Discussion Paper

Multilevel time series modelling of antenatal care coverage in Bangladesh at disaggregated administrative levels

Sumonkanti Das, Jan van den Brakel, Harm Jan Boonstra and Stephen Haslett

July 29, 2021

Abstract

Multilevel time series (MTS) models are applied to estimate trends in time series of antenatal care coverage at several administrative levels in Bangladesh, based on repeated editions of the Bangladesh Demographic and Health Survey (BDHS) within the period 1994-2014. MTS models are expressed in an hierarchical Bayesian framework and fitted using Markov Chain Monte Carlo simulations. The models account for varying time lags of three or four years between the editions of the BDHS and provide predictions for the intervening years as well. It is proposed to apply cross-sectional Fay-Herriot (FH) models to the survey years separately at district level, which is the most detailed regional level. Time series of these small domain predictions at the district level and their variance-covariance matrices are used as input series for the MTS models. Spatial correlations among districts, random intercept and slope at the district level, and different trend models at district level and higher regional levels are examined in the MTS models to borrow strength over time and space. Trend estimates at district level are obtained directly from the model outputs, while trend estimates at higher regional and national levels are obtained by aggregation of the district level predictions, resulting in a numerically consistent set of trend estimates.

keywords
Generalized variance function, Hierarchical Bayesian approach, MCMC simulation, Small area estimation, Demographic and Health Surveys

1 Introduction

Bangladesh Demographic and Health Survey (BDHS) has been widely used for estimating national and division level indicators on fertility, family planning, child mortality, child health, maternal health, and nutrition of children and women. In the sampling design of BDHS, administrative units lower than the division, such as districts and sub-districts, are not accounted for. Consequently sample sizes are too small to estimate any indicator under division level with standard design-based estimators. Over the time period 1994-2014 seven surveys have been conducted, providing time series of direct estimates at the national level and division level on aforementioned indicators to monitor progress in declining maternal and neonatal mortality in Bangladesh. However, for optimal allocation of resources and policy making, reliable statistical information at the more detailed regional level of districts is required. Since sample sizes of the BDHS are too small to produce this information using design-based inference methods, we develop multilevel time series (MTS) models as a form of small area estimation, as an alternative.

Small area estimation refers to a class of model based estimation procedures that improve upon the accuracy of direct domain estimates by increasing the effective sample size in each separate domain with sample information observed in other domains or preceding reference period. This is often referred to as borrowing strength over space or time, respectively. See Rao and Molina (2015) for a comprehensive overview of small area estimation methodology.

The BDHS is conducted repeatedly with varying time lags of 3 or 4 years between two consecutive surveys. Seven editions for the period of 1994 until 2014 are included in this study. In this paper multivariate MTS models are developed to produce reliable trend estimates of antenatal care (ANC) coverage at district level as well as division and national levels. These MTS models are developed in an hierarchical Bayesian framework and fitted using Markov Chain Monte Carlo simulations. The advantage of a multivariate time series approach is that it takes advantage of all available information by modelling cross-sectional and temporal correlations among districts and reference periods. The models are defined at an annual frequency and therefore properly account for the varying time lags of 3 or 4 years between the subsequent survey occasions. On top of that the MTS models provide predictions for the years without survey data.

As a first step direct estimates along with variance estimates are calculated from the cross-sectional data of the seven BDHS surveys at the most detailed regional level of districts for two related response variables: whether no or at least four antenatal consults have been received, abbreviated as ANC0 and ANC4. Two different approaches are followed. In the first approach, district-level direct estimates and standard errors are used as input for the MTS model. Since the estimated standard errors are unstable for some surveyed domains with few observations, these standard errors are first smoothed by means of a Generalized Variance Function (GVF) method (Wolter (2007)).

A drawback of the first approach is that additional auxiliary information, available from two censuses, cannot be included in the MTS models, since the censuses are conducted with intervals of ten years. This implies that the same values of auxiliary information, available from a particular census are used in two or even three subsequent editions of the BDHS conducted after this census but preceding the next census. This would have a negative impact on period-to-period estimates. This problem is circumvented in the second approach. Here separate cross-sectional Fay-Herriot (FH) models (Fay and Herriot, 1979) for each survey occasion are developed using the direct estimates at the district level and their smoothed standard errors. The census auxiliary information is used to improve the FH models. Subsequently the cross-sectional FH estimates are used as input in the MTS model as in the first approach. Note that the cross-sectional FH predictions for a particular survey year are correlated. The MTS models account for this correlation by using the full variance-covariance matrices of the cross-sectional FH predictions as input for the MTS model.

The MTS models borrow strength over time and space in several ways. Cross-sectional relations are modelled using fixed effects as well as district-level random intercepts and slopes, either independent or correlated. Spatial correlations among districts are also considered. Smooth and local level trends at district, division and national level are used to model temporal and cross-sectional correlations. Instead of defining a full correlation matrix between the trend disturbance terms at the district level, trends are defined at the division level, so they are shared by all underlying districts. Deviations from this overall trend are modelled with trends at the district level. This is a parsimonious way of modelling cross-sectional relations between districts (Boonstra and van den Brakel, 2019). Trend estimates at the district level are obtained directly from the model outputs, while trends at division and national levels are obtained by aggregation of the district level predictions. The MTS models account for these different aggregation levels of interest. The advantage of producing estimates for higher aggregation levels by aggregating predictions from the most detailed regional level is that all publication tables are numerically consistent by definition. Estimates for districts for the non-surveyed years and districts not covered in the surveyed years are also predicted based on the estimated time series models.

The MTS models developed in this paper are extensions of the FH model. Earlier accounts of multilevel time-series models extending the Fay-Herriot model to borrow strength over both time and space, include Rao and Yu (1994); Datta et al. (1999); You et al. (2003); You (2008); Pfeffermann and Burck (1990); Pfeffermann and Tiller (2006) and Boonstra and van den Brakel (2019).

The remainder of this article is organized as follows. In Section 2 the need for reliable low regional statistical information to evaluate Sustainable Development Goals related to maternal and neonatal mortality in Bangladesh is described. Section 3 briefly describes the data sources and the computation of direct estimates and variance estimates from the BDHS survey data, along with transformations of direct estimates and the Generalized Variance Function (GVF) approach for smoothing the variance estimates, which both improve model fitting. Section 4 describes the hierarchical Bayesian time series multilevel modelling framework. The models selected for ANC0 and ANC4 are presented in Section 5, along with a brief discussion of the model building process. Section 6 provides a discussion on the trend estimates based on the developed models, and some model evaluation results are illustrated in Section 7. The paper concludes with a discussion in Section 8.

2 Need for reliable regional statistics on maternal and neonatal mortality in Bangladesh

Bangladesh has made remarkable progress in reducing the maternal mortality ratio (MMR) and neonatal mortality rate (NMR) following the target of Millennium Development Goals (MDGs) 4 and 5. However, both the indicators MMR (170 per 100,000 live births WHO et al. (2014)) and NMR (28 per 1000 live births NIPORT (2015a)) are still reasonably high compared to the Sustainable Development Goals (SDG) of reducing MMR to 70 per 100,000 live births and NMR to 12 deaths per 1,000 live births in Bangladesh BBS (2020). Poor utilization of maternal health services such as antenatal care (ANC), skilled birth attendance (SBA) at delivery, and postnatal care (PNC) (NIPORT (2016)), is considered as one of the major reasons for these high mortality rates. Receiving sufficient ANC during pregnancy is important since it also increases usage of SBA and PNC (Mrisho et al. (2009)).

The most recent household survey indicates that the majority of pregnant women (75%) in Bangladesh receive ANC from medically trained providers. However, the proportion of women that receive WHO-recommended 4^+ ANC is much less at 37% (BBS and UNICEF (2019)). These data suggest that Bangladesh lags behind in reaching the national target of 50% 4^+ ANC utilization by the year 2016. Thus, to address this gap and to meet the target of Sustainable Development Goal (SDG) 3 of increasing 4+ ANC coverage to 98% by 2030 NIPORT (2015b), the country needs a comprehensive strategy and specific milestones. National level trends of ANC coverage indicate that the proportion of women having no ANC care (ANC0) improved to only 17.2% in 2019 from 85% in 1994, while the proportion of women who obtained at least four ANC (ANC4) increased to 37% in 2019 from 6% in 1994. The improvement of the indicators over this period varies by division. The most marked improvement is observed for the *Khulna* division where ANC0 and ANC4 shifted from about 70% and 5% to about 12% and 40%, respectively. The poorest development has been observed in *Sylhet* division.

The facilities for ANC services vary considerably within Bangladesh. There are community clinics and family welfare centers at the union level (also NGO clinics), upazila health complexes at sub-district level and district and tertiary hospitals at district level. Moreover, the access to private doctors varies according to the level of urbanization as well as the distance between the district/sub-district and the corresponding Metropolitan cities, particularly the capital city *Dhaka*. This inequality in the access to ANC is also explicitly visible at the division level. At disaggregated administrative levels such as district and sub-district, it can be expected that inequalities are even larger. There are, however, no studies that confirm this hypothesis, mainly because sufficient detailed survey data at those levels are not available. Recent evidence from disaggregated level studies on poverty, child nutrition and morbidity indicate high levels of inequality at both district and sub-district levels (Haslett and Jones (2004); Haslett et al. (2014); Das et al. (2020); Hossain et al. (2020)).

3 Data sources and input estimates

3.1 Data sources

Since 1993-94 the BDHS has been conducted under the authority of the National Institute of Population Research and Training (NIPORT) of the Ministry of Health and Family Welfare (MOHFW) to evaluate existing health and social programs and to design new strategies for improving the health status of the country's women and children. Until 2018, eight BDHS surveys have been conducted: in 1993-94, 1996-97, 2000, 2004, 2007, 2011, 2014 and 2017-18. In this study, the survey data over the period 1994-2014 have been used since the district level location of the surveyed clusters is not disclosed in the most recent BDHS 2017-18. Over the period of 1994-2014, three Population and Housing Censuses have been conducted, in 1991, 2001 and 2011. Full census data are not available, but only 10% of Census 1991 data, 10% of Census 2001 data and 5% of Census 2011 data are publicly available from IPUMS-International (https://international.ipums.org). A number of district-level contextual variables have been generated and used in the development of cross-sectional FH models to produce input estimates for the MTS models.

3.2 Direct estimates

The variables analysed in this paper are ANC0 and ANC4. Bangladesh is divided into 7 sub-national regions, called divisions. These divisions are further divided into 64 districts, which is the most detailed regional level considered in this study. As a first step, estimates and variance estimates of the two target variables at the district level are obtained from each survey year's unit-level data using the standard design-based direct survey estimator (hereafter denoted by DIR), where the survey weights are used to account for the sampling design and for non-response.

In this study, reproductive age ever-married women who had given a birth within the last three years before a survey year are considered as the target population. Since in the census population such pregnancy related information are not available, area-specific population size are estimated by the number of reproductive age ever-married women available in three Census data. Division and district specific population size are shown in Figure S.1.

The BDHS uses a two-stage stratified sample of households. The strata are formed from divisions and sub-divisions according to their urban-rural characterization. In the first stage, primary sampling units (PSUs or clusters) are selected with probability proportional to PSU size (the number of households) from each of the strata independently. In the second stage, a complete household listing is carried out in all selected PSUs and then about 30 households are selected from each PSU using systematic sampling. Note that division membership of the districts changed over time for some districts. In this paper the current clustering of districts within divisions is used for the entirely observed time period. The response rates among eligible women have been over 95% in all BDHS years. Though the sample size of the ever-married women is greater than 10,000 in all the surveys, in this study only the ever-married women who had a child birth in the three years preceding the survey year are considered, and therefore sample sizes are smaller. At the district level, mean sample sizes vary between 60 and 114, with some districts having less than 10 or even no observed women.

Sampling weights are calculated based on selection probabilities. These weights are then adjusted for household and individual non-response.

The direct estimate for the population proportion in a certain domain i for survey year t is computed as

$$\hat{\hat{Y}}_{it} = \frac{\sum_{j \in s_{it}} w_{ijt} y_{ijt}}{\sum_{j \in s_{it}} w_{ijt}}, \qquad (1)$$

where y is the response variable of interest, s_{it} is the set of ever-married women in domain i for which y is observed in year t, and w_{ijt} is the survey weight for person j living in area i in year t. Note that the weights w_{ijt} are scaled such that the sum over the weights in the sample is equal to the net sample size. The corresponding variance estimates are approximated as

$$var(\hat{\hat{Y}}_{it}) = \frac{1}{n_{it}(n_{it}-1)} \sum_{j \in s_{it}} w_{ijt} (y_{ijt} - \hat{\hat{Y}}_{it})^2, \qquad (2)$$

where n_{it} is the number of ever-married women observed in domain i at the survey year t. Initially, the variance was approximated by calculating the variance among the estimated PSU totals as if they were selected by using stratified sampling with replacement, known as the ultimate sampling unit variance approximation. This resulted in zero variance estimates for a few domains. Variance approximation (2) avoids these zero variance estimates, and otherwise results in variance estimates comparable with the initial approximation where PSUs were assumed to be selected with replacement.

The direct estimates and their estimated variances are computed for the survey years 1993-1994, 1996-97, 2000, 2004, 2007, 2011, and 2014. In our first MTS model, denoted MTS-I, these direct estimates are used as the input series. In a second approach, these direct estimates are first modelled using cross-sectional FH models, as explained in the next Subsection.

3.3 Fay-Herriot estimates

An issue with the MTS-I model is the use of census data as auxiliary variables in the MTS model. Because the time gap between two subsequent censuses is 10 years whereas the BDHS is conducted every 3 or 4 years, the census covariates remain the same until the new census data are available. It is anticipated that including these census data as covariates in the MTS-I models will bias estimates of trends and period-to-period changes. One way to take advantage of the census information is to model the direct estimates at the district level in separate cross-sectional FH models using relevant contextual variables extracted from the census data. It is also expected that the use of on-time available census auxiliary variables in repetitive cross-sectional FH models may affect regression coefficients and the accuracy of model predictions of the dependent variable, but not the predictions of the dependent variable itself. Compared to the direct estimates used in MTS-I, these cross-sectional FH models also provide better estimates by already borrowing some strength over districts.

The cross-sectional FH estimates and their standard errors are used as input to another MTS model, denoted by MTS-II. The fixed and random effect components used for developing survey-specific cross-sectional FH models are shown in Supplementary Tables S.1 and S.2. For all the models, district-specific random effects are assumed to follow a normal distribution. These Tables also show the transformation used to develop the FH models. Non-normal models are considered for the random effects (Laplace and horseshoe) and the sampling error (t-distribution) as alternatives for the normal distribution. This, however, did not improve the model fit. The cross-sectional FH estimates are correlated due to their common fixed effect components. Therefore, a third MTS model, denoted MTS-III, is developed using cross-sectional FH estimates and their full covariance matrix as input.

3.4 Generalized Variance Functions (GVF)

In the FH and MTS models, the variance estimates of the direct estimates are largely treated as fixed given quantities. Since these variance estimates can be very noisy due to the detailed estimation level, they are smoothed using Generalized Variance Functions (GVF) before using them in the FH and MTS models. Note that a district without sample information is considered as missing and is therefore not considered in the model development approach. The cross-sectional FH model can produce estimates and standard errors for these out-of-sample domains. These synthetic estimates are, however, not used in the development of the MTS-II and MTS-III models so as to allow a better comparison with the MTS-I model.

The GVF are regression models that relate the variance estimates to a few predictors such as sample size, survey design variable, and point estimates, Wolter (2007), Chapter 7. Now let $\hat{Y}_{it}$ denote the direct estimates for ANC0 or ANC4 for district i in year t. Also, let $se(\hat{Y}_{it})$ denote standard errors of $\hat{Y}_{it}$. For both target variables, we use the same GVF smoothing model

$$\log se(\hat{Y}_{it}) = \alpha + \beta \log \tilde{Y}_{it} + \gamma \log(m_{it} + 1) + \delta Division + \epsilon_{it}\,, \tag{3}$$

where m_{it} is the number of sampling units contributing to district i in year t and $Division$ is a categorical variable with 7 levels.

Since we cannot trust the direct estimates for very small m_{it}, the $\tilde{Y}_{it}$ on the right hand side of (3) are simple smoothed estimates

$$\begin{aligned} \tilde{Y}_{it} &= \lambda_{it}\hat{Y}_{it} + (1-\lambda_{it})\bar{Y}_{d[i]t}\,, \\ \lambda_{it} &= \frac{m_{it}}{m_{it}+1}\,, \end{aligned} \tag{4}$$

where $\bar{Y}_{d[i]t}$ denotes the mean for division d (d=1 to 7) to which district i belongs, in year t.

The regression errors ϵ_{it} are assumed to be independent and normally distributed with a common variance parameter σ^2. The GVF models are fitted only to districts with non-zero standard errors of the direct estimates. The predicted (smoothed) standard errors based on the fitted models are

$$se_{\text{pred}}(\hat{Y}_{it}) = \exp\left(\hat{\alpha} + \hat{\beta} \log \tilde{Y}_{it} + \hat{\gamma} \log(m_{it} + 1) + \hat{\delta} Division + \hat{\sigma}^2/2\right), \quad (5)$$

where $\hat{\sigma}$ is 0.03 for ANC0 and 0.003 for ANC4 coverage respectively. The R-squared model fit measures for both models are quite high 0.79 for ANC0 and 0.99 for ANC4. Note that the exponential back-transformation in (5) includes a bias correction, which in this case has only a small effect. The same approach has been implemented for getting smoothed standard errors for an individual survey year to develop the corresponding cross-sectional FH model.

3.5 Transformations of input series

The direct estimates and their standard errors for ANC0 and ANC4 display strong positive dependence. This may cause some problems in fitting the models, including MCMC simulation convergence problems. Therefore, a square root transformation is chosen as a variance stabilizing transformation, see Sakia (1992). Log and log-ratio transformations are also considered but the square root transformation appears to be the most appropriate variance stabilizing transformation, at least for ANC4. The square root transformation for ANC4 reduces the correlation between point estimates and their standard errors of the input series, reduces heterogeneity, improves the convergence of the MCMC simulation, and reduces the skewness of proportion data if they take values close to the lower boundary of zero.

For ANC0, the square root (SQRT) transformation is used for the year specific cross-sectional FH models in 2011 and 2014 only. In the other years, no transformation is applied. In all three MTS models, no transformation is applied for ANC0 since SQRT transformation for the input series shows some dependence between estimates and standard errors (see Supplementary Figure S.2). For ANC4 the SQRT transformation is applied for all cross-sectional FH models. The square root transformation is also found appropriate for the input series of all three MTS models for ANC4 (see Supplementary Figure S.2).

Let $\tilde{Y}_{it} = \sqrt{(\hat{Y}_{it} + \varepsilon)}$ denote the square root transformed direct estimates, where ε is a small number necessary (0.005) because for some districts direct estimates equal zero. It is understood that ε is added only if the direct estimate for a domain is equals zero. Using a first order Taylor approximation it can be shown that

$$se(\tilde{Y}_{it}) = se(\hat{Y}_{it})/(2\sqrt{\hat{Y}_{it} + \varepsilon}).$$

If the GVF smoothing model (3) is applied to the standard errors of the untransformed direct estimates, then the standard errors for domains with a very small number of sampling units can become unreasonably large due to the linearisation approximation. This issue is avoided by applying the GVF to the standard errors of the transformed estimates, i.e. $se(\tilde{Y}_{it})$. A comparison of the smoothed standard errors generated from GVF model on original scale and those from GVF model on SQRT scale for ANC4 are shown in Supplementary Figure S.3.

4 Time series multilevel modelling

In this study, direct estimates and their standard errors are available for the survey years 1994, 1997, 2000, 2004, 2007, 2011 and 2014. To account for the varying time-lags of 3 or 4 years between the subsequent survey years, the MTS models are defined at an annual frequency at the most detailed regional level of the 64 districts. With a time span of 21 years, there are 1344 domain-year combinations. With seven available survey years, the model is fitted to the 448 domain-year observations. The years between two subsequent surveys are defined as missing in the model. In this way the period-to-period evolution of the trend is specified correctly and the model provides predictions for the missing domain-year combinations.

For convenience let us now denote by $\hat{Y}_{it}$ the input series for the time series models for either ANC0 or ANC4 in year t and domain i. This can be the untransformed direct estimates, the SQRT transformed direct estimates or the model predictions obtained with the cross-sectional Fay Herriot models. Here domain index i runs from 1 to $M_d = 64$ and t from 1 to T corresponding to the years 1994 to 2014. We further combine these estimates into a vector $\hat{Y} = (\hat{Y}_{11}, \ldots \hat{Y}_{M_d 1}, \ldots \hat{Y}_{1T}, \ldots \hat{Y}_{M_d T})'$, a vector of dimension $M = M_d T$.

4.1 Model structure

The multilevel models considered take the general linear additive form

$$\hat{Y} = X\beta + \sum_{\alpha} Z^{(\alpha)} v^{(\alpha)} + e\,, \qquad (6)$$

where X is a $M \times p$ design matrix for a p-vector of fixed effects β, and the $Z^{(\alpha)}$ are $M \times q^{(\alpha)}$ design matrices for $q^{(\alpha)}$-dimensional random effect vectors $v^{(\alpha)}$. Here the sum over α runs over several possible random effect terms at different levels, such as local level and smooth trends at district and division levels, white noise at the most detailed level of the M domains, etc. This is explained in more detail below. In formula (6) $e = (e_{11}, \ldots, e_{M_d 1}, \ldots e_{M_d T})'$ denotes, depending on the input series, the sampling errors of the direct estimates or the prediction errors of the Fay Herriot model. The errors are taken to be normally distributed as $e \sim N(0, \Sigma)$ where $\Sigma = \bigoplus_{t=1}^{T} \Sigma_t$. If the input series are the untransformed direct estimates, then Σ_t is the covariance matrix for the untransformed direct estimates observed in year t. If the input series are transformed, then Σ_t is the covariance matrix for the transformed direct estimates, as described in Subsection 3.5. If the input series are the predictions based on the cross-sectional FH models, then Σ_t contains the estimated mean squared errors of the FH predictions. In the latter case a diagonal version that ignores the correlations between the domain predictions is considered as well as a full covariance matrix that also accommodates the correlations between domain predictions. Recall from Subsection 3.2 that changes in division membership of the districts are ignored in model MTSI. Assuming a diagonal matrix for the cross-sectional FH predictions under model MTSII also implies that the covariances between FH predictions are ignored. Note that this might affect the efficiency of the estimation procedure, but the point estimates obtained under both models are still unbiased.

Based on the distribution of sampling errors e, the likelihood function can be defined as

$$p\left(\hat{Y}|\eta, \Sigma\right) = N\left(\hat{Y}|\eta, \Sigma\right)\,, \qquad (7)$$

where $\eta = X\beta + \sum_\alpha Z^{(\alpha)} v^{(\alpha)}$ is the linear predictor. For the errors e a Student-t distribution instead of the normal distribution can be considered to give smaller weight to more outlying observations following West (1984).

The fixed effect part of η can contain components like an intercept, a linear trend, main effects for division and district and possibly the second-order interactions for linear trends and division or district. The vector β of fixed effects is assigned a normal prior $p(\beta) = N(0, 100I)$, which is only very weakly informative as a standard error of 10 is very large relative to the scales of the (transformed) direct estimates and the covariates used.

The second term on the right hand side of (6) consists of a sum of contributions to the linear predictor by random effects or varying coefficient terms. The random effect vectors $v^{(\alpha)}$ for different α are assumed to be independent, but the components within a vector $v^{(\alpha)}$ are possibly correlated to accommodate temporal or cross-sectional correlation. To describe the general model for each vector $v^{(\alpha)}$ of random effects, we suppress superscript α in what follows for notational convenience.

Each random effects vector v is assumed to be distributed as

$$v \sim N(0, A \otimes V)\,, \tag{8}$$

where V and A are $d \times d$ and $l \times l$ covariance matrices, respectively, and $A \otimes V$ denotes the Kronecker product of A with V. The total length of v is $q = dl$, and these coefficients may be thought of as corresponding to d effects allowed to vary over l levels of a factor variable. If, e.g., V corresponds to division, then V defines $d = 7$ different random effects that correspond to the 7 categories of division. If subsequently A corresponds to time, then $l = 21$ years. In that case each of the 7 effects can vary over its 21 levels (years in this case). Each random effect generated for a division × year combination is shared by all districts belonging to that division in that particular year.

The covariance matrix A describes the covariance structure among the levels of the factor variable, and is assumed to be known. Instead of covariance matrices, precision matrices $Q_A = A^{-1}$ are actually used, because of computational efficiency (Rue and Held, 2005). The covariance matrix V for the d varying effects can be parameterized in one of three different ways: (i) a fully parameterized covariance matrix, (ii) a diagonal matrix with unequal diagonal elements, and (iii) a diagonal matrix with equal diagonal elements. The scaled-inverse Wishart prior is used as proposed in O'Malley and Zaslavsky (2008) and recommended by Gelman and Hill (2007) when a full covariance matrix is assumed, while half-Cauchy priors are used for the standard deviations when the covariance matrix is assumed diagonal with equal or unequal elements. In case of diagonal variances, half-Cauchy priors are better default priors than the more common inverse gamma priors Gelman (2006).

The following random effect structures are considered in the model selection procedure:

1. Random intercepts for the M_d domains. In this case $A = I_{M_d}$ and V is a scalar variance parameter. This implies $v_{it} = v_i,\ \forall t$ and $v_i \sim N(0, \sigma_I^2)$.
2. First or second order random walks at different aggregation levels. A first order random walk or local level trend at district level is defined as $v_{it} = L_{it}$ with $L_{it} = L_{i,t-1} + \eta_{it}$ and $\eta_{it} \sim N(0, \sigma_{R1,i}^2)$. A second order random walk or smooth trend model at district level is defined as $v_{it} = L_{it}$ with $L_{it} = L_{i,t-1} + R_{i,t-1}$, $R_{it} = R_{i,t-1} + \eta_{it}$ and $\eta_{it} \sim N(0, \sigma_{R2,i}^2)$. Both kind of trends can be defined similarly at the division or national level. See Rue and Held (2005) for the specification of the precision matrix Q_A for first and second order random walks. A full covariance matrix

for the trend innovations can be considered to allow for cross-sectional besides temporal correlations, or a diagonal matrix with different or equal variance parameters to allow for temporal correlations only. In the case of equal variances $\sigma^2_{R1,i} = \sigma^2_{R1}$, $\forall\, i$ and $\sigma^2_{R2,i} = \sigma^2_{R2}$, $\forall\, i$. First and second order random walk components at district level are denoted below by $RW1\_District$ and $RW2\_District$ respectively. At division level they are denoted by $RW1\_Division$ and $RW2\_Division$.

3. The first order random walks as used in our models cannot capture an overall level as the corresponding random effects are constrained to sum to zero over time. Similarly, the second order random walks cannot capture both level and linear trend. This means that level and linear trend must be accommodated by other model terms, as either fixed or random effects. District-level intercepts have already been discussed under item 1. To also include linear trends by district, this component can be extended to random intercepts and slopes linear in time. In that case V can be either a 2×2 general covariance matrix
$$V = \begin{pmatrix} \sigma_I^2 & \rho_{IS} \\ \rho_{IS} & \sigma_S^2 \end{pmatrix},$$
accounting for correlations between intercepts and slopes, or a diagonal matrix with diagonal elements σ_I^2 and σ_S^2 the variances of the radom intercept and slopes respectively. This model component is referred to as $RIS\_District$ below.
4. Spatial random effects: random intercepts varying over the spatial location of districts following an intrinsic conditional autoregressive (ICAR) model (Besag and Kooperberg, 1995), defined as $v_i \mid v_{-i} \sim N((\Sigma_{i' \in nb(i)} v_{i'})/a_i, \sigma^2_{Sp}/a_i)$ for each spatial effect conditional on the others. Here $nb(i)$ is the set of domains neighbouring domain i and a_i the number of domains neighbouring domain i. See Rue and Held (2005) for the specification of the precision matrix Q_A. This spatial component is referred to later by $Spatial\_District$.
5. White noise: to allow for random unexplained variation, white noise at the most detailed domain-by-year level can be included. In this case $A = I_M$ and V a scalar variance parameter. This implies $v_{it} \sim N(0, \sigma_W^2)$.

We also investigated generalisations of (8) to non-normal distributions of random effects by implementing Student-t, horseshoe prior (Carvalho et al., 2010) and Laplace distribution ((Tibshirani, 1996; Park and Casella, 2008)). These alternative distributions have fatter tails allowing for occasional large effects. However, these distributions did not improve results for the considered target variables. Therefore the normal distribution is used for all random effect components.

4.2 Model estimation

The models are fitted using Markov Chain Monte Carlo (MCMC) sampling, in particular the Gibbs sampler (Geman and Geman, 1984; Gelfand and Smith, 1990). See Boonstra and van den Brakel (2018) for a specification of the full conditional distributions. The models are run in R (R Core Team, 2015) using package `mcmcsae` (Boonstra, 2021). The Gibbs sampler is run in parallel for three independent chains with randomly generated starting values. In the model building stage 1000 iterations are used, in addition to a 'burn-in' period of 100 iterations. This was sufficient for reasonably stable Monte Carlo estimates of the model parameters and trend predictions. For the selected model we use a longer run of 1000 burn-in plus 5000 iterations of which the draws of every fifth iteration are stored. This leaves $3 \times 1000 = 3000$ draws to compute estimates and standard errors. The convergence of the MCMC simulation is assessed using trace and autocorrelation plots as well as the Gelman-Rubin potential scale reduction factor (Gelman and Rubin, 1992), which diagnoses the mixing of the chains. For the longer simulation of the selected model all model parameters and model predictions have potential scale reduction factors below 1.01 and sufficient effective numbers of independent draws.

Many models of the form (6) have been fitted to the data. For the comparison of models using the same input data we use the Widely Applicable Information Criterion or Watanabe-Akaike Information Criterion (WAIC) (Watanabe, 2010, 2013) and the Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002). We also compare the models graphically by their model fits and trend predictions at three aggregation levels.

5 Selected models and model prediction

5.1 MTS model for ANC0

No transformation for the input series of the direct estimates or the FH estimates is considered. The model parameters in (6) are separated in fixed and random effects. The following simple fixed effects components are included in the finally selected model for MTS-I, MTS-II, and MTS-III:

$$1 + Division + yr.c + Division * yr.c \tag{9}$$

Here $yr.c$ denotes the centered and scaled quantitative year variable. The random effects part of the three models is shown in Table 5.1. If multiple varying effects are modeled, then there is a choice between scalar, diagonal or full covariance matrix V in (8). For variation over time, second order random walks $RW2\_Division$ and $RW2\_District$ were finally selected. White noise components are considered but not included in the final model since it did not further improve the model fit.

Model Component	Formula **V**	Variance Structure	Factor **A**	Number of Effects
RIS_District	$1 + yr.c$	full	$District$	128
RW2_Division	$Division$	scalar	RW2(yr)	147
RW2_District	$District$	scalar	RW2(yr)	1344
Spatial_District	1	scalar	Spatial($District$)	64

Table 5.1 Summary of the random effect components for the selected time series multilevel model for ANC0. The second and third columns refer to the varying effects with covariance matrix V in (8), whereas the fourth column refers to the factor variable associated with A in (8). The last column contains the total number of random effects for each component.

5.2 MTS model for ANC4

The square-root transformation is applied to the input series of the direct and FH estimates of ANC4 for models MTS-I, MTS-II, and MTS-III. For MTS-I the GVF is applied to the transformed standard errors to obtain the variance matrix Σ. Applying the GVF model on the original scale produces unstable, higher standard errors for some domains with small sample size, particularly for the districts of Chittagong Hill Track region for which either very small or no sample information were available. For the fixed effect component a factor variable called "Region" has been created based on the degree of urbanization following Rahman et al. (2019). The variable has four levels; 1 for three big cities Dhaka, Chittagong and Gazipur, 2 for other nine regional big cities (Barisal, Bogra, Comilla, Khulna, Mymensing, Narayanganj, Rajshahi, Rangpur, Sylhet), 3 for three hilly districts (Bandarban, Khagrachhari and Rangamati) and 4 for the remaining districts. This variable mainly helped to adjust the estimates for the three hilly districts which have very few (even no) information in the considered seven surveys. The final model has the following fixed effects components:

$$1 + Division + yr.c + Region \tag{10}$$

The interaction between "Division" and "yr.c" (like in the ANC0 model) was found to be insignificant in the ANC4 model. The random effect components for ANC4 model shown in Table 5.2 are very similar to those used for the model of ANC0 (shown in Table 5.1). A local level trend instead of smooth trend at division level (RW1_Division in Table 5.2) has been considered since the smooth trend component (RW2_Division, as in Table 5.1) resulted in some bias in the national and divisional trends. Also, the model with RW1_Division component gives better scores for the information criteria compared to the model with RW2_Division component. White noise components are considered but not included in the final model since it did not further improve the model fit.

Model Component	Formula **V**	Variance Structure	Factor **A**	Number of Effects
RIS_District	$1 + yr.c$	full	$District$	128
RW1_Division	$Division$	scalar	RW1(yr)	147
RW2_District	$District$	scalar	RW2(yr)	1344
Spatial_District	1	scalar	Spatial($District$)	64

Table 5.2 Summary of the random effect components for the selected multilevel time series model for ANC4. The second and third columns refer to the varying effects with covariance matrix V in (8), whereas the fourth column refers to the factor variable associated with A in (8). The last column contains the total number of random effects for each term.

5.3 Trend estimation

Trend estimates are computed based on the MCMC simulation results. In a first step, for each MCMC replicate, an M-dimensional vector containing predictions at the most detailed level of all year-district combinations is computed as

$$\eta^{(r)} = X\beta^{(r)} + \sum_{\alpha} Z^{(\alpha)} v^{(\alpha,r)}, \tag{11}$$

where superscript (r) indexes the retained MCMC draws. Note that $\eta^{(r)}$ also includes predictions for the years without survey observations. Since a square root transformation was applied to the ANC4 series, initially the following back-transformation for the vectors $\eta^{(r)}$ was considered:

$$\theta^{(r)} = (\eta^{(r)})^2 + (se(\tilde{Y}_{it}))^2 . \tag{12}$$

The second term on the right hand side is a (relatively small) bias correction using the transformed and smoothed standard errors. The bias correction stems from the fact that the design expectation of the direct estimates can be written as

$$E(\hat{Y}) = E((\tilde{Y})^2) = E((\eta + \tilde{E})^2) = \eta^2 + var(\tilde{E}),$$

where $\tilde{E}$ is the vector of sampling errors after transformation, assumed to be normally distributed with standard errors $se(\tilde{Y}_{it})$. A difficulty with the data at hand is that the bias correction can only be applied to the survey years, since standard errors are only available for those years. Applying the bias correction only for the survey years distorts the trend estimates, as illustrated in Supplementary Figures S.4, S.5, S.6. In case of MTS-I model, the impact of this bias correction is most clear for those domains with zero direct estimates particularly for *Chittagong* hilly districts. The impact of the bias correction is less in case of MTS-II and MTS-III models since the estimated standard errors of the FH estimates are already smoothed enough and consistent. However, at national and division levels this bias correction causes some overestimation in some survey years for all the trends based on the MTS models. Therefore, the bias correction for the SQRT transformation is not applied in the trend estimates but only used in the calculation of cross-sectional FH estimates.

Trend estimates at the most detailed level of districts for all years are obtained by taking the mean over the MCMC replications $\eta^{(r)}$. Standard errors for the trend estimates are obtained by taking the standard deviation of the MCMC replicates. Trends at the divisional and national levels are obtained as a weighted average of the trend estimates at the most detailed regional level of districts, using the number of ever-married women as a weighting variable.

6 Results

The trends of ANC0 and ANC4 shown in the Figures consist of five types of estimates with their approximate 95% confidence intervals: (i) weighted direct estimates (DIR) at the surveyed year (black error-bar line), (ii) cross-sectional FH estimates (FH) at the surveyed year (green error-bar line), (iii) estimates based on MTS model using the DIR estimates and their smoothed standard errors as input (red line) referred to as MTS-I, (iv) estimates based on MTS model using the FH estimates and their standard errors as input (green line) referred to as MTS-II and (v) estimates based on MTS model using the FH estimates and their variance-covariance matrix as input (blue line) referred to as MTS-III hereafter.

6.1 ANC0

Variance estimates of the random components for the three time series models for ANC0 are given in Table 6.1.

The national level trends of ANC0 are shown in Figure 6.1. The figure shows that the DIR and cross-sectional FH estimates are very similar at the survey years with approximately equal 95% CI. This can be expected for figures at the national level, since the gain in precision obtained with a small area prediction model with respect to a direct estimator becomes smaller as the sample size increases. During the initial period 1994-2000, the national level trend based on the MTS-I model (red line) follows the DIR and cross-sectional FH estimates, while the trends based on MTS-II and MTS-III models are slightly higher. For the period 2004-2010, the trend based on MTS-I model is slightly higher than the trends based on MTS-II and MTS-III models. The differences are, however, very small.

Model	$\hat{\sigma}_I^{(3)}$ (SE)	$\hat{\sigma}_S^{(3)}$ (SE)	$\hat{\rho}_{IS}^{(3)}$ (SE)	$\hat{\sigma}_{Sp}^{(3)}$ (SE)	$\hat{\sigma}_{R2}^{(2)}$ (SE)	$\hat{\sigma}_{R2}^{(3)}$ (SE)
MTS-I	0.083 (0.013)	0.054 (0.007)	0.168 (0.171)	0.068 (0.032)	0.019 (0.003)	0.024 (0.002)
MTS-II	0.069 (0.012)	0.033 (0.004)	0.254 (0.180)	0.071 (0.028)	0.020 (0.003)	0.013 (0.002)
MTS-III	0.062 (0.013)	0.027 (0.013)	0.227 (0.201)	0.067 (0.030)	0.020 (0.003)	0.009 (0.001)

Table 6.1 Posterior means of standard deviation parameters of random components of MTS-I, MTS-II, MTS-III models for ANC0. Superscripts (2) **and** (3) **refer to division and district levels, resp.**

The trends at division level, shown in Figure 6.2, indicate that the trends under MTS-I are very similar to those based on MTS-II and MTS-III models with some small exceptions in Dhaka, Khulna and Rajshahi divisions. The differences in Dhaka and Khulna division may cause most of the differences in the national level trends.

The trends based on the MTS-II and MTS-III models are almost identical at national and division levels. This is supported by the estimated variance components of the division-level smooth-trend random component under the developed two models ($\hat{\sigma}_{(RW2)}^{(2)}$: about 0.020). However, there are more substantial differences in the trends under MTS-II and MTS-III at the district level, see Figures 6.3 and 6.4. Plots for all districts are given in Appendix S.1. The trends based on the MTS-III model are smoother than those based on the MTS-II model, which is a result of the smaller values of the estimated variance component $\hat{\sigma}_{(RW2)}^{(3)}$ under MTS-III (see Table 6.1).

The trends at the district level have a tendency to follow the pattern of their respective division level trend shown in Figure 6.2. This is particularly the case for domains with a relatively small number of observations such as districts *Bandarban*, *Khagrachnari* and *Rangamati* in Figure 6.3 that belong to *Chittagong* division in Figure 6.2. To reduce this tendency, an MTS model was developed by removing smooth trend component $RW2\_Division$ at division level in Table 5.1. This, however, resulted in highly smooth unrealistic trends at the national and divisional levels. In a similar way, to examine the need for a spatial component, MTS models were developed with and without considering the spatial component ($Spatial\_District$ in Table 5.1). It is observed that the spatial component slightly improves the estimates for those districts with small or zero sample sizes. See for example the trends of *Bandarban* and *Rangamati* districts of *Chittagong* division.

The MTS-I model shows upward trends for some districts during the period of 1994-2000. These developments are unplausible from a subject matter point of view and are nicely corrected by the cross-sectional FH estimates based MTS-II and MTS-III models. See for example *Noakhali*, *Bandarban*, *Rangamati*, *Narayanganj*, *Rajbari*, and *Narail* districts in Figure 6.3. Some districts have volatile trends according to the DIR estimates and MTS-I model during the whole period mainly due to variation in the sample size. See for example, *Bandarban*, *Bhola*, *Khagrachhari*, *Kishoreganj* and *Rangamati* in Figure 6.3, *Chapai Nababganj*, *Feni*, *Jhalokati*, *Joypurhat*, and *Pabna* districts in Figure 6.4. From a subject matter point of view a smooth decreasing trend for ANC0 coverage is expected. In particular the turning points that are visible in several districts arround 2007 and 2011 are not expected. The trends based on the MTS-II and

MTS-III models ignore most of these volatilities and show reasonable smooth trends for these districts and are therefore more realistic compared to MTS-I, since they suppress this unexpected increase. Nevertheless the fit of all three models are compatible with the observed data. MTS-II appears to be a nice compromise between models I and III.

In most cases models MTS-II and MTS-III behave similarly. However, model MTS-III, which accounts for correlation among the cross-sectional FH estimates, overestimates ANC0 for some districts (such as *Chapai Nababganj*, *Lalmonirhat* and *Shariatpur* districts in Figure 6.4) and also slightly underestimate the trend in some districts (such as *Khagrachari*, *Rangamati*, and *Shirajganj* districts in Figure 6.3) compared to the cross-sectional FH estimates. Again MTS-II seeks a compromise between smooth trends under MTS-III and more volatile trends under MTS-I in most of the districts and appears to be the preferred model for estimating trends of ANC0.

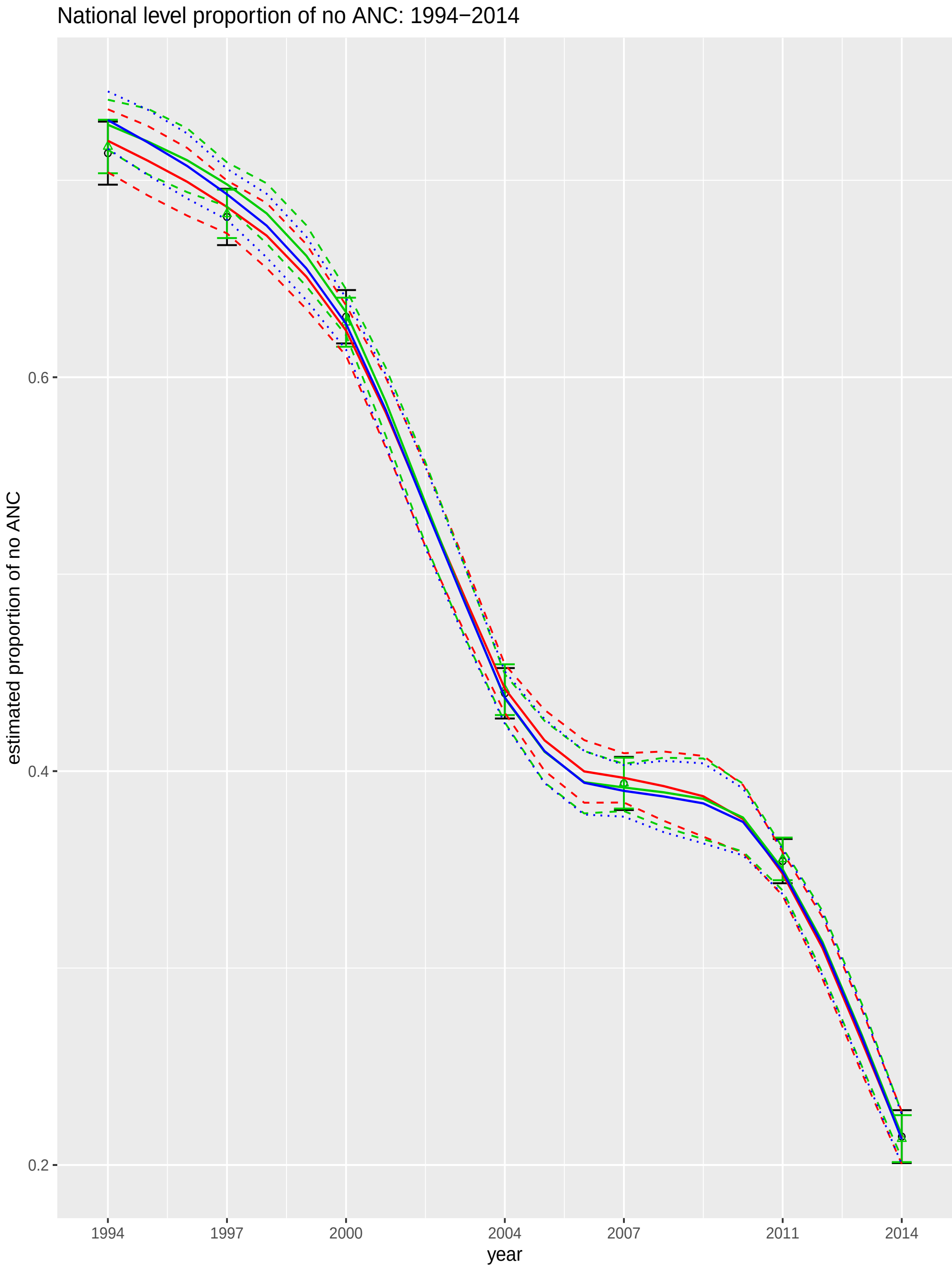

Figure 6.1 National level trends of ANC0 in Bangladesh: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

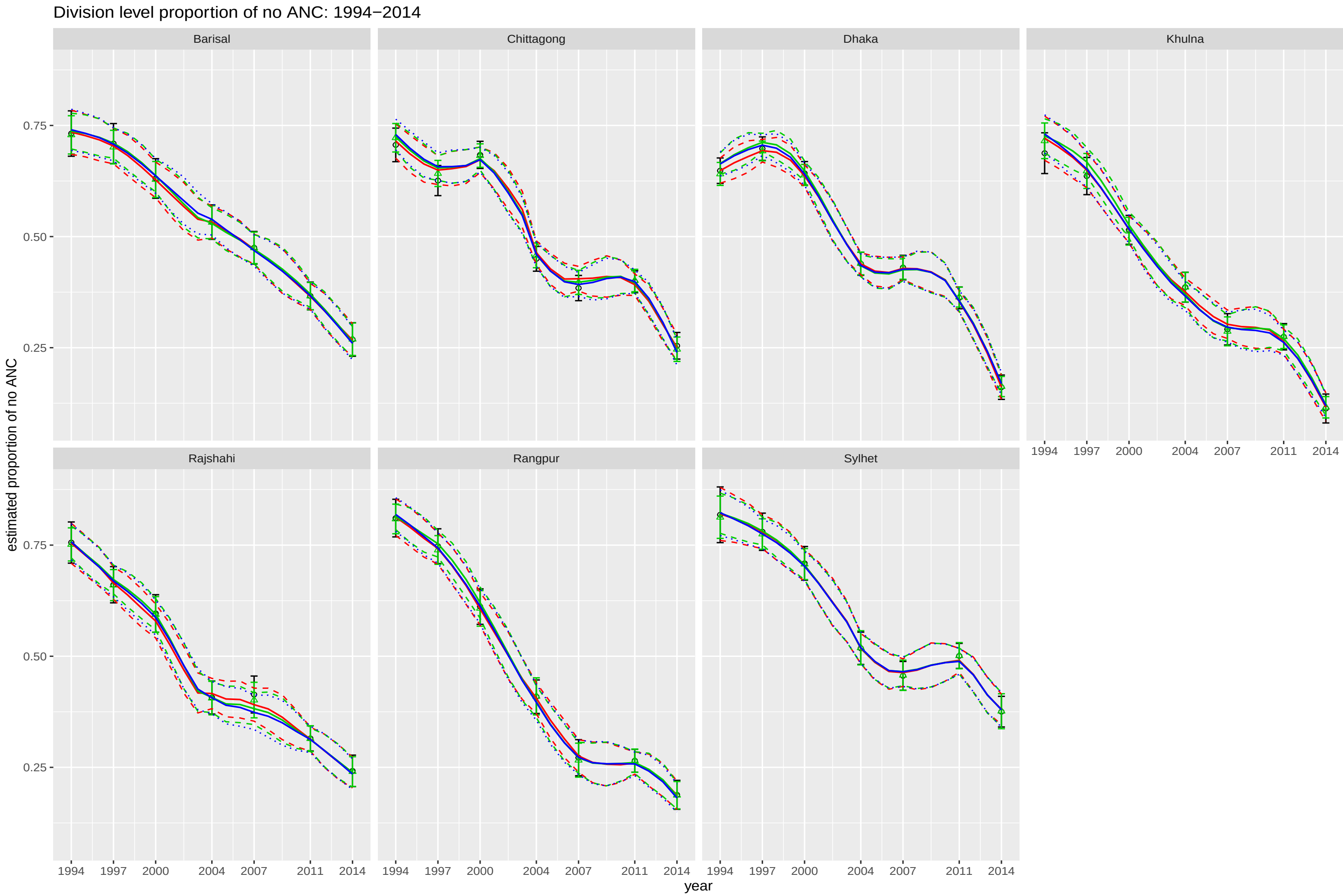

Figure 6.2 Division level trends of ANC0 in Bangladesh: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

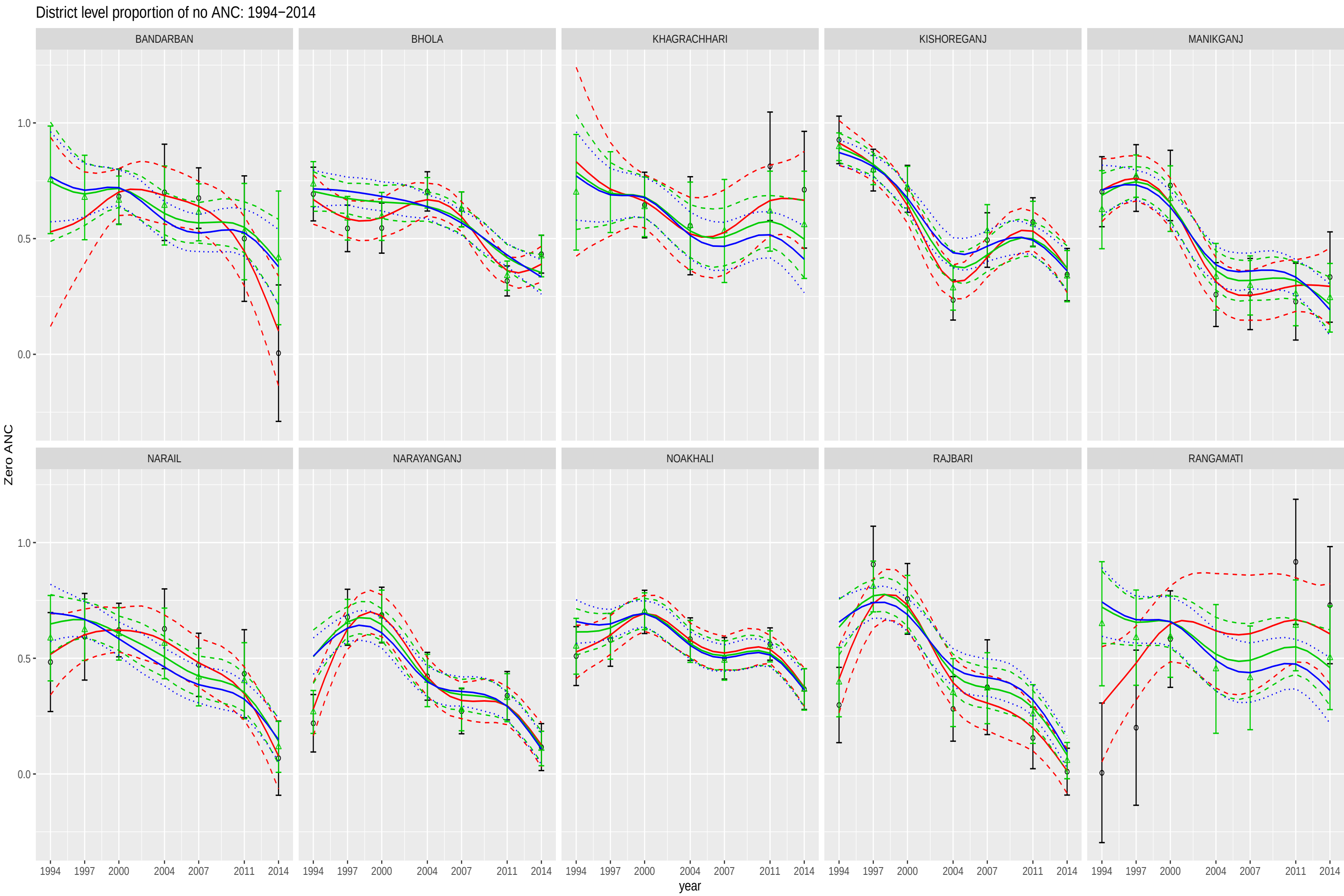

Figure 6.3 District level trends of ANC0 in Bangladesh: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

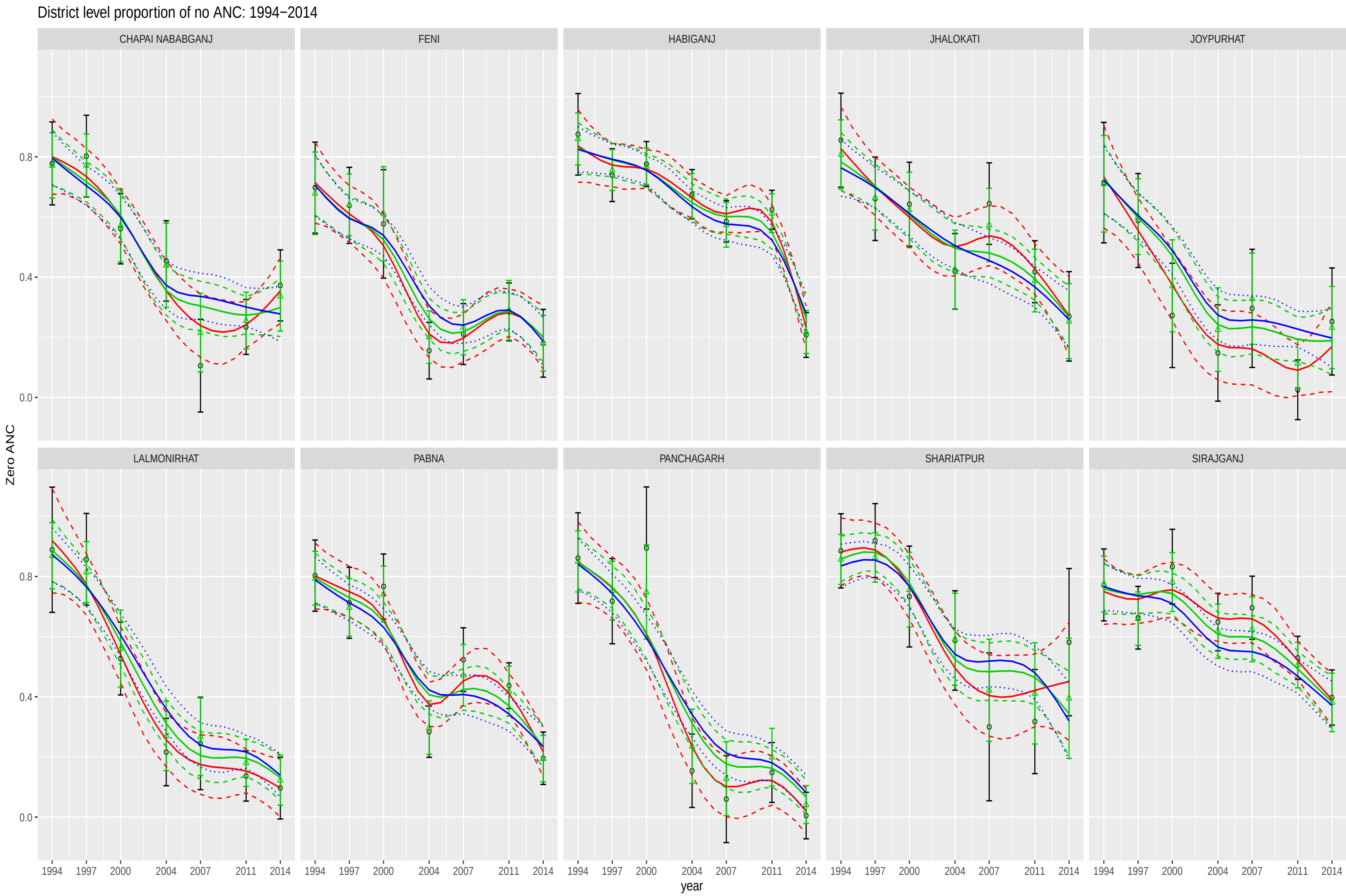

Figure 6.4 District level trends of ANC0 in Bangladesh: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

Model	$\hat{\sigma}_I^{(3)}$ (SE)	$\hat{\sigma}_S^{(3)}$ (SE)	$\hat{\rho}_{IS}^{(3)}$ (SE)	$\hat{\sigma}_{Sp}^{(3)}$ (SE)	$\hat{\sigma}_{R1}^{(2)}$ (SE)	$\hat{\sigma}_{R2}^{(3)}$ (SE)
MTS-I	0.060 (0.010)	0.033 (0.005)	0.428 (0.178)	0.047 (0.026)	0.012 (0.004)	0.009 (0.001)
MTS-II	0.046 (0.007)	0.022 (0.003)	0.501 (0.162)	0.035 (0.018)	0.016 (0.003)	0.004 (0.001)
MTS-III	0.038 (0.006)	0.018 (0.006)	0.522 (0.165)	0.027 (0.016)	0.014 (0.003)	0.002 (0.001)

Table 6.2 Posterior means of standard deviation parameters of random components of MTS-I, MTS-II, MTS-III models for ANC4. Superscripts (2) **and** (3) **refer to division and district levels, resp.**

6.2 ANC4

Variance estimates of the random components for the three time series models for ANC4 are given in Table 6.2.

The national level trend of ANC4 shown in Figure 6.5 shows a linear upward increase from 6% in 1994 to about 31% in 2014. Like ANC0, the DIR and cross-sectional FH estimates of ANC4 are very similar at the survey years with approximately equal 95% CI. Trends estimated from the MTS-I (red line), MTS-II (green line) and MTS-III (blue line) show very similar patterns. Compared to the DIR and cross-sectional FH estimates, the trend of MTS-I is slightly lower in 2007 and 2014. Trends under MTS-II and MTS-III in survey year 2011 are somewhat larger compared to the DIR and cross-sectional FH estimates. The trends at division level are shown in Figure 6.6. The three MTS models give very similar trend estimates. Some differences occur in *Chittagong*, *Dhaka* and *Rangpur* divisions. With MTS-I the trend is slightly higher compared to the DIR and FH estimates for *Rangpur* division over the 1994-2000 period. For MTS-II and MTS-III, the trend is somewhat higher in *Rajshahi* division during 2011-2014 period compared to the DIR and FH estimates. All three MTS models show slightly bow-shaped 95% CI bands in between two subsequent survey years, which indicates slightly higher uncertainty during the non-survey years compared to the survey years.

Although the trends based on MTS-II and MTS-III are almost identical at national and division levels, the estimated variance components of both model differ considerably as follows from Table 6.2. These differences lead to substantial differences in the trend estimates at the district level for MTS-II and MTS-III. Plots for some of the districts are provided in Figures 6.7 and 6.8. In Appendix S.2 plots for all districts are given. Similar to the trends of ANC0, the MTS-III model based trends of ANC4 are smoother than those based on the MTS-II model. The smaller variance components of MTS-III also result in narrower confidence bands compared to MTS-II.

The trend estimates under MTS-I are volatile and show unexpected downward trends for some districts, see for example *Bhola* and *Pirojpur* districts of Barisal division, *Gazipur*, *Kishoreganj* and *Manikganj* of *Dhaka* division, *Bogra*, *Chapai Nababganj* and *Rajshahi* districts of *Rajshahi* division, and *Habiganj* district of *Sylhet* division in Figure 6.7. From a subject matter point of view, such strong movements and turning points are not expected for ANC4 coverage. Therefore it appears that MTS-I follows the DIR estimates too strongly. The trends based on the MTS-II model generally ignore these volatilities and show reasonably smooth trends for these districts. The MTS-III model shows even smoother trends for some of these districts, as for example *Bogra* and *Habiganj* districts in Figure 6.7, and *Mymensingh*, and *Sylhet* districts in Figure 6.8.

The main difficulty arises for the three hilly districts of *Chittagong* division, i.e. *Khagrachhari*, *Rangamati*, and *Lakshmipur* (the first two districts are plotted in Figure 6.8). The MTS-I model shows very poor trend estimates for ANC4 over the whole period mainly due to the erratic DIR estimates, which are either zero or highly inconsistent in most of the surveys. The cross-sectional FH estimates are more robust and consequently the MTS-II and MTS-III models show reasonable upward trends for ANC4. It is expected that women residing in urbanized and better socioeconomic areas are supposed to receive more ANC visits compared to those residing in rural and poor socioeconomic areas. The MTS-I model shows in some districts lower and in other districts higher than expected trend estimates over the whole time period. For example, the trend obtained with MTS-I for *Narsingdi* in Figure 6.8, which is a highly urbanized district of *Dhaka* division, is lower than expected. Similarly the trend under MTS-I *Munshiganj* in Figure 6.8, which is a less urbanized district of *Dhaka* is higher than expected. Similarly the trend estimates under MTS-I are over the whole period higher than expected in *Meherpur* district of *Khulna* division, *Lalmonirhat* and *Panchagarh* districts of *Rangpur* division. See Figure 6.8. The trend estimates under MTS-II and MTS-III seem more plausible because the cross-sectional FH estimates appear to be more realistic than the DIR estimates.

Overall, as in the case of ANC0, MTS-II is a good compromise between MTS-I and MTS-III.

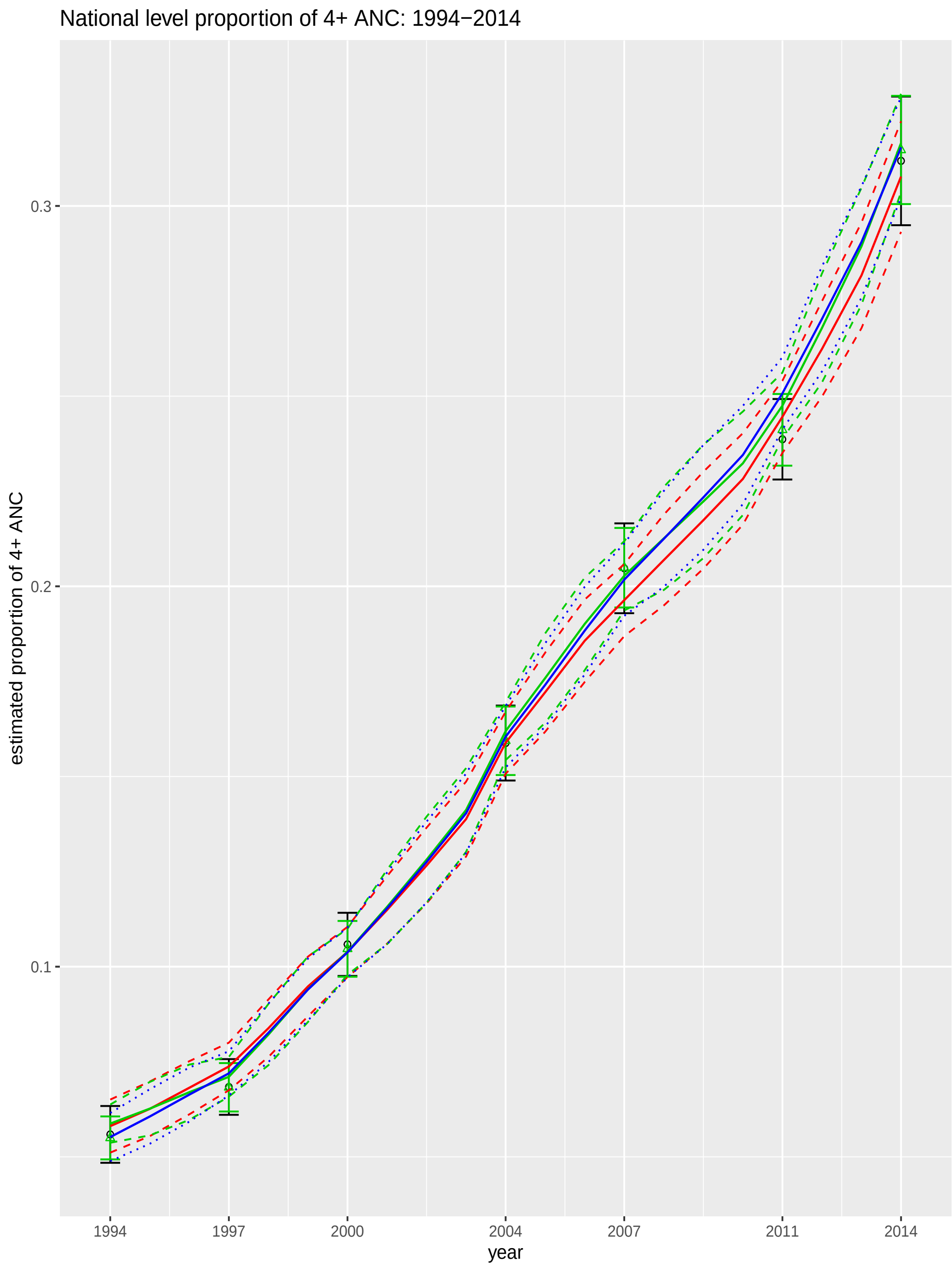

Figure 6.5 National level trends of ANC4 in Bangladesh: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

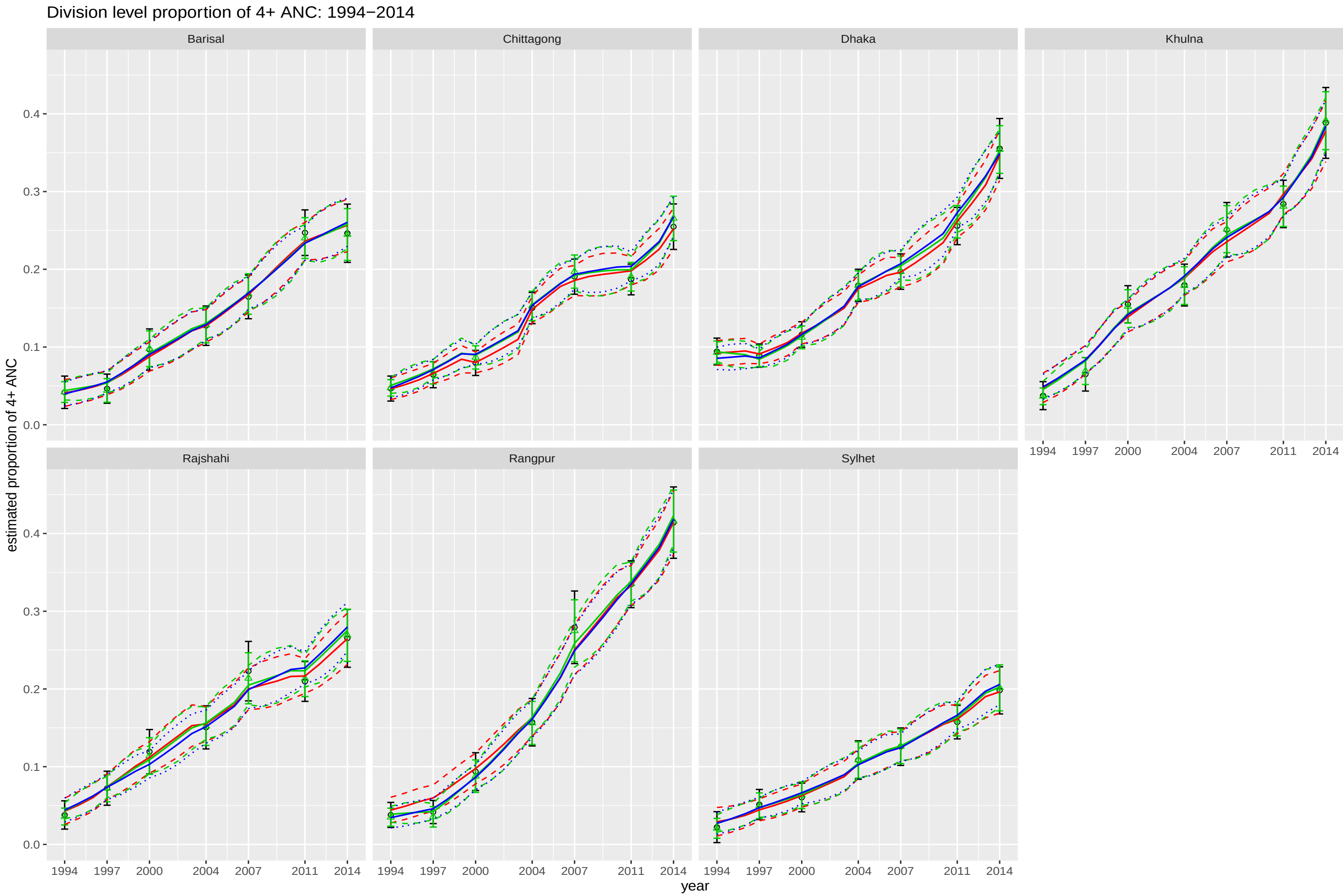

Figure 6.6 Division level trends of ANC4 in Bangladesh: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

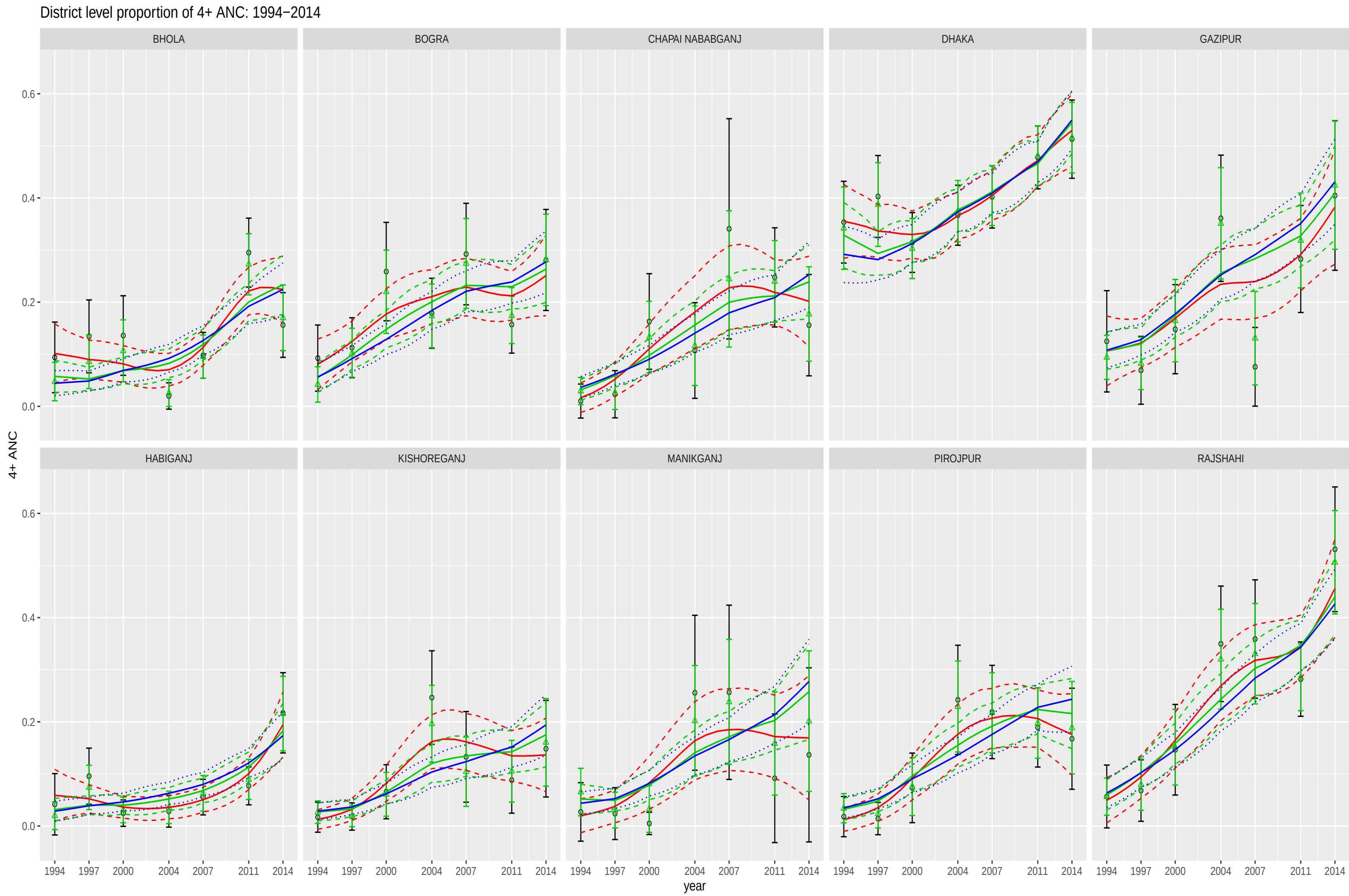

Figure 6.7 District level trends of ANC4 in Bangladesh: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

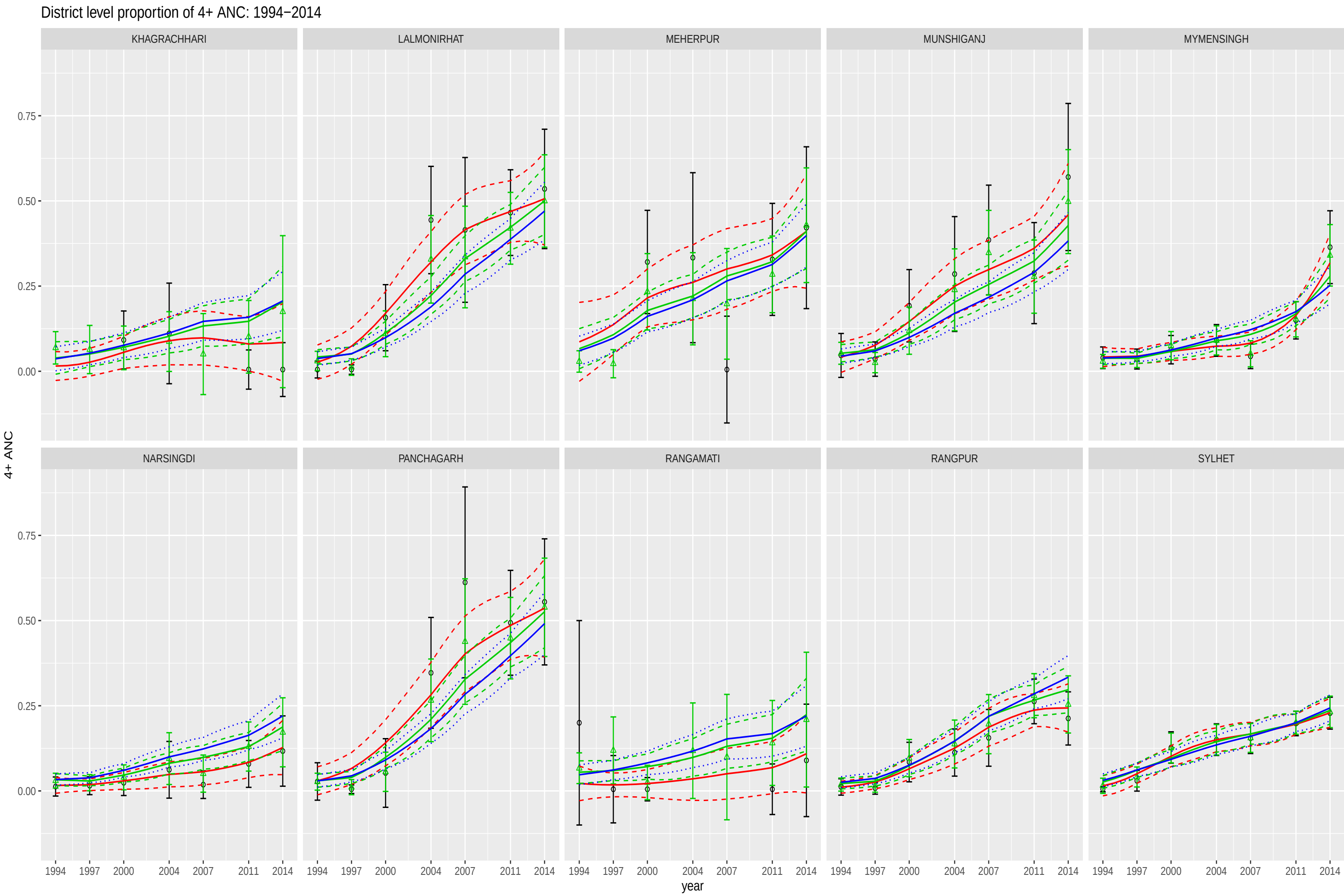

Figure 6.8 District level trends of ANC4 in Bangladesh: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

7 Model assessment

In this study, models were selected based on the WAIC, DIC and graphical comparisons of their trend predictions at three hierarchical levels.

In addition to these model diagnostics, three discrepancy measures are defined to evaluate and compare the time-series multilevel models. The first two measures are the Relative Bias (RB) and Absolute Relative Bias (ARB), which express the differences between model estimates and direct estimates, as percentage of the latter. For a given model, RB_{it} and ARB_{it} for domain i and (survey) year t are defined as

$$RB_{it} = \frac{(\hat{\theta}_{it} - \hat{Y}_{it})}{\hat{Y}_{it}} \times 100\% . \tag{13}$$

$$ARB_{it} = \frac{|\hat{\theta}_{it} - \hat{Y}_{it}|}{\hat{Y}_{it}} \times 100\% . \tag{14}$$

with $\hat{\theta}_{it}$ the model prediction and $\hat{Y}_{it}$ the direct estimate. The third discrepancy measure is the Relative Reduction of the Standard Errors (RRSE), which measures the percentage of reduction in standard error of the model-based estimates compared to the direct estimates, i.e.,

$$RRSE_{it} = 100\% \times (se(\hat{Y}_{it}) - se(\hat{\theta}_{it}))/se(\hat{Y}_{it}) . \tag{15}$$

These three discrepancy measures are calculated at national, division and district (i.e. most detailed) levels. The distributions of these measures are presented in terms of the minimum value, 1st quartile (Q_1), median, mean, 3rd quartile (Q_3) and maximum value.

Additionally, observed coverage rate (CR expressed in %) for 95% confidence interval of the considered cross-sectional FH and MTS models are calculated at division and district levels by identifying whether the estimated 95% confidence interval (CI) of $\hat{\theta}_{it}$ contains the direct estimates ($\hat{Y}_{it}$). Coverage at the district level is the percentage of district by year combinations (about 7×64 domains) where the direct estimate is included in the CI of $\hat{\theta}_{it}$. Coverage at the division level is the percentage of division by year combinations (7×7 domains) where the direct estimate is included in the CI of $\hat{\theta}_{it}$. Coverage rates are defined in a similar way for each survey year by averaging over all available districts in one particular survey year. Finally coverage is calculated for each division seperately by averaging over the 7 survey years.

The distributions of the RB_{it} (13), ARB_{it} (14) and $RRSE_{it}$ (15) for three administrative levels are provided in Tables 7.1, 7.2, and 7.3 for ANC0 and ANC4 for the cross-sectional FH, MTS-I, MTS-II, and MTS-III models. Table 7.1 shows that FH and MTS-I models provide lower mean RB for ANC0 and ANC4 at all three levels, while MTS-II provides slightly lower mean RB compared to MTS-III model at the district level. The ARB distributions in Table 7.2 show that the performance of MTS-II is in between MTS-I and MTS-III for all administrative levels except the national level for ANC4. The ARB values are the lowest for the cross-sectional FH model. It is also observed that the ARB increases as the domain sample size becomes smaller. Table 7.3 shows that MTS-II has the highest RRSE values at national and division levels, while at district level this model shows slightly lower RRSE than the MTS-III model for both ANC0 and ANC4. The variance reduction increases as the domain sample sizes become smaller. The reason that standard errors for the trends at national and division level under MTS-II are smaller than MTS-III is because under MTS-II the covariances between the cross-sectional FH predictions at the district level in the input series are ignored. These covariances are predominantly positive and therefore the standard errors of trends at aggregated levels are higher and more realistic under MTS-III. The higher RB, ARB and RRSE values for models MTS-II and MTS-III are a consequence of the more smooth trends obtained under both models. Small variances under smooth trends imply a larger amount of bias with respect to the direct estimates. As discussed in Section 6, these trends are more plausible compared to the cross-sectional FH model and MTS-I model, since from a subject matter point of view a smooth decreases for ANC0 and increase for ANC4 are expected.

This conclusion is confirmed by the CR values shown in Table 7.4. The CRs for the cross-sectional FH models are too high, indicating that the FH predictions tend too much to the direct estimates. The CR levels are reasonably good for MTS-I, substantially lower for MTS-II and the lowest for MTS-III. The lower coverage rates of MTS-II and MTS-III at the district level is reflected by the corresponding higher ARB and higher RRSE. These findings show that MTS-I model predictions are more volatile and tend to the direct estimates, MTS-III model predictions are highly smoothed, and MTS-II model predictions seem like a reasonable compromise between MTS-I and MTS-III model predictions, particularly at the district level.

Parameter	Aggregation Level	Model	Min.	Q_1	Median	Mean	Q_3	Max.
ANC0	Nation	FH	-0.48	-0.20	0.23	0.08	0.37	0.47
		MTS-I	-1.84	-0.68	0.54	-0.05	0.73	0.88
		MTS-II	-1.16	-0.55	0.29	0.41	1.19	2.43
		MTS-III	-1.53	-0.80	-0.57	-0.02	0.60	2.35
	Division	FH	-0.68	-0.48	-0.36	0.05	0.50	1.31
		MTS-I	-0.99	-0.50	-0.31	0.05	0.64	1.41
		MTS-II	-0.77	0.04	0.15	0.59	1.08	2.50
		MTS-III	-1.44	-0.37	0.13	0.15	0.89	1.35
	District	FH	-8.77	-1.72	0.14	0.31	1.67	12.41
		MTS-I	-10.35	-1.24	-0.49	-0.66	0.30	1.87
		MTS-II	-7.87	-1.15	0.77	1.25	2.89	18.34
		MTS-III	-10.05	-2.63	0.89	1.34	3.91	21.43
ANC4	Nation	FH	-1.65	-0.62	0.07	-0.07	0.65	1.04
		MTS-I	-4.09	-1.60	0.05	1.00	3.19	7.88
		MTS-II	-1.85	0.27	1.98	1.91	3.80	5.07
		MTS-III	-2.00	-1.35	1.06	1.11	3.10	5.23
	Division	FH	-1.33	-0.60	-0.13	0.24	0.43	3.47
		MTS-I	-1.17	-0.25	-0.04	-0.07	0.32	0.59
		MTS-II	-0.50	0.68	1.18	1.55	1.70	5.39
		MTS-III	-2.08	0.31	0.73	1.24	1.92	5.58
	District	FH	-17.83	-4.85	0.40	2.08	6.78	64.77
		MTS-I	-16.32	-3.80	-0.56	-0.42	2.98	15.57
		MTS-II	-22.00	-5.30	0.57	4.57	12.47	84.31
		MTS-III	-29.92	-8.23	0.57	6.12	14.07	124.63

Table 7.1 Summary statistics of relative bias (RB, in %) at different aggregation levels for the SAE estimates of ANC0 and ANC4

Parameter	Aggregation Level	Model	Min.	Q_1	Median	Mean	Q_3	Max.
ANC0	Nation	FH	0.04	0.27	0.42	0.34	0.46	0.48
		MTS-I	0.26	0.63	0.75	0.87	0.99	1.84
		MTS-II	0.29	0.44	0.58	1.05	1.59	2.43
		MTS-III	0.49	0.61	0.96	1.18	1.61	2.35
	Division	FH	0.39	0.50	0.65	0.90	1.20	1.84
		MTS-I	0.48	0.66	0.78	1.39	2.13	2.90
		MTS-II	0.79	0.96	1.56	1.78	2.14	3.88
		MTS-III	1.00	1.14	1.41	1.88	2.43	3.61
	District	FH	1.08	2.73	4.17	5.12	5.84	15.94
		MTS-I	1.48	3.93	6.58	7.53	9.02	26.67
		MTS-II	3.15	6.46	10.31	11.32	14.50	33.01
		MTS-III	4.15	8.65	12.54	13.49	16.98	38.16
ANC4	Nation	FH	0.07	0.25	0.92	0.76	1.08	1.65
		MTS-I	0.05	1.60	2.47	3.09	4.00	7.88
		MTS-II	0.97	1.68	1.98	2.71	3.80	5.07
		MTS-III	1.06	1.19	1.46	2.46	3.53	5.23
	Division	FH	0.98	1.40	1.71	1.87	2.06	3.47
		MTS-I	1.96	3.06	4.31	4.07	4.64	6.82
		MTS-II	2.18	3.66	4.68	4.33	5.07	6.00
		MTS-III	3.66	4.60	5.36	5.27	5.63	7.46
	District	FH	1.93	7.64	12.91	14.29	17.60	64.77
		MTS-I	3.86	14.27	18.72	20.61	28.10	53.45
		MTS-II	7.07	19.47	26.22	28.51	35.88	84.31
		MTS-III	8.62	21.36	29.32	33.13	41.00	124.63

Table 7.2 Summary statistics of absolute relative bias (ARB, in %) at different aggregation levels for the SAE estimates of ANC0 and ANC4

Parameter	Aggregation Level	Model	Min.	Q_1	Median	Mean	Q_3	Max.
ANC0	Nation	FH	-0.65	4.03	8.10	8.00	12.72	15.01
		MTS_I	-0.03	1.35	4.01	3.67	5.82	7.33
		MTS_II	4.07	7.90	13.71	12.89	17.47	21.68
		MTS_III	-3.52	1.02	3.30	3.69	7.15	9.67
	Division	FH	2.99	5.82	7.56	7.03	8.64	9.75
		MTS-I	2.66	4.00	5.32	5.16	6.65	6.84
		MTS-II	8.47	12.74	13.70	13.12	14.34	15.53
		MTS-III	3.30	4.71	5.21	5.47	6.12	8.16
	District	FH	-1.60	7.17	10.20	9.98	12.04	21.61
		MTS-I	7.91	15.16	17.81	18.06	21.15	27.47
		MTS-II	12.60	27.84	34.08	33.80	38.46	48.53
		MTS-III	19.48	32.61	38.40	37.79	41.55	52.71
ANC4	Nation	FH	8.58	11.22	11.66	13.71	14.49	24.32
		MTS_I	6.64	12.16	14.60	14.75	18.50	20.66
		MTS_II	17.79	22.87	23.56	25.12	27.99	32.75
		MTS_III	10.33	16.58	19.45	18.15	21.04	22.04
	Division	FH	11.08	11.80	14.07	14.23	16.39	18.08
		MTS-I	11.82	14.31	14.46	15.78	18.18	19.17
		MTS-II	20.32	24.96	27.39	26.34	28.15	30.45
		MTS-III	15.49	20.37	21.75	21.72	24.51	25.05
	District	FH	0.34	11.62	16.77	17.63	22.60	38.62
		MTS-I	17.79	27.84	30.48	30.93	33.65	43.40
		MTS-II	29.58	43.37	46.86	48.10	54.96	66.75
		MTS-III	35.63	48.88	51.75	52.94	59.31	70.35

Table 7.3 Summary statistics of relative reduction of standard errors (RRSE in %) at different aggregation levels for the SAE estimates of ANC0 and ANC4

Parameter	Model	Year wise CR at District Level							Overall CR by Level	
		1994	1997	2000	2004	2007	2011	2014	District	Division
ANC0	FH	100.00	98.33	100.00	100.00	100.00	98.36	100.00	99.53	100.00
	MTS-I	100.00	90.00	93.44	88.52	93.22	98.36	100.00	94.81	100.00
	MTS-II	88.33	63.33	70.49	67.21	71.19	75.41	91.53	75.10	95.92
	MTS-III	83.33	53.33	50.82	52.46	61.02	55.74	79.66	62.22	95.92
ANC4	FH	98.15	98.28	100.00	100.00	100.00	100.00	90.20	98.36	100.00
	MTS-I	87.04	84.48	68.33	76.27	81.97	96.72	100.00	84.58	95.92
	MTS-II	44.44	51.72	50.00	52.54	62.30	65.57	76.47	57.55	97.96
	MTS-III	44.44	41.38	40.00	38.98	50.82	55.74	72.55	48.70	97.96

Table 7.4 Observed coverage rate (CR in %) of the model predictions for 95% confidence interval at district and division levels as well as district level by survey years for the SAE estimates of ANC0 and ANC4

8 Discussion

In this study, multilevel time-series models have been developed for two response variables related to antenatal care coverage in Bangladesh using only seven editions of the Bangladesh Demographic and Health Survey (BDHS) over the period of 1994-2014. Time gaps between two consecutive survey years vary between 3 and 4 years. Time series models are defined at an annual frequency where years without a survey edition are treated as missing. In this way, the model accounts for the varying time gaps between the subsequent editions of the BDHS. An additional advantage of modelling the observed time series in this way is that the time series models produce predictions in the years without sample surveys. Estimates for the percentage of women receiving no antenatal consult (ANC0) and the percentage of women receiving at least 4 antenatal consults (ANC4) are produced at three regional levels, namely the national level, a break down in 7 divisions and a breakdown in 64 districts.

In a first approach, year-domain-specific direct estimates and their standard errors are used as input in the MTS model, after applying a variance function to smooth the standard errors. Trends obtained under this model, denoted MTS-I, are rather volatile since the trend estimates tend to follow the direct estimates. Another drawback of the MTS-I model is that the use of auxiliary information from two available censuses results in less realistic evolutions of the trend estimates, since the same values for the auxiliary information from a particular census are used in two or three subsequent editions of the survey. Therefore in a second approach, cross-sectional Fay-Herriot (FH) models are developed for the available survey year data where the available census data are considered as auxiliary variables. These FH estimates and their standard errors are used as input in the MTS model instead of the direct estimates. This results in more accurate input series for the time series model and the use of the auxiliary information from the census in the cross-sectional FH models doesn't compromise the evolution of the trends. For this second approach, two types of MTS models are developed considering either a diagonal or a full covariance matrix for the cross-sectional FH estimates. These models are referred to as MTS-II and MTS-III, respectively.

The models are developed at the most detailed regional level of districts. Division and national level trends are estimated by aggregating predictions of the district level trends. In this way, figures at different aggregation levels are numerically consistent by definition.

The observed time series are very short (seven years only). It was therefore important to develop parsimonious time series models. This is achieved by defining trends at the division level that are shared by all underlying districts. Deviations from the overall trends of the individual districts are modelled as district specific trends. This avoids specifying full covariance matrices for the 64 trends at the most detailed level of districts. Specifying trend components at the level of the divisions is also necessary to obtain accurate aggregated predictions for the divisions and the national level. Further model regularization was considered by specifying global-local priors. This, however, did not further improve the model fits. Therefore standard normal priors are used in this application.

In small area estimation, it is often considered to benchmark the domain estimates to the direct estimates at the national level as an attempt to reduce the bias in the model based domain predictions. In this application the trend estimates at the national level under the MTS models are already very close to the direct estimates, see Figures 6.1 and 6.5. Note that good fits at higher aggregation levels are obtained by adding trend components at the division level in the MTS models. An additional benchmark is therefore not considered in this application.

To summarize, the main contribution of this study is to produce more reliable and consistent estimates at different regional levels by proposing a time series modelling technique that accounts for the issues of varying survey time gaps, survey and census time gaps, and domain-specific small sample sizes. In addition, the study contributes to the SAE literature by showing how the MTS model using FH estimates as input (instead of direct estimates) helps to obtain reliable trend estimates. All the developed time series models provide estimates with better accuracy. The MTS-II model, however, seems to provide most plausible trends for both response variables, particularly at the district level, by compromising volatility in the trends under the MTS-I model and flatness in the trends under the MTS-III model. This choice is supported by the fact that these variables are likely to be relatively smooth over time. Fitting these models to the series of ANC0 and ANC4 is therefore certainly suitable in concept. This also justifies the interpolation of the trends for the years without sample surveys.

Because MTS-II ignores the predominantly positive correlations between the cross-sectional FH input series, the standard errors of the trends at aggregated levels are actually too small. Since MTS-III accounts for these correlations, the standard errors for national and division trends are larger but also more realistic.

One limitation of this study is related to the bias correction for the square root transformation that is applied to ANC4. The bias correction can only be applied to the trend estimates in the survey years. This results in awkward increases of point estimates if the sampling error is not smoothed enough, particularly for the domains with small sample size. This hampers estimation of period-to-period changes between survey years and non-survey years. Therefore the bias correction is only applied to the cross-sectional FH models and not to the MTS models.

For ANC0, the national level shows a downward trend. The decline in the trend temporarily stopped during 2004-2011. The trend of ANC4 shows steady increase over the considered study period. Division level trends for ANC0 show a steady decline for all the divisions except *Dhaka*, *Chittagong* and *Sylhet* divisions. The trends for these three divisions remained stable during the period of 2004-2011 which mainly causes the flat trend at the national level of ANC0. On the other hand, at the division level ANC4 shows almost linear upward trends for most of the divisions except *Dhaka* and *Chittagong*. The greatest improvement is observed for *Khulna* and *Rangpur* divisions where the trends of ANC4 reach to more than 40% in 2014. District-level trends help to identify highly vulnerable districts in terms of the two considered response variables. Though the national level trend of ANC0 declines to about 21% in 2014, a few districts get below 10% (*Dhaka*, *Jhenaidaha*, and *Meherpur*) while a considerable number of districts still have ANC0 higher than 35% (*Bhola*, *cox's Bazar*, *Kishoregonj*, *Noakhali*, *Sunamganj*, *Sirajgonj*, and three *Chittagong* hill tract districts). For ANC4, a few districts have estimates above 50% (*Dhaka*, *Nilphamari*, and *Panchagarh*) and most of the districts with high ANC0 have ANC4 estimates less than 20%. These district level trends might help policy makers to focus on vulnerable hotspots where both ANC0 and ANC4 indicators are still poor. Obviously, detailed level trends might help policy makers to take actions for reducing disaggregated level inequalities in the race of achieving SDG goals.

Acknowledgement

We wish to thank Measure Evaluation and National Institute of Population Research and Training (NIPORT) for making the BDHS data publicly available. In addition, IPUMS deserves thanks for providing the access to the sample data of Bangaldesh Census 1991, Census 2001, and Census 2011. The views expressed in this paper are those of the authors and do not necessarily reflect the policy of Statistics Netherlands.

References

BBS (2020). SDG Tracker: Bangladesh Development Mirror. https://www.sdg.gov.bd/page/thirty_nine_plus_one_indicator/5#1. [Online; accessed 18-December-2020].

BBS and UNICEF (2019). Progotir Pathey Bangladesh: Bangladesh Multiple Indicator Cluster Survey 2019. Technical report, Bangladesh Bereau of Statistics (BBS).

Besag, J. and C. Kooperberg (1995). On conditional and intrinsic autoregression. *Biometrika 82*(4), 733–746.

Boonstra, H. J. (2021). *mcmcsae: Markov Chain Monte Carlo Small Area Estimation*. R package version 0.6.0.

Boonstra, H. J. and J. van den Brakel (2018). Hierarchical bayesian time series multilevel models for consistent small area estimates at different frequencies and regional levels. Statistics Netherlands discussion paper, December 4, 2018.

Boonstra, H. J. and J. van den Brakel (2019). Estimation of level and change for unemployment using multilevel and structural time-series models. *Survey Methodology 45*(3), 395–425.

Carvalho, C. M., N. G. Polson, and J. G. Scott (2010). The horseshoe estimator for sparse signals. *Biometrika 97*(2), 465–480.

Das, S., B. Kumar, and L. A. Kawsar (2020). Disaggregated level child morbidity in Bangladesh: An application of small area estimation method. *PLoS ONE 15*(5), e0220164.

Datta, G., P. Lahiri, T. Maiti, and K. Lu (1999). Hierarchical bayes estimation of unemployment rates for the states of the u.s. *Journal of the American Statistical Association 94*(448), 1074–1082.

Fay, R. and R. Herriot (1979). Estimates of income for small places: An application of james-stein procedures to census data. *Journal of the American Statistical Association 74*(366), 269–277.

Gelfand, A. and A. Smith (1990). Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association 85*, 398–409.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis 1*(3), 515–533.

Gelman, A. and J. Hill (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gelman, A. and D. Rubin (1992). Inference from iterative simulation using multiple sequences. *Statistical Science 7*(4), 457–472.

Geman, S. and D. Geman (1984). Stochastic relaxation, gibbs distributions and the bayesian restoration of images. *IEEE Trans. Pattn Anal. Mach. Intell. 6*, 721–741.

Haslett, S. and G. Jones (2004). Local Estimation of Poverty and Malnutrition in Bangladesh. Technical report, Bangladesh Bureau of Statistics and United Nations World Food Programme.

Haslett, S., G. Jones, and M. Isidro (2014). Local Estimation of Poverty and Malnutrition in Bangladesh. Technical report, United Nations World Food Programme and International Fund for Agricultural Development.

Hossain, M. J., S. Das, and H. Chandra (2020). Disaggregate level estimates and spatial mapping of food insecurity in Bangladesh by linking survey and census data. *PLoS ONE 15*(4), e0230906.

Mrisho, M., B. Obrist, J. A. Schellenberg, R. A. Haws, A. K. Mushi, H. Mshinda, M. Tanner, and D. Schellenberg (2009). The use of antenatal and postnatal care: perspectives and experiences of women and health care providers in rural southern Tanzania. *BMC pregnancy and childbirth 9*(1), 10.

NIPORT (2015a). Bangladesh Demographic and Health Survey 2014: Key Indicators. Technical report, National Institute of Population Research and Training.

NIPORT (2015b). Bangladesh Demographic and Health Survey 2014: Policy Briefs. Technical report, National Institute of Population Research and Training.

NIPORT (2016). Bangladesh Demographic and Health Survey 2014. Technical report, National Institute of Population Research and Training.

O'Malley, A. and A. Zaslavsky (2008). Domain-level covariance analysis for multilevel survey data with structured nonresponse. *Journal of the American Statistical Association 103*(484), 1405–1418.

Park, T. and G. Casella (2008). The bayesian lasso. *Journal of the American Statistical Association 103*(482), 681–686.

Pfeffermann, D. and L. Burck (1990). Robust small area estimation combining time series and cross-sectional data. *Survey Methodology 16*, 217–237.

Pfeffermann, D. and R. Tiller (2006). Small area estimation with state-space models subject to benchmark constraints. *Journal of the American Statistical Association 101*, 1387–1397.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rahman, M. S., H. Mohiuddin, A.-A. Kafy, P. K. Sheel, and L. Di (2019). Classification of cities in Bangladesh based on remote sensing derived spatial characteristics. *Journal of Urban Management 8*(2), 206–224.

Rao, J. and I. Molina (2015). *Small Area Estimation*. Wiley-Interscience.

Rao, J. and M. Yu (1994). Small area estimation by combining time series and cross-sectional data. *The Canadian Journal of Statistics 22*, 511–528.

Rue, H. and L. Held (2005). *Gaussian Markov Random Fields: Theory and Applications*, Volume 56. Chapman and Hall/CRC.

Sakia, R. (1992). The box-cox transformation technique: a review. *The Statistician 41*, 169–178.

Spiegelhalter, D., N. Best, B. Carlin, and A. van der Linde (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B 64*(4), 583–639.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.

Watanabe, S. (2010). Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research 11*, 3571–3594.

Watanabe, S. (2013). A widely applicable bayesian information criterion. *Journal of Machine Learning Research 14*, 867–897.

West, M. (1984). Outlier models and prior distributions in bayesian linear regression. *Journal of the Royal Statistical Society. Series B (Methodological)*, 431–439.

WHO, UNICEF, and Others (2014). Trends in maternal mortality: 1990 to 2013: estimates by who, unicef, unfpa, the world bank and the united nations population division. Technical report, World Health Organization.

Wolter, K. (2007). *Introduction to Variance Estimation*. Springer.

You, Y. (2008). An integrated modeling approach to unemployment rate estimation for sub-provincial areas of canada. *Survey Methodology 34*(1), 19–27.

You, Y., J. Rao, and J. Gambino (2003). Model-based unemployment rate estimation for the canadian labour force survey: A hierarchical bayes approach. *Survey Methodology 29*(1), 25–32.

Appendix
Supplementary Materials

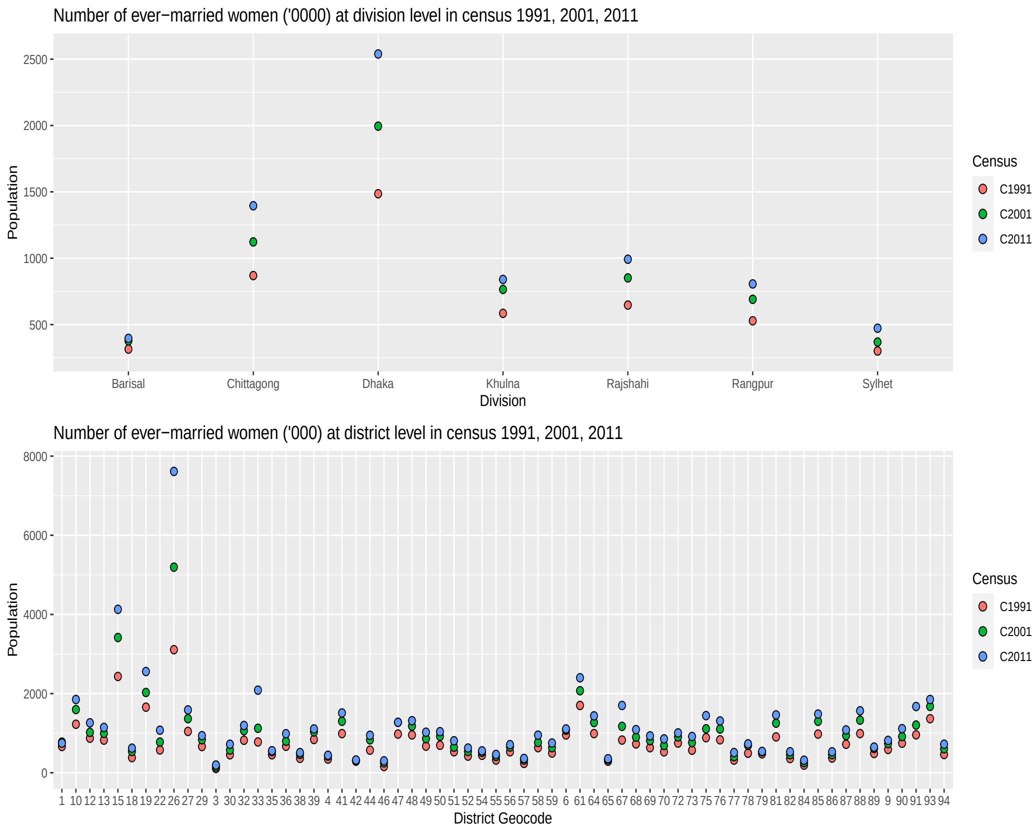

Figure S.1 Population size of ever-married women at division and district level for the three census years

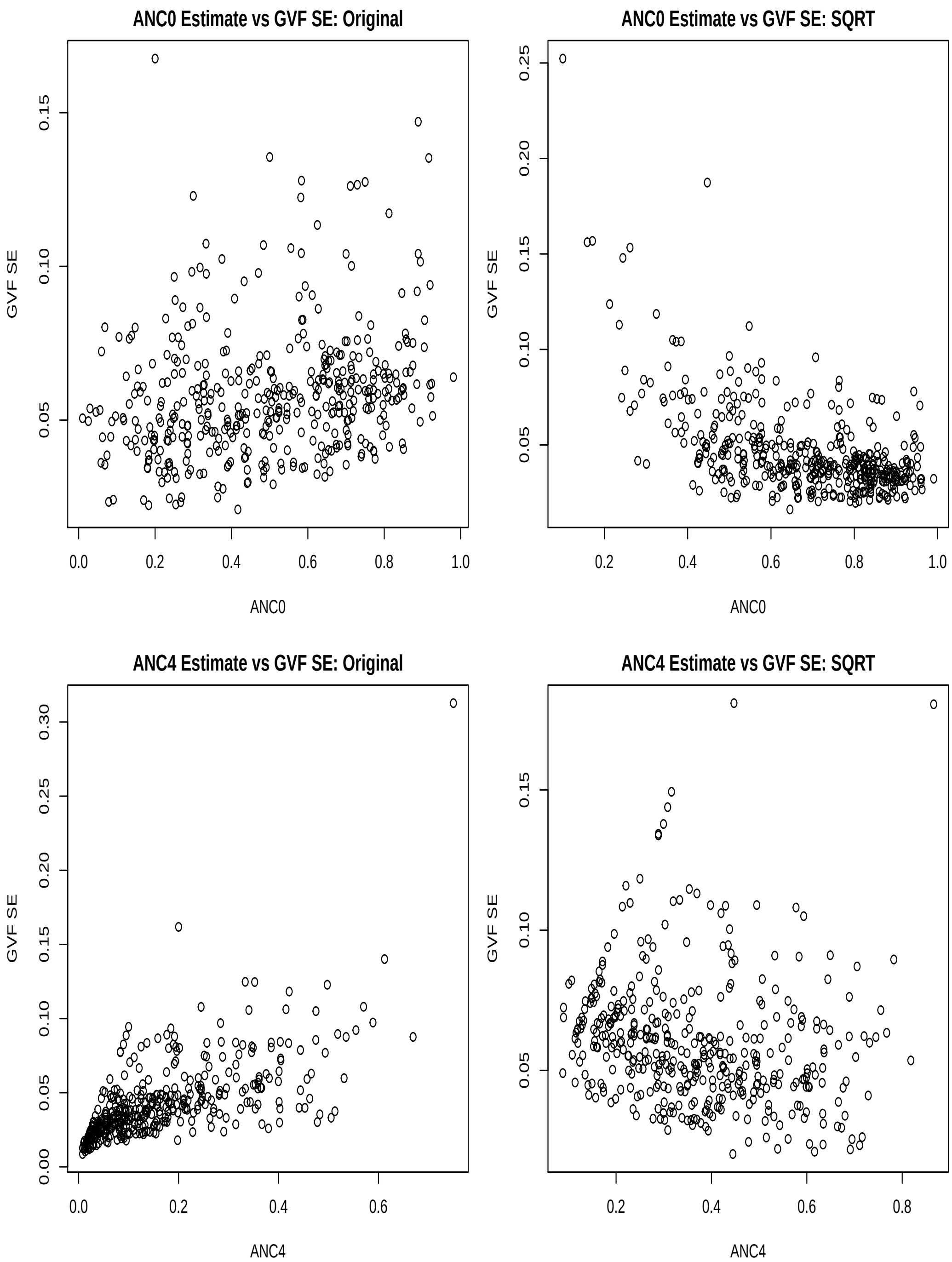

Figure S.2 Relation between point estimates and (smoothed) standard erros for ANC0 and ANC4 on the original and square root (SQRT) transformations. Each point corresponds to a direct estimate and its (smoothed) standard error for a certain domain and year. Note that the sample sizes are very diverse among the districts.

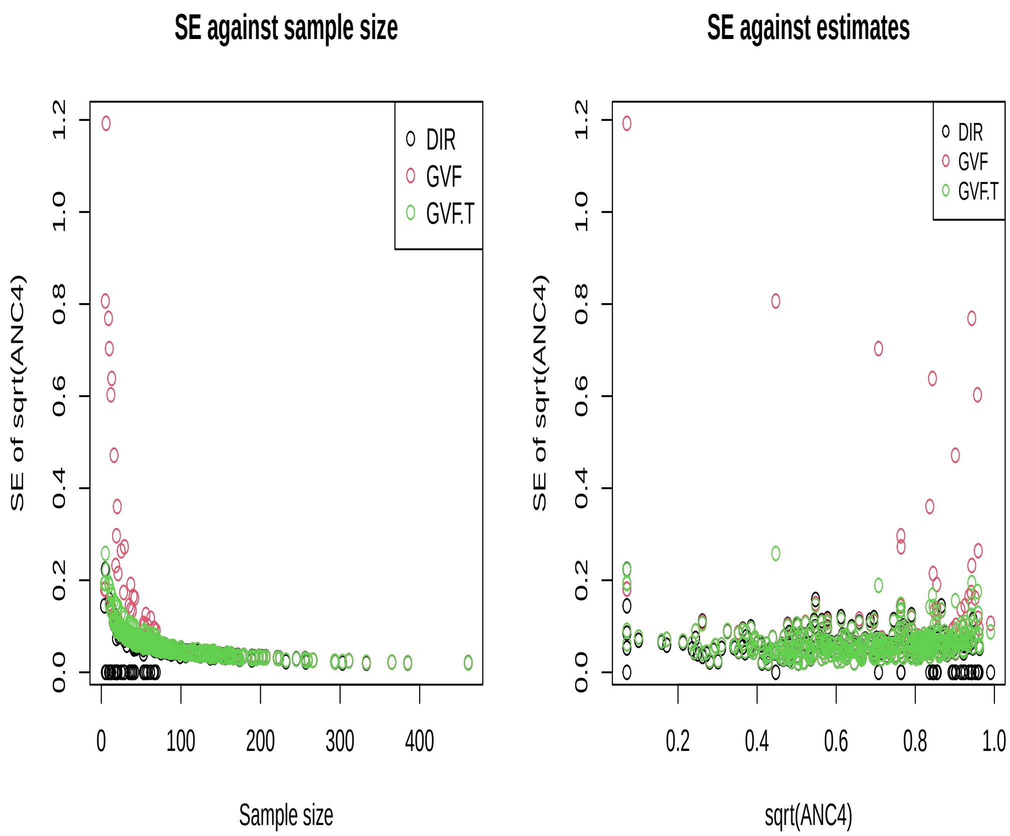

Figure S.3 Standard errros of square-root transformed ANC4 (i.e. sqrt(ANC4)) calculated based on direct estimates (DIR), GVF model using original scale (GVF) and GVF model using transformed scale (GVF.T) standard errors. Each point corresponds to a standard error for a certain domain and year.

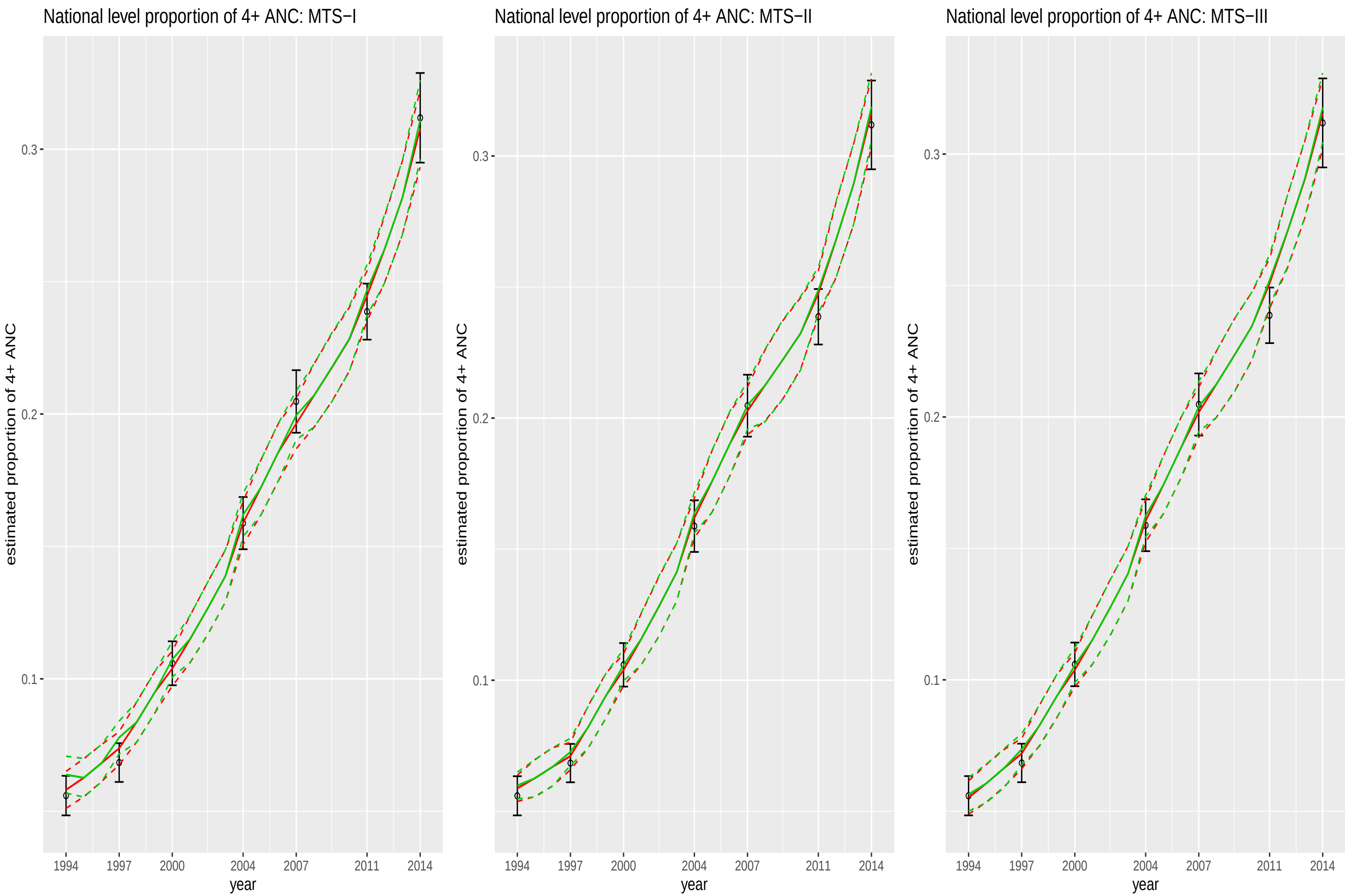

Figure S.4 National level trends of ANC4 in Bangladesh: (i) DIR (black error-bar line), (ii) MTS (red line) without bias correction, and (iii) MTS (green line) with bias correction.

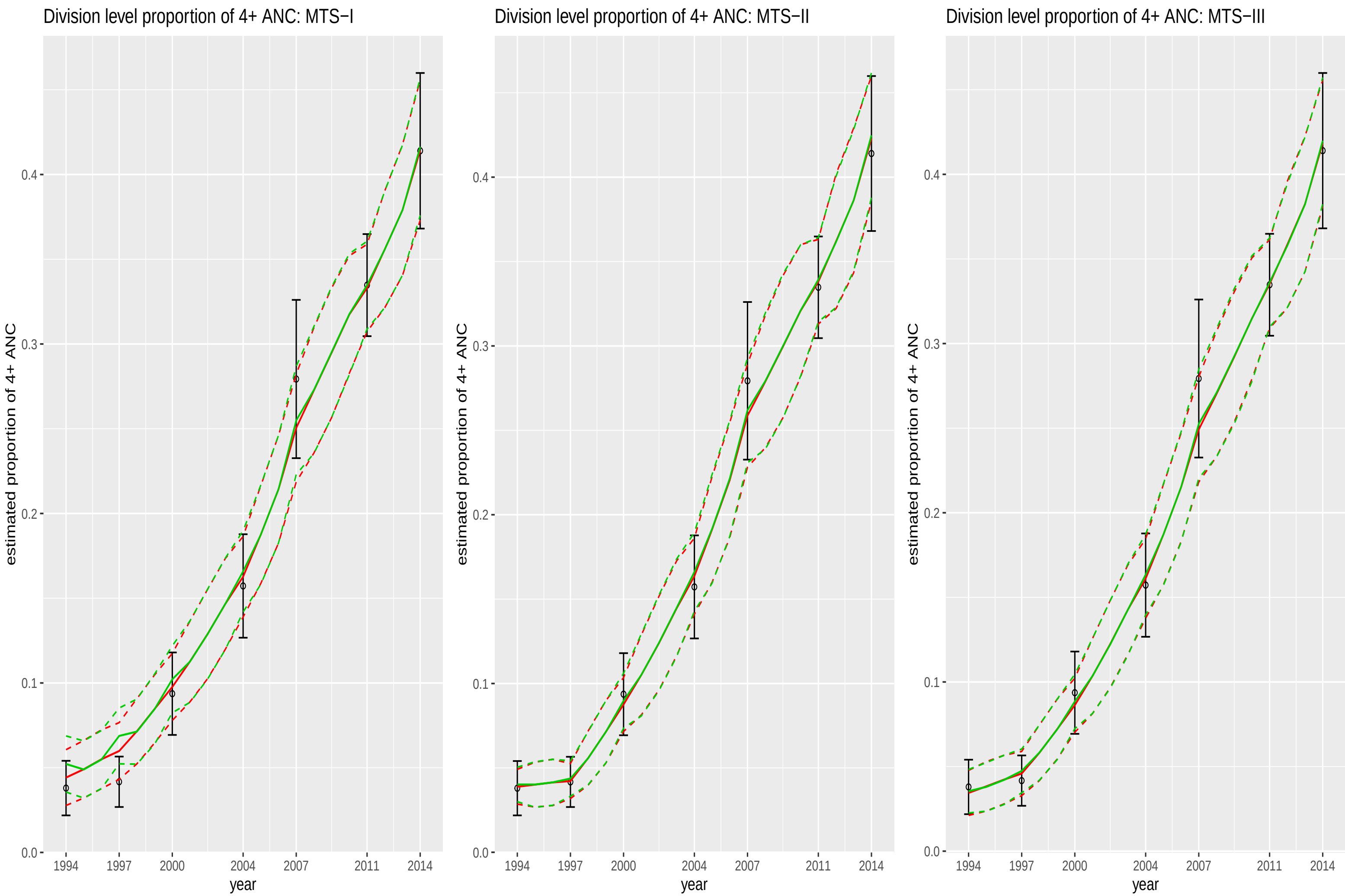

Figure S.5 Division level trends of ANC4 in *Rangpur* division: (i) DIR (black error-bar line), (ii) MTS (red line) without bias correction, and (iii) MTS (green line) with bias correction.

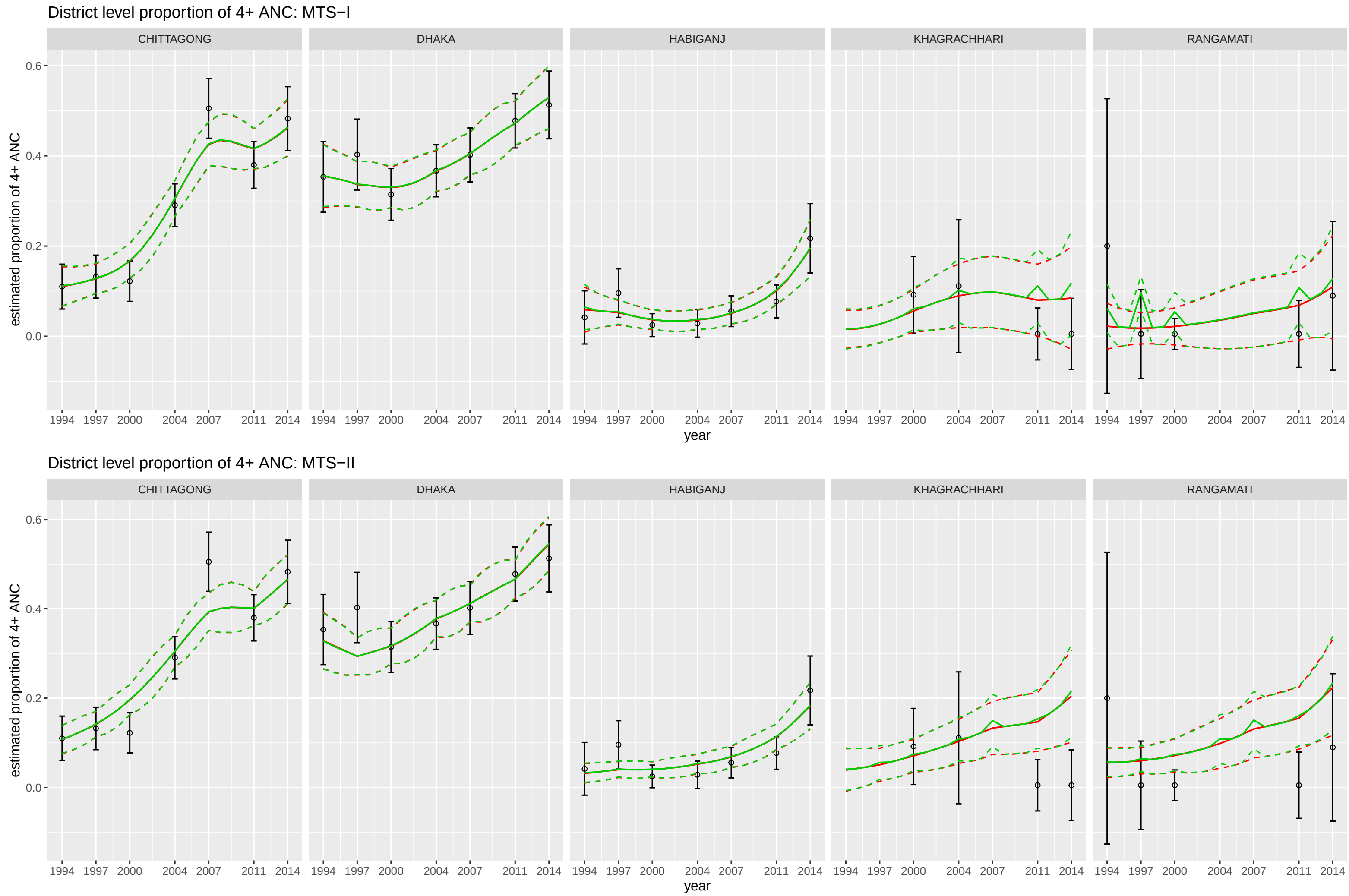

Figure S.6 District level trends of ANC4: (i) DIR (black error-bar line), (ii) MTS (red line) without bias correction, and (iii) MTS (green line) with bias correction.

S.1 Trends of ANC0 in Bangladesh districts

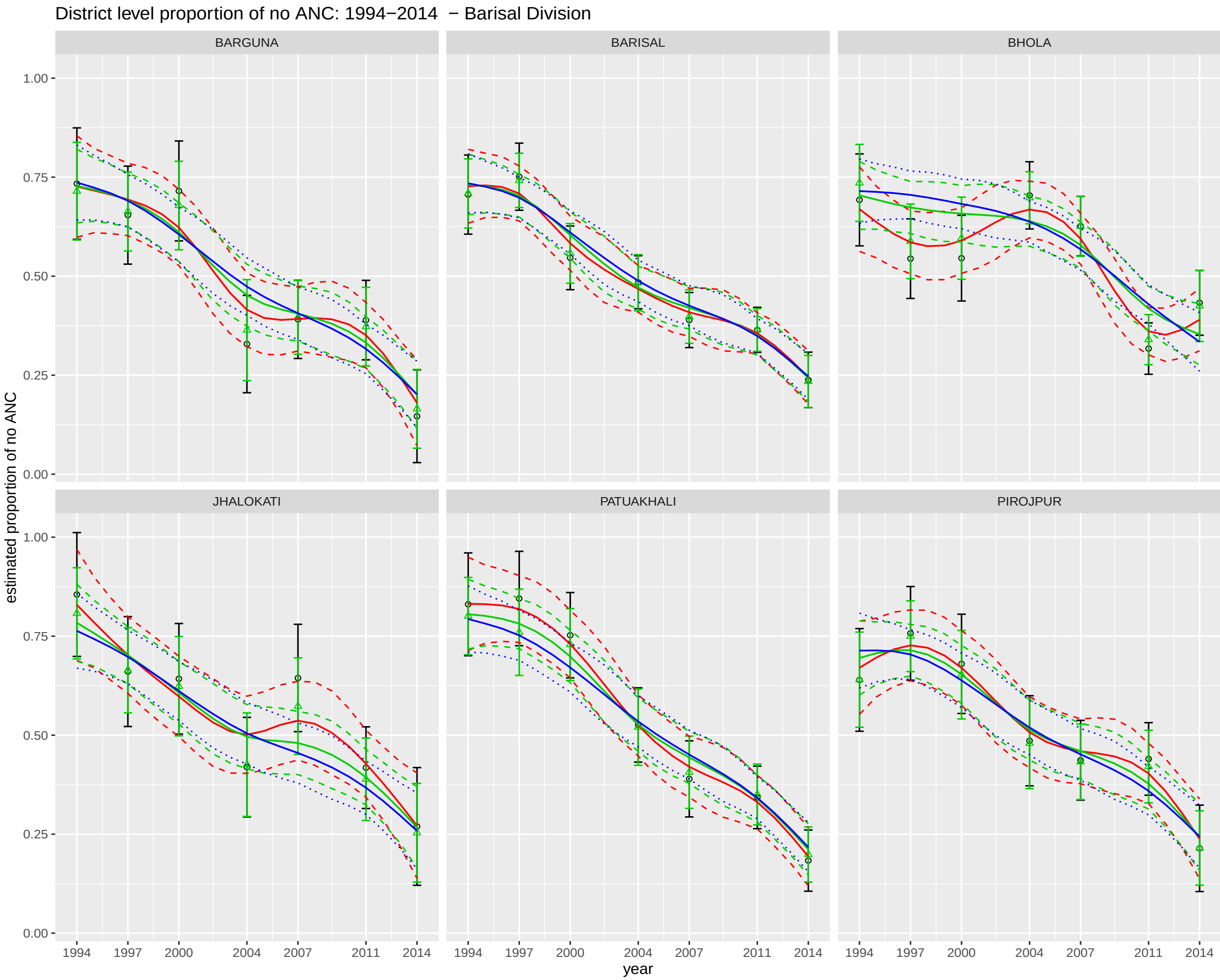

Figure S.7 District level trends of ANC0 in Bangladesh - *Barisal* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

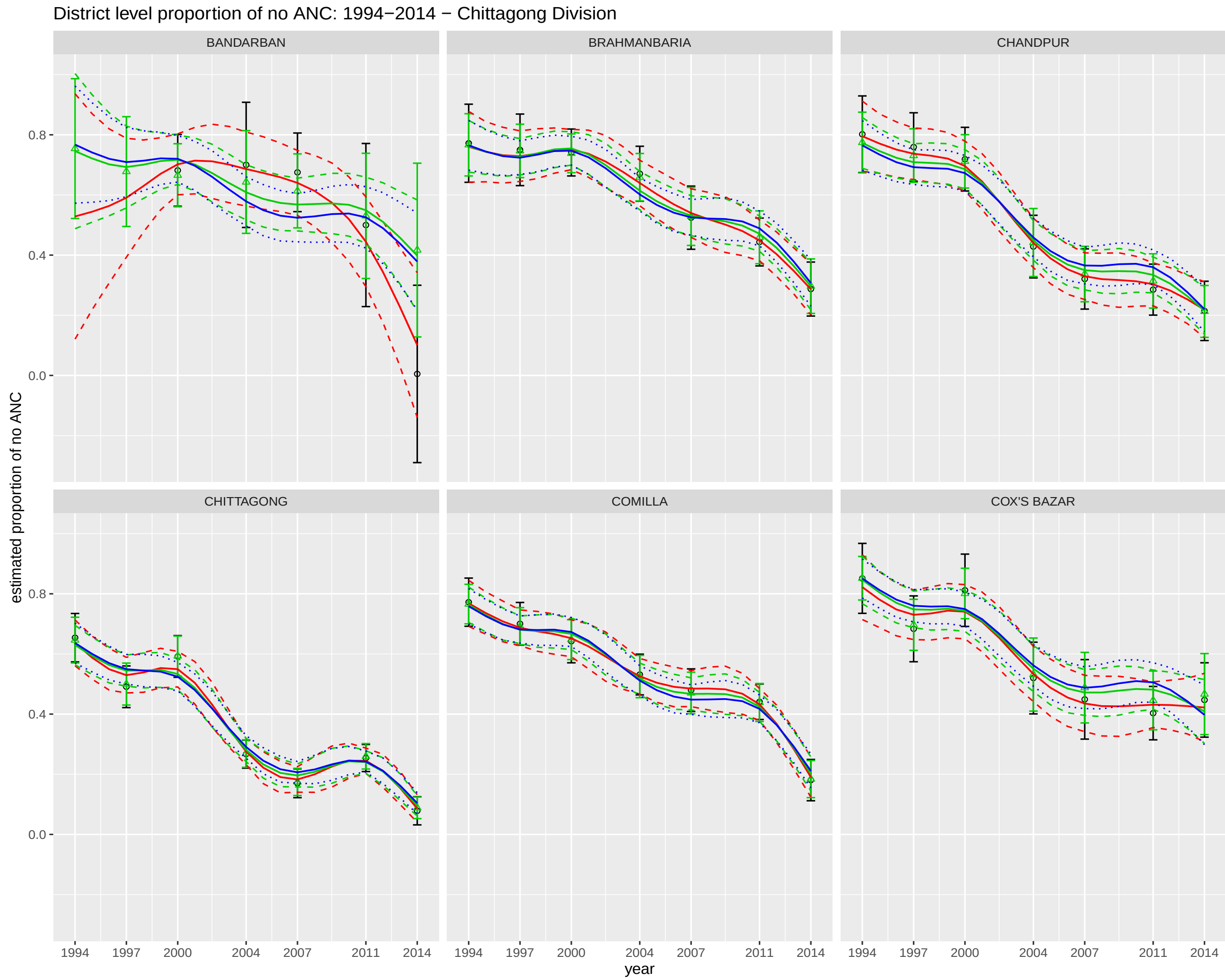

Figure S.8 District level trends of ANC0 in Bangladesh - *Chittagong* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

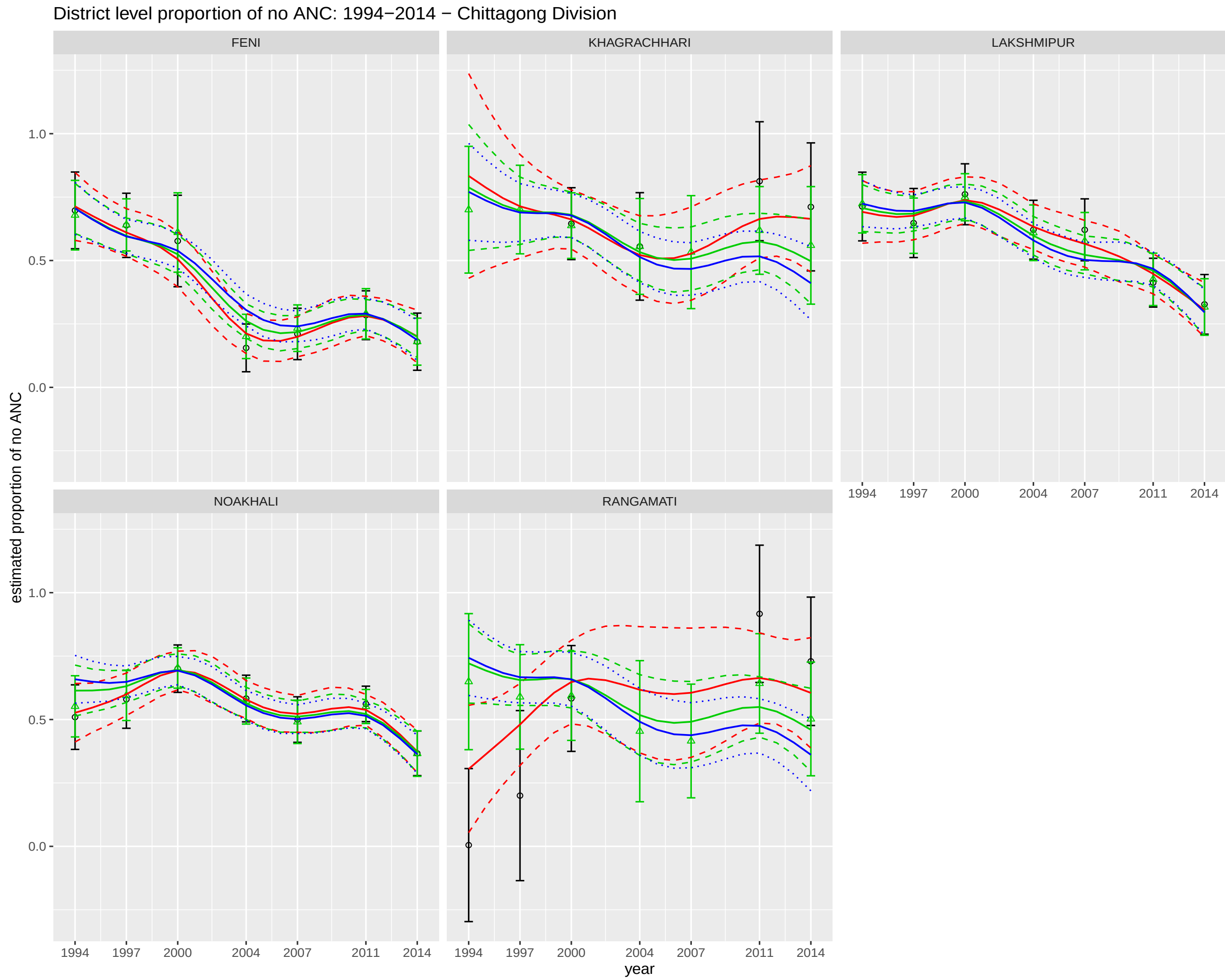

Figure S.9 District level trends of ANC0 in Bangladesh - *Chittagong* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

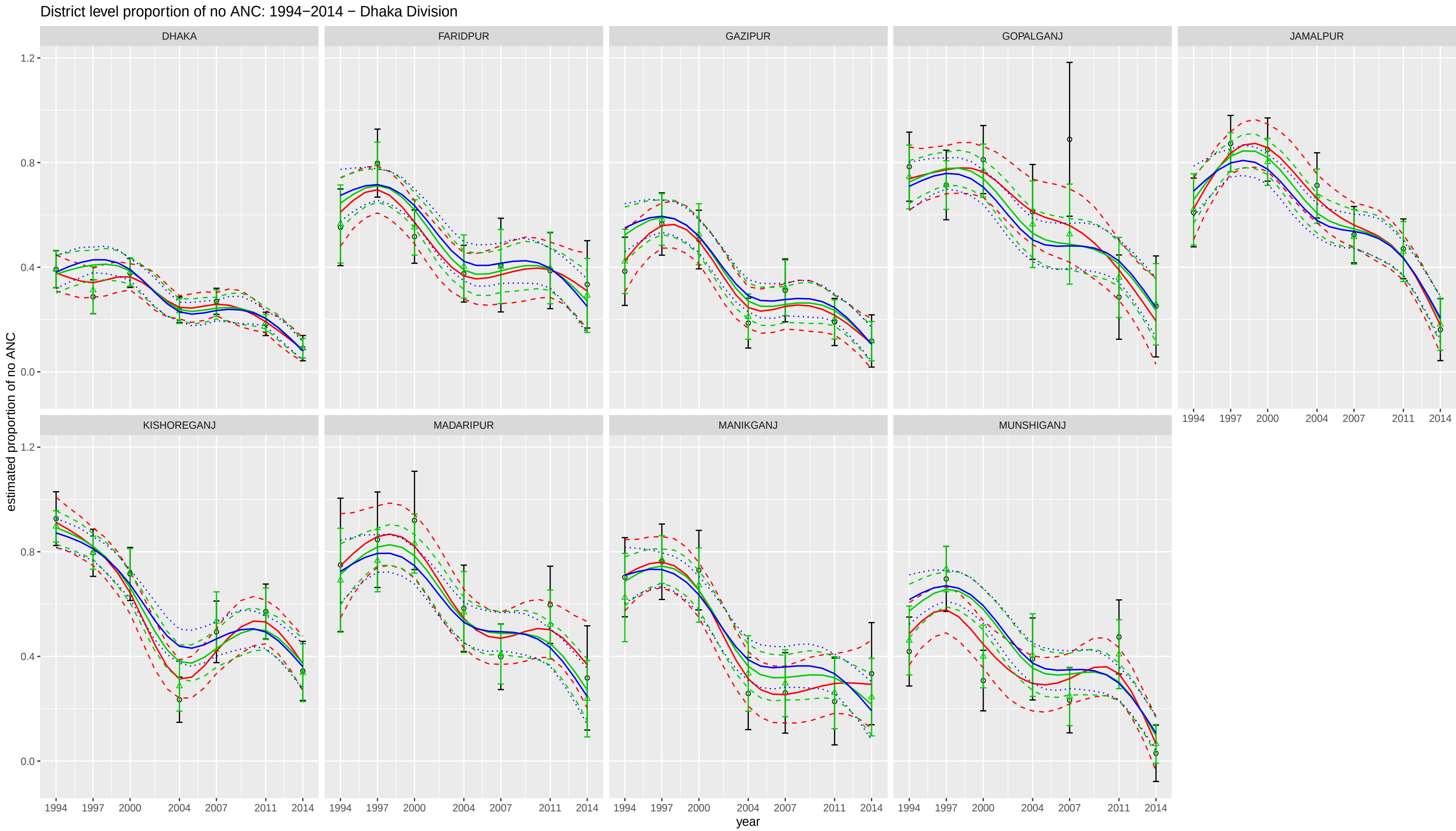

Figure S.10 District level trends of ANC0 in Bangladesh - *Dhaka* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

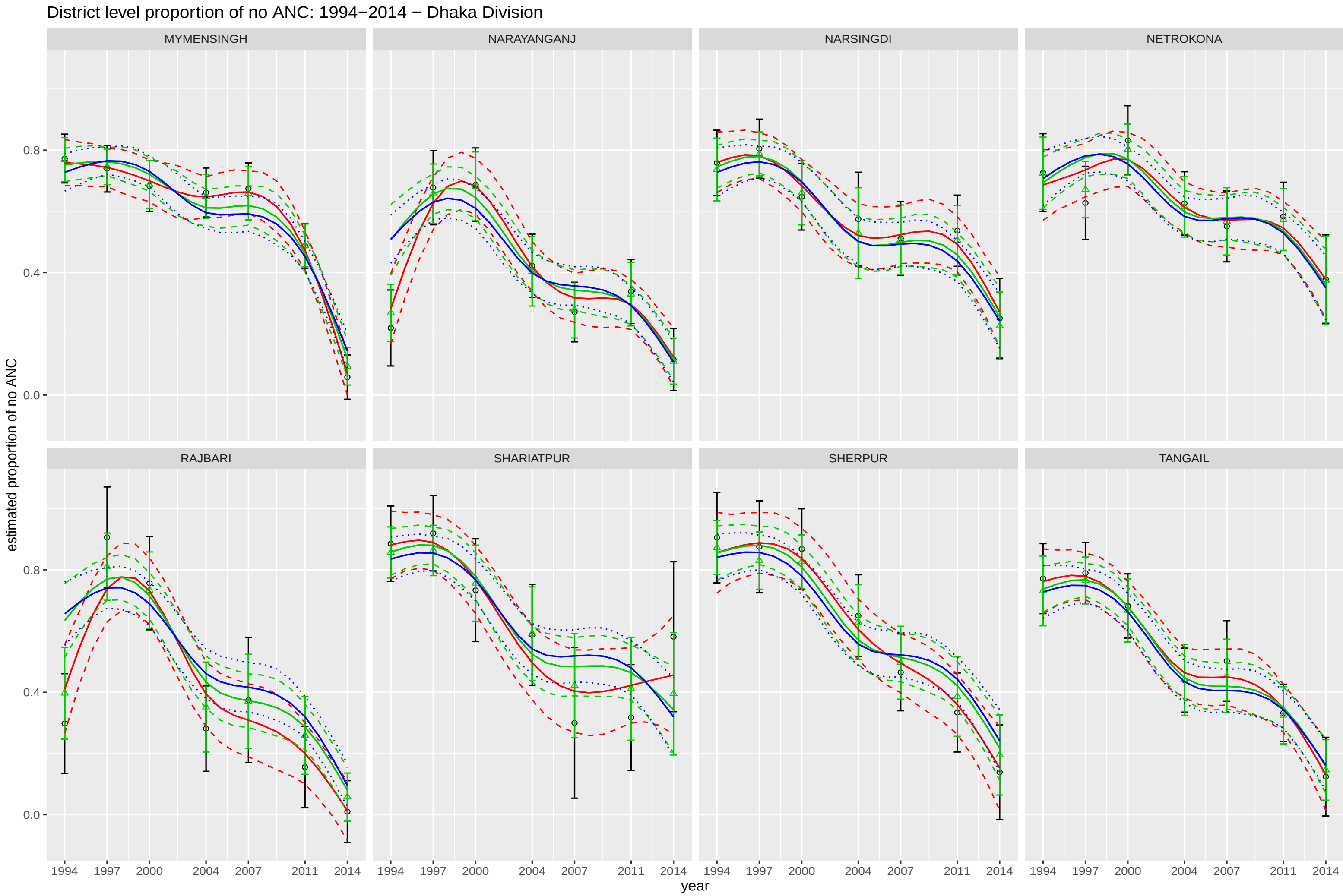

Figure S.11 District level trends of ANC0 in Bangladesh - *Dhaka* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

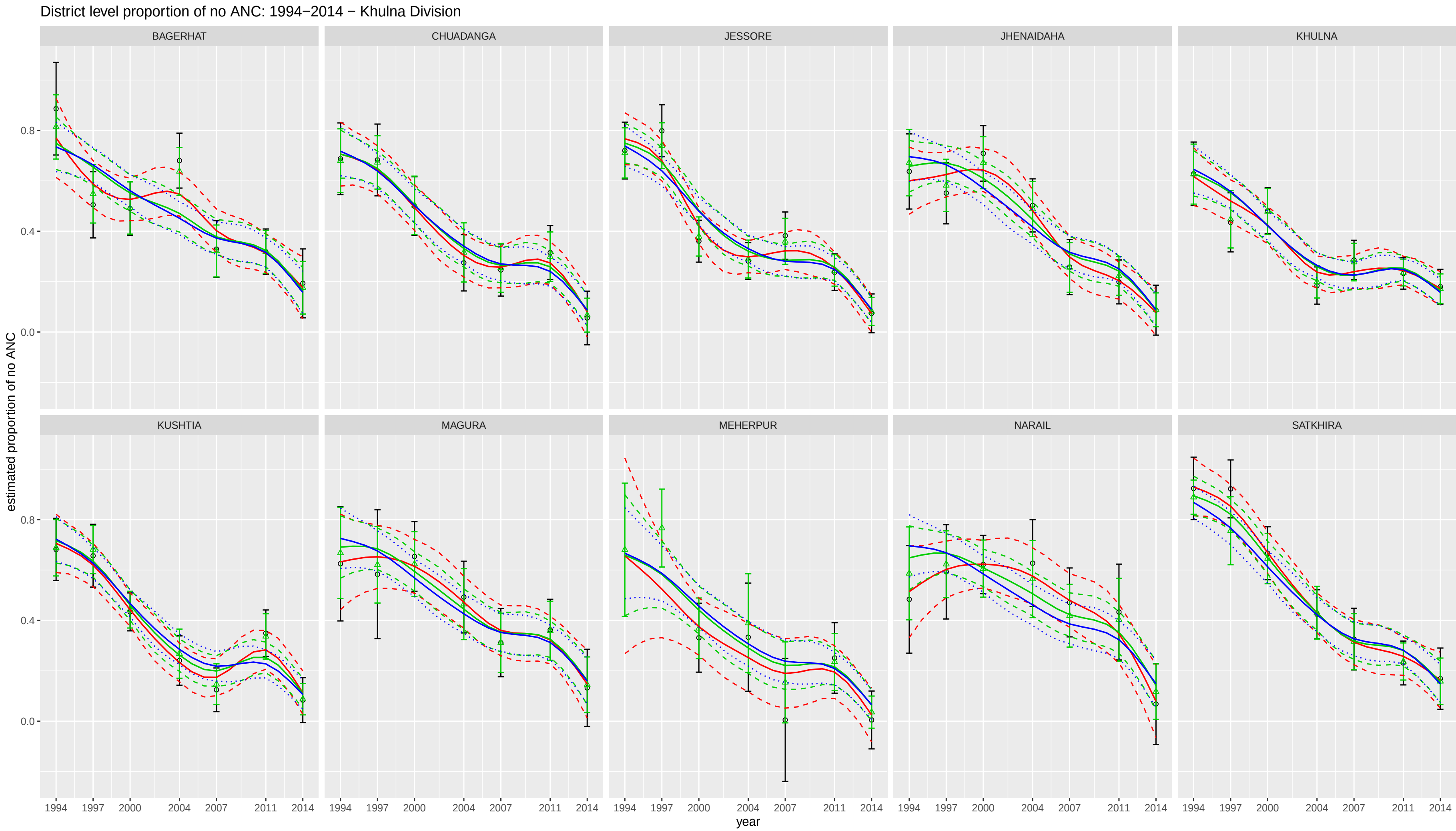

Figure S.12 District level trends of ANC0 in Bangladesh - *Khulna* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

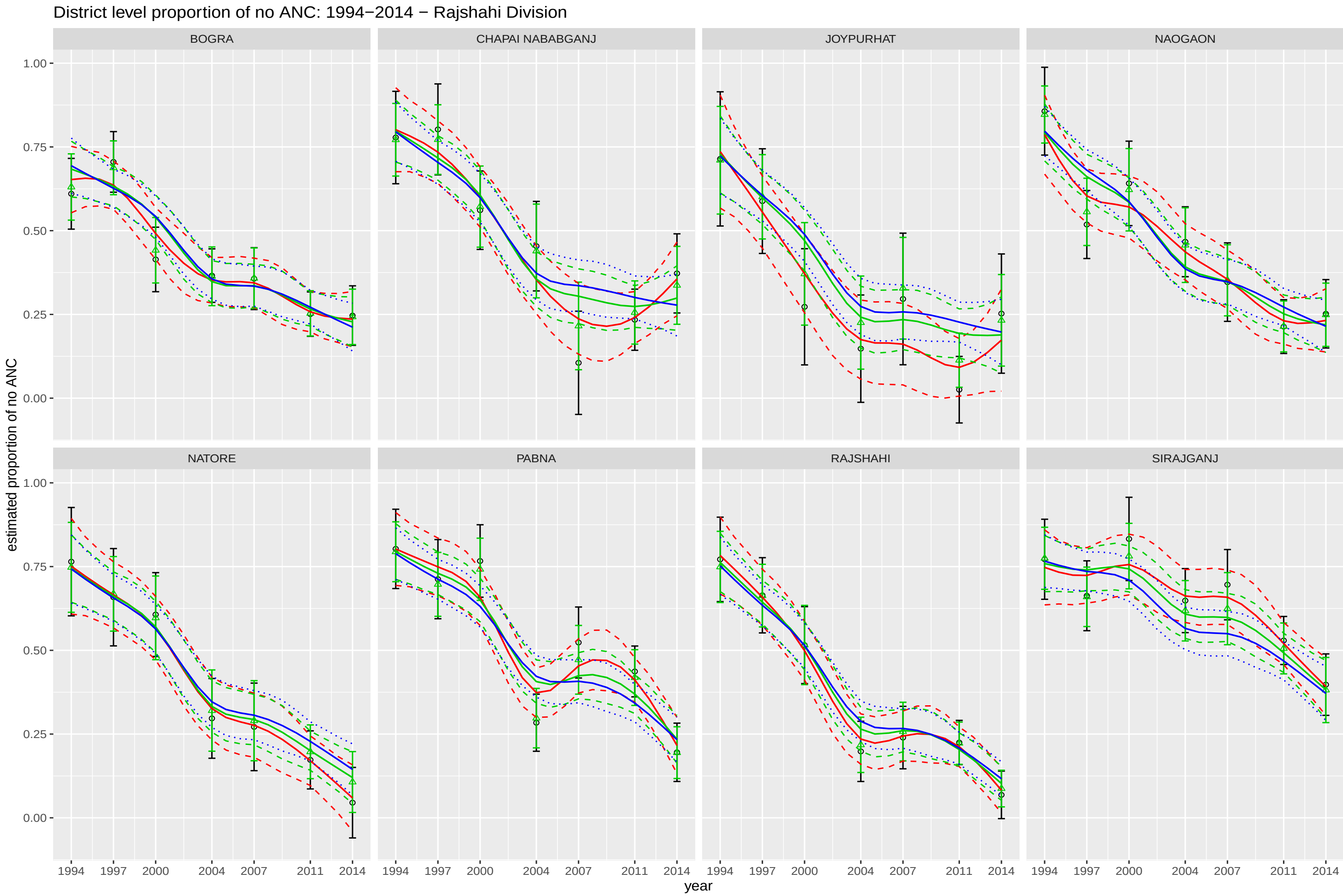

Figure S.13 District level trends of ANC0 in Bangladesh - *Rajshahi* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

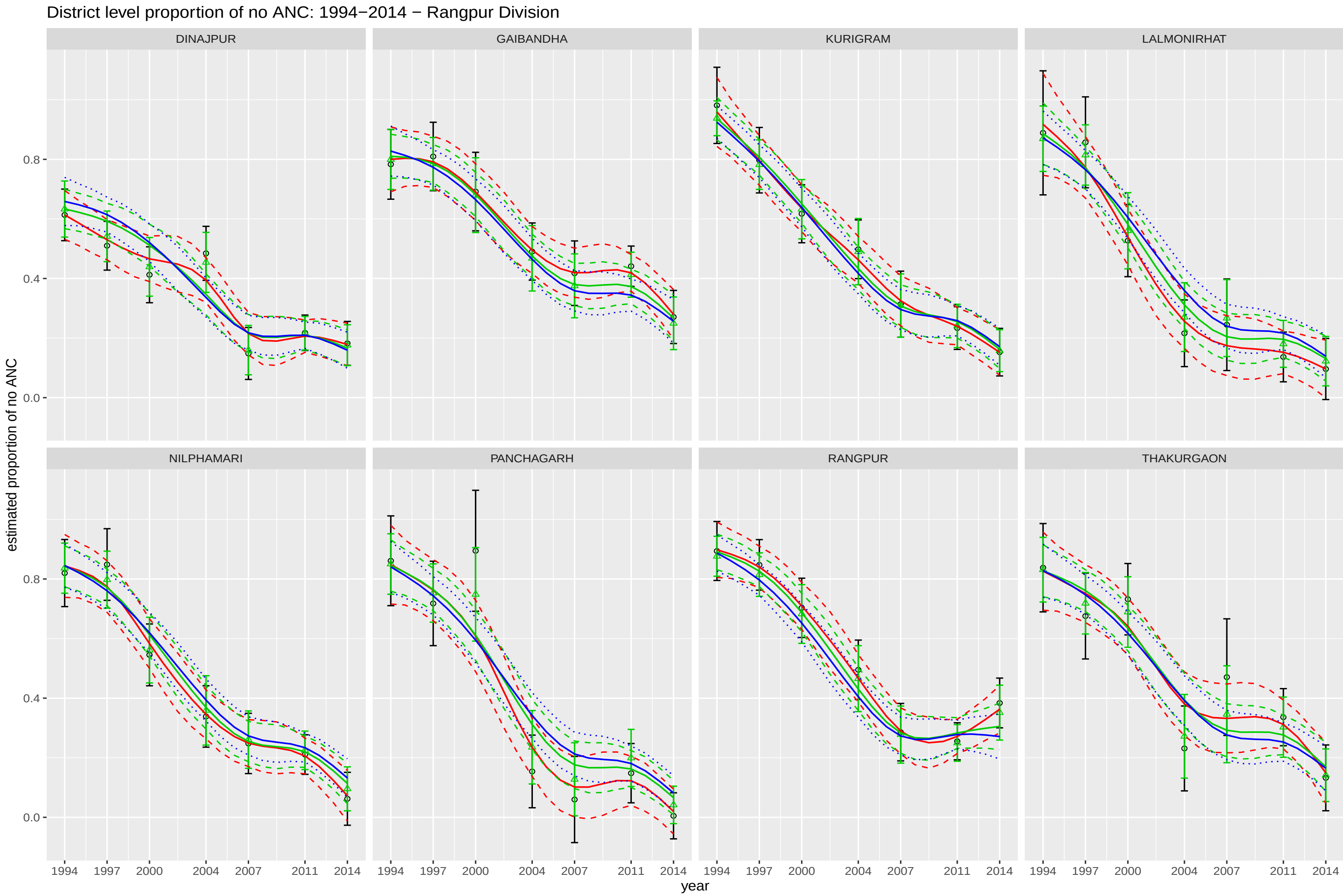

Figure S.14 District level trends of ANC0 in Bangladesh - *Rangpur* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

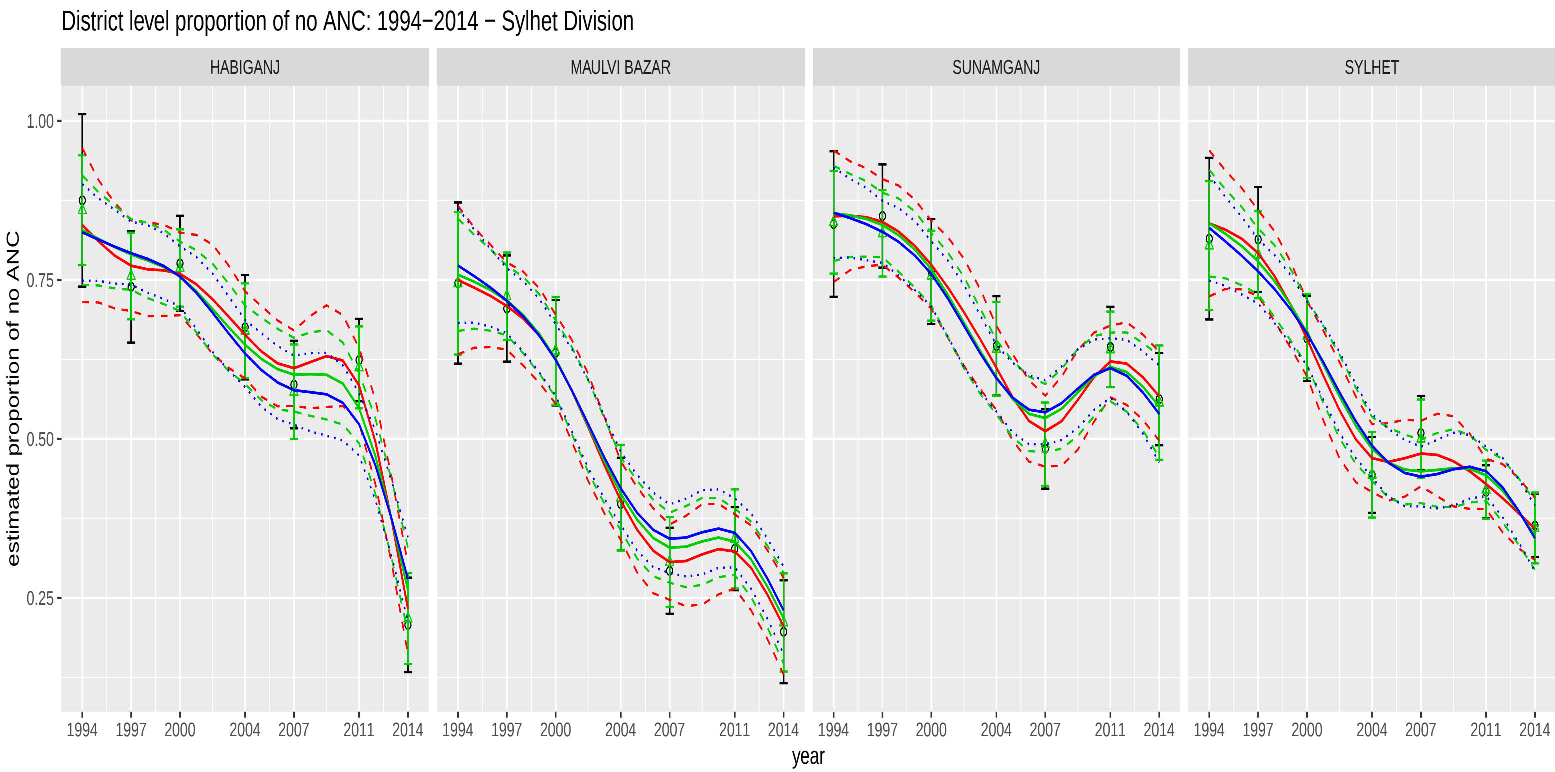

Figure S.15 District level trends of ANC0 in Bangladesh - *Sylhet* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

S.2 Trends of ANC4 in Bangladesh districts

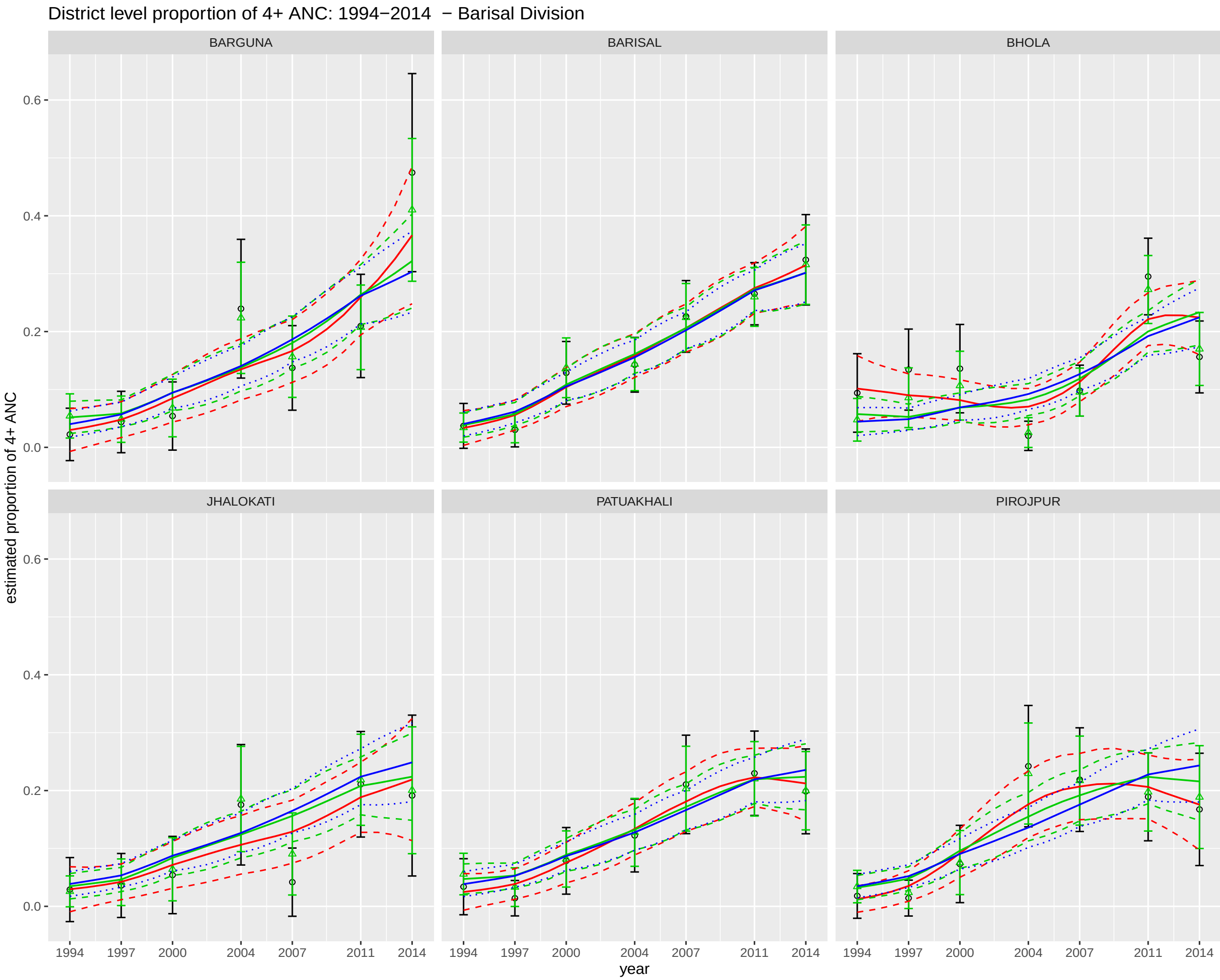

Figure S.16 District level trends of ANC4 in Bangladesh - *Barisal* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

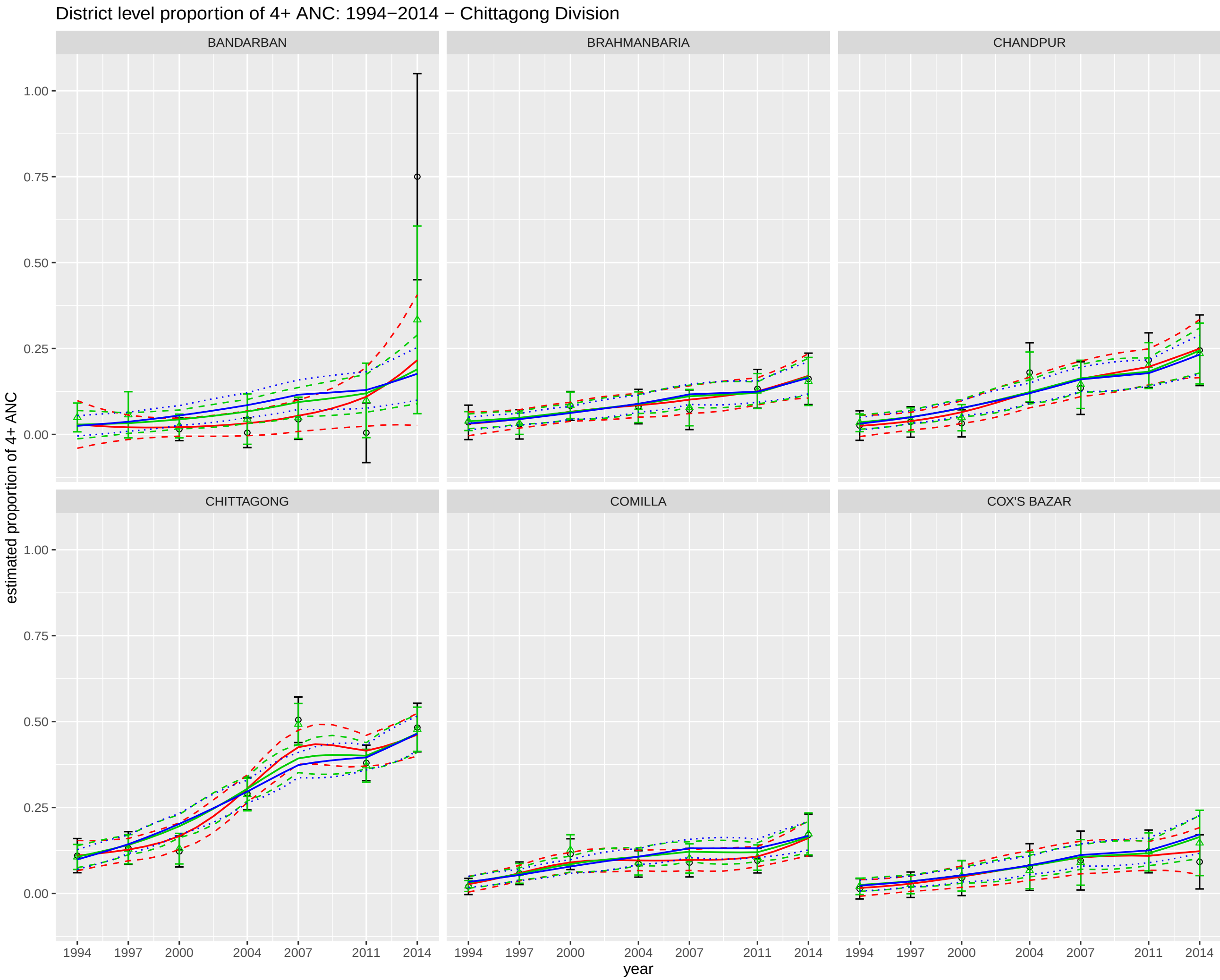

Figure S.17 District level trends of ANC4 in Bangladesh - *Chittagong* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

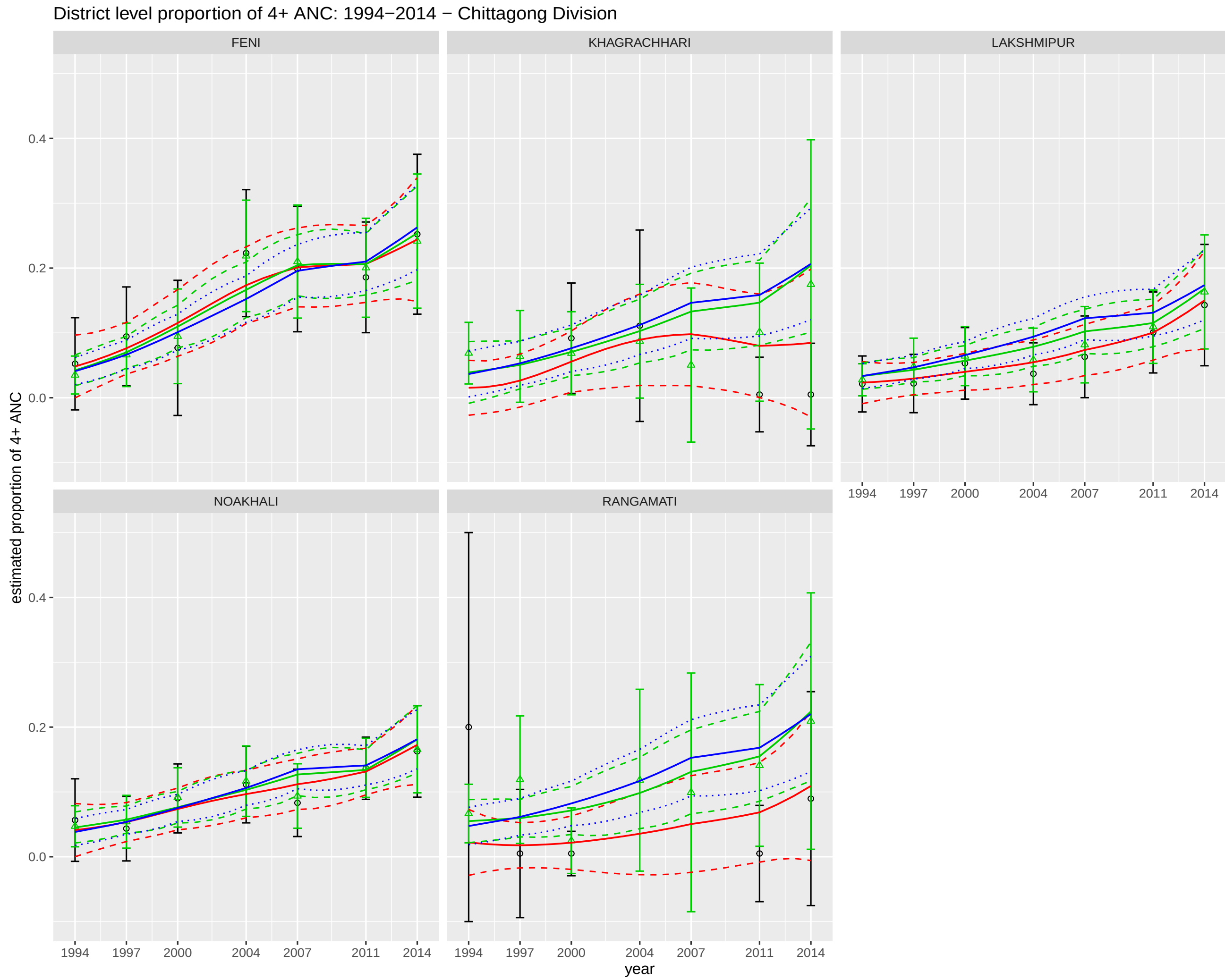

Figure S.18 District level trends of ANC4 in Bangladesh - *Chittagong* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

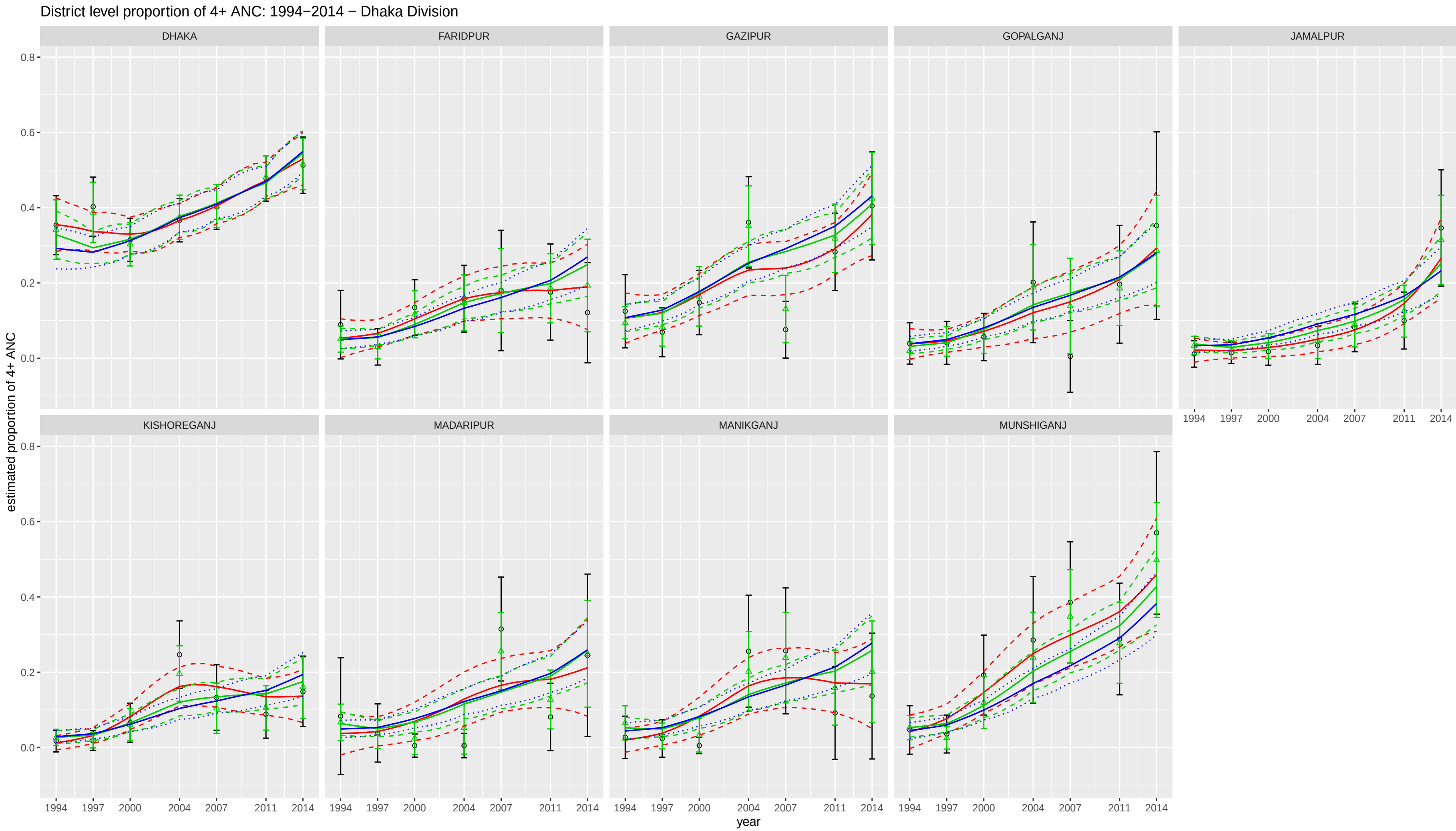

Figure S.19 District level trends of ANC4 in Bangladesh - *Dhaka* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

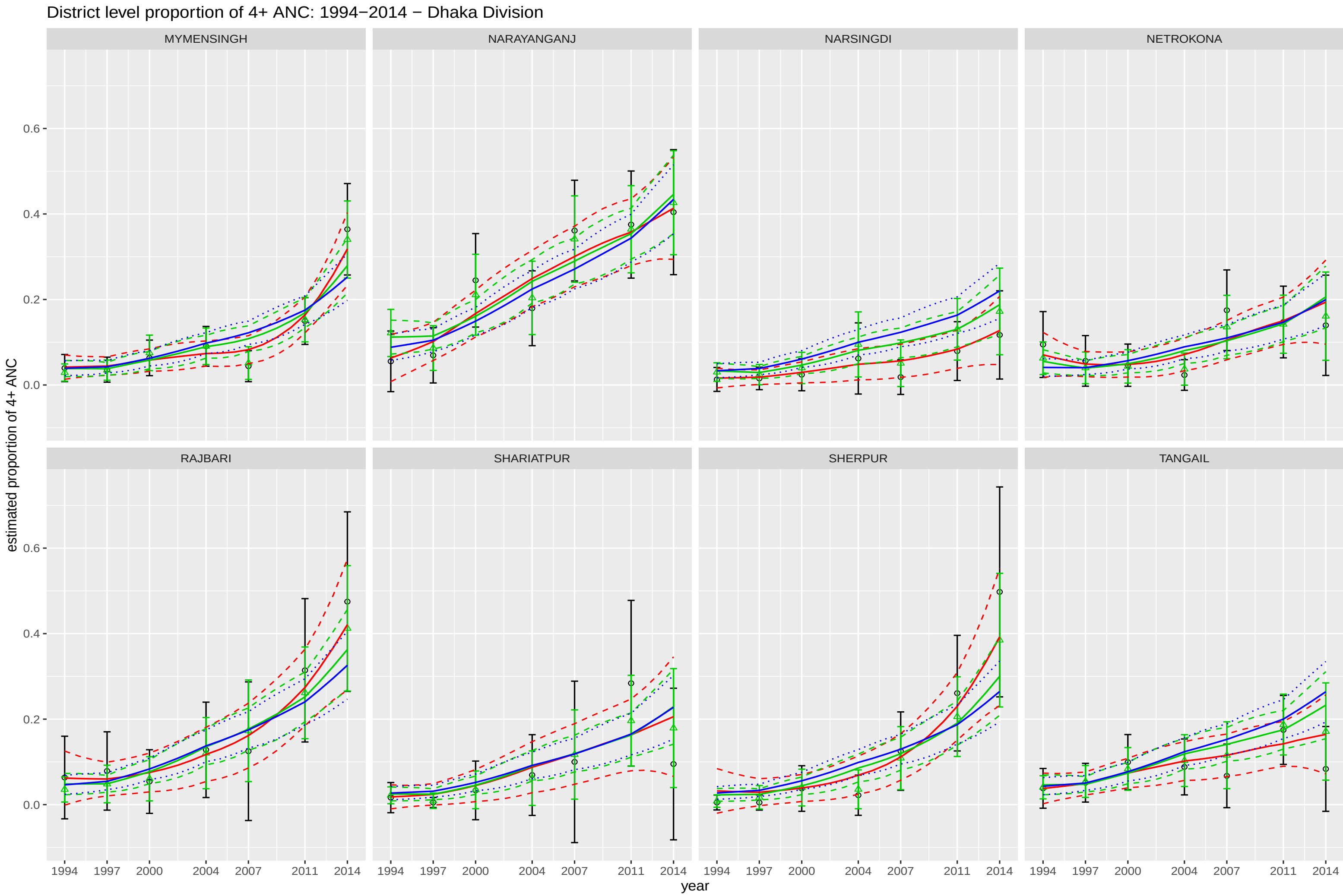

Figure S.20 District level trends of ANC4 in Bangladesh - *Dhaka* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

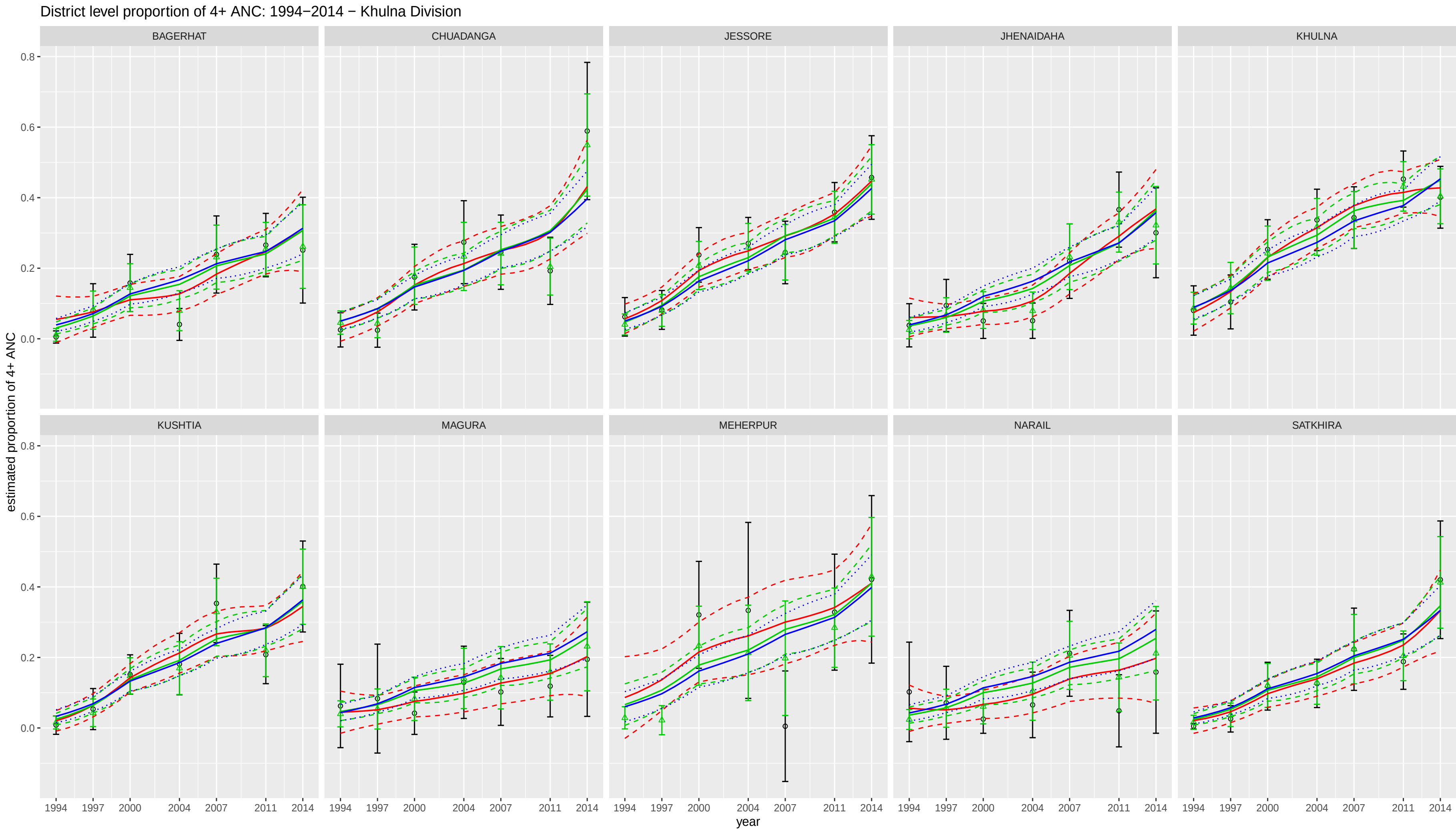

Figure S.21 District level trends of ANC4 in Bangladesh - *Khulna* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

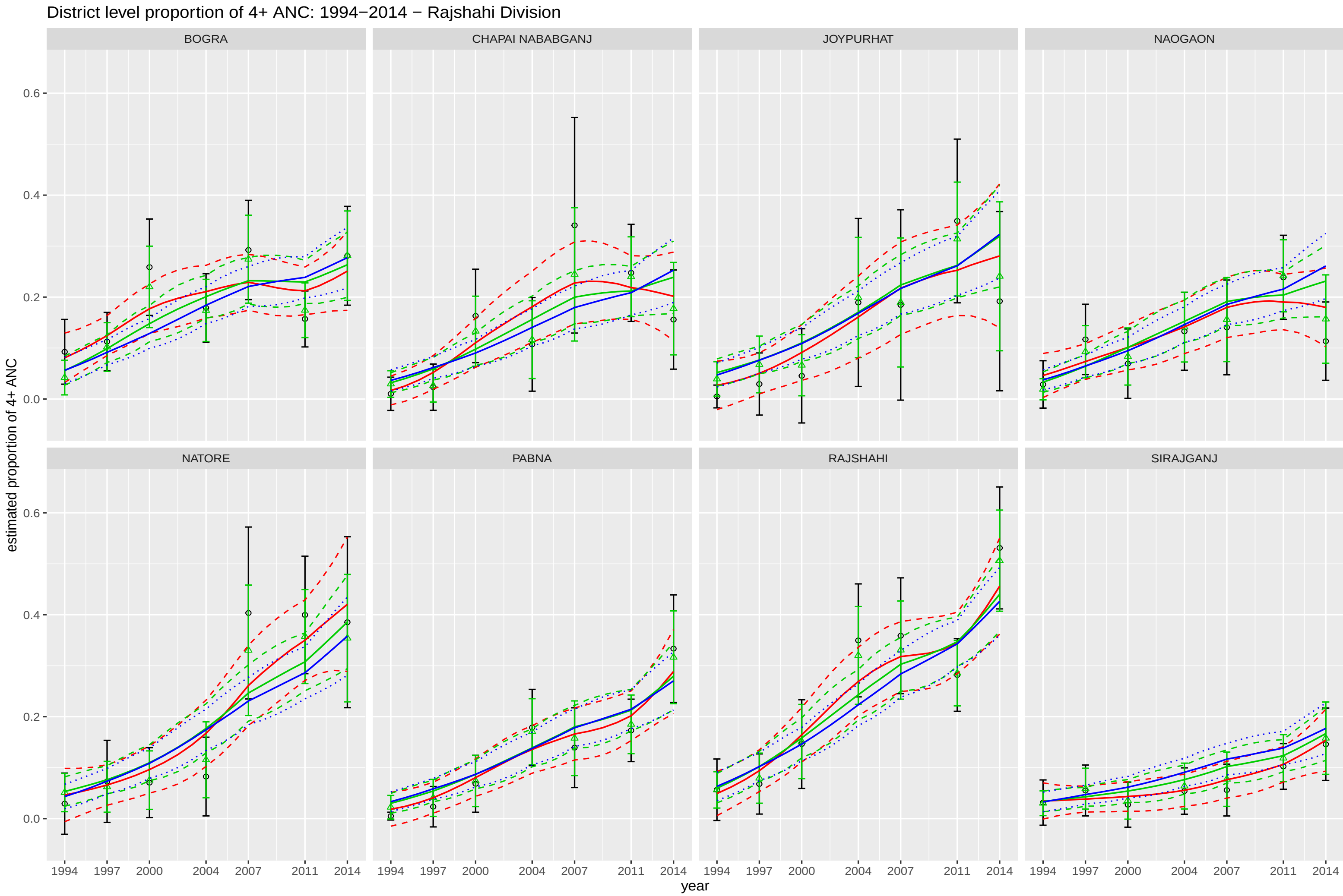

Figure S.22 District level trends of ANC4 in Bangladesh - *Rajshahi* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

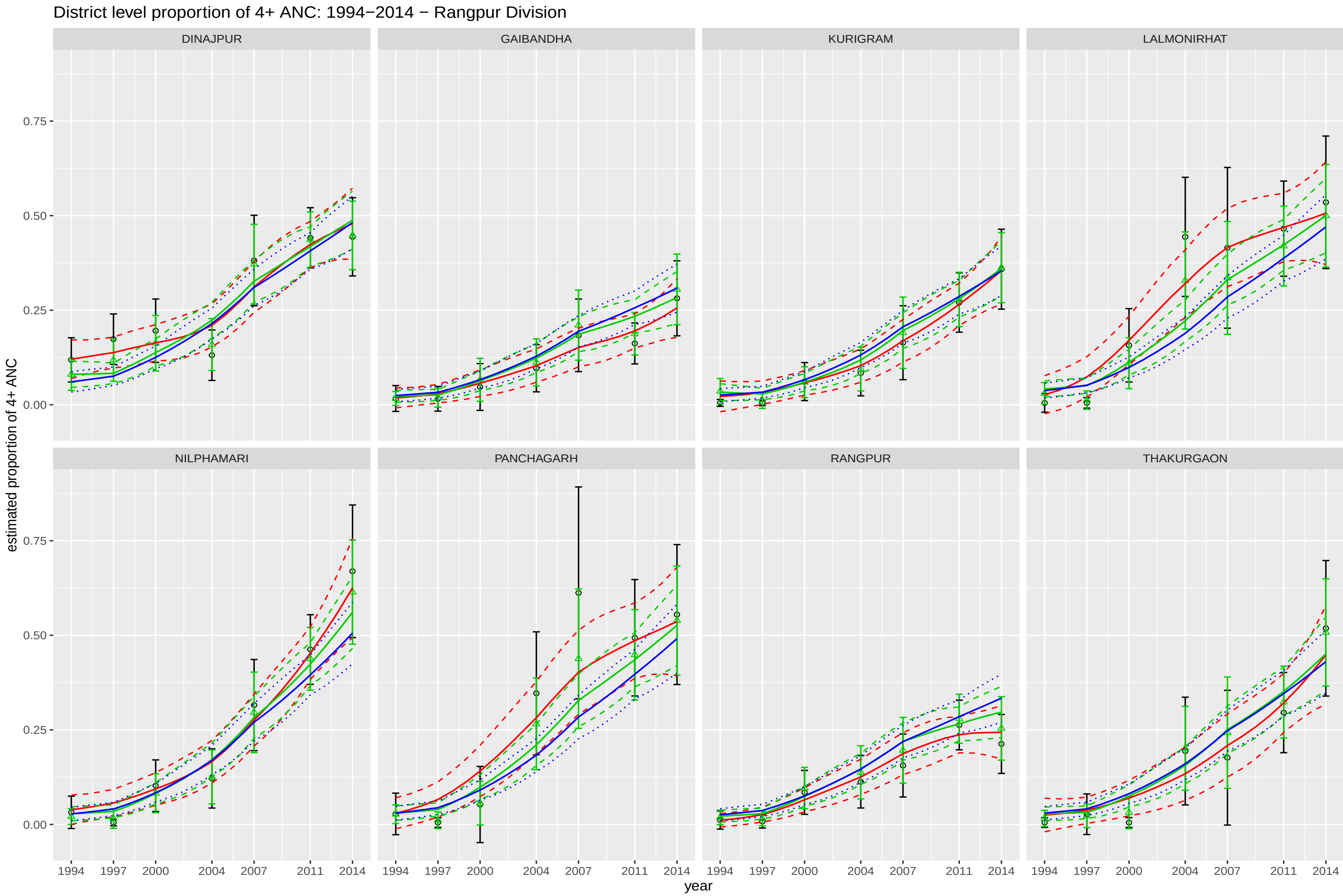

Figure S.23 District level trends of ANC4 in Bangladesh - *Rangpur* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

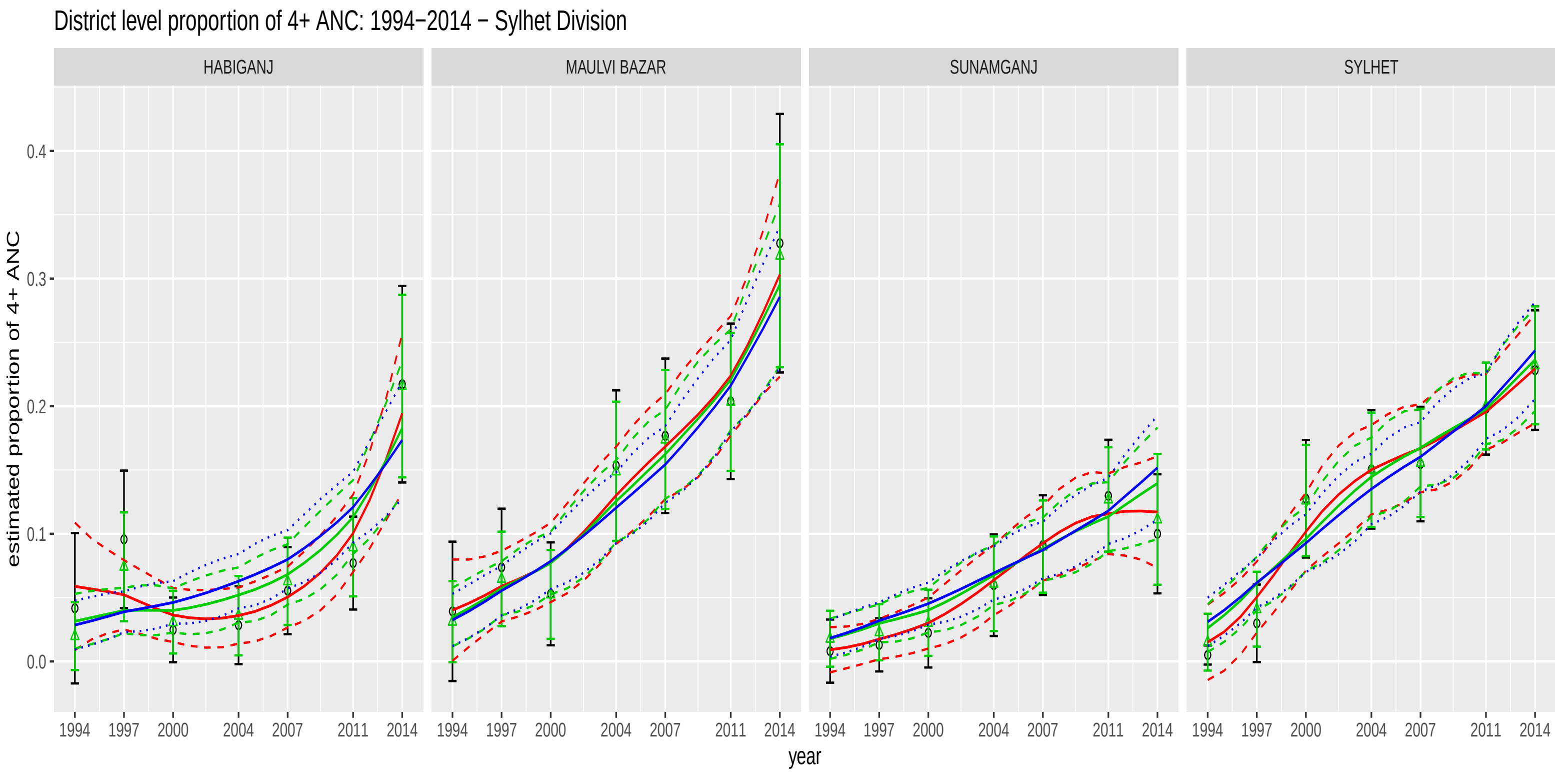

Figure S.24 District level trends of ANC4 in Bangladesh - *Sylhet* Division: (i) DIR (black error-bar line), (ii) cross-sectional FH (green error-bar line), (iii) MTS-I (red line), (iv) MTS-II (green line) and (v) MTS-III (blue line).

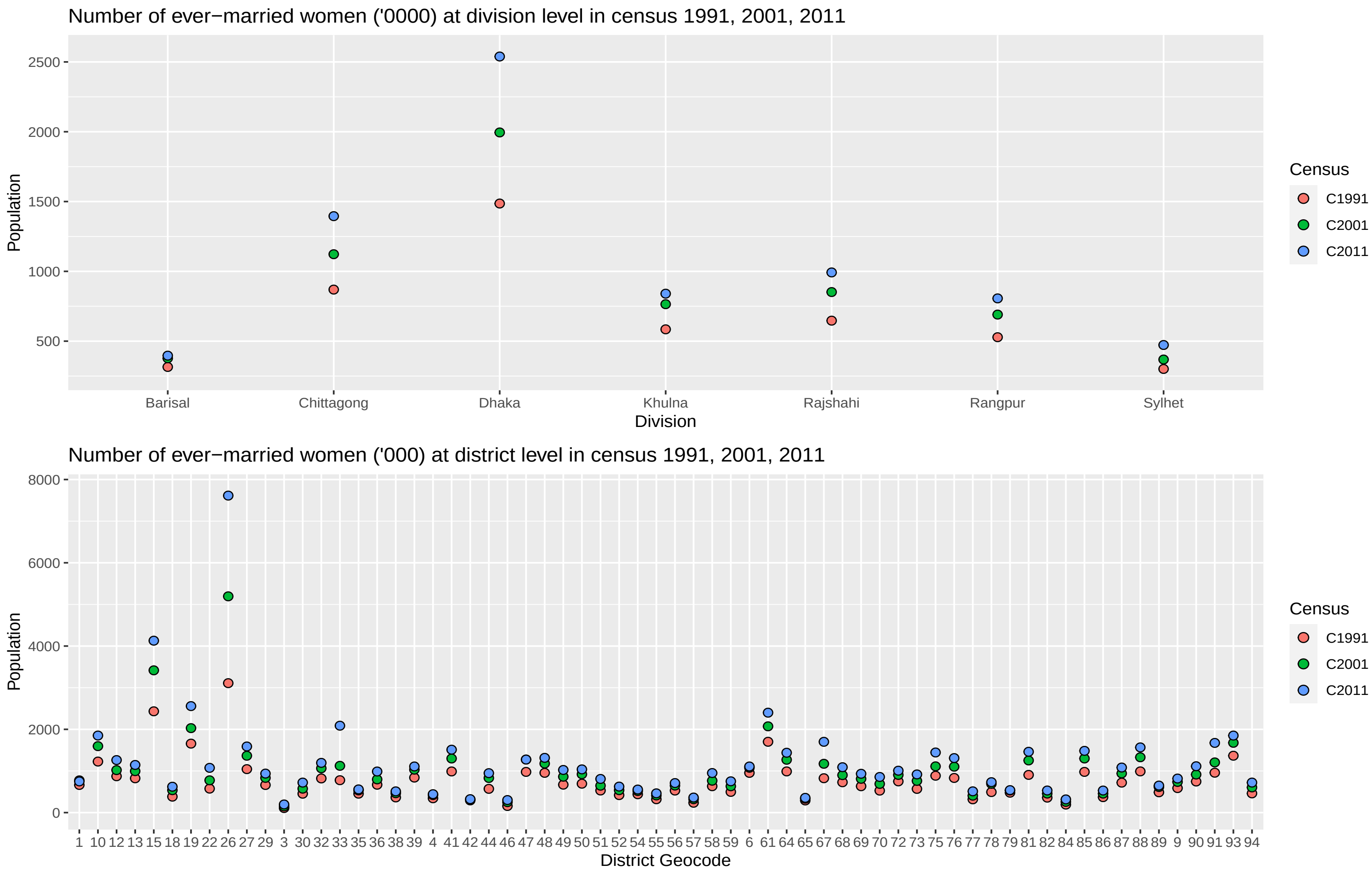

Figure S.25 **Number of ever-married women at division and district level in Census 1991, 2001, and 2011.**

Survey Year	Transformation	Fixed Effects	Random Effect	Census Data
1994	No	$1 + Division + P\_HH\_High\_Edu\_W + P\_Ur\_HH\_2^{+}\_U5$	RI: District level Random Intercept	1991
1997	No	$1 + Division + P\_MW\_Sec\_Edu + P\_HH\_Sec\_Edu\_Head$	RI	1991
2000	No	$1 + Division + P\_Ru\_HH\_U5) + P\_W$	RI	1991
2004	No	$1 + Division + P\_HH\_U5\_Sec\_Edu\_W + P\_MW$	RI	2001
2007	No	$1 + Division + P\_U5 + P\_Ru\_HH\_Size\_4^{+}$	RI	2001
2011	SQRT	$1 + Division + sqrt(P\_Ru\_HH\_U5) + sqrt(P\_MW)$	RI	2011
2014	SQRT	$1 + Division + sqrt(P\_Ur\_HH\_2^{+}\_U5) + sqrt(P\_Ru\_HH\_Elec)$	RI	2011

Table S.1 Fixed and Random effects of survey-year specific FH models for ANC0.

Survey Year	Transformation	Fixed Effects	Random Effect	Census Data
1994	SQRT	$1 + Division + sqrt(P\_Ru\_U5) + sqrt(P\_HH\_2^{+}\_U5) + sqrt(P\_Ur\_HH\_2^{+}\_U5)$	RI: District level Random Intercept	1991
1997	SQRT	$1 + Division + sqrt(P\_HH\_U5\_Sec\_Edu\_W) + log(P\_HH\_Sec\_Edu\_Head)$	RI	1991
2000	SQRT	$1 + Khulna + Region + sqrt(P\_MW_{P}rim_{E}du) + sqrt(P\_HH\_W_{I}lli_{E}du)$	RI	1991
2004	SQRT	$1 + Division + sqrt(P\_HH\_U5\_Prim\_Edu\_W) + sqrt(P\_W)$	RI	2001
2007	SQRT	$1 + Rangpur + Region + sqrt(P\_U5) + sqrt(P\_Ru\_HH\_Size\_4^{+})$	RI	2001
2011	SQRT	$1 + Rangpur + Chittagong + sqrt((P\_HH\_U5\_Sec\_Edu\_W) + sqrt(P\_W)$	RI	2011
2014	SQRT	$1 + Rangpur + Chittagong + Rajshahi + Region + sqrt(P\_W) + sqrt(P\_Ru\_HH\_Sing\_Mot)$	RI	2011

Table S.2 Fixed and Random effects of survey-year specific FH models for ANC4.

Variable	Definition
Division	Barishal, Chittagong, Dhaka, Khulna, Rajshahi, Rangpur, Sylhet
Region	(1) Densely populated Dhaka, Chittagong and Gazipur districts (2) 9 regional districts with big cities, (3) 3 hilly districts (Bandarban, Khagrachhari and Rangamati), (4) 49 other districts (less urbanized areas)
Chittagong	Chittagong Division?
Dhaka	Dhaka Division?
Khulna	Khulna Division?
Rangpur	Rangpur Division?
Rajshahi	Rajshahi Division?
P_U5	Proportion of Under-5 children
P_W	Proportion of women aged 15-49 years
P_MW	Proportion of married women aged 15-49 years
$P\_MW\_Prim\_Edu$	Proportion of married women aged 15-49 years having primary education
$P\_MW\_Sec\_Edu$	Proportion of married women aged 15-49 years having at least secondary education
$P\_HH\_No\_Edu\_W$	Proportion of household (HH) with illiterate women aged 15-49 years
$P\_HH\_Prim\_Edu\_W$	Proportion of household (HH) with primary educated women aged 15-49 years
$P\_HH\_High\_Edu\_W$	Proportion of household (HH) with higher educated women aged 15-49 years
$P\_HH\_Sec\_Edu\_Head$	Proportion of HH with at least secondary educated HH head
$P\_Ru\_HH\_4^+$	Proportion of rural HH of size 4 and more
$P\_Ru\_HH\_Elec$	Proportion of rural HH with electricity
$P\_Ru\_HH\_Sing\_Moth$	Proportion of rural HH with single mother
$P\_HH\_U5\_Sec\_Edu\_W$	Proportion of HH having under-5 children and women aged 15-49 years having at least secondary education
$P\_HH\_2^+\_U5$	Proportion of HH with 2 or more under-5 children
$P\_Ru\_HH\_U5$	Proportion of rural HH with under-5 children
$P\_Ru\_HH\_2^+\_U5$	Proportion of rural HH with 2 or more under-5 children
$P\_Ur\_HH\_2^+\_U5$	Proportion of urban HH with 2 or more under-5 children

Table S.3 District level contextual variables generated from 10% Census 1991, 10% Census 2001, and 5% Census 2011 data for ANC0

Colophon

Publisher
Statistics Netherlands
Henri Faasdreef 312, 2492 JP The Hague
www.cbs.nl

Prepress
Statistics Netherlands, Grafimedia

Design
Edenspiekermann

Information
Telephone +31 88 570 70 70, fax +31 70 337 59 94
Via contact form: www.cbs.nl/information