A Bayesian Coherent Framework for Forecasting Fertility by Age and Education

Abstract

Fertility forecasts play a central role in shaping public policy, influencing decisions across education, labor markets, and social protection. In Colombia, strong differences in fertility by educational attainment point to structural inequalities and shifting reproductive behavior yet most forecasting approaches overlook these dynamics. This paper investigates whether accounting for coherence across education groups improves the stability and relevance of fertility projections. It applies a Bayesian hierarchical modelling framework to evaluate the effect of coherence assumptions on forecast outcomes. The findings show that models which enforce consistency in trends across subpopulations produce more stable and interpretable projections, without masking important educational differences. Forecasts that incorporate educational gradients are also more aligned with recent demographic shifts and policy needs. These results highlight the risks of using overly simplistic or independent models in the presence of social stratification and demonstrate the added value of coherent modelling for long-term planning. By improving how inequality is represented in fertility forecasts, this research contributes to a more accurate and inclusive basis for decision making in population policy. The approach is broadly applicable to other settings where demographic behaviors vary systematically between socioeconomic groups.

Introduction

Fertility forecasting is a key component of demographic analysis, shaping policies in education, labour markets, and social security systems (Bongaarts, 2002). In developing regions, where demographic transitions occur rapidly and socioeconomic inequalities are widespread, accurate fertility projections are essential for designing and evaluating long-term public policies (Cleland, 2009). These projections help governments plan for future service demands, allocate resources efficiently, and anticipate shifts in population structure that affect economic growth and the sustainability of social support systems. However, forecasting fertility remains a particularly complex task due to its sensitivity to a wide range of social, economic, cultural, and policy-related factors (Brass, 1974; Booth, 2006). Compared to mortality, which typically follows more stable and predictable trajectories, fertility is more responsive to short-term behavioural changes and institutional dynamics.

As a result, turning points such as fertility declines, recoveries, or plateaus are difficult to anticipate using standard extrapolative methods. Although theories like the second demographic transition (SDT) offer valuable insights into long term fertility behaviour, they provide limited operational guidance for forecasting purposes. The SDT framework posits that beyond the initial fertility decline associated with modernisation, low fertility levels persist due to ideational changes linked to individual autonomy, gender equality, and new family forms Lesthaeghe (2010). While these theories have been widely applied in Europe, their relevance to Latin America remains debated, as regional fertility transitions have often reflected persistent social inequalities, diverse family arrangements, and distinct policy contexts(Esteve et al., 2012; Cavenaghi and Lesthaeghe, 2020). Technical challenges further complicate the task; for instance, the influence of tempo effects on period based indicators like the total fertility rate (TFR) can obscure underlying fertility quantum unless explicitly adjusted for (Bongaarts and Feeney, 1998).

One critical determinant of fertility that has gained prominence in recent decades is educational attainment. Higher levels of education are consistently associated with delayed childbearing and lower completed fertility, highlighting the importance of incorporating education into forecasting models (Martín, 1995; Lutz et al., 2014). Yet, this level of disaggregation introduces its own set of challenges. Subpopulation-specific fertility estimates particularly by education are often incomplete, inconsistently reported, or unavailable over long time series, especially in low- and middle-income countries (Schmertmann et al., 2014). In the absence of robust data and validated modelling approaches, official projections frequently adopt scenario-based methods involving high, medium, and low variants. While useful for communicating uncertainty, these scenarios are typically non-probabilistic and rely on expert judgment rather than formal statistical inference. Recent developments in stochastic and Bayesian methods offer promising avenues for improving fertility forecasting by quantifying uncertainty and enabling coherent projections across educational subgroups (Gerland et al., 2014). Nonetheless, considerable challenges remain, particularly in terms of

data quality, model transparency, and the integration of structural behavioural determinants. Addressing these issues is essential for producing reliable and policy-relevant fertility projections, especially in contexts where demographic change is both rapid and uneven (Booth, 2006).

Against this backdrop, forecasting fertility by educational attainment requires not only disaggregated estimates, but also models that preserve internal consistency across subgroups. While approaches such as that of Batyra et al. (2023) enable the analysis of fertility differentials by education, they typically forecast each subgroup independently, without enforcing coherence in shared age-specific or temporal trends. This may lead to inconsistencies when aggregating fertility rates, particularly when structural patterns such as declining fertility across all groups are ignored. Although robustness checks and sensitivity analyses are commonly applied, the absence of an integrated modelling framework limits comparability across subpopulations. These methodological gaps underscore the need for forecasting approaches that accommodate subgroup heterogeneity while ensuring demographic coherence, particularly in settings marked by educational disparities.

Building on probabilistic methods, Bayesian hierarchical models have gained prominence due to their ability to integrate prior demographic knowledge and account for interdependencies across subpopulations (Raftery et al., 2014). Unlike traditional time-series models that treat fertility rates as independent trajectories, Bayesian approaches impose a structured framework that allows for variations while ensuring coherence in demographic forecasts (Schmertmann et al., 2014). Recent studies have introduced coherent Bayesian methods that maintain consistency across demographic groups, improving forecast reliability (Wiśniowski et al., 2015). These models are particularly relevant in settings where data inconsistencies exist, as they allow for structured borrowing of information across related subpopulations (Alkema et al., 2011). Despite these advancements, the application of Bayesian coherent methods to fertility forecasting by education level remains an emerging field, with limited empirical validation. This paper builds on these developments by evaluating how different levels of coherence in Bayesian hierarchical models influence the accuracy and applicability of fertility projections.

This paper compares three Bayesian hierarchical fertility forecasting models, each designed to capture different levels of coherence across educational subpopulations. The first model estimates fertility trends independently for each education group, treating them as unrelated processes. The second model introduces structural coherence by enforcing consistency in age-specific fertility responses, while allowing for heterogeneity across education levels. The third model extends this further by imposing coherence in both age-specific and period-specific trends, reflecting shared temporal dynamics across groups. By systematically evaluating these models, the study aims to identify which specification offers the best trade-off between subgroup flexibility and overall model coherence in forecasting fertility by educational attainment.

This paper is structured as follows. The methodology presents a bayesian hierarchical framework for fertility forecasting, comprising three models: (i) a

non-coherent model that estimates fertility separately by education level; (ii) a coherent model that enforces consistency in age-specific fertility declines across education groups; and (iii) a fully coherent model that ensures consistency in both age-specific and period-specific fertility trends. The section data and model implementation describes the fertility data used in the analysis and outlines the estimation procedure. The section results evaluates model performance based on predictive accuracy. The paper concludes by summarizing the main contributions and implications of the study.

Methodology: Bayesian Hierarchical Framework for Fertility Forecasting

This study extends the Bayesian Lee-Carter model developed by Wiśniowski et al. (2015) into a hierarchical framework tailored for forecasting fertility rates across educational groups. The key innovation lies in modelling these subpopulations jointly rather than independently, allowing for shared patterns in both the age-specific shape and the temporal evolution of fertility. Instead of fitting entirely separate models for each education level, the hierarchical structure enables partial pooling of information capturing common demographic trends while still allowing for group-specific differences. This approach ensures internal consistency across education-specific forecasts and improves the coherence of aggregated fertility projections. In doing so, it addresses a key limitation of independent forecasting approaches, which may overlook structural similarities and lead to inconsistencies when comparing or summing fertility rates across subgroups.

Extending the Lee-Carter Model to Fertility Forecasting

The Lee-Carter model (Lee, 1993), originally designed for mortality forecasting in Lee and Carter (1992), has been widely adapted for fertility modeling. It decomposes fertility rates into an age-specific component and a time-varying index of fertility trends. However, when estimated independently for different educational group inconsistencies may arise due to the underlying demographic relationships.

To address these limitation this paper applies a Bayesian hierarchical framework, incorporating prior information and ensuring coherence across subpopulations. The approach follows three model specifications:

Non-coherent Bayesian Lee-Carter model, estimating each education level independently

The Non-Coherent Lee-Carter Model for forecasting fertility rates (Fertility Model A) is formulated within a Hierarchical Bayesian Model (HBM) framework. It extends the original Lee-Carter model to flexibly capture fertility dynamics using diverse data sources, particularly for subpopulations defined

by educational attainment (Wiśniowski et al., 2015). This specification builds on the Bayesian extensions of the models introduced by Lee (1993) and further developed by Wiśniowski et al. (2015) to accommodate heterogeneity in demographic behaviours across strata.

Figure 1 presents the structure of the Non-Coherent Hierarchical Bayesian Lee-Carter Model used to estimate fertility rates by educational attainment. The model separately forecasts fertility for four educational categories: no formal education ($e = 0$), primary education ($e = 1$), secondary education ($e = 2$), and post-secondary education ($e = 3$). The hierarchical structure enables the specification of prior distributions and the modelling of stochastic components, allowing for flexible estimation of subgroup-specific trends without enforcing coherence across educational groups.

The graphical representation consists of three primary node types: variable nodes (circles), constant nodes (curved squares), and data nodes (squares). The variable nodes denote parameters that are estimated during the iterative computational process whereas constant nodes represent fixed quantities that inform prior distributions or baseline assumptions. Data nodes indicate observed fertility. The model integrates both stochastic and deterministic relationship visually distinguished by different arrow types. Solid arrows represent stochastic dependencies capturing uncertainty and probabilistic variation in fertility estimates across time and educational groups. Dotted arrows indicate deterministic dependencies ensuring internal coherence in model constraints and structural assumptions.

For each educational category e, the core parameters include $\sigma_{logf_e}(x,t)$, representing the variance of log fertility rates over age x and time t, and $logf_e(x,t)$, denoting the logarithm of fertility rates at specific ages and periods. The hierarchical Bayesian framework also includes a latent component, denoted as $h_e(x,t)$, which combines the effects of age, education, and time. Specifically, $h_e(x,t)$ is expressed as a linear combination of the Lee–Carter parameters a_x^e, b_x^e, and k_t^e, representing the baseline age–education pattern, the sensitivity of fertility to age-specific changes, and the overall time–education trend, respectively.

Fertility Model A is structured into six hierarchical layers: the data layer (two stages), the process layer, and the prior layer (divided into three stages). For each e represents educational attainment ($e = 0$ for no formal education, $e = 1$ for primary education, $e = 2$ for secondary education, and $e = 3$ for post-secondary education), Model A is described by the following mathematical expressions consisting of seven hierarchical layers:

- **Data Layer (First Stage):**

$$B_{x,t}^e \sim \text{Poisson}\left(f_{x,t}^e \cdot E_{x,t}^e\right). \tag{1}$$

- **Data Layer (Second Stage):**

$$\log f_{x,t}^e \sim \mathcal{N}\left(\alpha_x^e + \beta_x^e \kappa_t^e, \Sigma^f\right). \tag{2}$$

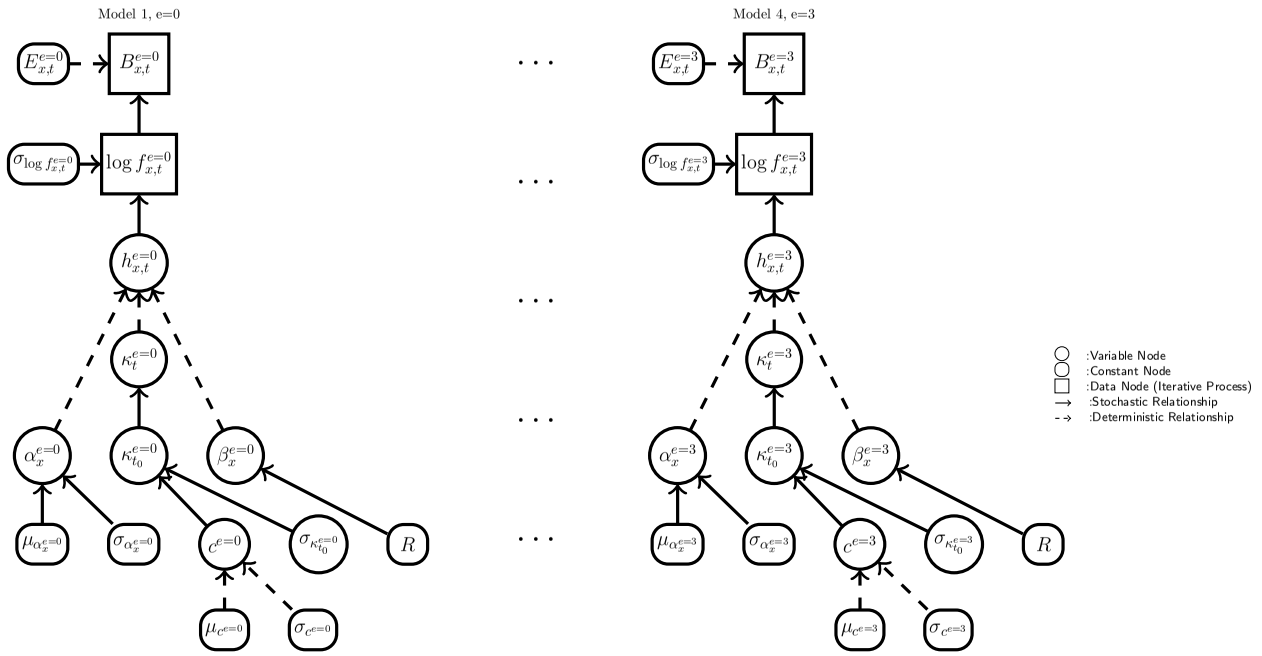

Figure 1: Fertility Model A: Non-Coherent Hierarchical Bayesian Lee-Carter Models for forecasting fertility rates by educational attainment independently (0 = No formal education, 1 = Primary education, 2 = Secondary education, 3 = Post-secondary education). Circles represent variable nodes, curved squares represent constant nodes, and squares represent data nodes where the iterative process occurs. Solid arrows represent stochastic relationships, while dotted arrows indicate deterministic relationships.

- **Process Layer:**

$$h^e_{x,t} = \alpha^e_x + \beta^e_x \cdot \kappa^e_t. \tag{3}$$

- **Prior Layer (First Stage):**

$$\alpha^e_x \sim \mathcal{N}\left(f_{a^e_x}, \sigma_{a^e_x}\right), \tag{4}$$

$$\tag{5}$$

$$\beta^e_x \sim \text{Dirichlet}(\alpha_\beta), \tag{6}$$

$$\tag{7}$$

$$\kappa^e_t \sim \mathcal{N}\left(c^e + \kappa^e_{t-1}, \sigma^2_{\kappa^e_t}\right). \tag{8}$$

- **Prior Layer (Second Stage):**

$$\kappa^e_{t_0} \sim \mathcal{N}\left(c^e, \sigma^2_{\kappa^e_{t_0}}\right). \tag{9}$$

- **Prior Layer (Third Stage):**

$$c^e \sim \mathcal{N}\left(f_{c^e}, \sigma_{c^e}\right), \tag{10}$$

$$\tag{11}$$

$$\sigma_{\kappa^e_t} \sim \text{Exponential}(\lambda_{\kappa^e_t}). \tag{12}$$

The model specification presented in equations (1)–(12) follows the structure of the Lee–Carter framework, adapted here for fertility forecasting within a hierarchical Bayesian setting. In the first stage (equation (1)), the observed number of births $B^e_{x,t}$ at age x, time t, and education level e is assumed to follow a Poisson distribution with mean equal to the product of the fertility rate $f^e_{x,t}$ and the population exposure $E^e_{x,t}$. The second stage (equation (2)) models the logarithm of fertility rates as a normal distribution centered on a linear predictor composed of three key parameters: the term α^e_x represents the baseline log-fertility rate at age x for individuals with education level e, capturing the average age pattern of fertility across time; the coefficient β^e_x reflects the sensitivity of fertility at age x (within education group e) to changes over time, effectively weighting the influence of temporal variation; and the parameter κ^e_t captures the overall level of fertility at time t for education level e, representing period-specific changes that are common across all ages within the group. Together, these components allow the model to account for both age-specific fertility profiles and their evolution over time within each educational subgroup. This specification captures both the baseline fertility pattern by age and education (α^e_x) and its temporal evolution through the interaction between β^e_x and κ^e_t.

Equation (3) defines the latent process $h^e_{x,t}$, which represents the systematic component of the model and links age, period, and education effects through a linear combination of these parameters. The prior layers (equations (8)–(12))

complete the hierarchical structure by assigning probability distributions to the parameters. The age-specific intercepts α_x^e follow normal priors, the sensitivity parameters β_x^e follow Dirichlet priors to ensure normalization, and the time indexes κ_t^e evolve according to a random-walk process with drift c^e and variance $\sigma^2_{\kappa_t^e}$. Hyperpriors are placed on c^e and $\sigma_{\kappa_t^e}$ to allow for uncertainty in the overall level and variability of temporal change. Together, these equations define a coherent Bayesian implementation of the Lee–Carter model that jointly captures age-, period-, and education-specific fertility dynamics.

By integrating the data, process and prior layer the posterior distribution of Model A parameters α_x^e, β_x^e, and κ_t^e, is derived as being proportional to the product of the likelihood and prior distribution as outlined in Equation 13.

$$\begin{aligned} p\left(\alpha_x^e, \beta_x^e, \kappa_t^e \mid \log f_{x,t}^e\right) \propto & \prod_{t \in T} \prod_{x \in X} \left[p\left(\log f_{x,t}^e \mid h_{x,t}^e\right)\right] \cdot p\left(h_{x,t}^e \mid \alpha_x^e, \beta_x^e, \kappa_t^e\right) \cdot p\left(\alpha_x^e\right) \\ & \cdot p\left(\beta_x^e\right) \cdot p\left(\kappa_t^e \mid c^e\right) \cdot p\left(\kappa_t^e \mid \sigma_k^e\right) \cdot p\left(c^e\right) \cdot p\left(\sigma_k^e\right) \\ = & \prod_{t \in T} \prod_{x \in X} \left[p\left(\log f_{x,t}^e \mid f\left(\alpha_x^e, \beta_x^e, \kappa_t^e\right)\right)\right] \\ & \cdot p\left(\alpha_x^e\right) \cdot p\left(\beta_x^e\right) \cdot p\left(\kappa_t^e \mid c^e\right) \cdot p\left(\kappa_t^e \mid \sigma_k^e\right) \cdot p\left(c^e\right) \cdot p\left(\sigma_k^e\right). \end{aligned} \tag{13}$$

In this context, $p(\cdot)$ signifies a given probability distribution, whereas $f(\cdot)$ denotes a functional dependency within the process layer. More specifically, the function is expressed as $f\left(\alpha_x^e, \beta_x^e, \kappa_t^e\right) = h_{x,t}^e$, where $h_{x,t}^e$ is determined by the sum of α_x^e, β_x^e, and κ_t^e.

While this independence enables flexibility in accounting for heterogeneous fertility patterns it may also introduce discrepancies when aggregating forecasts across educational groups. To further refine fertility projection the model extends the Bayesian Lee-Carter framework by incorporating stochastic age-period interactions. This enhancement enables robust estimation of future fertility rates under different demographic scenarios. The iterative structure, as indicated in the data node facilitates convergence of posterior distributions through Markov Chain Monte Carlo (MCMC). Additionally, the hierarchical specification ensures that model parameters remain identifiable and interpretable within the broader demographic forecasting context.

To address the limitations of the non-coherent specification, the next model introduces coherence across educational groups in the age-specific dimension. This is achieved by enforcing a shared structure for the age sensitivity parameter, thereby aligning fertility trends across strata.

Coherent Bayesian model in β across educational level enforcing consistency in age-specific fertility trends.

Similar to the model described in equation 13, age-specific fertility trends over time are modeled coherently across all educational groups. This specification

imposes coherence on the age-specific sensitivity parameter β, resulting in a common rate of change in fertility across educational levels.

The standard Lee-Carter model is given by:

$$\log f_{x,t}^e = \alpha_x^e + \beta_x^e \kappa_t^e + \epsilon_{x,t}^e.$$

To introduce coherence across educational attainment, the β parameter is constrained as:

$$\beta_x^0 = \beta_x^1 = \beta_x^2 = \beta_x^3,$$

where $e = 0, 1, 2, 3$ represents the educational levels. This constraint allows the model to reflect consistent fertility trends across educational groups for each sex.

The hierarchical framework is visually depicted in Figure 2. Similar to Figure 1 in this representation, circles indicate variable nodes, while rounded-corner squares represent deterministic or constant nodes. The squares illustrate the iterative structure across age and time dimensions. Solid arrows indicate stochastic dependencies, whereas dashed lines denote deterministic relationships.

The structure of the model illustrated in Figure 2 is composed of the following hierarchical layers:

- **Data Layer (First Stage):**

$$B_{x,t}^e \sim \text{Poisson}\left(f_{x,t}^e \cdot E_{x,t}^e\right). \tag{14}$$

- **Data Layer (Second Stage):**

$$\log f_{x,t}^e \sim \mathcal{N}\left(\alpha_x^e + \beta_x \kappa_t^e, \Sigma^f\right). \tag{15}$$

- **Process Layer:**

$$h_{x,t}^e = \alpha_x^e + \beta_x \cdot \kappa_t^e. \tag{16}$$

- **Prior Layer (First Stage):**

$$\alpha_x^e \sim \mathcal{N}\left(f_{a_x^e}, \sigma_{a_x^e}\right), \tag{17}$$

$$\tag{18}$$

$$\beta_x \sim \text{Dirichlet}(\alpha_\beta), \tag{19}$$

$$\tag{20}$$

$$\kappa_t^e \sim \mathcal{N}\left(c^e + \kappa_{t-1}^e, \sigma_{\kappa_t^e}^2\right). \tag{21}$$

- **Prior Layer (Second Stage):**

$$\kappa_{t_0}^e \sim \mathcal{N}\left(c^e, \sigma_{\kappa_{t_0}^e}^2\right). \tag{22}$$

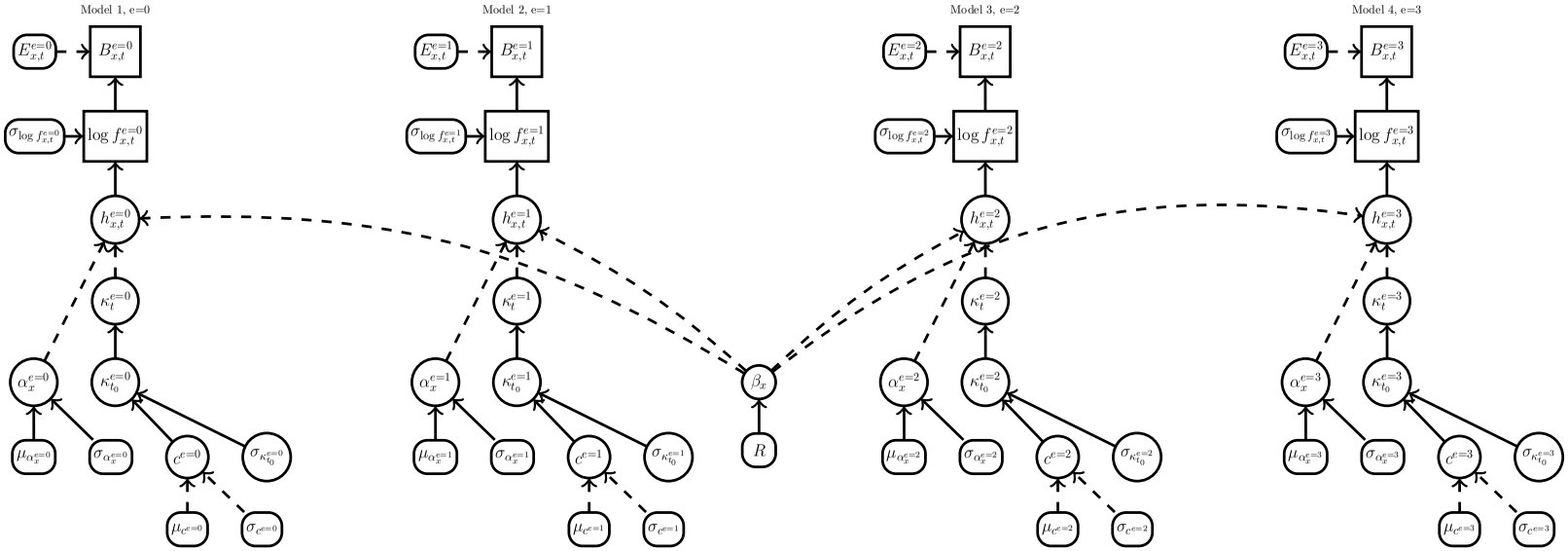

Figure 2: Fertility Model B: Coherent Lee–Carter Model in β across educational attainment for forecasting (0 = No formal education, 1 = Primary education, 2 = Secondary education, 3 = Post-secondary education).

- **Prior Layer (Third Stage):**

$$c^e \sim \mathcal{N}\left(f_{c^e}, \sigma_{c^e}\right), \tag{23}$$

$$\tag{24}$$

$$\sigma_{\kappa_t^e} \sim \text{Exponential}(\lambda_{\kappa_t^e}). \tag{25}$$

The posterior distribution of Model B, which incorporates data, process, and prior layers, is derived for each subpopulation defined by educational attainment. The parameters α_x^e, β_x, and κ_t^e follow a posterior distribution that is proportional to the product of the likelihood and prior distributions, as formulated in Equation 26.

$$\begin{aligned}
&p\Big(\alpha_x^{e=0}, \ldots, \alpha_x^{e=3}, \beta_x, \kappa_t^{e=0}, \ldots, \kappa_t^{e=3} \mid \log f_{x,t}^{e=0}, \ldots, \log f_{x,t}^{e=3}\Big) \\
&\quad \propto \prod_{t \in T} \prod_{x \in X} \left[p\left(\log f_{x,t}^{e=0} \mid f\left(\alpha_x^{e=0}, \beta_x, \kappa_t^{e=0}\right)\right) \cdots p\left(\log f_{x,t}^{e=3} \mid f\left(\alpha_x^{e=3}, \beta_x, \kappa_t^{e=3}\right)\right) \right] \\
&\qquad \cdot \prod_{e=0}^{3} p\left(\alpha_x^e\right) \cdot p\left(\beta_x\right) \\
&\qquad \cdot \prod_{e=0}^{3} p\left(\kappa_t^e \mid c^e\right) \cdot p\left(\kappa_t^e \mid \sigma_k^e\right) \\
&\qquad \cdot \prod_{e=0}^{3} p\left(c^e\right) \cdot p\left(\sigma_k^e\right).
\end{aligned} \tag{26}$$

Here, $p(\cdot)$ denotes a specified probability distribution, while $f(\cdot)$ represents a functional relationship within the process layer. In particular, the function is defined as $f\left(\alpha_x^e, \beta_x, \kappa_t^e\right) = h_{x,t}^e = \alpha_x^e + \beta_x + \kappa_t^e$.

The following model proposes a more general framework of coherence taking into account the velocity of the changes over the time and the change into the time across the subpopulations.

Coherent Bayesian model in both β and κ across educational level ensuring alignment in both age-specific and period-specific fertility trends.

The Bayesian framework presented here introduces coherence simultaneously in the age-specific sensitivity parameter (β) and the period-specific fertility component (κ), ensuring consistency in both dimensions across educational attainment levels. By constraining these parameters, the model preserves the structural integrity of fertility trends, aligning both the rate of fertility change across ages and the temporal evolution of fertility levels among different educational groups.

The mathematical representation of this coherence is defined by the constraints:

$$\beta_x^0 = \beta_x^1 = \beta_x^2 = \beta_x^3,$$

$$\kappa_t^0 = \kappa_t^1 = \kappa_t^2 = \kappa_t^3.$$

This formulation assumes that the variations in fertility over time, captured by κ_t are aligned across all subpopulations, and the differential impact of age on fertility, represented by β_x, evolve consistently across all educational subgroups. Under these assumptions, the coherent Bayesian Lee-Carter model for fertility forecasting is expressed as:

$$\log f_{x,t}^e = \alpha_x^e + \beta_x \kappa_t + \epsilon_{x,t}^e,$$

where α_x^e represents the educational-specific baseline fertility level, and $\epsilon_{x,t}^e$ denotes the residual error term capturing unstructured variability.

This joint coherence model addresses a key limitation in independent fertility forecasting across educational groups, where variations in β_x and κ_t could lead to divergent trends, even if underlying demographic forces remain similar. By enforcing consistency, this approach ensures that fertility improvements are modeled in a demographically plausible manner, avoiding artificial disparities between educational subgroups.

The hierarchical structure of Fertility Model C consists of the following layers:

- **Data Layer (First Stage):**

$$B_{x,t}^e \sim \text{Poisson}\left(fx, t^e \cdot E_{x,t}^e\right). \tag{27}$$

- **Data Layer (Second Stage):**

$$\log f_{x,t}^e \sim \mathcal{N}\left(\alpha_x^e + \beta_x \kappa_t, \Sigma^\mu\right). \tag{28}$$

- **Process Layer:**

$$m_{x,t}^e = \alpha_x^e + \beta_x \cdot \kappa_t. \tag{29}$$

- **Prior Layer (First Stage):**

$$\alpha_x^e \sim \mathcal{N}\left(fa_x^e, \sigma_{a_x^e}\right), \tag{30}$$

$$\tag{31}$$

$$\beta_x \sim \text{Dirichlet}(\alpha_\beta), \tag{32}$$

$$\tag{33}$$

$$\kappa_t \sim \mathcal{N}\left(c + \kappa_{t-1}, \sigma_{\kappa_t}^2\right). \tag{34}$$

- **Prior Layer (Second Stage):**

$$\kappa_{t_0} \sim \mathcal{N}\left(c, \sigma_{\kappa_{t_0}}^2\right). \tag{35}$$

- **Prior Layer (Third Stage):**

$$c \sim \mathcal{N}\left(fc, \sigma_c\right), \tag{36}$$

$$\tag{37}$$

$$\sigma_{\kappa_t} \sim \text{Exponential}(\lambda_{\kappa_t}). \tag{38}$$

This hierarchical structure is graphically represented in Figure 3 for educational groups. Similarly to Figures 1 and 2, in this figure, circles represent variable nodes, while rounded-corner squares denote deterministic or constant nodes. The squares illustrate the repeated structure across age and time. Solid arrows represent stochastic dependencies, and dashed lines indicate deterministic relationships.

The posterior distribution of the model parameters is proportional to the product of the likelihood and prior distributions:

$$\begin{aligned} p\left(\alpha_x^{e=0}, \cdots, \alpha_x^{e=3}, \beta_x, \kappa_t \mid \{\log f_{x,t}^e\}\right) \propto & \prod_{t \in T} \prod_{x \in X} \prod_{e=0}^{3} p\left(\log f_{x,t}^e \mid f\left(\alpha_x^e, \beta_x, \kappa_t\right)\right) \\ & \cdot \prod_{e=0}^{3} p\left(\alpha_x^e\right) \cdot p\left(\beta_x\right) \\ & \cdot p\left(\kappa_t \mid c\right) \cdot p\left(\kappa_t \mid \sigma_\kappa\right) \cdot p\left(c\right) \cdot p\left(\sigma_\kappa\right). \end{aligned} \tag{39}$$

In this setting, $p(\cdot)$ represents a specified probability distribution, while $f(\cdot)$ denotes the functional relationship within the process layer. Specifically, this function is defined as $f\left(\alpha_x^e, \beta_x, \kappa_t\right) = h_{x,t}^e = \alpha_x^e + \beta_x + \kappa_t$.

Forecasting of Fertility Rates

The proposed framework comprises three models—A, B, and C—each representing a different level of structural coherence across educational groups. Model A estimates fertility trends independently for each subgroup, without imposing any coherence constraints. Model B introduces coherence in the age-specific sensitivity parameter (β) by modeling it jointly across educational groups. Model C further extends this structure by also introducing coherence in the period-specific component (κ), thereby ensuring alignment in both age-specific and temporal fertility patterns across subpopulations.

In Model A, the kappa parameter is forecasted independently as a random walk with drift. Model B introduces coherence by incorporating a multivariate normal structure:

$$\begin{bmatrix} \kappa_t^0 \\ \kappa_t^1 \\ \kappa_t^2 \\ \kappa_t^3 \end{bmatrix} \sim \text{MVN}_4 \left(\begin{bmatrix} c_0^0 \\ c_0^1 \\ c_0^2 \\ c_0^3 \end{bmatrix} + \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \\ c_{41} & c_{42} & c_{43} & c_{44} \end{bmatrix} \begin{bmatrix} \kappa_{t-1}^0 \\ \kappa_{t-1}^1 \\ \kappa_{t-1}^2 \\ \kappa_{t-1}^3 \end{bmatrix}, \Sigma_\kappa \right) \tag{40}$$

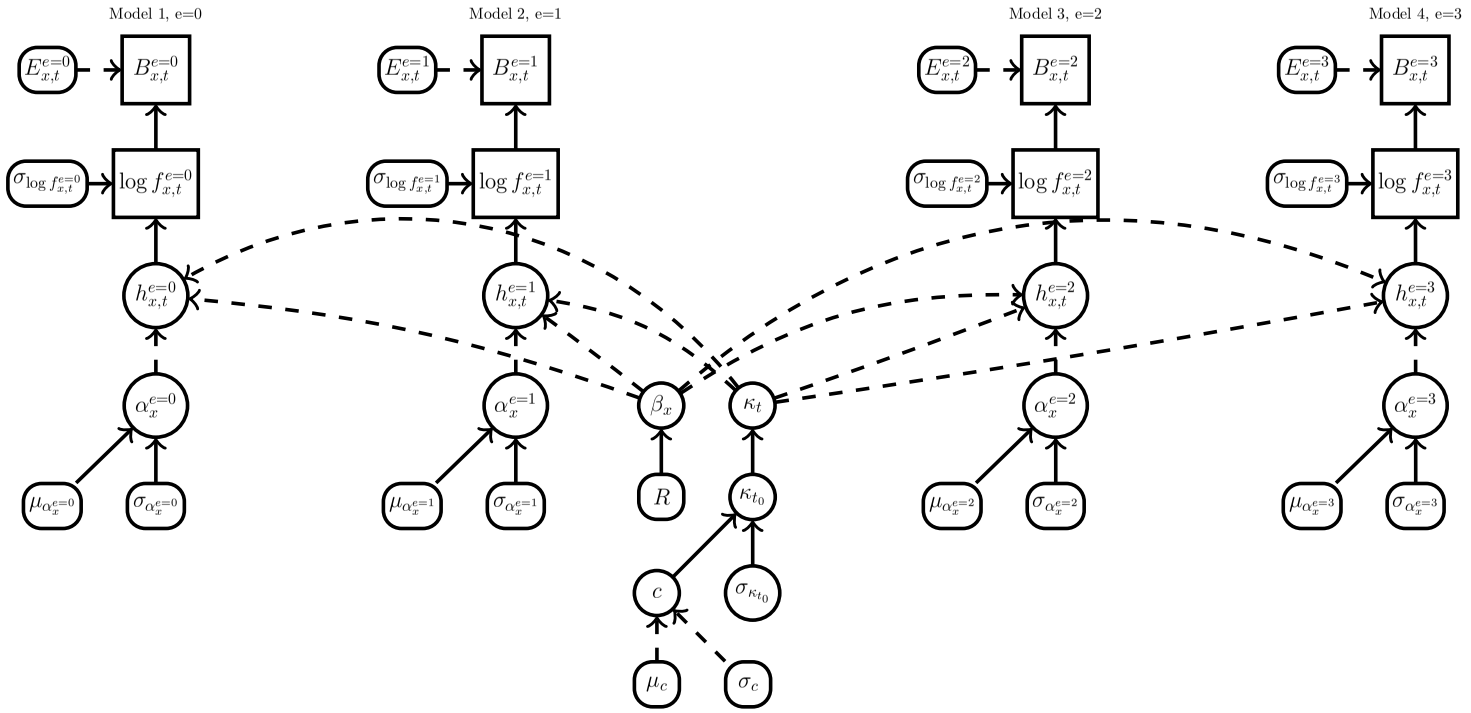

Figure 3: Fertility Model C: Coherent Lee–Carter Model in β and κ across educational attainment groups.

Model C further extends this structure by modeling kappa parameters coherently across all educational groups but the relation is similar to Model A. This coherence ensures that variations in fertility trends are captured systematically across different levels of education, allowing for a more unified projection framework.

In these equations (40), the drift vector c_0 accounts for systematic temporal changes, while the covariance matrix Σ_κ captures the relationships among different demographic subgroups. By incorporating these dependencies, the model not only reflects immediate variations but also accounts for lagged effects across educational categories, enhancing the accuracy of long-term fertility projections.

The term MVN_4 denotes a four-dimensional multivariate normal distribution governed by the precision matrix Σ_κ. The elements c_{ij} (where $i, j = 1, \ldots, 4$) define the autoregressive structure, ensuring that fertility dynamics evolve according to a VAR(4) process. Additionally, the drift terms c_0 embed underlying long-term trends, strengthening the model's capability to generate reliable forecasts across educational groups.

Prior Specification

This section outlines the prior distributions used in the hierarchical Bayesian models for forecasting fertility rates. Weakly informative priors are specified to allow the data to dominate inference while providing regularisation in the estimation process. The models differ in the level of structural coherence imposed across educational groups, but share a common prior structure for core parameters. The age-specific effects (α_x^e) are assigned normal priors to capture baseline fertility patterns, while the sensitivity parameters (β_x) and temporal components (κ_t^e) follow priors appropriate to the coherence assumptions in each model. When parameters are shared across groups (e.g. in Models B and C), hierarchical priors are employed to model their joint distribution. The full specification of priors is assigned as follows:

$$
\begin{aligned}
\theta_i &\sim \mathcal{N}(\mu_\theta, \sigma_\theta), \quad \mu_\theta = 0, \quad \sigma_\theta = 10 \\
\beta_i &\sim \text{Dirichlet}(\alpha_\beta), \quad \alpha_\beta = 1 \\
\tau_i &\sim \text{Exponential}(\lambda_\tau), \quad \lambda_\tau = 0.1 \\
\delta_i &\sim \mathcal{N}(\mu_\delta, \sigma_\delta) \\
\omega_k &\sim \text{Exponential}(\lambda_\omega), \quad \lambda_\omega = 0.1 \\
\Psi_i &\sim \text{Exponential}(\lambda_\Psi), \quad \lambda_\Psi = 1
\end{aligned}
$$

The intercept terms θ_i are assigned a Normal prior with moderate variance to accommodate variability in baseline fertility levels. The beta parameters, β_i, follow a Dirichlet distribution to ensure non-negativity and sum-to-one constraints, making them suitable for fertility share modeling. The variance components τ_i, ω_k, and Ψ_i follow Exponential priors to regulate dispersion and prevent overfitting, while the drift term δ_i allows for gradual shifts in fertility trends.

Model Evaluation and Validation

To evaluate the performance of the proposed Bayesian fertility models, predicted fertility rates are compared with observed values. The primary metric used is the root mean square error (RMSE), which measures the deviation between the logarithm of observed and estimated fertility rates. RMSE is calculated separately for each educational group to assess predictive accuracy at the subgroup level. An overall RMSE is then obtained by averaging across groups, providing a summary measure for model comparison. The model with the lowest overall RMSE is considered to offer the best predictive performance. The RMSE is defined as:

$$RMSE_{LC} = \sqrt{\frac{1}{N} \sum_{x} \sum_{t} (\log f^{e}_{x,t} - \log \hat{f}^{e}_{x,t})^2} \tag{41}$$

where $f^{e}_{x,t}$ and $\hat{f}^{e}_{x,t}$ denote the observed and estimated fertility rates for age x, time t, and education level e, respectively. The total number of observations is N, with summation over all ages and periods.

For further validation, we conduct an ex-post evaluation using a truncated dataset from 1998 to 2013. Forecasts generated from this period are compared to actual fertility rates beyond 2013, providing insights into long-term predictive accuracy. Additionally, visual diagnostics, such as fertility trajectories and residual analysis, complement numerical assessments by identifying systematic patterns and areas where model refinements may be necessary.

For further validation, we conduct an ex-post evaluation using a truncated dataset from 1998 to 2013. Forecasts generated from this period are compared to actual fertility rates beyond 2013, providing insights into long-term predictive accuracy. Additionally, visual diagnostics, such as fertility trajectories and residual analysis, complement numerical assessments by identifying systematic patterns and areas where model refinements may be necessary. These evaluations ensure that the reconstructed fertility rates are robust and reliable for subsequent demographic analyses. After applying the different models described, this validation is used to assess the accuracy of the fertility rates derived from multiple data sources.

Data and Model Implementation

This study relies on data from administrative registers and reconstructed population estimates to compute fertility rates disaggregated by age, sex, and educational attainment for the period 1998 to 2018. Birth counts were obtained from the administrative records of DANE and classified using the ISCED 2011 framework into four educational levels: no formal education, primary education completed, secondary education completed, and post-secondary education. The corresponding population denominators were reconstructed following the methodology described in Appendix A.

The methodology also draws on insights from Ellison et al. (2023), who highlight the importance of reconciling different data sources to ensure demographic consistency. By integrating multiple datasets including census-derived estimates and vital registration data, the approach captures distinct fertility behaviors among different population subgroups. This ensures that fertility rates remain aligned with observed demographic processes particularly by linking births directly to mothers through household survey data. The combination of these data sources enhances the robustness of fertility estimate preventing inconsistencies that may arise when relying on a single data source.

Period Versus Cohort modeling and forecast

Fertility rates are constructed from two complementary data sources, responding to the absence of official annual population estimates disaggregated simultaneously by age, sex, and education. The first data source is a population reconstruction developed by Sanchez-Segura (2025), which applies a Holt-filter variant of the multistate approach to estimate educational composition over time using census and survey benchmarks. This reconstruction generates consistent annual estimates of population exposure by age, sex, and education for the period 1998–2018 (see Appendix A). The second data source consists of administrative birth records from the Colombian national statistical office (DANE), which provide detailed information on the age and educational attainment of mothers. These records are used to compute age- and education-specific fertility rates for the same period. Although coverage and quality of educational reporting in birth records are not uniform across all years, quality checks and harmonisation procedures were applied to ensure internal consistency and comparability across time.

While period-based fertility indicators such as the total fertility rate (TFR) -we plan to present the TFR for the time of the conference- are known to be influenced by tempo effects particularly in settings with delayed childbearing the use of period-based modelling is justified in this study for both practical and substantive reasons. First, reliable cohort-parity-specific data disaggregated by education are not available, making cohort-based approaches infeasible in the Colombian context. Second, the aim of this study is to generate short- to medium-term forecasts that are responsive to recent trends in educational composition and fertility behaviour. Period models offer a more timely and flexible foundation for this purpose. Although tempo distortions are acknowledged as a limitation particularly their potential to misrepresent fertility quantum the benefits of period-based estimation outweigh these concerns in data-constrained environments. Forecast accuracy is evaluated using internal validation, through holdout samples and comparisons between predicted and observed fertility rates by education group, using root mean square error (RMSE) and predictive interval coverage as key metrics.

To illustrate these fertility trends, Figure 4 presents the logarithm of fertility rates by age group, educational attainment, and year, with the color gradient ranging from dark blue (1998) to yellow (2018). The four panels correspond

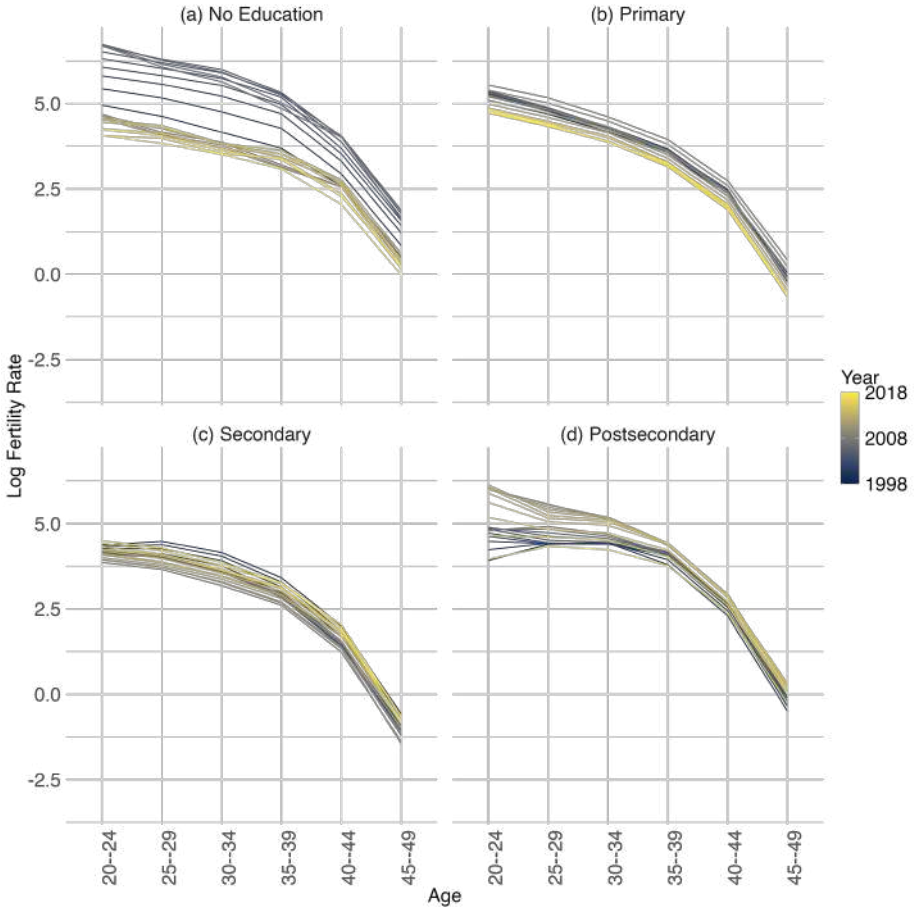

Figure 4: Log of fertility rates by age (x-axis), educational attainment (first row: no formal education and primary; second row: secondary and post-secondary), and year (color gradient from dark blue representing 1998 to yellow representing 2018) based on the population reconstruction of Sanchez-Segura (2025).

to different educational levels: no education, primary, secondary, and post-secondary education. The first row contains the no education and primary education groups, while the second row corresponds to secondary and post-secondary education. The y-axis represents the logarithm of fertility rates, and the x-axis denotes five-year age groups.

At the beginning of the period, in 1998, the highest fertility rates are observed among individuals with no formal education, followed by those with primary education. The secondary and post-secondary education groups exhibit comparatively lower fertility rates. Across all education levels, the age group 20–24 records the highest fertility rates, followed by 25–29, with a decline in older age groups. By 2018, fertility rates have decreased across all education levels, with more pronounced reductions among younger cohorts. While the general shape of the fertility-age profile remains similar over time, the decline in fertility is most evident in the lower education groups. The largest reduction in fertility rates between 1998 and 2018 is observed among individuals with no formal education, indicating a substantial shift in reproductive patterns within this group.

A closer examination of each educational category reveals distinct patterns of fertility decline. The first panel, corresponding to individuals with no formal education, shows that in 1998, fertility rates peak at ages 20–24, followed by a decline from age 30 onwards. The transition from dark blue to yellow illustrates a reduction in fertility rates across all age groups by 2018, with the most significant declines occurring at younger ages.

Similarly, the second panel, representing individuals with primary education, exhibits a fertility peak at ages 20–24 and a steady decline at older ages. However, by 2018, a reduction in fertility is observed across all age groups, though the magnitude of change is slightly lower compared to the no education group. This pattern suggests a consistent decline in fertility across lower education levels, with variations in the extent of change over time.

The third panel, corresponding to secondary education, shows a fertility peak at ages 20–24 in 1998, followed by a sharp decline beyond age 30. Compared to lower education groups, fertility rates in this category are lower at all ages. By 2018, a further reduction in fertility rates is observed, particularly in younger age groups. The fourth panel, representing post-secondary education, exhibits a distinct fertility-age profile, with the peak fertility rate occurring at ages 25–29 instead of 20–24. In 1998, fertility rates in this group are the lowest among all education levels. By 2018, further declines are observed, particularly among younger women, with fertility rates becoming more concentrated in the 25–34 age range.

A comparison across educational levels shows a negative association between educational attainment and fertility levels. In 1998, fertility rates are highest among individuals with no formal education and lowest among those with post-secondary education. The peak fertility age shifts from 20–24 in lower education groups to 25–29 in the post-secondary education group. By 2018, the general decline in fertility is evident across all groups, with reductions in early-age fertility more pronounced among secondary and post-secondary education groups,

while overall fertility reductions are greater in lower education groups.

The trends indicate a divergence in fertility patterns based on educational attainment. The decline in fertility is more substantial among individuals with no formal and primary education, while secondary education exhibits a slower rate of decline. The post-secondary group, in contrast, shows a more pronounced shift in fertility age distribution, with a relative increase in fertility at older ages compared to previous years. These variations suggest differences in reproductive timing and fertility behavior across educational groups, reflecting distinct socio-economic and demographic factors influencing childbearing decisions.

Model Implementation

The fertility models proposed in this paper were estimated using the No-U-Turn Sampler (NUTS), an adaptive variant of the Hamiltonian Monte Carlo (HMC) algorithm introduced by Hoffman and Gelman (2014). NUTS is particularly appropriate for high-dimensional hierarchical Bayesian models, as it adaptively selects the path length during sampling, thereby enhancing both convergence and computational efficiency.

All models were implemented in the Julia programming language using the Turing.jl probabilistic programming framework, which supports flexible specification and inference for Bayesian models. Posterior samples were drawn using multiple chains, and standard convergence diagnostics including the potential scale reduction factor ($\hat{R}$) and effective sample size were evaluated to ensure the reliability of the estimates.

All model code, data processing scripts, and replication materials are publicly available at: `https://github.com/anfesanz/Coherent_Fertility`.

Results

The results of this paper provide a comprehensive assessment of fertility forecasting models based on Bayesian hierarchical frameworks. Using the No-U-Turn Sampler (NUTS) for Bayesian inference, the model parameters are estimated to ensure coherent posterior distributions of fertility rates across educational levels. A key feature of the approach is the imposition of coherence on β across educational strata, maintaining consistent fertility trends and reducing the likelihood of implausible divergences in forecasts among subpopulations. Unlike non-coherent models, which estimate β independently for each educational group, the coherent Bayesian approach enforces a structured relationship that enhances stability and interpretability in the projections. The following sections detail the model's performance, beginning with an evaluation of model fit, followed by an assessment of both out-of-sample and in-sample forecasts. A comparative analysis of different model specifications further elucidates the predictive accuracy across demographic subgroups. Additionally, Appendix B provides an in-depth examination of the model's goodness of fit, including parameter estimation diagnostics and posterior evaluations for the Bayesian models applied in

Table 1: Model fit evaluation: Root Mean Square Error (RMSE) of log fertility rates, stratified by education level and Models A, B, and C (1998–2018).

Education Level	**Model A**	**Model B**	**Model C**
No Formal Education	0.116	0.123	0.025
Primary	0.132	0.113	0.133
Secondary	0.112	0.135	0.043
Post-Secondary	0.159	0.148	0.083

this paper.

Model Fit Evaluation (1998–2018)

After estimating the Bayesian hierarchical individual and coherent models, the RMSE of the estimations was calculated against the observed values. Table 1 presents this comparison for each education level, while Table 2 provides an overall RMSE estimation by model to facilitate model comparison.

Table 1 displays the RMSE for three models (Model A, Model B, and Model C) across different education levels: No Formal Education, Primary, Secondary, and Post-Secondary. The RMSE quantifies the prediction error of a model, with lower values indicating better performance. The table includes three model structures: a single-group model that estimates parameters separately for each education level, a model that assumes a common beta coefficient across education levels, and a model that incorporates both beta and kappa coefficients to account for education level heterogeneity.

The RMSE values indicate that Model C consistently outperforms Models A and B across most education levels. This suggests that incorporating both beta and kappa coefficients to account for education level improves model accuracy compared to models that do not incorporate coherence across levels.

For individuals without formal education, Model C (0.025) achieves a significantly lower RMSE than Model A (0.116) and Model B (0.123). Among those with primary education, Model B (0.113) has the lowest RMSE, followed closely by Model A (0.132) and Model C (0.133). For individuals with secondary education, Model C (0.043) again demonstrates the lowest RMSE, while Model A (0.112) and Model B (0.135) exhibit higher errors. In the post-secondary education group, Model C (0.083) yields the lowest RMSE, followed by Model B (0.148) and Model A (0.159).

These results suggest that a model incorporating both beta and kappa coefficients (Model C) provides better predictive accuracy across most education levels.

Table 2 presents the average RMSE across all education levels for each model. Model C (0.071) achieves the lowest RMSE, suggesting superior performance in capturing fertility rate trends. In contrast, Models A and B both exhibit higher average RMSE values (0.130), indicating less accurate fits to the observed data.

Table 2: Model fit evaluation: Root Mean Square Error (RMSE) of log fertility rates by model, from 1998 to 2018.

Model	**RMSE**
Model A	0.130
Model B	0.130
Model C	0.071

Overall, Model C, which incorporates both beta and kappa coefficients, demonstrates the lowest overall error. This highlights its effectiveness in modeling log fertility rates compared to Models A and B, which do not incorporate coherence across education levels. The superior performance of Model C suggests that accounting for parameter consistency across education groups improves predictive accuracy.

Model Fit to Observed Data in 2018

This section presents a comparative analysis of model fits for fertility rates in 2018, stratified by education level. Figure 5 displays the results of three models (Model A, Model B, and Model C, as described in the previous section) across four educational strata: no formal education, primary education, secondary education, and post-secondary education. The x-axis represents age groups, while the y-axis represents the logarithm of fertility rates. Observed data are depicted using black lines, model fits using grey lines, and the 95 percent credible intervals are shaded in grey.

Model A exhibits a consistent trend across all educational levels. The logarithm of fertility rates declines with increasing age, with slight variations between education groups. The fit closely follows the observed data, particularly for ages 25 to 40, where the credible interval remains narrow. Beyond age 40, the model shows increased uncertainty, reflected in the widening of the grey-shaded area.

For individuals with no formal education, Model A aligns with the observed data, though minor deviations emerge at older ages. The primary education group follows a similar pattern, with relatively small discrepancies. In the secondary and post-secondary education groups, the fit remains robust but shows slight overestimation at younger ages and underestimation at older ages. These patterns suggest that Model A provides a relatively stable representation of fertility rates across educational levels.

Building on the observations from Model A, Model B provides a slightly different fit. The general declining trend in fertility rates with age is maintained across education groups, but the model exhibits greater uncertainty beyond age 35, as indicated by a wider credible interval. The Bayesian fit follows the observed data more closely in some cases, particularly for the secondary education group, but diverges slightly for those with no formal education and post-secondary education.

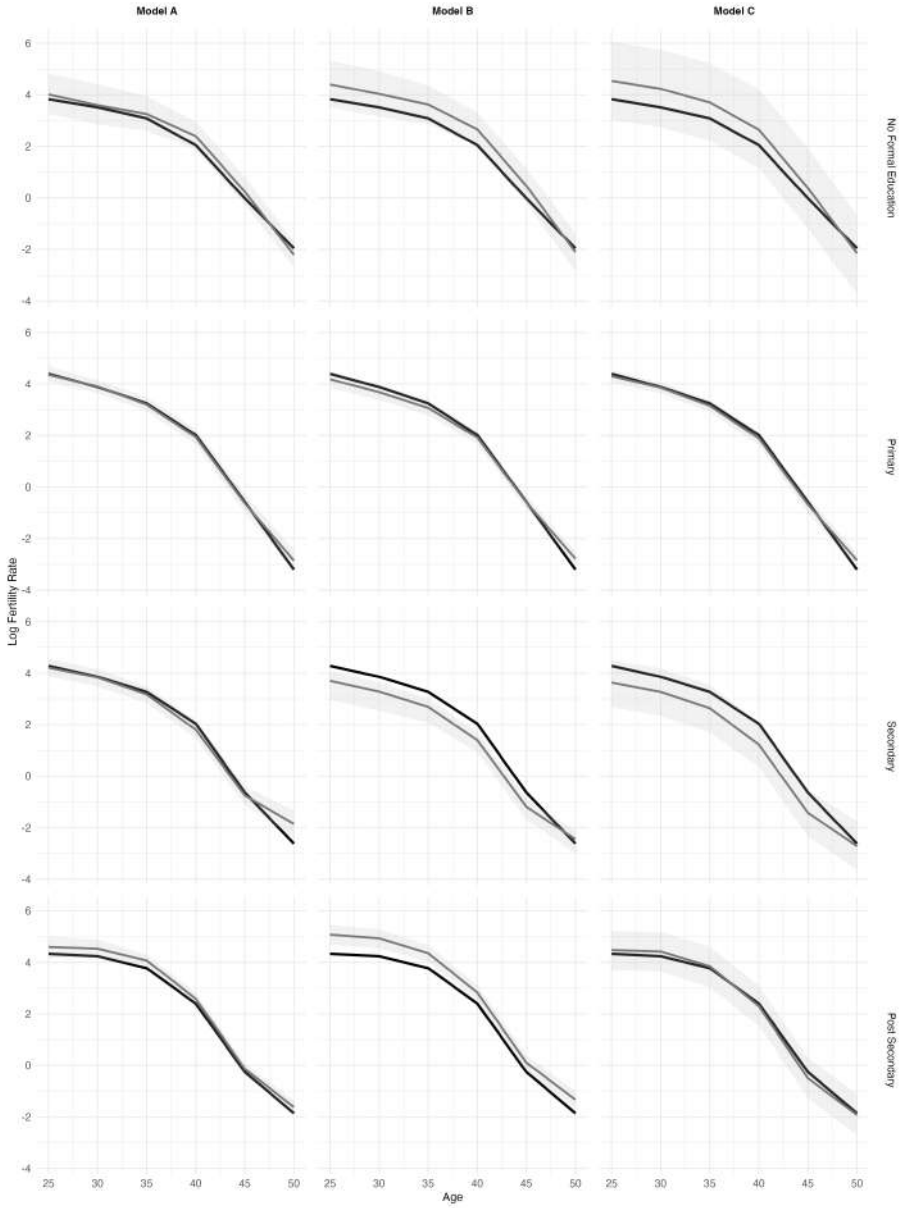

Figure 5: Model fits for 2018, stratified by education level: no formal education (row 1), primary education (row 2), secondary education (row 3), and post-secondary education (row 4). Three models are displayed: Model A (column 1, non coherent), Model B (column 2, β coherent), and Model C (column 3, β and κ coherent). The x-axis represents age groups, and the y-axis represents the logarithm of fertility rates. Black lines indicate observed data, grey lines show the Bayesian fit, and the grey shaded area represents the 95% credible interval.

In the no formal education and primary education groups, Model B tends to overestimate fertility rates at younger ages, leading to a small but noticeable deviation. The fit improves in middle ages (30–40) but diverges again at later ages. In the secondary education group, Model B aligns well with the observed data, reducing overestimation at younger ages compared to Model A. However, in the post-secondary education group, the model slightly underestimates fertility rates across all ages. These variations highlight differences in the capacity of Models A and B to capture fertility trends accurately.

Extending the comparison, Model C follows a similar pattern to Models A and B but with notable distinctions. The fit closely follows the observed data at younger ages. However, in contrast to Model A, the credible intervals remain relatively wide across all age groups, indicating greater uncertainty in model estimates.

For individuals with no formal education, Model C follows the observed data but slightly underestimates fertility rates at ages 25–35. The primary education group shows a reasonable fit, but like Model B, Model C slightly overestimates fertility rates at younger ages. In contrast, for the secondary and post-secondary education groups, the model tends to underestimate fertility rates beyond age 35, with a more noticeable divergence compared to Model A and Model B. These results suggest that Model C introduces greater variability in the estimates across age groups.

A direct comparison between Model A and Model B reveals differences in uncertainty and fit precision. The primary distinction lies in the width of the credible intervals and variations in the fit at different ages. Model B exhibits slightly larger uncertainty, particularly beyond age 35, whereas Model A maintains narrower credible intervals. Model B aligns better with the secondary education group but diverges from the observed data in the no formal education and post-secondary education groups, where Model A provides a more stable fit. These findings indicate that Model A may be preferable in cases where greater precision is required.

Comparing Model A and Model C, both models exhibit similar overall trends. However, Model C introduces higher uncertainty across all age groups. While Model A maintains a more consistent fit with narrower credible intervals, Model C follows the observed data more closely at younger ages but diverges at older ages. Model A provides more stable estimates, whereas Model C reflects greater variability in the underlying data. This distinction suggests that Model C may be more sensitive to fluctuations in fertility rates at younger ages.

Model B and Model C share several tendencies, particularly in their overestimation of fertility rates at younger ages and underestimation at older ages for certain education levels. However, Model C exhibits wider credible intervals across all ages, while Model B maintains a more precise fit in the middle age range. The performance of both models is relatively similar, though Model B appears to manage uncertainty more effectively. These observations suggest that Model B may be more reliable in settings where middle-age fertility estimates are of particular interest.

Overall, Model C provides the most stable fit with the least uncertainty, par-

ticularly in middle-age groups. Model B and Model C exhibit similar strengths and weaknesses but introduce greater uncertainty at later ages, particularly in education groups with higher variability in fertility rates. The selection of the most appropriate model depends on the priority given to stability versus the ability to capture variability at younger ages. These findings underscore the importance of selecting a model based on the specific requirements of fertility analysis across education levels.

To complement the visual assessment of model fit, Table 3 reports the credible interval spread ratios, calculated as the ratio between the 95th percentile and the median of the posterior distribution of forecasted log fertility rates for the period 2014–2018. The results are disaggregated by education level and model. Across all education groups, Model A generally yields wider credible intervals compared to Models B and C, particularly for females with no formal education (1.47) and post-secondary education (1.25). Model B produces narrower intervals for most groups relative to Model A but shows slightly higher spread ratios for females with secondary education (1.09). Model C exhibits the lowest spread ratios in three out of four education levels, most notably among females with post-secondary education (0.65), indicating reduced posterior dispersion in those forecasts. Differences in spread ratios across models reflect variation in the extent of posterior uncertainty introduced by model structure and coherence assumptions.

Table 3: Credible interval spread ratios (95th percentile / median) of the posterior distribution of log fertility rate forecasts for 2014–2018 (trained on 1998–2013 data) education, and model.

	Model A	Model B	Model C
Females, No Formal Education	1.47	1.24	1.36
Females, Primary	0.88	0.89	0.79
Females, Secondary	0.99	1.09	1.00
Females, Post-Secondary	1.25	1.04	0.65

Forecasts of Fertility Rates in 2028

This section presents a comparative analysis of fertility rate forecasts for the year 2028, based on three different models: Model A, Model B, and Model C. An evaluation of forecast accuracy over the full forecasting period will be provided as an appendix. The forecasts are disaggregated by education level, including no formal education, primary education, secondary education, and post-secondary education. Figure 6 displays the projected fertility rates, with age groups on the x-axis and the logarithm of fertility rates on the y-axis. Black lines represent the observed data from 2018, grey lines indicate the Bayesian forecasts, and the grey shaded areas denote the 95% credible intervals.

Model A projects a consistent decline in fertility rates with increasing age across all education levels. The credible intervals remain relatively narrow for

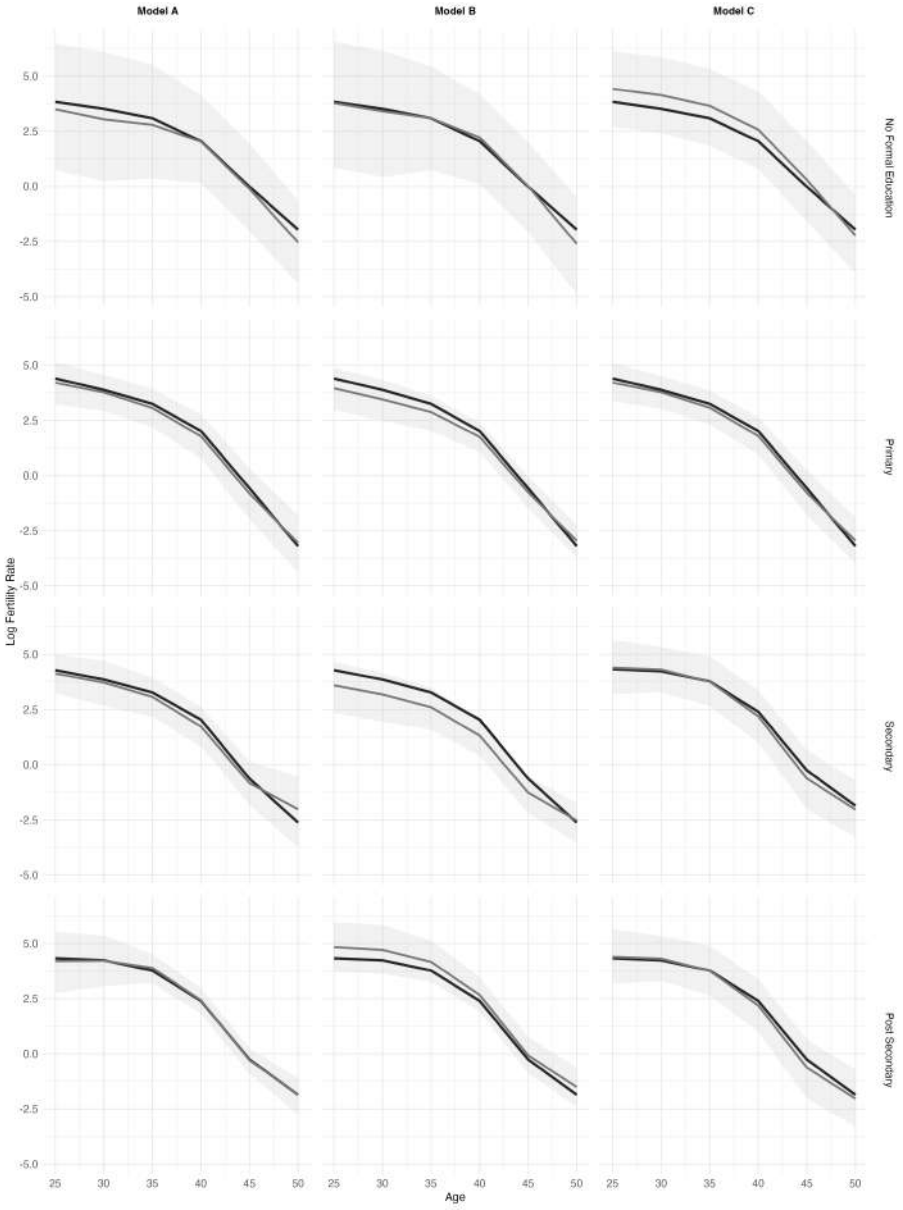

Figure 6: Fertility forecasts for 2028, stratified by education level: no formal education (row 1), primary education (row 2), secondary education (row 3), and post-secondary education (row 4). Three models are displayed: Model A (column 1, non coherent), Model B (column 2, β coherent), and Model C (column 3, β and κ coherent). The x-axis represents age groups, and the y-axis represents the logarithm of fertility rates. Black lines indicate observed data from 2018, grey lines show the Bayesian forecast, and the grey shaded area represents the 95% credible interval.

younger ages but widen progressively beyond age 40, indicating increased uncertainty in the forecast. The model appears to align closely with observed data from 2018, particularly for individuals with primary and secondary education. However, for individuals with no formal education and post-secondary education, minor deviations are observed, especially at older ages where uncertainty is more pronounced.

Model B follows a similar overall pattern to Model A but exhibits wider credible intervals, particularly beyond age 35. The projected fertility rates align with the observed data from 2018 for the secondary education group but tend to slightly overestimate fertility at younger ages and underestimate it at older ages for the primary and post-secondary education groups. The forecasts for individuals with no formal education show increased uncertainty, with wider credible intervals compared to Model A, suggesting greater variability in the underlying fertility patterns.

Model C also predicts a decreasing trend in fertility rates with age, but with some notable distinctions. The credible intervals remain wider across all age groups compared to Models A and B, reflecting greater uncertainty in the forecasts. While the model aligns well with observed data at younger ages, it tends to underestimate fertility rates beyond age 35 for all education groups. The projections for individuals with no formal education show the highest level of uncertainty, with a wider spread of credible intervals.

A comparison between Model A and Model B reveals that both models capture the general trend of fertility decline with age. However, Model B exhibits greater uncertainty, as indicated by its wider credible intervals. Model A provides a more stable forecast, particularly for individuals with primary and secondary education, whereas Model B demonstrates greater variability, especially for individuals with no formal education and post-secondary education. While both models align with observed data from 2018, Model A appears to maintain a more consistent fit at younger ages.

When comparing Model A and Model C, Model C introduces greater uncertainty across all age groups, as indicated by wider credible intervals. While Model A provides a more stable forecast, Model C follows the observed data more closely at younger ages but diverges at older ages, particularly for the secondary and post-secondary education groups. The choice between these models depends on whether stability (Model A) or flexibility in capturing variations at younger ages (Model C) is prioritized.

Model B and Model C exhibit similar tendencies, particularly in their overestimation of fertility rates at younger ages and underestimation at older ages for certain education groups. However, Model C shows the widest credible intervals across all ages, suggesting greater uncertainty in its projections. Model B maintains a slightly more precise fit in the middle-age range but still exhibits greater variability than Model A. The performance of both models is relatively similar, though Model B appears to offer slightly more stability than Model C.

Model A provides the most stable fertility forecast with the least uncertainty, particularly in middle-age groups. In contrast, Model B and Model C display similar strengths and weaknesses but introduce greater uncertainty at older

ages, especially in education groups with higher variability in fertility rates. The choice of the most appropriate model depends on balancing stability with flexibility in capturing variations at younger ages.

In-Sample Forecasting Performance (1998–2018)

To evaluate model performance, Figures 7–10 present the age-specific logarithm of fertility rates across different education levels, along with in-sample forecasts for 2014–2018. These forecasts provide insight into how well each model replicates observed trends and their associated levels of uncertainty.

This section provides a comparative analysis of model-based forecasts for age-specific fertility rates (ASMR) from 1998 to 2018, with a focus on in-sample forecasting for 2014–2018. The analysis is structured by six age groups: 25–29, 30–34, 35–39, 40–44, 45–49, and 50–54. Observed data are represented by black lines, while Model A, Model B, and Model C are indicated by red dashed, dark blue dotted, and light blue dash-dotted lines, respectively. The shaded areas denote the credible intervals for each model's forecasts, providing insight into the uncertainty associated with each prediction.

Model A follows the observed data closely across all age groups, with a gradual increase in log fertility rates until approximately 2012, followed by a decline in most age groups. The model captures the long-term trend well, though some deviations occur in the later years. The credible intervals remain relatively narrow, indicating lower uncertainty in the estimates. However, slight underestimations are observed for the 50–54 age group in the later years.

Model B introduces a different pattern, showing greater fluctuations in fertility rates. While the model aligns with observed data in earlier years, deviations become more pronounced from 2010 onwards, particularly for the 25–29 and 30–34 age groups. The model's credible intervals are slightly wider than those of Model A, reflecting increased uncertainty. In the 50–54 age group, Model B tends to overestimate fertility rates compared to observed data.

Model C exhibits the widest credible intervals among the three models, indicating higher uncertainty in its forecasts. The model generally follows the observed trend but demonstrates more pronounced deviations in the middle age groups (35–39 and 40–44), where fertility rates appear overestimated. Additionally, Model C captures some of the short-term fluctuations more effectively than the other models, particularly for the 45–49 age group, but at the cost of increased uncertainty.

Comparing Model A and Model B, both models capture the long-term fertility trend, but Model B exhibits greater variability in the forecasts. Model A provides a more stable fit with narrower credible intervals, whereas Model B appears more sensitive to short-term fluctuations. In the 50–54 age group, Model B tends to overestimate fertility rates more than Model A.

Model C introduces higher uncertainty across all age groups compared to Model A. While Model A provides more stable estimates with narrower credible intervals, Model C follows the observed data more closely at certain points,

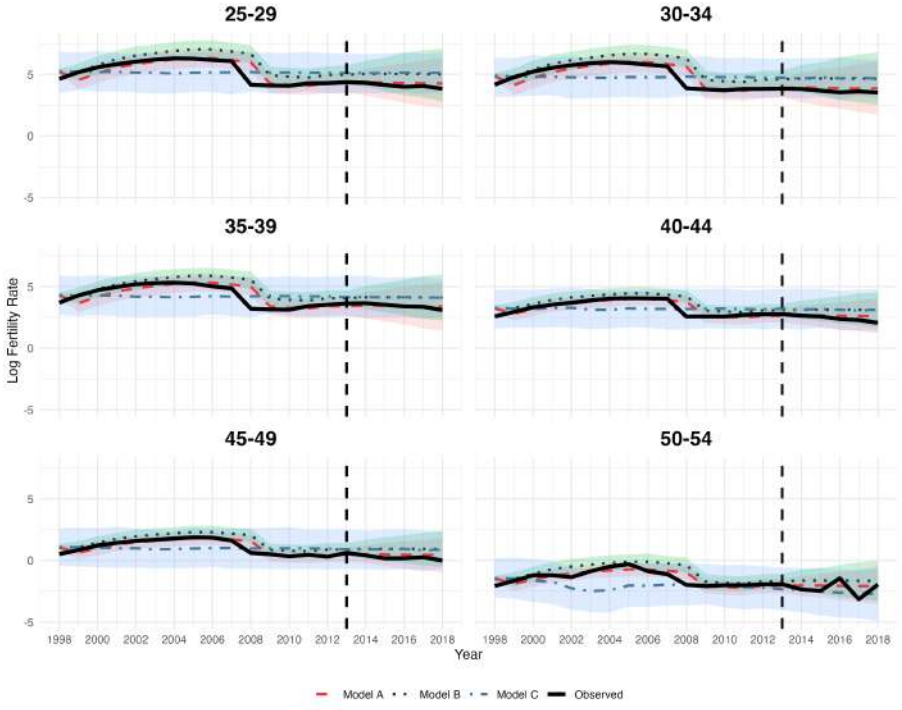

Figure 7: Age-specific logarithm of fertility rates (1998–2018) with in-sample forecasting for 2014–2018 for individuals with no formal education. Black lines represent observed data, red dashed lines correspond to Model A, dark blue dotted lines to Model B, and light blue dash-dotted lines to Model C.

particularly in the younger age groups (25–29 and 30–34). However, at older ages, Model C tends to deviate more significantly from observed fertility rates.

Model B and Model C share some similarities in capturing fluctuations, but Model C introduces wider credible intervals, indicating greater uncertainty. Model B tends to follow observed trends more closely in the middle age groups, while Model C provides a more flexible fit but with increased uncertainty. In the 50–54 age group, Model C underestimates fertility rates compared to Model B.

Overall, Model A provides the most stable forecasts with the least uncertainty, particularly in the middle-age groups. Model B and Model C exhibit similar strengths and weaknesses but introduce greater uncertainty at older ages, particularly in education groups with higher variability in fertility rates. The selection of the most appropriate model depends on the balance between stability and the ability to capture short-term fluctuations.

Figure 8 displays the results, where black lines represent observed data, red dashed lines correspond to Model A, dark blue dotted lines to Model B, and light blue dash-dotted lines to Model C. The shaded areas indicate the credible intervals for each model's forecasts.

Model A closely follows the observed data across all age groups, displaying a stable pattern in log fertility rates with minor fluctuations. The model captures long-term trends effectively, particularly for the 25–34 and 35–39 age groups. The credible intervals remain narrow, indicating lower uncertainty in the estimates. However, in the 50–54 age group, Model A tends to slightly underestimate fertility rates in later years.

Model B exhibits a pattern similar to Model A but with slightly wider credible intervals, particularly in the 35–39 and 40–44 age groups. The model generally aligns with observed data but demonstrates small deviations from 2010 onward, particularly in the younger age groups (25–29 and 30–34). The credible intervals widen further in the 50–54 age group, indicating increased uncertainty compared to Model A.

Model C produces forecasts with the widest credible intervals among the three models, reflecting higher uncertainty. The model follows observed trends closely in earlier years but diverges slightly in the post-2010 period. Model C captures some fluctuations more effectively than Models A and B, particularly in the 45–49 age group, but at the cost of increased uncertainty. The divergence is more evident in the 50–54 age group, where Model C underestimates fertility rates.

Comparing Model A and Model B, both models effectively capture the long-term fertility trend, though Model B exhibits slightly larger uncertainty. Model A provides a more stable fit with narrower credible intervals, while Model B appears more responsive to short-term variations. Model B exhibits minor overestimation in some younger age groups and underestimation in older groups.

Model C introduces greater uncertainty across all age groups compared to Model A. While Model A maintains narrower credible intervals, Model C follows observed data more closely at certain points, particularly in younger age groups. However, at older ages, Model C shows more pronounced deviations

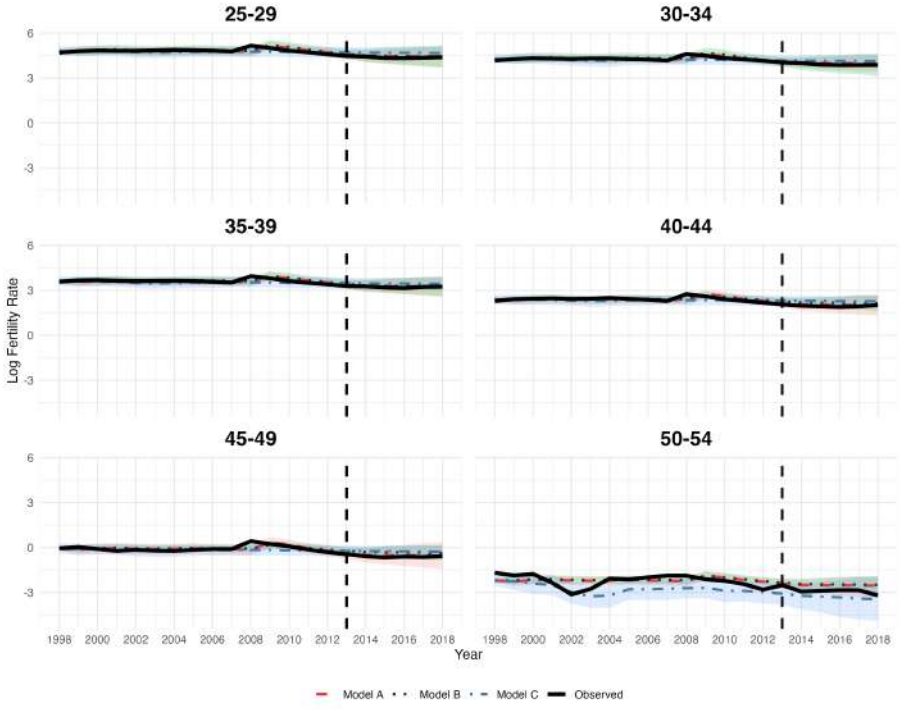

Figure 8: Age-specific logarithm of fertility rates (1998–2018) with in-sample forecasting for 2014–2018 for individuals with primary education. Black lines represent observed data, red dashed lines correspond to Model A, dark blue dotted lines to Model B, and light blue dash-dotted lines to Model C.

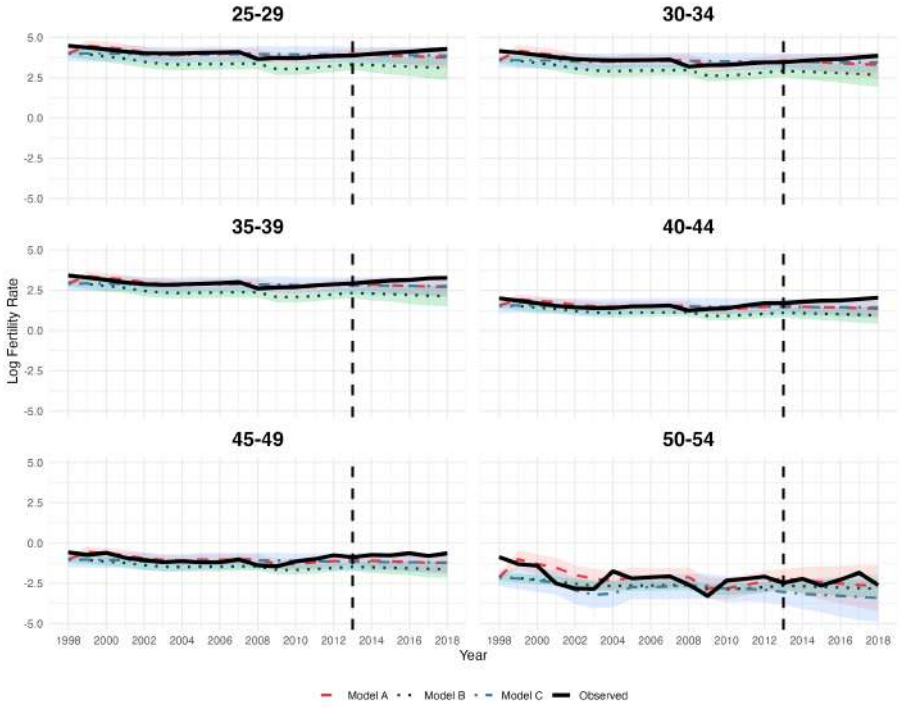

Figure 9: Age-specific logarithm of fertility rates (1998–2018) with in-sample forecasting for 2014–2018 for individuals with secondary education. Black lines represent observed data, red dashed lines correspond to Model A, dark blue dotted lines to Model B, and light blue dash-dotted lines to Model C.

from observed fertility rates.

Model B and Model C share some similarities in capturing fluctuations, but Model C introduces wider credible intervals, indicating greater uncertainty. Model B aligns more closely with observed data in middle age groups, whereas Model C provides more flexibility in its forecasts but at the expense of stability. In the 50–54 age group, Model C underestimates fertility rates more significantly compared to Model B.

Overall, Model A provides the most stable forecasts with the least uncertainty across age groups. Model B and Model C exhibit similar trends but introduce greater uncertainty, particularly in the older age groups. The choice of the most appropriate model depends on the balance between stability and flexibility in capturing fertility trends.

Figure 9 displays the results, where black lines represent observed data, red dashed lines correspond to Model A, dark blue dotted lines to Model B, and light blue dash-dotted lines to Model C. The shaded areas indicate the credible intervals for each model's forecasts.

Model A closely follows the observed data across all age groups, showing a stable pattern in log fertility rates. The model captures long-term trends effectively, particularly for the 25–34 and 35–39 age groups. The credible intervals remain narrow, indicating lower uncertainty in the estimates. However, in the 50–54 age group, Model A tends to slightly underestimate fertility rates in later years.

Model B produces forecasts similar to Model A but with slightly wider credible intervals, particularly in the 35–39 and 40–44 age groups. The model generally aligns with observed data but exhibits minor deviations from 2010 onward, particularly in the younger age groups (25–29 and 30–34). The credible intervals widen further in the 50–54 age group, indicating increased uncertainty compared to Model A.

Model C generates forecasts with the widest credible intervals among the three models, reflecting higher uncertainty. The model follows observed trends closely in earlier years but diverges slightly in the post-2010 period. Model C captures some fluctuations more effectively than Models A and B, particularly in the 45–49 age group, but at the cost of increased uncertainty. The divergence is more evident in the 50–54 age group, where Model C underestimates fertility rates.

Comparing Model A and Model B, both models effectively capture the long-term fertility trend, though Model B exhibits slightly larger uncertainty. Model A provides a more stable fit with narrower credible intervals, while Model B appears more responsive to short-term variations. Model B exhibits minor overestimation in some younger age groups and underestimation in older groups.

Model C introduces greater uncertainty across all age groups compared to Model A. While Model A maintains narrower credible intervals, Model C follows observed data more closely at certain points, particularly in younger age groups. However, at older ages, Model C shows more pronounced deviations from observed fertility rates.

Model B and Model C share some similarities in capturing fluctuations, but Model C introduces wider credible intervals, indicating greater uncertainty. Model B aligns more closely with observed data in middle age groups, whereas Model C provides more flexibility in its forecasts but at the expense of stability. In the 50–54 age group, Model C underestimates fertility rates more significantly compared to Model B.

Overall, Model A provides the most stable forecasts with the least uncertainty across age groups. Model B and Model C exhibit similar trends but introduce greater uncertainty, particularly in the older age groups. The choice of the most appropriate model depends on the balance between stability and flexibility in capturing fertility trends.

This section presents a comparative analysis of model-based forecasts for age-specific fertility rates (ASMR) from 1998 to 2018, with in-sample forecasting for the period 2014–2018. The analysis is stratified by six age groups: 25–29, 30–34, 35–39, 40–44, 45–49, and 50–54. Figure 10 displays the results, where black lines represent observed data, red dashed lines correspond to Model A, dark blue dotted lines to Model B, and light blue dash-dotted lines to Model C.

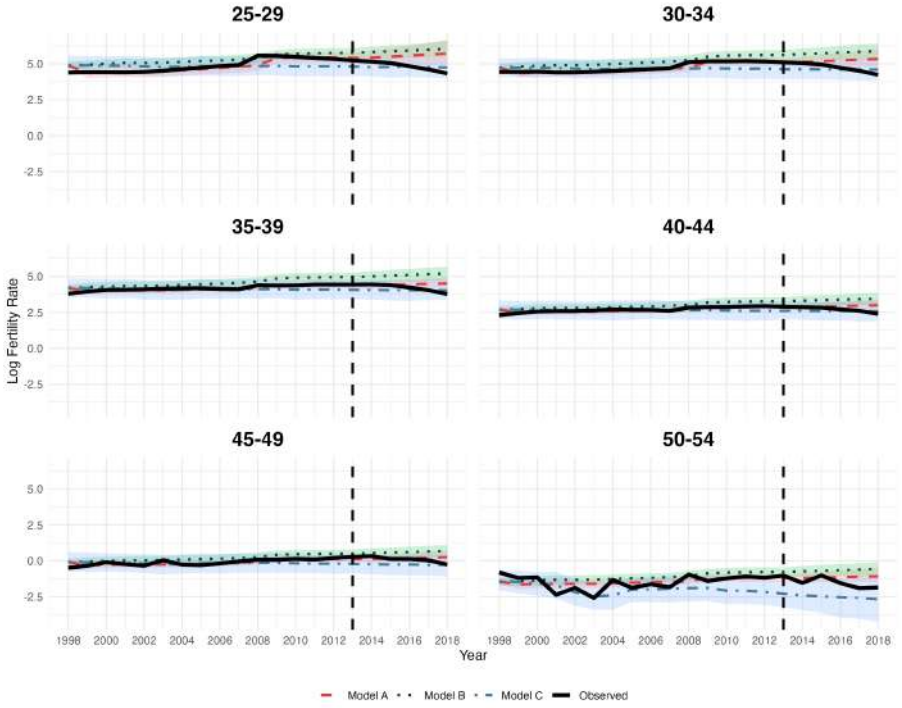

Figure 10: Age-specific logarithm of fertility rates (1998–2018) with in-sample forecasting for 2014–2018 for individuals with post-secondary education. Black lines represent observed data, red dashed lines correspond to Model A, dark blue dotted lines to Model B, and light blue dash-dotted lines to Model C.

The shaded areas indicate the credible intervals for each model's forecasts.

Model A closely follows observed data across all age groups, maintaining a relatively stable pattern in log fertility rates with minimal fluctuations. The model effectively captures long-term trends, particularly for the 25–34 and 35–39 age groups. The credible intervals remain narrow, indicating lower uncertainty in estimates. However, for the 50–54 age group, Model A shows a slight underestimation of fertility rates in later years.

Model B follows a similar trend to Model A but with slightly wider credible intervals, especially in the 35–39 and 40–44 age groups. The model aligns well with observed data but displays minor deviations after 2010, particularly in younger age groups (25–29 and 30–34). The credible intervals widen further in the 50–54 age group, reflecting increased uncertainty compared to Model A.

Model C produces forecasts with the widest credible intervals among the three models, indicating higher uncertainty. While it follows observed trends closely in earlier years, it diverges slightly in the post-2010 period. Model C captures some fluctuations more effectively than Models A and B, particularly

Table 4: In-Sample Forecast Evaluation: Root Mean Square Error (RMSE) of log fertility rates by model in the final year (2018) of the in-sample forecast period (2014–2018).

Model	**RMSE**
Model A	32.11
Model B	55.18
Model C	6.77

in the 45–49 age group, but at the cost of increased uncertainty. The divergence is more evident in the 50–54 age group, where Model C underestimates fertility rates.

Comparing Model A and Model B, both models effectively capture the long-term fertility trend, though Model B exhibits slightly greater uncertainty. Model A provides a more stable fit with narrower credible intervals, while Model B appears more responsive to short-term variations. Model B exhibits minor overestimation in some younger age groups and underestimation in older groups.

Model C introduces greater uncertainty across all age groups compared to Model A. While Model A maintains narrower credible intervals, Model C follows observed data more closely at specific points, particularly in younger age groups. However, at older ages, Model C deviates more significantly from observed fertility rates.

Model B and Model C share similarities in capturing fluctuations, but Model C introduces wider credible intervals, indicating greater uncertainty. Model B aligns more closely with observed data in middle age groups, whereas Model C provides greater flexibility in forecasts but at the expense of stability. In the 50–54 age group, Model C underestimates fertility rates more significantly compared to Model B.

Overall, Model A provides the most stable forecasts with the least uncertainty across age groups. Model B and Model C exhibit similar trends but introduce greater uncertainty, particularly in older age groups. The choice of the most appropriate model depends on the balance between stability and flexibility in capturing fertility trends.

In-Sample Forecast Coherence

Building on the assessment of model fit, the evaluation of forecast accuracy provides further insight into model performance. To quantify the ability of each model to replicate observed fertility rates, Table 4 presents the Root Mean Square Error (RMSE) for the final year (2018) of the in-sample forecast period (2014–2018). The RMSE measures the deviation between estimated and observed log fertility rates, where lower values indicate a better model fit. This comparison allows for an assessment of each model's forecasting precision within the observed data range.

Model C produces the lowest RMSE (6.77), indicating the closest alignment

with observed fertility rates in 2018. This suggests that Model C provides the most accurate predictions among the three models, effectively capturing temporal fertility trends. Model A, with an RMSE of 32.11, exhibits moderate predictive accuracy. Although it aligns reasonably well with observed data, the error magnitude is substantially higher than that of Model C. In contrast, Model B yields the highest RMSE (55.18), reflecting the greatest deviation from observed values. This result indicates that the forecasted fertility rates from Model B diverge more significantly from empirical data, potentially due to increased variability or structural differences in its parameterization.

The variation in RMSE values reflects differences in how each model represents fertility patterns. The lower RMSE for Model C suggests that its parameterization better captures observed trends, whereas Models A and B introduce higher forecast uncertainty. These findings emphasize the importance of selecting an appropriate model for fertility forecasting, as accuracy depends on the methodological approach employed.

Extending the evaluation beyond RMSE values, a broader comparison of model forecasts provides additional insights into their performance.

Models A and B demonstrate similar long-term trends in fertility forecasts, with both capturing the overall trajectory observed in historical data. However, Model B exhibits slightly wider credible intervals, particularly for older age groups, indicating greater uncertainty in its projections. Model A provides a more stable forecast, with narrower intervals and a closer alignment with observed data for most age groups. Model B, in contrast, appears more responsive to short-term fluctuations, which may contribute to increased variability in its estimates. While both models effectively track observed trends, Model A maintains greater stability, making it preferable for long-term projections, whereas Model B may be better suited to capturing temporal variations.

Comparing Model A and Model C reveals notable differences in uncertainty estimation. Model C consistently presents the widest credible intervals, suggesting a higher degree of uncertainty in its forecasts. While Model C follows observed trends relatively well in younger age groups, it diverges more significantly in older age groups, where it tends to underestimate fertility rates. Model A, by contrast, maintains a more stable fit with narrower credible intervals, making it more reliable for long-term forecasting. Model C's greater flexibility allows it to capture fluctuations better in specific cases, but at the cost of increased uncertainty, particularly in later years.

Models B and C share some similarities, particularly in their treatment of short-term fluctuations. Both models exhibit wider credible intervals than Model A, with Model C showing the greatest level of uncertainty. Model B generally aligns more closely with observed data across all age groups, whereas Model C demonstrates greater variability in its predictions. In older age groups, Model C's forecasts tend to be lower than those of Model B, suggesting a potential underestimation of fertility rates. While Model B offers a balance between trend stability and responsiveness, Model C's flexibility in capturing variability is accompanied by a higher degree of forecast uncertainty.

Overall, Model A provides the most stable and reliable forecasts, with narrow

credible intervals and a strong alignment with observed trends. Model B, while slightly more uncertain, captures short-term variations more effectively. Model C, with the widest credible intervals, introduces the most flexibility but at the expense of increased forecast uncertainty, particularly in older age groups. The selection of the most appropriate model depends on the priority assigned to stability versus adaptability in capturing fertility trends.

Conclusion

This paper has examined the application of Bayesian hierarchical models in forecasting fertility rates in Colombia, disaggregated by educational attainment. By systematically comparing three models with different levels of coherence—(i) a non-coherent Bayesian model, (ii) a partially coherent model enforcing consistency in age-specific fertility trends, and (iii) a fully coherent model ensuring alignment in both age-specific and period-specific trends the analysis has demonstrated the advantages of structured modeling in demographic forecasting.

The findings indicate that coherent Bayesian approaches substantially enhance predictive accuracy and internal consistency, addressing the demographic heterogeneity associated with educational stratification. The fully coherent Bayesian model outperformed the other two by maintaining internal coherence while allowing for the necessary flexibility in fertility projections across subpopulations. Moreover, fertility trends in Colombia exhibit a pronounced negative correlation with educational attainment. Women with lower educational levels maintain higher fertility rates, with their peak childbearing years occurring earlier, whereas those with post-secondary education display lower fertility levels and a postponement in childbearing age. These patterns emphasize the role of education in shaping reproductive behaviour and highlight the importance of incorporating structured coherence in fertility forecasting models to enhance policy relevance.

The paper sought to determine how different Bayesian hierarchical models influence the accuracy and coherence of fertility projections in Colombia. The results confirm that structured forecasting frameworks significantly improve predictive accuracy, particularly in contexts where data quality and consistency are challenging. By enforcing coherence in age-specific and period-specific fertility trends, the paper demonstrated that these models provide more reliable demographic projections, which are crucial for informing evidence-based policymaking.

In addition to the methodological advancements, the paper examined how educational attainment influences fertility trends in Colombia. The findings reveal a non-linear relationship between education and fertility that deviates from the commonly assumed inverse association. Fertility rates declined more substantially among women without formal education than among those with only primary education. Among women with secondary education, fertility levels remained relatively stable, while an increase was observed among women with postsecondary education. These results suggest that the effect of education on

reproductive behaviour is more complex than traditionally theorised, and may reflect a combination of delayed childbearing, shifting socio-economic aspirations, and improved access to healthcare and support services among highly educated women. By capturing these differentiated fertility trajectories across educational groups, the Bayesian hierarchical models employed in this paper provide a nuanced and policy-relevant framework for future demographic forecasting.

Beyond its methodological contributions, this paper has practical implications for demographic research and policy development. The theoretical advancements reinforce the importance of coherent Bayesian models in enhancing the accuracy of fertility projections while maintaining demographic plausibility. By integrating coherence constraints, this approach ensures consistency across demographic subgroups while allowing for meaningful variations in fertility behaviours. The empirical application to Colombia serves as a case paper for other developing nations experiencing similar fertility differentials by education.

From a policy perspective, reliable fertility forecasts are essential for planning education systems, labour markets, and social welfare programs. The paper's findings provide valuable insights for policymakers, enabling more precise resource allocation and the development of targeted interventions aimed at addressing demographic disparities. Understanding fertility differentials by education can help in designing education-based reproductive health programs, improving maternal and child health outcomes, and balancing demographic and economic development objectives.

Despite the robustness of the methodology, certain limitations must be acknowledged. First, the availability and quality of fertility data posed challenges, as inconsistencies across census data, household surveys, and birth records required harmonization techniques. While these adjustments helped improve data reliability, residual inconsistencies may still affect projections. Second, although Bayesian hierarchical models incorporate uncertainty, they rely on historical trends and may not fully capture abrupt demographic shifts resulting from policy changes, economic downturns, or public health crises.

Additionally, while this paper focuses on educational attainment as the primary stratification variable, fertility behaviour is influenced by a broader range of socioeconomic factors, including income levels, employment stability, urbanization, and access to reproductive healthcare. Future research should consider integrating these factors into forecasting models to provide a more holistic understanding of fertility trends.

Given these limitations, several avenues for future research emerge. A promising extension involves incorporating economic indicators such as income distribution, employment trends, and social mobility into Bayesian hierarchical fertility models. These factors significantly influence reproductive decisions and could further enhance forecasting accuracy. Another important research direction is the integration of policy variables into fertility forecasting. Given that government policies on education, reproductive health, and social security impact fertility behaviour, incorporating these elements into Bayesian models would improve their applicability to real-world scenarios.

Further, comparative studies applying similar Bayesian hierarchical methods to other countries with varying demographic contexts would help validate the generalizability of this approach. Additionally, advancements in computational methods provide opportunities for refining fertility forecasting models. Future research could explore the integration of machine learning techniques with Bayesian modeling to enhance predictive accuracy and computational efficiency. Hybrid approaches combining probabilistic modeling with deep learning methods may offer new insights into fertility trends and their long-term implications.

Fertility forecasting remains an essential component of demographic research, particularly in regions undergoing rapid social and economic transformations. This paper highlights the importance of developing sophisticated forecasting models capable of accurately capturing demographic heterogeneity and generating reliable projections that inform policy decisions. As fertility patterns continue to evolve, the need for rigorous, data-driven methodologies will only increase.

By refining Bayesian hierarchical fertility forecasting methods, this research contributes to a deeper understanding of population dynamics and demographic change. The findings emphasize the continued relevance of fertility forecasting in shaping policies related to education, labour markets, and social security. Moving forward, interdisciplinary collaboration and methodological advancements will be key in ensuring that fertility projections remain robust, adaptable, and relevant to policymakers and researchers alike.

A Appendix A - Population Reconstruction

This paper integrates multiple data sources. The population at risk is reconstructed using censuses, household surveys, and administrative birth records from 1998 to 2018. A multistate demographic approach, refined in the Appendix A, ensures consistency in estimating fertility rates by age and education. Additionally, fertility estimates are validated against alternative demographic sources, as discussed in the Appendix A, to confirm their reliability. By leveraging multiple datasets, this paper accounts for potential reporting inconsistencies and measurement errors, improving the accuracy of fertility forecasting models.

Building on previous efforts to improve fertility forecasting, this paper employs a methodology that reconciles multiple data sources to reconstruct the population in Colombia (presented in the Appendix A). The approach integrates census data, household survey and vital statistics to generate a coherent demographic foundation for fertility analysis. By systematically linking these source the methodology ensures that fertility estimates remain consistent with observed demographic pattern improving the accuracy and representativeness of the constructed population at risk.

The administrative birth records exhibit high coverage and completeness; however, information on maternal education is occasionally missing or inconsistently reported. Standardisation procedures and quality checks were applied to

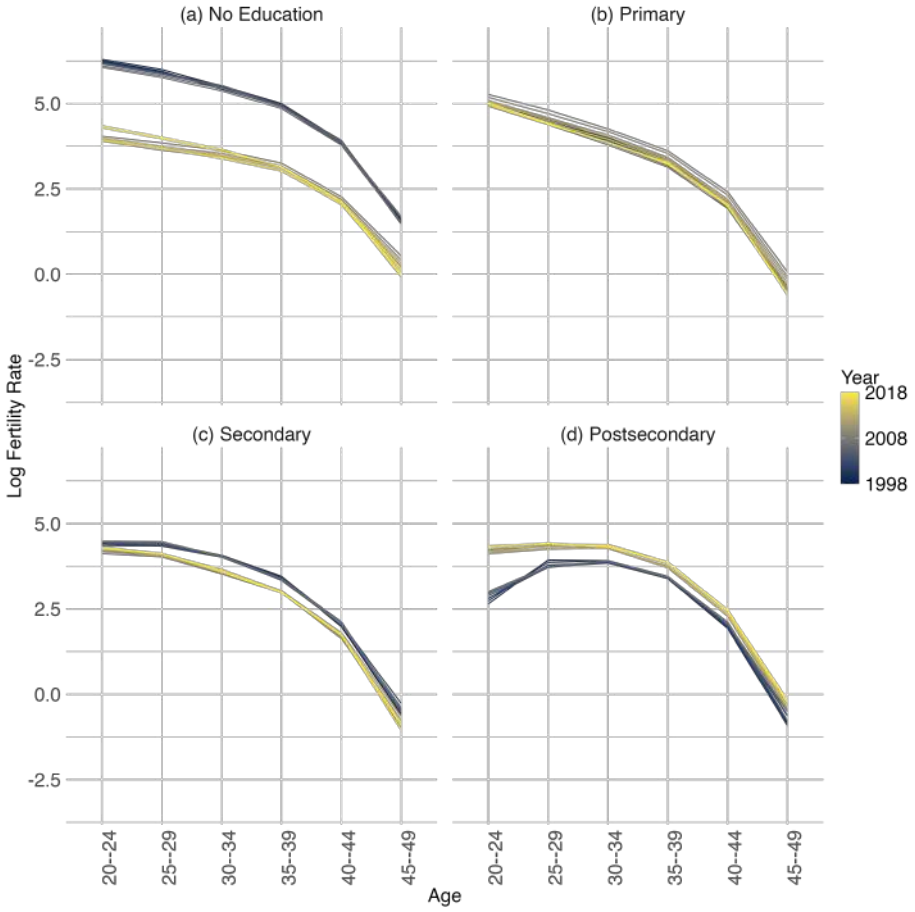

Figure 11: Log of fertility rates by age (x-axis), educational attainment (first row: no formal education and primary; second row: secondary and post-secondary), and year (color gradient from dark blue representing 1998 to yellow representing 2018) based on the WCDE reconstruction.

ensure internal consistency across years and subgroups.

This section describes the fertility data used for estimation. Fertility rates are reconstructed from historical birth records disaggregated by age and education level (see details in Appendix A), household surveys, and national censuses in Colombia). The data span the period 1998–2018 and are aligned with fertility data sources to ensure consistency in demographic projections. Two versions of the population estimates were compared to assess the consistency of transitions between years, given that the Holt Filter method modification produces more accurate and credible fertility rates (see details in the appendix A).

B Appendix B - Inference Evaluation

Lee Carter Parameters

A detailed presentation of each Lee–Carter parameter for all models is provided in the appendix (see Appendix Section B). However, given the central role of the κ parameter in forecasting, this section focuses on its estimation across models to provide insight into future fertility trends.

Across Models A, B, and C, the average log fertility rates exhibit similar trends for each education level. However, the κ parameter captures the overall temporal trend between 1998 and 2018, distinguishing the forecasting behavior of the models.

To further analyze the individual Lee-Carter parameter estimations for each education level, the paper period is divided into two segments: 1998–2007 and 2008–2018. For Model A, the first period exhibits a substantial increase in the average log fertility rates for women without formal education. This increase surpasses that of all other educational levels, peaking in 2004 before beginning a decline. During the same period, the average log fertility rates for women with postsecondary education show a slight increase, while rates for those with primary education remain stable, and those with secondary education experience a decrease.

For the period 2008–2018, women with secondary education exhibit the lowest fertility rates, followed by those without formal education, then women with primary education, and finally, women with postsecondary education who have the highest rates. By 2018, fertility rates for women with postsecondary education stabilize, with a decrease from 4 to 0. Women with primary education experience the most significant decrease (from 3 to -2), followed by those without formal education (from -2 to -4). Interestingly, women with secondary education show an increase in fertility rates (from -6 to -2), suggesting a compositional shift within this group.

Overall, the rate of change in fertility rates by the end of the period follows an educational gradient: women with postsecondary education experience the most rapid change, followed by those with secondary education, then women with primary education, with the slowest change occurring among women without formal education.

Model B exhibits a trend similar to Model A in the estimation of the κ parameter across all educational levels. In contrast, Model C differentiates the two periods 1998–2007 and 2008–2018 in the estimation of the κ parameter by modeling the parameters β and κ coherently across all education levels. During the first period (1998–2007), the log fertility rate remains stable until 2007, after which a sharp peak occurs. This is followed by a decreasing trend until 2016, after which the log fertility rate begins to rise again.

References

Alkema, L., Raftery, A. E., Gerland, P., et al. (2011). Probabilistic projections of the total fertility rate for all countries. *Demography*, 48(3):815–839.

Batyra, E., Leone, T., and Myrskylä, M. (2023). Forecasting of cohort fertility by educational level in countries with limited data availability: The case of brazil. *Population Studies*, 77(2):179–195.

Bongaarts, J. (2002). The end of the fertility transition in the developing world. *Population and Development Review*, 28(3):419–443.

Bongaarts, J. and Feeney, G. (1998). On the quantum and tempo of fertility. *Population and Development Review*, 24(2):271–291.

Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting*, 22(3):547–581. Accessed via ScienceDirect at https://www.sciencedirect.com/science/article/pii/S016920700600046X.

Brass, W. (1974). Perspectives in population prediction: Illustrated by the statistics of england and wales. *Journal of the Royal Statistical Society. Series A (General)*, 137(4):532–583.

Cavenaghi, S. and Lesthaeghe, R. (2020). Demographic transformations and inequality: The second demographic transition in latin america and the caribbean. In *Demographic Transformations and Inequality: The Second Demographic Transition in Latin America and the Caribbean*, pages 1–28. Springer.

Cleland, J. (2009). Contraception in historical and global perspective. *Best Practice & Research Clinical Obstetrics & Gynaecology*, 26(6):793–807.

Ellison, J., Kuang, B., Christison, S., Berrington, A., and Kulu, H. (2023). Estimating the 2011 total fertility rate for england & wales and scotland using alternative data sources. Working Paper 106, ESRC Centre for Population Change, Southampton, UK. ESRC Centre for Population Change Connecting Generations Working Paper.

Esteve, A., Lesthaeghe, R., and López-Gay, A. (2012). The latin american cohabitation boom, 1970–2007. *Population and Development Review*, 38(1):55–81.

Gerland, P., Raftery, A. E., Ševčíková, H., Li, N., Gu, D., Spoorenberg, T., Alkema, L., Fosdick, B. K., Chunn, J., Lalic, N., Bay, G., Buettner, T., Heilig, G. K., and Wilmoth, J. (2014). World population stabilization unlikely this century. *Proceedings of the National Academy of Sciences*, 111(33):11897–11902.

Hoffman, M. D. and Gelman, A. (2014). The no-u-turn sampler: Adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research*, 15(1):1593–1623.

Lee, R. D. (1993). Modeling and forecasting the time series of us fertility: Age distribution, range, and ultimate level. *International Journal of Forecasting*, 9(2):187–202.

Lee, R. D. and Carter, L. R. (1992). Modeling and forecasting u.s. mortality. *Journal of the American Statistical Association*, 87(419):659–671.

Lesthaeghe, R. (2010). The unfolding story of the second demographic transition. *Population and Development Review*, 36(2):211–251.

Lutz, W., Butz, W. P., and K.C., S. (2014). *World Population and Human Capital in the Twenty-First Century*. Oxford University Press.

Martín, T. C. (1995). Women's education and fertility: Results from 26 demographic and health surveys. *Studies in Family Planning*, 26(4):187–202.

Raftery, A. E., Alkema, L., and Gerland, P. (2014). Bayesian population projections for the united nations. *Statistical Science*, 29(1):58–68.

Sanchez-Segura, A. F. (2025). Historical population reconstruction for colombia by age, sex, and education, 1998–2018. https://doi.org/10.48420/29309810. Data set.

Schmertmann, C., Zagheni, E., Goldstein, J. R., and Myrskylä, M. (2014). Bayesian forecasting of cohort fertility. *Journal of the American Statistical Association*, 109:500–513.

Wiśniowski, A., Smith, P. W. F., Bijak, J., et al. (2015). Bayesian population forecasting: Extending the lee-carter method. *Demography*, 52(3):1035–1059.