A Systematic Assessment of National Under-5 Mortality Rate by Place of Residence for 109 Countries

Fengqing Chao* Danzhen You† Lucia Hug† Jon Pedersen‡ Hernando Ombao* Leontine Alkema§

April 20, 2020

Full Paper for PAA 2020¶

Abstract

The progress in reducing the under-5 mortality rate (U5MR) since 1990 has been remarkable but uneven within countries between urban and rural populations. While trend U5MR estimates have been published frequently, an analysis of U5MR by urban and rural area has not been available across countries over time. In this paper, we provide annual estimates of U5MR among urban residence for 109 countries from 1990 to 2018 using a Bayesian time series model and assess the corresponding uncertainty. The analyses are based on an extensive database complied from surveys (including DHS, MICS, RHS, PAPFAM, PAPCHILD), censuses, and vital registration system. We present results for selected countries and identify country-years with the highest disparities in U5MR between urban area and national level.

*Computer, Electrical and Mathematical Sciences and Engineering Division, 4700 King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia.

†Division of Data, Analytics, Planning and Monitoring, United Nations Children's Fund, New York, NY, USA

‡Fafo, Oslo, Norway

§Department of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts, Amherst, USA

¶The project described is solely the responsibility of the authors and does not necessarily represent the official views of the United Nations.

Contents

List of Tables

List of Figures

1 Introduction

Under-5 mortality rate (U5MR) is defined as the probability of dying before reaching age 5. The progress in reducing the U5MR since 1990 has been remarkable – by the year 2018 the U5MR is about 41% of that in 1990 [1]. Since the era of Sustainable Development Goal (SDG), leaving no one behind and reducing disparities across populations become a priority [2]. The disparity in U5MR between subgroups withint a country can be masked if we only focus on the national-level U5MR. Therefore, in order to identify the most disadvantaged and vulnerable children within a country, it is important to study how U5MR disparity changes overtime. Previous studies have provided evidence on U5MR disparity between girls and boys [3] and by household economic status [4]. This project aims to contribute to the literature of U5MR disparity within countries overtime on the dimension of resident status.

Historically, given the self-governed nature of many Western cities, the demography and economy in western urban areas were different from their rural surroundings [5, 6]. As one of the principal demographic characteristics, the western cities have a higher mortality rate than in their rural surroundings. The excessive mortality in urban than in rural during the medieval and early modern periods was largely due to the frequent and severe epidemics [7]. The urban-rural difference of U5MR during this period could be more extreme than the overall mortality for all ages since the children under age of 5 are more sensitive to the environmental differences than the rest of the population. It is only after the early modern Europe do we see a reverse of the urban-rural mortality disparity that the better access to the most advanced health care facilities in cities than in rural areas improved the survival chances more to urban population than rural population.

Estimates of U5MR (or infant mortality) by residence, or disparities in U5MR between urban and rural areas, have been published previously [7, 8, 9, 10, 11, 12]. Before this study, the Health Equity Monitor (HEM) by WHO [13] provided the most comprehensive information on U5MR disaggregated by place of residence for country-years with available data from Demographic and Health Surveys, Multiple Indicator Cluster Surveys, and Reproductive Health Surveys, including 73 countries[1]. However, no studies provided model-based time trends with comparable annual estimates covering 1990–2018.

In this paper, we estimate country-specific U5MR from urban residents for 109 countries from 1990 to 2018. We implement a Bayesian time series model, accounting for the varying levels of uncertainty associated with different observations, and assess the uncertainty related to outcomes. In this extended abstract, we summarize the proposed method and present preliminary results.

2 Data

The residence-specific under-5 mortality rate indicates the probability for children born in either urban or rural area of dying before age of five years. There are 528 data points available from 109 countries for U5MR in urban area. The data used in this systematic analysis are observed ratios of urban to national-level U5MR from various sources: vital registration systems, sample registration systems, household surveys, and censuses. Figure 1 gives an overview of the data availability for urban U5MR.

The U5MR among urban residents from vital registration and sample registration and surveillance data are calculated by standard life-table methods. For data based on full birth histories from Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS), selected surveys from the Pan-Arab Programme on Family Health (PAPFAM) and Pan Arab Project for Child Development (PAPCHILD), and selected Reproductive Health Surveys (RHS), we calculated the residence-specific U5MR in the 5 years before the survey to reduce the effect of urban-rural migration over time and potential recall bias. For surveys and censuses that collected summary birth histories, indirect methods are used to estimate the residence-

[1]Latest HEM data are available at: http://www.who.int/gho/health_equity/outcomes/under5_mortality/en/

specific U5MR. In particular, for MICS data, an indirect method based on time since first birth [14] is used to calculate mortality based on births and deaths that occurred to mothers who had their first births five to nine years before the time of the survey. For summary birth history data collected from DHS, we calculate mortality by using an indirect method based on age of mother [15] aged 25 to 29 through 45 to 49. When microdata or tabulations are not available, we use estimates from published surveys and census reports.

Urban and rural definition In DHS, MICS, RHS, PAPFAM and PAPCHILD surveys, there is an indicator to determine whether the household interviewed is in urban or rural areas. These surveys do not specifically define the urban/rural residency in a surveyed country, but rather adopt the urban-rural definition of that country. For data from vital registration systems, the urban/rural residency status follows the criteria provided by the corresponding national statistical offices. Due to the complex nature in which the urbanization progresses, it is challenging to define urban areas using a universal set of rules [16]. The urban/rural definition differ across countries and/or evolve over time [17]: some are population size/density based, some are infrastructure based, many have unclear definition. Although there have been attempts of using satellite data to identify city areas across the globe [18], the approach only became available after the late 2000's, and it is difficult or impossible to employ definitions generated by such data to current survey datasets. We do not adjust or exclude data based on the urban-rural definition, but accept the country definitions.

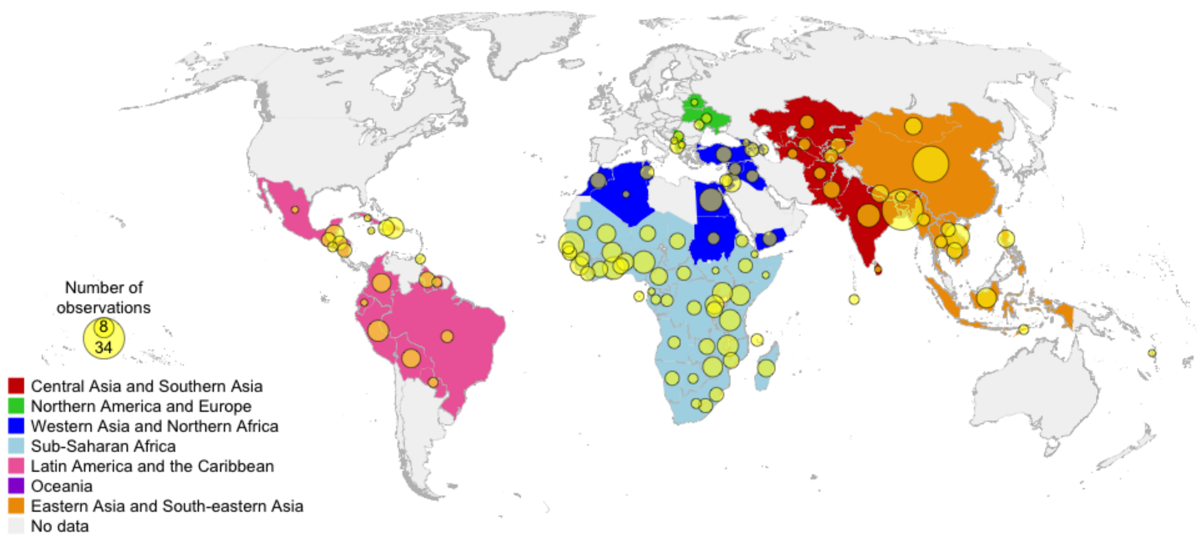

Figure 1: **Data availability for urban U5MR.** Countries are coloured by regions. Circle size is proportional to the number of observations available for each country.

3 Summary of estimation method

We implement a statistical model to estimate levels and trends in urban U5MR from 1990 to 2018. The input data of the model are the observed ratio of urban U5MR to the national-level U5MR from empirical data, the estimates of national-level U5MR (i.e. urban and rural areas combined) published by the UN Inter-agency Group for Child Mortality Estimation (IGME) in 2019 [1], and the proportion of population residing in urban area published by the UN World Urbanization Prospects revision 2018 [16]. Details for the statistical model are provided in appendix 6.2, 6.3, and 6.4 and summarized in this section.

We model the country-year-specific ratio of urban to national-level U5MR by a time series model on the log-scale. The time series model follows an AR(1) structure. The mean of the time series model at each country-year is assumed to follow a multivariate linear regression model with two country-year-specific covariates: (i) the proportion of population residing in urban area; and (ii) the national-level U5MR. The coefficients of the two covariates are modeled on country level in order to account for the difference in the definition of urban/rural residence and the difference in the rate of urbanization across countries.

Estimates for ratios of urban to national U5MR are based on all available data in a country. The data quality model incorporate stochastic variance for SRS/VR data and sampling variance for survey data. By including the variance terms, less informative observations are down-weighted compared with more informative observations. We fit the model in the open source software `R 3.5.1` [19] and `R`-packages `R-INLA` [20].

4 Preliminary Results

4.1 Overview of countries with outlying urban-national U5MR disparity

Figure 2 illustrates 25 countries with ratios of urban to national-level U5MR in 2018 significantly below 1. In 2018, the lowest ratio of urban to national-level U5MR is estimated to be in China at 0.51 (90% uncertainty interval [0.43; 0.60]), indicating just about half chance of dying for children age under-5 for those who live in the urban area compare to the whole country. From 1990 to 2018, there is a significant positive increase in the China ratio at 0.16 [0.08; 0.25]. Hence, compare to the ratio of urban to national U5MR back in 1990 at 0.35 [0.32; 0.38], the disparity in U5MR between the urban and national level has been narrowed since 1990.

Cambodia and Mali are the only other 2 countries (besides China) that with a change in ratio between 1990 and 2018 that is significantly different from zero among the 25 countries. In contract to China, both Cambodia and Mali have significant decreases in their ratios (i.e. further away from 1): the ratio in Cambodia decreases from 0.78 [0.63; 0.96] in 1990 to 0.53 [0.39; 0.73] in 2018, while the ratio in Mali decreases from 0.77 [0.70; 0.85] in 1990 to 0.56 [0.43; 0.72] in 2018. For both countries, their disparity in U5MR between urban residences and the whole countries are growing significantly wider between 1990 and 2018.

Among the 25 countries, five of them have increases in their point estimates between 1990 and 2018 and become closer to 1: China, Morocco, Peru, Senegal, and Timor Leste. This implies that the disparities between the urban and national U5MR have become narrower between 1990 and 2018 for the five countries.

4.2 Illustration for country estimates

Figure 3 illustrates country-specific results from 1990 to 2018 for three selected countries: (i) the ratio of urban to national-level U5MR and urban U5MR; and (ii) urban U5MR. China is a country with empirical data only from Sampling Vital Registration system (SVR). The estimated ratio of urban to national-level U5MR in China is significantly smaller than 1 for the entire observation period. Given the results for the ratios, the estimated urban U5MR is significantly lower than their corresponding national U5MR over time. Bangladesh has a mix of empirical data from SVR and survey. The estimated ratio of urban to national U5MR is below 1 but increases steadily over time especially after 2000 and fluctuates just below 1 after 2010. Hence, the urban U5MR in Bangladesh before 2000 is estimated to be much lower than the national level and the difference between the urban and national U5MR narrowed after 2000. Tanzania is a country with survey data as the only data source. Its estimated ratio of urban to national-level U5MR fluctuates around 1 throughout the whole observation period. In consequence, the estimated urban U5MR are at similar levels to the national level for most of the period.

4.3 Urban-national U5MR disparity and urbanization

Figure 4 presents a snapshot of the model results of the year 2018. The ratio of urban to national U5MR is plotted against the proportion of urban among total population, alongside with the corresponding urban U5MR for each country-year. The corresponding information on infant mortality rate (IMR) from developed countries (Japan[2] and United Kingdom[3]) are added in the same plot[4].

In general, as the proportion of urban population approach 100%, the ratio of urban to national U5MR should converge to 1. Meanwhile, as a country becomes more urbanized (using the proportion of urban population as an proxy), the urban U5MR decreases.

Although the results from most countries fall within this described general pathway, exceptions remains. On one hand, we observe a set of countries with a relatively high proportion of urban population together with a low ratio of urban to national U5MR: China, Georgia, Ukraine, Thailand, Mali, Lao PDR, and Cambodia. On the other hand, another set of countries with outlying results are with high ratio of urban to national U5MR and high proportion of urban population: Montenegro, Cuba, Sao Tome and Principe, Trinidad and Tobago, and Jamaica.

5 Discussion

To our best knowledge, our study is the first systematic analysis to provide an assessment of levels and trends of resident inequalities in U5MR with a wide coverage of developing countries. Our findings confirmed evidence from previous studies that the gaps of survival and accessing healthcare resources between the under-five children between residential status have remained or even worsen for some countries since 1990 even though the overall healthcare situation has improved. In addition, we identified countries with unexpectedly high resident U5MR disparity among countries with low proportion of urban population and among countries that are highly urbanized. Further research should focus on explaining the outlying disparities across countries. More efforts from international organizations should be drawn to the children aged under-five from those groups with outlying disparities.

[2]Data source for Japan: Statistics of Japan, available at: `https://www.e-stat.go.jp/en`.

[3]Data source for United Kingdom: Office for National Statistics UK, available at: `https://www.ons.gov.uk`.

[4]IMR for the urban and total population are computed directly from vital registration data. Most of the under-5 deaths in developed countries occur within the first year of new born children. Hence, we plot IMR information in order to approximate the disparity in the U5MR from developed countries.

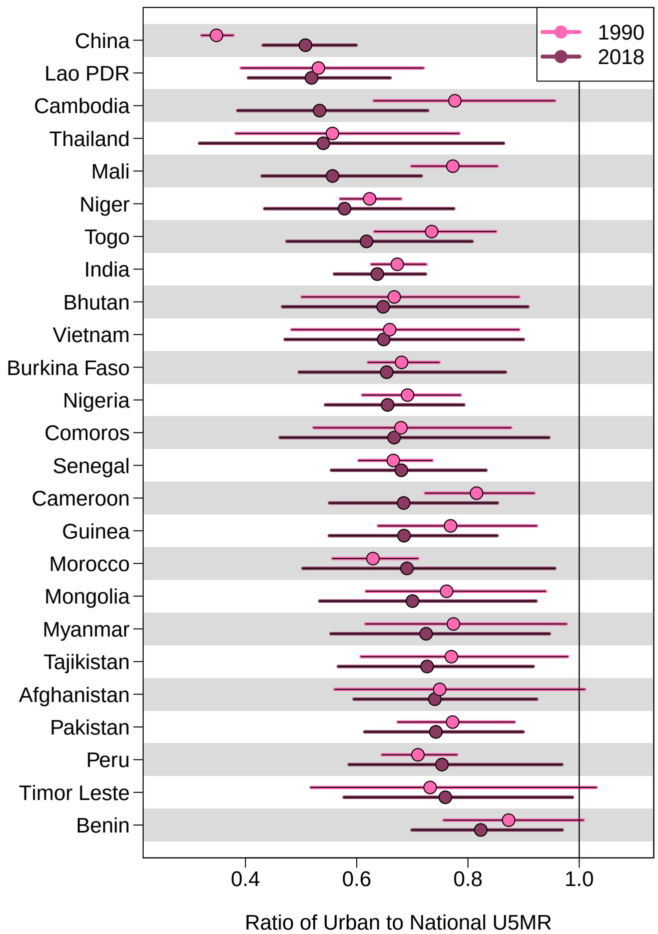

Figure 2: **Countries with the lowest ratios of urban to national U5MR in 2018.** Countries are ordered by increasing point estimates for the year 2018. Countries are selected if their 90% uncertainty intervals of the ratios in 2018 are statistically significantly below 1. Dots are point estimates. Horizontal lines are 90% uncertainty intervals. Results for the year 1990 and 2018 are differentiated by colors.

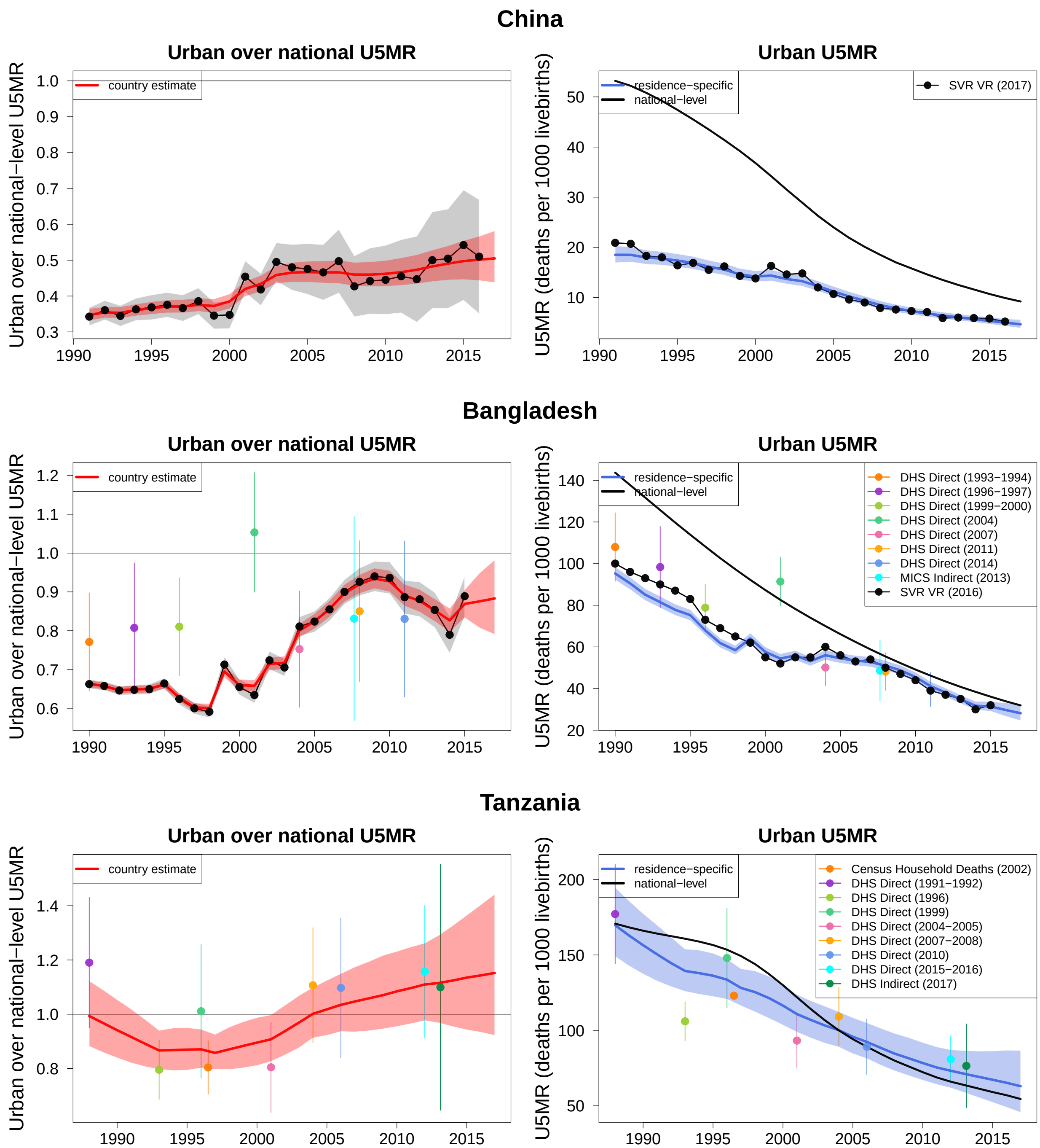

Figure 3: **Urban U5MR results over time for selected countries.** Observations from different data series are differentiated by colors. The vertical line at observations represent the sampling or stochastic errors associated with the observations. Curves and shaded areas with same color: point estimates and corresponding 90% uncertainty intervals. Left column: ratio of urban to national-level U5MR; right column: urban U5MR is in blue, and the point estimates for the national-level U5MR are in black curves (taken from the UN IGME 2019 results [1]).

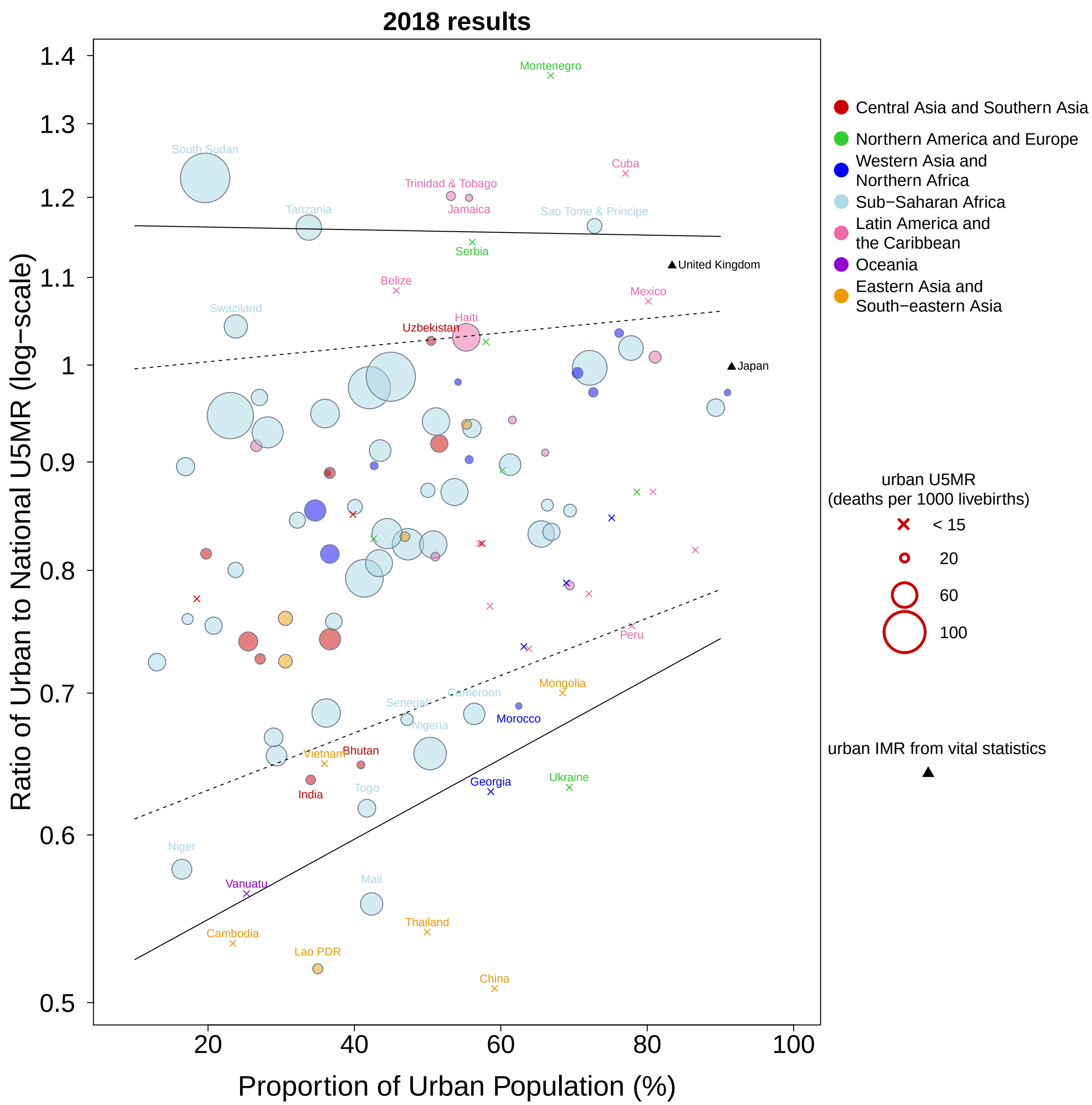

Figure 4: **2018 results overview.** Countries are coloured by regions. Circle size is proportional to the number of observations available for each country. Black triangle dots are the ratio of urban to total infant mortality rate (IMR) from developed countries. The two solid lines are the 95% bounds, and the two dashed lines are the 80% bounds. Both sets of bounds are smoothed by the Friedman's SuperSmoother approach [21, 22].

6 Appendix

6.1 Data sources

An overview of the data sources included in the database is in Table 1.

Data source type	# obs.
DHS Direct	305
Others Direct	68
Indirect	69
VR/SRS	59
total	**528**

Table 1: **Distribution of observations by source type.** Observations of urban U5MR are grouped by source types. "Direct" refers to observation obtained from full birth histories while "Indirect" refers to observations obtained from summary information and demographic methods. DHS: Demographic and Health Survey; MICS: Multiple Indicator Cluster Survey; SRS: Sample Registration System; VR: Vital Registration.

6.2 Model for urban and rural under-5 mortality rate

Our goal is to estimate the urban U5MR $Q_{c,t}$ for country c, year t. The urban U5MR is assumed to relate to the national-level U5MR as follows:

$$Q_{c,t} = Q_{\text{total},c,t} \cdot R_{c,t},$$

where $R_{c,t}$ is the ratio of urban to national-level U5MR. $Q_{\text{total},c,t}$ denotes the national-level U5MR. For each country-years, $Q_{\text{total},c,t}$ is known value based on the UN IGME estimates published in 2019 [1].

The ratio of urban to national-level U5MR $R_{c,t}$ is modeled on the log-scale. For $c = 1, \cdots, C$:

$$\begin{aligned}
R_{c,t} &= \exp\{V_{c,t}\}, \text{ for } t = 1, \cdots, T, \\
V_{c,t} &= \mu_{c,t} + \epsilon_{c,t}, \text{ for } t = 1, \\
\epsilon_{c,t} &\sim \mathcal{N}(0, \frac{\sigma_\epsilon^2}{1-\rho^2}), \text{ for } t = 1, \\
V_{c,t} &= \rho(V_{c,t-1} - \mu_{c,t}) + \mu_{c,t} + \epsilon_{c,t}, \text{ for } t = 2, \cdots, T, \\
\epsilon_{c,t} &\sim \mathcal{N}(0, \sigma_\epsilon^2), \text{ for } t = 2, \cdots, T.
\end{aligned}$$

$V_{c,t} = \log(R_{c,t})$ is modeled as an auto-regressive time series model of order 1 (AR(1)). The AR(1) model has a global variance of distortion parameter σ_ϵ^2, a global autoregressive parameter ρ, and a country-year-specific level parameter $\mu_{c,t}$. Each $\mu_{c,t}$ is modeled as a multivariate linear regression function. For $c = 1, \cdots, C$ and $t = 1, \cdots, T$, we assume:

$$\mu_{c,t} = \alpha_c \cdot x_{c,t} + \beta_c \cdot y_{c,t}.$$

The explanatory variables for the regression function are: (i) $x_{c,t}$, the proportion of population residing in urban area on log scale; and (ii) $y_{c,t}$, the national-level U5MR on log-scale (i.e. $y_{c,t} = \log(Q_{\text{total},c,t})$). The country-specific regression coefficients for $x_{c,t}$ and $y_{c,t}$ are α_c and β_c respectively.

6.3 Data model

Sampling errors are accounted for in the data model for data of the ratio of urban (or rural) to national-level U5MR. For the i-th observed ratio r_i from country $c[i]$ in year $t[i]$, we assume:

$$v_i \sim \mathcal{N}(V_{c[i],t[i]}, \sigma_i^2), \text{ for } i = 1, \cdots, n,$$

where $v_i = \log(r_i)$ and σ_i^2 is the sampling variance for the i-th observation. If the data is from full birth history data series, its corresponding sampling variance is obtained using the Jackknife repeated replication method used by DHS Sampling and Household Listing Manual [23].

Computing The model estimation is implemented in the open source softwares `R 3.5.1` [19] and `R`-packages `R-INLA` (Integrated Nested Laplace Approximations) [20].

6.4 Model summary

Table 2 summarizes the notation and indices for this project:

Symbol	**Description**
i	Indicator for observation, $i = 1, \cdots, n$, where $n = 528$.
t	Indicator for year, $t = 1, \cdots, T$. $t = 1$ corresponds to the year 1990 and $t = T$ to the year 2018.
c	Indicator for country, $c = 1, \cdots, C$, where $C = 109$.
$R_{c,t}$	The true ratio of urban U5MR to national-level U5MR for country c in year t.
$V_{c,t}$	$V_{c,t} = \log(R_{c,t})$
r_i	i-th observed ratio of urban U5MR to national-level U5MR for country $c[i]$ in year $t[i]$.
v_i	$v_i = \log(r_i)$
σ_i^2	The i-th sampling variance for v_i.
$\mu_{c,t}$	country-year-specific mean for the AR(1) time series model.
ρ	Autoregressive parameter for AR(1) time series model.
σ_ϵ^2	Variance of distortion terms in AR(1) time series model.
$x_{c,t}$	The log of proportion of population residing in urban area for country c in year t.
$y_{c,t}$	The national-level U5MR on log-scale for country c in year t.
α_c	Regression coefficient for $x_{c,t}$ for country c.
β_c	Regression coefficient for $y_{c,t}$ for country c.

Table 2: **Notation summary.**

Model for urban and rural under-5 mortality rate For $c = 1, \cdots, C$:

$$\begin{aligned}
Q_{c,t} &= Q_{\text{total},c,t} \cdot R_{c,t}, \text{ for } t = 1, \cdots, T \\
R_{c,t} &= \exp\{V_{c,t}\}, \text{ for } t = 1, \cdots, T \\
V_{c,t} &= \rho(V_{c,t-1} - \mu_{c,t}) + \mu_{c,t} + \epsilon_{c,t}, \text{ for } t = 2, \cdots, T \\
V_{c,t} &= \mu_{c,t} + \epsilon_{c,t}, \text{ for } t = 1 \\
\mu_{c,t} &= \alpha_c \cdot x_{c,t} + \beta_c \cdot y_{c,t}, \text{ for } t = 1, \cdots, T \\
\epsilon_{c,t} &\sim \mathcal{N}(0, \sigma_\epsilon^2), \text{ for } t = 2, \cdots, T \\
\epsilon_{c,1} &\sim \mathcal{N}(0, \frac{\sigma_\epsilon^2}{1-\rho^2}), \text{ for } t = 1.
\end{aligned}$$

Data model

$$v_i \sim \mathcal{N}(V_{c[i],t[i]}, \sigma_i^2), \text{ for } i = 1, \cdots, n.$$

Prior distributions

$$\begin{aligned}
\alpha_c &\sim \mathcal{U}(-5, 5), \text{ for } c = 1, \cdots, C, \\
\beta_c &\sim \mathcal{U}(-5, 5), \text{ for } c = 1, \cdots, C, \\
\rho &\sim \mathcal{PC}_{\text{cor1}}(0.8, 0.5), \\
\sigma_\epsilon &\sim \mathcal{PC}_{\text{prec}}(1, 0.01).
\end{aligned}$$

where $\mathcal{PC}()$ refers to Penalized Complexity priors [24].

References

1 UNICEF. UN Inter-agency group for Child Mortality Estimation. Levels & Trends in Child Mortality. Report 2019. New York;.

2 UN. Transforming our world: the 2030 Agenda for Sustainable Development. New York: United Nations; 2015. Available at: `https://sustainabledevelopment.un.org/`.

3 Alkema L, Chao F, You D, Pedersen J, Sawyer CC. National, regional, and global sex ratios of infant, child, and under-5 mortality and identification of countries with outlying ratios: a systematic assessment. The Lancet Global Health. 2014;2(9):e521–e530.

4 Chao F, You D, Pedersen J, Hug L, Alkema L. National and regional under-5 mortality rate by economic status for low-income and middle-income countries: a systematic assessment. The Lancet Global Health. 2018;6(5):e535–e547.

5 Pirenne H. Medieval Cities, their Origins and the Revival of Trade. Princeton: Princeton University Press; 1926.

6 Mols R. Introduction à la démographie historique des villes d'Europe du XIVe au XVIIIe siècle: Les résultats. vol. 2. Louvain: Publications Universitaires de Louvain; 1954.

7 Woods R. Urban-rural mortality differentials: an unresolved debate. Population and development review. 2003;29(1):29–46.

8 Sastry N. What explains rural-urban differentials in child mortality in Brazil? Social science & medicine. 1997;44(7):989–1002.

9 Van de Poel E, O'donnell O, Van Doorslaer E. What explains the rural-urban gap in infant mortality: household or community characteristics? Demography. 2009;46(4):827–850.

10 Saikia N, Singh A, Jasilionis D, Ram F. Explaining the rural–urban gap in infant mortality in India. Demographic Research. 2013;29:473–506.

11 Bhatia M, Ranjan M, Dixit P, Dwivedi L. Mind the gap: temporal trends in inequalities in infant and child mortality in India (1992–2016). SSM-Population Health. 2018;.

12 Fotso JC. Urban–rural differentials in child malnutrition: trends and socioeconomic correlates in sub-Saharan Africa. Health & Place. 2007;13(1):205–223.

13 Hosseinpoor AR, Bergen N, Schlotheuber A, Victora C, Boerma T, Barros AJ. Data resource profile: WHO health equity monitor (HEM). International journal of epidemiology. 2016;45(5):1404–1405e.

14 Hill K. Indirect estimation of child mortality. Tools for demographic estimation. 2013;p. 148–164.

15 You D, Hug L, Ejdemyr S, Idele P, Hogan D, Mathers C, et al. Global, regional, and national levels and trends in under-5 mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN Inter-agency Group for Child Mortality Estimation. The Lancet. 2015;386(10010):2275–2286.

16 United Nations, Department of Economic and Social Affairs, Population Division. World Urbanization Prospects: The 2018 Revision; 2018. Available at `https://population.un.org/wup/`.

17 Alkema L, Jones GW, Lai CU. Levels of urbanization in the world's countries: testing consistency of estimates based on national definitions. Journal of Population Research. 2013;30(4):291–304.

18 Schneider A, Friedl MA, Potere D. A new map of global urban extent from MODIS satellite data. Environmental Research Letters. 2009;4(4):044003.

19 R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2018. Available from: `https://www.R-project.org/`.

20 Rue H, Martino S, Chopin N. Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations (with discussion). Journal of the Royal Statistical Society B. 2009;71:319–392.

21 Friedman JH. SMART User's Guide. Laboratory for Computational Statistics. Stanford University Technical Report No 1. 1984;.

22 Friedman JH. A variable span smoother. Stanford Univ CA Lab for Computational Statistics; 1984.

23 ICF International. Demographic and Health Survey Sampling and Household Listing Manual. Calverton, Maryland, U.S.A.; 2012. ISBN 3-900051-07-0.

24 Simpson D, Rue H, Riebler A, Martins T, Sørbye S. Penalising model component complexity: A principled, practical approach to constructing priors. Statistical Science. 2017;32(1):1–28.