var.est))}

# getting the estimates #
get.est(glm80)
```

9.2. *Choosing Prior for RW1 Random Effect.*

```
#
# (R,1/R) is the range of the residual odds ratios
# gives a=d/2 where d = degrees of freedom of marginal Student's t

R <- 0.5; d <- 1; a <- d/2; p <- 0.025; b <-(log(R))^2*d/(2*qt(p,df=d)^2)

# Gives a=0.5 and b=0.001488
c(a,b)
1/sqrt(qgamma(p=c(0.025,0.975),a,b))
1/sqrt(qgamma(p=c(0.9,0.1),a,b)) # 0.03316518 0.4341181

# Check range using simulation #

nsamp <- 100000
tausamp <- rgamma(nsamp,a,b)
Usamp <- rnorm(nsamp,mean=0,sd=1/sqrt(tausamp))
quantile(exp(Usamp),p=c(0.025,0.5,0.975))

# The adjacency matrix for 6 years
m2 = c(1,2,2,2,2,1)
# create the adjacency #
adj2<-c(2,
        1,3,
        2,4,
        3,5,
        4,6,
        5)

make.Q <- function(num.neighbors, neighbors, omega.sq = 1){
  n <- length(num.neighbors)
```

```
  mat <- matrix(0, ncol = n, nrow = n)
  diag(mat) <- num.neighbors
  mat[cbind(rep(1:n, num.neighbors), neighbors)] <- -1
  mat/omega.sq
}
vars.Q <- function(eigenvalues,eigenvectors){
  margsum <- 0
  nloop <- length(eigenvalues)-1
  for (i in 1:nloop){
    ev <- eigenvectors[,i]
    margsum <- margsum + ev %*% t(ev)/ eigenvalues[i]
  }
  margvars <- diag(margsum)
  margvars
}
#
sim.Q <- function(Q){
  eigenQ <- eigen(Q)
  rankQ <- qr(Q)$rank
  sim <- as.vector(eigenQ$vectors[,1:rankQ] %*%
                     matrix(
                       rnorm(rep(1, rankQ), rep(0, rankQ),
                       1/sqrt(eigenQ$values[1:rankQ])),
                       ncol = 1))
  sim
}
#
Q <- make.Q(m2, adj2, 1)
eigentemp <- eigen(Q)
eigenvaluesQ <- eigentemp$values
eigenvectorsQ <- eigentemp$vectors
rankQ <- qr(Q)$rank # 5
margy <- mean(vars.Q(eigenvaluesQ,eigenvectorsQ))
#
nsamp <- 5000
astar <- a; bstar <- b/margy
c(astar,bstar)
taustarsamp <- rgamma(nsamp,astar,bstar)
Ustarsamp <- matrix(nrow=nsamp,ncol=6)
for (i in 1:nsamp){
```

```
  Qstar <- Q*taustarsamp[i]
  Ustarsamp[i,] <- sim.Q(Qstar)
}
quantile(exp(Ustarsamp),p=c(0.025,0.5,0.975)) # 0.5285916 0.9999869 1.8503786
```

9.3. *Choosing Prior for RW2 Random Effect.*

```
R <- 0.5; d <- 1; a <- d/2; p <- 0.025; b <-(log(R))^2*d/(2*qt(p,df=d)^2)
# Gives a=0.5 and b=0.002857959 # [0.5,2]
c(a,b)
1/sqrt(qgamma(p=c(0.025,0.975),a,b))
1/sqrt(qgamma(p=c(0.9,0.1),a,b)) # 0.03316518 0.4341181

#
# Check range using simulation
#

nsamp <- 100000
tausamp <- rgamma(nsamp,a,b)
Usamp <- rnorm(nsamp,mean=0,sd=1/sqrt(tausamp))
quantile(exp(Usamp),p=c(0.025,0.5,0.975))

# The adjacency matrix for 6 years
Q<-matrix(0,nrow=6,ncol=6)
Q[1,c(1,2,3)]<-c(1,-2,1)
Q[2,c(1,2,3,4)]<-c(-2,5,-4,1)
Q[3,c(1,2,3,4,5)]<-c(1,-4,6,-4,1)
Q[4,c(2,3,4,5,6)]<-c(1,-4,6,-4,1)
Q[5,c(3,4,5,6)]<-c(1,-4,5,-2)
Q[6,c(4,5,6)]<-c(1,-2,1)


vars.Q <- function(eigenvalues,eigenvectors){
  margsum <- 0
  # make sure - value represents rank deficiency
  nloop <- length(eigenvalues)-2
  for (i in 1:nloop){
    ev <- eigenvectors[,i]
    margsum <- margsum + ev %*% t(ev)/ eigenvalues[i]
  }
```

```
  margvars <- diag(margsum)
  margvars
}
#
sim.Q <- function(Q){
  eigenQ <- eigen(Q)
  rankQ <- qr(Q)$rank
  sim <- as.vector(eigenQ$vectors[,1:rankQ] %*%
                     matrix(
                       rnorm(rep(1, rankQ), rep(0, rankQ), 1/sqrt(eigenQ$values[1:ran
                       ncol = 1))
  sim
}
#

eigentemp <- eigen(Q)
eigenvaluesQ <- eigentemp$values
eigenvectorsQ <- eigentemp$vectors
rankQ <- qr(Q)$rank # 4
margy <- mean(vars.Q(eigenvaluesQ,eigenvectorsQ))
#
nsamp <- 5000
astar <- a; bstar <- b/margy
c(astar,bstar)
# astar<-0.500000000
# bstar<-0.001530466

taustarsamp <- rgamma(nsamp,astar,bstar)
Ustarsamp <- matrix(nrow=nsamp,ncol=6)
for (i in 1:nsamp){
  Qstar <- Q*taustarsamp[i]
  Ustarsamp[i,] <- sim.Q(Qstar)
}
quantile(exp(Ustarsamp),p=c(0.025,0.5,0.975))
```

9.4. *Choosing Prior for ICAR Random Effect.*

```
#
# (R,1/R) is the range of the residual odds ratios
# gives a=d/2 where d = degrees of freedom of marginal Student's t
```

```
R <- 0.5; d <- 1; a <- d/2; p <- 0.025; b <-(log(R))^2*d/(2*qt(p,df=d)^2)
# Gives a=0.5 and b=0.001488

1/sqrt(qgamma(p=c(0.9,0.1),a,b)) # 0.03316518 0.4341181

# Check range using simulation #

nsamp <- 100000
tausamp <- rgamma(nsamp,a,b)
Usamp <- rnorm(nsamp,mean=0,sd=1/sqrt(tausamp))
quantile(exp(Usamp),p=c(0.025,0.5,0.975))

# The adjacency matrix for Tanzania without Zanzibar
Amat<-as.matrix(read.table("adj_regions_nozanz.txt"))
m2 = apply(Amat,1,sum)

# create the adjacency list #
nums<-c(1:21)

adj2<-NULL

for(i in 1:21){

  adj2<-c(adj2,nums[as.numeric(Amat[i,])==1])
}

make.Q <- function(num.neighbors, neighbors, omega.sq = 1){
  n <- length(num.neighbors)
  mat <- matrix(0, ncol = n, nrow = n)
  diag(mat) <- num.neighbors
  mat[cbind(rep(1:n, num.neighbors), neighbors)] <- -1
  mat/omega.sq
}
vars.Q <- function(eigenvalues,eigenvectors){
  margsum <- 0
  nloop <- length(eigenvalues)-1
  for (i in 1:nloop){
    ev <- eigenvectors[,i]
    margsum <- margsum + ev %*% t(ev)/ eigenvalues[i]
  }
```

```
  margvars <- diag(margsum)
  margvars
}
#
sim.Q <- function(Q){
  eigenQ <- eigen(Q)
  rankQ <- qr(Q)$rank
  sim <- as.vector(eigenQ$vectors[,1:rankQ] %*%
                     matrix(
                       rnorm(rep(1, rankQ), rep(0, rankQ),
                       1/sqrt(eigenQ$values[1:rankQ])),
                       ncol = 1))
  sim
}
#
Q <- make.Q(m2, adj2, 1)
eigentemp <- eigen(Q)
eigenvaluesQ <- eigentemp$values
eigenvectorsQ <- eigentemp$vectors
rankQ <- qr(Q)$rank # 20
margy <- mean(vars.Q(eigenvaluesQ,eigenvectorsQ))
#
nsamp <- 5000
astar <- a; bstar <- b/margy
c(astar,bstar)
taustarsamp <- rgamma(nsamp,astar,bstar)
Ustarsamp <- matrix(nrow=nsamp,ncol=21)
for (i in 1:nsamp){
  Qstar <- Q*taustarsamp[i]
  Ustarsamp[i,] <- sim.Q(Qstar)
}
quantile(exp(Ustarsamp),p=c(0.025,0.5,0.975)) # 0.5152361 0.9999245 1.9420017
```

9.5. *Using INLA to fit the models.*

```
#### --- Setting priors --- ####
# range of [0.5,2] #
a.iid<-0.5
b.iid<-0.001488
```

```
a.rw2<-0.5
b.rw2<-0.002857959

a.icar<-0.5
b.icar<-0.003602143

# the formula #
mod11<-logit.est~f(survey, param=c(a.iid,b.iid))
+f(survey.area, param=c(a.iid,b.iid))
+f(survey.time param=c(a.iid,b.iid))
+f(region.unstruct,param=c(a.iid,b.iid))
+f(region.struct, graph=Amat,model="besag",param=c(a.icar,b.icar))
+f(time.struct,model="rw2",param=c(a.rw2,b.rw2))
+f(time.unstructparam=c(a.iid,b.iid))
+f(time.area, param=c(a.iid,b.iid))

inla.fit <- inla(mod11,
   family = "gaussian", control.compute=list(dic=T,mlik=T,cpo=T),
   data =exdat,
   control.predictor=list(compute=TRUE),
   control.family=list(hyper=list(prec=list(initial=log(1),fixed=TRUE))),
   scale=logit.prec)

summary(inla.fit)
```

References.

Binder, D. (1983). On the variances of asymptotically normal estimators from complex surveys. *International Statistical Review*, **51**, 279–292.

Lohr, S. (2009). *Sampling: Design and Analysis*. Cengage Learning.

Pedersen, J. and Liu, J. (2012). Child mortality estimation: Appropriate time periods for child mortality estimates from full birth histories. *PLoS Medicine*, **9**(8).

Roberts, G., Rao, N., and Kumar, S. (1987). Logistic regression analysis of sample survey data. *Biometrika*, **74**(1), 1–12.