

Artificial Intelligence–Enabled Modeling of Childcare Service Demand and Policy Efficiency in Inner Mongolia Autonomous Region, China: A Comprehensive Multi-Method Demographic Analysis

Qingyan Meng^{1*}

¹School of Educational Science, Jining Normal University, Ulanqab, Inner Mongolia 012000, China.

*Corresponding Author: Qingyan Meng Email: mengqyd@163.com

Received: May 19, 2026 Accepted: August 23, 2026

Abstract: Demographic contraction is a process that is rapidly unfolding in Inner Mongolia Autonomous Region (IMAR), and is characterized by declining fertility, increasing mortality, negative natural population growth, and a decreasing childcare-age population. These trends present significant challenges for planning childcare, early childhood education and regional fertility policy. This study proposes a forecasting and policy evaluation model based on artificial intelligence and an official panel data set of 2021–2025 with 36 indicators from the Inner Mongolia Statistical Yearbook, China Statistical Yearbook and National Bureau of Statistics communiqués. The framework combines statistical trend analysis, 6 machine-learning regressors, SHAP explainability, BiLSTM forecasting with Monte Carlo dropout, policy-efficiency indices and 5 policy scenarios through 2030. The birth rate for IMAR decreased from 6.26‰ (2021) to 4.28‰ (2025), whereas the natural population growth rate went from 3.76‰ (2021) to –4.58‰ (2025). The number of births used as a proxy for annual childcare demand shrank from 15.0 wan in 2023 to 10.2 wan in 2025, while the population aged 0-2 decreased from 40.4 wan in 2023 to 35.4 wan in 2025. The BiLSTM model further predicts a decrease in births to 9.02 wan by 2028. The machine learning model with the highest performance was Gradient Boosting (RMSE = 1.557 wan; MAPE = 11.33%), which considered birth rate, natural growth rate, and the intensity of childcare as the most important demand predictors, as identified by the SHAP values. A combination of lower policy efficiency and weaker policy-to-outcome conversion has led to the lowest scores for Policy Efficiency Score and Fertility-Friendly Development Index in 2025. A comprehensive intervention that combines a childcare subsidy, an improvement in parental leave, housing policy and protection for women in the labor market can make a difference and bring the projected birth rate to 5.03‰ and the projected demand for childcare to 10.654 wan by 2030, according to scenario analysis. The results offer a basis for adaptive childcare infrastructure, optimising fertility policies, and demographic governance in quickly ageing and low-fertility areas.

Keywords: Childcare demand forecasting; Fertility Decline; BiLSTM Neural Network; SHAP Explainability; Machine Learning; Urbanization; Natural Population Growth; Pro-Natalist Policy; Early Childhood Education

1. Introduction

The global demographic landscape is undergoing a profound and historically unprecedented transformation. Across East Asia, fertility rates have fallen far below replacement level, driven by the interacting forces of rapid urbanization, rising educational attainment among women, escalating costs of child-rearing, delayed marriage, and the erosion of traditional family formation norms under the pressures of competitive labor markets and high-cost urban living environments [1,2]. China, the world's most populous nation for most of recorded history, has entered a phase of absolute population decline a

milestone reached in 2022 when the national population fell for the first time since the Great Leap Forward famine of the early 1960s. National annual births declined from 10.62 million in 2021 to 7.92 million in 2025, and the total fertility rate (TFR) has fallen to approximately 1.0 among the lowest ever recorded for a major economy despite the formal abolition of birth number restrictions in 2021 and the introduction of successive waves of pro-natalist incentive packages at national and provincial levels [4,5]. The persistence and indeed acceleration of fertility decline in the face of intensifying policy intervention has emerged as one of the defining governance challenges of contemporary China, with far-reaching structural consequences for labor supply, pension system sustainability, healthcare demand, housing markets, and critically for this study the childcare service ecosystem upon which future human capital formation depends [6,7].

While national-level demographic trends in China have attracted substantial and growing scholarly attention, sub-national demographic heterogeneity remains comparatively understudied, particularly in ethnic minority autonomous regions where demographic structures, policy environments, economic development trajectories, and socioeconomic conditions diverge markedly from Han-majority urban centers. This gap is important because it means that policies developed at the national level and infrastructure planning standards set at the national level may systematically misalign with the actual demographic situations of regional governments, resulting in a "shortfall of provision" in areas in which demand is highly localized but where the overall picture is of a child population decline or an "oversupply" of provision in areas of rapid decline. The challenge of this gap needs to be tackled by region-specific and data-driven analyses that can reflect the multi-driver dynamics of the evolution of childcare demands and the fertility-decline phenomenon at the sub-national level, which is nonlinear. The divergence is particularly acute in Inner Mongolia Autonomous Region (IMAR). IMAR is an area in northern China, bordering Mongolia and Russia, and with a total area of 1.183 million km² and a population of approximately 23.74 million (2,374 wan) persons in 2025, which has seen a deterioration in its population in all major indicators of vital statistics, to a greater degree than even China's alarming national trend. The region's birth rate of 4.28‰ in 2025 represents only 76.02% of the national birth rate index, and its natural population growth rate of -4.58‰ is more than double the national figure of -2.41‰ in absolute terms. Total population has been declining since 2022, falling at a statistically significant rate of -6.517 wan per year a trend driven by the compounding effects of falling births, rising deaths, and net outmigration associated with economic restructuring away from the region's historically dominant coal, pastoral, and heavy industrial sectors toward service-oriented urban economies [10]. These dynamics create a particularly challenging environment for childcare service planning: aggregate demand is contracting structurally, yet the spatial distribution of remaining demand is shifting rapidly toward urban centers as urbanization accelerates from 68.2% in 2021 to 71.48% in 2025 [11].

The demand for childcare in this study is operationalized using the number of births per year, while the Age 0–2 cohort population is an essential planning factor for regional governments tasked with providing nursery infrastructure and setting the workforce of early childhood education (ECE), funding transfers for family support programs and long-term demographic investment strategies. The spatial concentration of births is likely to lead to more childcare needs in urban areas, while also being a concern in rural areas because of the risk of oversized infrastructure. Accurate forecasting of this demand is important not only to prevent over-provision of infrastructure in rural areas in bad economies, but also ensure that adequate childcare capacity is provided in the urban areas. But conventional demographic projection methods such as linear extrapolation, cohort-component models or Leslie matrix models are not appropriate to account for the multi-driver and nonlinear nature of the fertility decline in the rapidly urbanizing regions with small historical time series and structural policy discontinuities in China's successive birth policy transitions [12]. For these reasons, there has been an increasing amount of research that has used machine learning and deep learning techniques to predict demographics, including national scale population predictions and city scale modelling of childcare demands [13,14]. In this context, there are several benefits of machine learning methods over traditional demographical methods. Gradient boosting algorithms have been found to be more effective especially in small sample regression problems, for their sequential error correction property, self-regulated smoothing property, and their capacity to detect nonlinear interaction between features without having to explicitly specify a model [15]. Bidirectional Long Short-Term Memory (BiLSTM) networks have shown particular promise for short

demographic time series by processing sequences in both forward and backward temporal directions, thereby capturing both long-run trend momentum and anomalous short-run reversals such as the temporary 2024 birth rate rebound observed in IMAR that unidirectional models systematically misrepresent [16]. SHapley Additive exPlanations (SHAP), grounded in cooperative game theory, provide theoretically principled, model-agnostic feature attribution that is increasingly adopted in policy-relevant ML applications as a bridge between predictive accuracy and interpretive transparency accessible to non-technical policymakers [17]. Despite these methodological advances, their integrated application to sub-national Chinese demographic analysis particularly in the context of childcare demand forecasting under data-scarce conditions and explicit policy efficiency evaluation remains limited in the published literature [1,13].

The policy efficiency dimension of this study is particularly novel. Existing studies of Chinese pro-natalist policy have predominantly focused on documenting the failure of incentive-based approaches to reverse fertility decline at the national level [4,5], with less attention paid to the regional variation in policy-to-outcome conversion efficiency and the structural determinants of that variation. The Policy Efficiency Score (PES) and Fertility-Friendly Development Index (FFDI) constructed in this study provide quantitative tools for evaluating not merely whether policy exists but whether it works a distinction of critical importance given that IMAR's policy pressure index reached its study-period maximum of 3.52 in 2025, the same year that produced the worst demographic outcomes on record [8,9]. This inverse relationship between policy intensity and demographic outcomes encapsulates what this study terms the "fertility policy efficiency paradox" a phenomenon with important implications for the design of next-generation pro-natalist interventions.

This study aims to answer three fundamental research questions: (1) Are there statistically significant demographic trends that are observed in IMAR between 2021 and 2025 and how do those trends compare to the national benchmarks for all key vital statistics and derived indicators? (2) What are the most important variables explaining when children demand childcare services and what are the relative performances of the competing ML models when the dataset is very small (5 observations per year)? What is the predicted path of demand for childcare under alternative policy scenarios with varying levels of support for fertility and varying urbanization pathways up to 2030, and what levels of policy efficiency are possible? (3) This research, which explicitly addresses these questions by applying an integrated framework that brings together the statistical rigor, explainability of ML, deep learning forecasting, and scenario simulation to understand the situation within the framework of the one Belt and One Road initiative in a demographically stressed region, will serve as a supplement to the growing body of research on data-driven demographic governance in such areas and will offer a replicable methodological template for similar sub-national scenarios around the world [10,11,14].

The rest of the paper is organized as follows. Each of these is described in detail in Section 2, with the full multi-method analytical framework described in Section 2.4. In Section 3, the results are given in a comprehensive way, where figures and tables appear where they are discussed. Section 4 discusses the study results in light of the literature and the implications for policy. Section 5 summarizes and outlines recommendations, limitations and suggestions for future research.

2. Materials and Methods

2.1. Study Area

The study area is Inner Mongolia Autonomous Region (IMAR; Chinese: 内蒙古自治区), a provincial-level autonomous region in northern China established in 1947. IMAR spans approximately 1.183 million km² the third largest provincial-level administrative unit in China and borders eight Chinese provinces and autonomous regions as well as Mongolia and Russia. The region's economy has historically been dominated by coal mining, pastoral agriculture, and heavy industry, though structural economic transformation toward services and renewable energy has accelerated since 2015. IMAR's population of 2,374 wan (2025) is distributed across 12 prefecture-level cities and leagues (aimags), with significant urban-rural heterogeneity. The Han ethnic majority comprises approximately 79% of the population, with Mongol (17%), Hui, Manchu, Daur, and other ethnic minorities accounting for the remainder. IMAR's demographic profile is shaped by decades of net outmigration of working-age population to eastern coastal

provinces, an aging population structure, and a historically lower fertility rate relative to national averages trends that have intensified markedly since 2020.

2.2. Dataset Description and Data Sources

This study constructed a longitudinal demographic dataset covering Inner Mongolia Autonomous Region and China at the national level for the period 2021–2025 to support artificial intelligence–based analysis of fertility dynamics, childcare service demand, and the demographic context of fertility-friendly policy development. The dataset was compiled exclusively from publicly available official statistical publications and contains five annual observations across 33 variables, encompassing original demographic indicators, national benchmark indicators, and transparently derived analytical features. No synthetic demographic observations were introduced, and all original and derived variables were retained separately to ensure full transparency and reproducibility.

For Inner Mongolia, the primary data sources were the 2021–2025 population and socioeconomic statistical bulletins published by the Inner Mongolia Autonomous Region Statistics Bureau. The collected variables include year-end resident population, urban and rural population, urbanization rate, male and female population, annual number of births, crude birth rate, annual deaths, crude death rate, and natural population growth rate. For example, Inner Mongolia recorded 15.0×10^4 births and a birth rate of 6.26‰ in 2021, compared with 10.2×10^4 births and a birth rate of 4.28‰ in 2025. The corresponding official annual figures for 2022, 2023, and 2024 were incorporated directly from the regional statistical bulletins without modification or imputation.

To establish a national comparison benchmark, corresponding demographic indicators for China as a whole were extracted from national economic and social development statistical communiqués and related official statistical reports published by the National Bureau of Statistics of China (NBS). These national indicators include total population, annual births, crude birth rate, annual deaths, crude death rate, and natural population growth rate. National birth rates declined from 7.52‰ in 2021 to 5.63‰ in 2025, while the natural population growth rate changed from +0.34‰ to –2.41‰ over the same period. Intermediate annual national statistics for 2022–2024 were obtained from the corresponding NBS annual communiqués.

The full structure of the dataset, organized by data component, main variables, and analytical role, is summarized in Table 1. As shown in Table 1, the dataset spans six conceptual data components: IMAR core demographics, population structure variables, national China benchmark indicators, fertility-gap indicators, trend indicators, childcare demand proxies, and standardized benchmark indices.

Table 1. Dataset Structure: Data Components, Main Variables, and Analytical Roles, Inner Mongolia–China Demographic Dataset, 2021–2025.

Data Component	Main Variables	Analytical Role
Inner Mongolia demographics	Population, births, birth rate, deaths, death rate, natural growth rate	Core regional predictors and outcome variables
Population structure	Urban/rural population, urbanization rate, male/female population	Structural determinants of childcare demand
National China benchmark	Population, births, birth rate, deaths, death rate, natural growth rate	National reference for comparative modeling
Fertility-gap indicators	Regional–national birth rate gap; natural growth rate gap	Measure of relative demographic pressure
Trend indicators	Birth growth rate, national birth growth rate, urbanization change	Capture temporal demographic momentum
Childcare demand proxies	Annual births; estimated population aged 0–2 years	Approximate potential childcare service demand
Benchmark indices	Birth rate index (China=100); population share; birth share	Standardized comparative features for ML modeling

The present dataset captures the demographic demand side of childcare service planning exclusively. The source documents do not directly provide childcare center capacity, childcare places, service utilization rates, public expenditure levels, subsidy intensity, staffing ratios, or other policy-input variables. Therefore, in the empirical analysis, demographic indicators are treated as the foundation for childcare demand modeling, while policy efficiency measures (PES, FFDI) are framed as demographic policy-pressure and relative-efficiency proxies rather than direct measures of childcare system performance. This limitation is acknowledged explicitly in the Discussion and Conclusion sections.

2.3. Variable Construction and Operationalization

In addition to the directly observed variables described in Section 2.2, several derived features were constructed to improve the suitability of the dataset for quantitative and AI-enabled modeling. All derivations are documented below with explicit formulae to ensure reproducibility.

Birth Rate Index (BRI): A standardized comparative index expressing IMAR's birth rate as a percentage of the national birth rate, computed annually as:

$$BRI_t = \frac{BR_t^{IM}}{BR_t^{CN}} \times 100 \quad (1)$$

where BR_t^{IM} denotes the crude birth rate of Inner Mongolia in year t (‰), and BR_t^{CN} denotes the corresponding national crude birth rate (‰). BRI values below 100 indicate that IMAR's birth rate falls below the national average; values declined from 83.24 (2021) to 76.02 (2025), confirming IMAR's progressively deteriorating relative fertility position.

Childcare Demand Proxy Age 0–2 Cohort: Because official sources do not directly report the annual population requiring childcare services, a childcare demand proxy for children aged 0–2 years was constructed from observed birth cohorts as:

$$CD_t = B_t + B_{t-1} + B_{t-2} \quad (2)$$

where B_t denotes the number of births (in wan, i.e., 10^4) in year t . This indicator approximates the potential size of the population requiring nursery and childcare services at any given point in time. The proxy explicitly assumes no adjustment for infant mortality or interregional migration a simplification that may introduce minor upward bias in demand estimates but is consistent with standard practice in regional childcare demand modeling under data-scarce conditions [13,14]. Because the available birth series begins in 2021, this three-cohort proxy is available only for 2023–2025, yielding values of 40.4 wan (2023), 38.6 wan (2024), and 35.4 wan (2025). A normalized childcare demand index with 2023=100 was further generated to facilitate temporal comparison.

Year-on-Year Change Variables: Year-on-year changes in births (wan) and urbanization rate (percentage points) were calculated to capture demographic momentum and the rate of structural change:

$$\Delta B_t = B_t - B_{t-1} \quad (3)$$

$$\Delta U_t = U_t - U_{t-1} \quad (4)$$

where U_t denotes the urbanization rate (%) in year t . These derived variables are available for 2022–2025 (n=4).

Fertility-Gap Indicators: The regional–national birth rate gap and natural growth rate gap were computed as simple differences:

$$Gap_{BR,t} = BR_t^{IM} - BR_t^{CN} \quad (5)$$

$$Gap_{NGR,t} = NGR_t^{IM} - NGR_t^{CN} \quad (6)$$

where NGR denotes the natural population growth rate (‰). Negative values indicate that IMAR's vital statistics are worse than the national average; the birth rate gap ranged from -1.19‰ to -1.39‰ across the study period (mean= -1.288‰).

2.4. Descriptive Statistical Analysis

Descriptive statistics were computed for all 36 variables, including: arithmetic mean, standard deviation (SD), minimum, maximum, and coefficient of variation ($CV\% = SD/\text{mean} \times 100$). $CV\%$ was used as the primary measure of relative variability across variables, enabling comparison of instability across indicators measured on different scales. Variables with $CV > 30\%$ were classified as exhibiting high structural instability. All computations were performed in Python 3.10 using NumPy 1.24 and Pandas 2.0.

2.5. Correlation Analysis

Pearson product-moment correlation coefficients (r) were computed for all pairwise combinations of key variables to quantify linear associations. Spearman rank correlation coefficients (ρ) were computed in parallel to capture monotonic associations robust to non-normality and outliers particularly important given the small sample size ($n=5$). Two-tailed significance tests were conducted for all correlation pairs, with p -values reported and significance classified at $\alpha=0.05$. For the primary research question of childcare demand prediction, all variables were tested as predictors of `childcare_demand_proxy_births_wan`, with results. Correlation matrices were visualized as annotated heatmaps using Seaborn 0.12.

2.6. Linear Trend Analysis

Ordinary Least Squares (OLS) linear regression was applied to each of nine key variables as a function of calendar year (independent variable), yielding annual slope estimates, intercepts, Pearson R , R^2 , p -values, standard errors, trend direction classifications, and statistical significance assessments at $\alpha=0.05$. For variables with $n<5$ observations (childcare intensity and Age 0–2 cohort, both $n=3$), trend statistics are reported but significance is not formally assessed due to insufficient degrees of freedom. Trend visualization was produced using Matplotlib 3.7, with 95% confidence bands computed from the OLS standard errors.

2.7. Machine Learning Models and Cross-Validation

Six supervised ML regression models were trained to predict the primary childcare demand proxy (`births_wan`) from the full feature set of available predictors. The models evaluated were: (1) Gradient Boosting Regressor (scikit-learn; $n_estimators=100$, $learning_rate=0.1$, $max_depth=3$); (2) Random Forest Regressor ($n_estimators=100$, $random_state=42$); (3) Ridge Regression ($\alpha=1.0$); (4) Lasso Regression ($\alpha=0.1$); (5) ElasticNet ($\alpha=0.1$, $l1_ratio=0.5$); and (6) Support Vector Regression (SVR; RBF kernel, $C=1.0$, $epsilon=0.1$). Feature engineering included standardization (zero mean, unit variance) applied to all input features prior to model training.

Given the extreme data scarcity ($n=5$ annual observations), Leave-One-Out Cross-Validation (LOO-CV) was employed as the sole feasible validation strategy. In LOO-CV, each of the five observations is held out once as the test set while the remaining four are used for training, yielding five test predictions from which aggregate performance metrics are computed. This approach maximizes training data utilization while providing an unbiased estimate of out-of-sample prediction error, though it is known to produce high variance in performance estimates for very small n [15]. Performance metrics reported include: Root Mean Squared Error (RMSE, wan), Mean Absolute Error (MAE, wan), R^2 (LOO-CV), and Mean Absolute Percentage Error (MAPE%). All ML computations were performed using scikit-learn 1.3.

2.8. Residual Diagnostics

Residual diagnostics were conducted for the best-performing ML model (Gradient Boosting) and the Random Forest model to validate statistical assumptions. Diagnostics included: (i) residuals vs. fitted values plot to assess homoscedasticity; (ii) normal Q-Q plot to assess residual normality; (iii) standardized residuals over time to identify outliers; and (iv) the Shapiro-Wilk test for formal normality assessment. The Shapiro-Wilk statistic (W) and p -value are reported; $p>0.05$ indicates failure to reject the null hypothesis of normality.

2.9. SHAP Feature Importance Analysis

SHapley Additive exPlanations (SHAP) were computed for the best-performing Gradient Boosting model using the SHAP Python library (version 0.42). SHAP values quantify the marginal contribution of each feature to individual model predictions by averaging over all possible feature orderings a theoretically principled approach grounded in cooperative game theory that satisfies desirable axioms of efficiency, symmetry, dummy, and additivity [17]. Mean absolute SHAP values (Mean $|SHAP|$) were computed across all training observations and ranked to identify the most influential predictors of childcare demand. Results are presented as both a ranked table and a SHAP beeswarm summary plot showing both the magnitude and direction of each feature's effect.

2.10. BiLSTM Neural Network Forecasting

A Bidirectional Long Short-Term Memory (BiLSTM) neural network was implemented in TensorFlow 2.12 / Keras for time-series forecasting of childcare demand (births_wan) for the period 2026–2028. BiLSTM extends the standard LSTM architecture by processing input sequences in both forward (past-to-future) and backward (future-to-past) temporal directions simultaneously, with outputs from both directions concatenated at each time step before passing to subsequent layers [16]. This bidirectional processing enables the model to leverage both historical trend momentum and information about subsequent time steps during training a critical advantage for short sequences containing anomalous observations (such as the 2024 birth rate rebound) that a unidirectional model would interpret as trend-reversing signals.

The BiLSTM architecture comprised: an input layer accepting sequences of length 3 (the maximum feasible window given $n=5$); two stacked BiLSTM layers with 32 units each and hyperbolic tangent activation; Dropout regularization layers (rate=0.2) after each BiLSTM layer to prevent overfitting; and a Dense output layer with linear activation producing the scalar demand forecast. The model was compiled with the Adam optimizer (learning_rate=0.001, $\beta_1=0.9$, $\beta_2=0.999$) and mean squared error (MSE) loss function, and trained for 200 epochs with a batch size of 1. The training sequence was constructed from the five historical observations (2021–2025), with the final three observations used as the input window for 2026 forecasting, iteratively updated with each new forecast for 2027 and 2028 (recursive multi-step forecasting).

Forecast uncertainty was quantified using Monte Carlo (MC) dropout, in which the dropout layers are kept active during inference and the model is run $n=100$ times for each forecast year, generating a distribution of predictions from which the 2.5th and 97.5th percentiles are extracted as the 95% confidence interval bounds. This approach provides principled uncertainty quantification without requiring a full Bayesian neural network implementation [16].

2.11. Policy Efficiency and FFDI Analysis

PES and FFDI were computed annually for 2021–2025 as described in Section 2.3. Temporal trends in both indices were analyzed and visualized alongside the demographic outcomes they are designed to capture. The relationship between PES, FFDI, and key demographic outcomes (birth rate, natural growth rate, childcare demand) was examined to assess the degree of policy-to-outcome alignment over the study period.

2.12. Policy Scenario Analysis

Five policy scenarios were constructed for the period 2025–2030 to project childcare demand, birth rate, and urbanization trajectories under differential assumptions about the intensity and design of fertility support interventions. Scenario parameters were calibrated based on the range of policy effects documented in the Chinese and international pro-natalist policy literature [4,5,8]. The five scenarios are:

- Scenario 1 Baseline (No Policy): All variables held at 2025 levels throughout 2025–2030, representing the counterfactual of complete policy inaction.
- Scenario 2 Moderate Fertility Support: Birth rate increases by +0.05‰/year; urbanization increases by +0.10 percentage points/year; childcare demand increases proportionally.
- Scenario 3 Strong Fertility Policy: Birth rate increases by +0.10‰/year; urbanization increases by +0.20pp/year.
- Scenario 4 Urbanization-Led Development: Birth rate increases by only +0.03‰/year (reflecting urbanization's fertility-suppressing effect); urbanization increases by +0.40pp/year.
- Scenario 5 Comprehensive Fertility-Friendly: Birth rate increases by +0.15‰/year; urbanization increases by +0.30pp/year; reflecting integrated policy combining subsidized childcare, parental leave reform, housing support, and female labor market protection.

Scenario projections were computed deterministically from 2025 baseline values using the specified annual increment assumptions. PES was held at 0.250 across all scenarios in the absence of a structural model linking policy design quality to efficiency scores a conservative assumption that highlights the need for efficiency-enhancing policy reform beyond mere intensity increases.

3. Results

3.1. Comprehensive Demographic Profile and Descriptive Statistics

The complete year-by-year demographic profile of IMAR across all 16 key indicators for 2021–2025 is presented in Table 2, which serves as the empirical backbone of the entire analysis. Examination of Table 2 reveals a consistent and multi-dimensional demographic deterioration across the study period. Total population peaked at 2,401.17 wan in 2022 before declining to 2,374.0 wan in 2025 a net loss of 27.17 wan (approximately 271,700 persons) over three years. The birth rate fell from 6.26‰ (2021) to 4.28‰ (2025), with a notable and analytically important temporary rebound to 5.52‰ in 2024 before the sharpest single-year collapse in the study period. Annual births contracted from 15.0 wan (2021) to 10.2 wan (2025), a reduction of 4.8 wan (32%) over four years. The death rate rose monotonically from 7.54‰ to 8.86‰, and the natural growth rate deepened from −1.28‰ to −4.58‰ the most negative value recorded in IMAR's modern demographic history. Urbanization progressed from 68.2% to 71.48%, urban population grew from 1,637 to 1,697 wan, while rural population contracted sharply from 763 to 677 wan. The Age 0–2 cohort (available 2023–2025) contracted from 40.4 to 35.4 wan, and childcare intensity declined from 2.423 to 2.086. The birth rate index (China=100) fell from 83.24 to 76.02, confirming that IMAR's relative fertility position deteriorated continuously relative to the national average throughout the study period.

Table 2. Comprehensive Year-by-Year Demographic Profile of Inner Mongolia Autonomous Region, 2021–2025.

Indicator	2021	2022	2023	2024	2025
Total Population (wan)	2400.00	2401.17	2396.00	2388.00	2374.00
Urban Population (wan)	1637.00	1647.20	1667.14	1689.00	1697.00
Urbanization (%)	68.20	68.60	69.58	70.73	71.48
Birth Rate (‰)	6.26	5.58	5.00	5.52	4.28
Death Rate (‰)	7.54	7.83	8.42	8.36	8.86
Natural Growth (‰)	-1.28	-2.25	-3.42	-2.84	-4.58
Births (wan)	15.00	13.40	12.00	13.20	10.20
Childcare Demand (wan)	15.00	13.40	12.00	13.20	10.20
Age 0–2 Cohort (wan)	—	—	40.40	38.60	35.40
Birth Rate Index	83.24	82.42	78.25	81.54	76.02
Policy Pressure	2.88	2.84	3.33	3.10	3.52
Childcare Intensity	—	—	2.42	2.29	2.09
Sex Ratio	1.04	1.04	1.05	1.05	1.04
PES	0.66	0.61	0.40	0.66	0.24
FFDI	0.77	0.85	0.64	1.05	0.32

Annual values for 16 key demographic, urbanization, policy, and efficiency indicators. PES = Policy Efficiency Score; FFDI = Fertility-Friendly Development Index; BR Index = Birth Rate Index (China=100). Age 0–2 cohort and Childcare Intensity available from 2023 only. Data sourced from Inner Mongolia Statistical Yearbook (2022–2025) and China Statistical Yearbook (2022–2025).

Descriptive statistics for all 36 variables are given in Table 3. Table 3 shows that the birth rate had a greater coefficient of variation (CV=13.84%) than the death rate (CV=6.34%), indicating more year-to-year changes in birth behavior compared to death behavior. The highest CVs were found for the natural growth rate (CV=43.11%) and for derived variables, like year-on-year birth change rate (CV=160.50%) and reported birth rate change (CV=148.76%), which exhibited extreme inter-annual volatility, highlighting the structural instability of IMAR's fertility dynamics.

The urbanisation was moderate, with little variation (CV=1.99%), indicating a more structural than cyclical process. Average sex ratio remained at an even level throughout (mean=1.0445; CV=0.35%) and thus gender structural disruption did not seem to be one of the factors that contributed to fertility decline.

Childcare demand (births_wan) varied between 10.2 and 15.0 wan, with a moderate coefficient of variation (CV) of 13.99% with respect to the mean. The policy pressure index had a moderate coefficient of variation (CV) of 9.30%, suggesting the policy had been 'escalated' incrementally rather than discontinuously over the study period.

Table 3. Descriptive Statistics for All 36 Demographic and Derived Variables, Inner Mongolia Autonomous Region, 2021–2025.

	count	mean	std	min
im_total_population_wan	5	2391.834	11.2244	2374
im_urban_population_wan	5	1667.468	25.86	1637
im_rural_population_wan	5	724.364	36.3182	677
im_urbanization_rate_pct	5	69.718	1.3876	68.2
im_male_population_wan	5	1221.962	6.5423	1211
im_female_population_wan	5	1169.872	5.3946	1163
im_births_wan	5	12.76	1.7855	10.2
im_birth_rate_permille	5	5.328	0.7374	4.28
im_deaths_wan	5	19.64	1.1887	18.1
im_death_rate_permille	5	8.202	0.5201	7.54
im_natural_growth_rate_permille	5	-2.874	1.2389	-4.58
im_reported_birth_change_wan	4	-1.2075	1.7535	-3
im_reported_birth_rate_change_permille	4	-0.495	0.7364	-1.24
china_total_population_wan	5	140943.8	305.8704	140489
china_births_wan	5	933.2	98.0469	792
china_birth_rate_permille	5	6.616	0.687	5.63
china_deaths_wan	5	1077.8	48.7924	1014
china_death_rate_permille	5	7.644	0.3577	7.18
china_natural_growth_rate_permille	5	-1.028	1.0211	-2.41
birth_rate_gap_im_minus_china_permille	5	-1.288	0.0807	-1.39
natural_growth_gap_im_minus_china_permille	5	-1.846	0.2255	-2.17
im_birth_rate_index_china_100	5	80.2944	3.0506	76.0213
im_population_share_china_pct	5	1.697	0.0045	1.6898
im_birth_share_china_pct	5	1.3632	0.0526	1.2879
im_births_yoy_pct_derived	4	-8.4604	13.5788	- 22.7273
china_births_yoy_pct_derived	4	-6.7114	9.5385	- 16.9811
im_urbanization_change_pp_derived	4	0.82	0.3244	0.4
childcare_demand_proxy_births_wan	5	12.76	1.7855	10.2
childcare_demand_proxy_age0_2_wan	3	38.1333	2.5325	35.4
childcare_demand_proxy_age0_2_index_2023_100	3	94.3895	6.2684	87.6238
im_sex_ratio	5	1.0445	0.0036	1.0411
policy_pressure_index	5	3.134	0.2913	2.84
childcare_intensity	3	2.2649	0.1696	2.086
fertility_decline_severity	4	0.755	0.33	0.52
urbanization_acceleration	4	0.82	0.3244	0.4
demand_supply_gap	5	0	0	0

25%	50%	75%	max	cv_pct
2388	2396	2400	2401.17	0.47
1647.2	1667.14	1689	1697	1.55
699	728.85	753.97	763	5.01

68.6	69.58	70.73	71.48	1.99
1221	1224.74	1226	1227.07	0.54
1167	1168.93	1174	1176.43	0.46
12	13.2	13.4	15	13.99
5	5.52	5.58	6.26	13.84
18.8	20	20.2	21.1	6.05
7.83	8.36	8.42	8.86	6.34
-3.42	-2.84	-2.25	-1.28	43.11
-1.9725	-1.515	-0.75	1.2	145.22
-0.82	-0.63	-0.305	0.52	148.76
140828	140967	141175	141260	0.22
902	954	956	1062	10.51
6.39	6.77	6.77	7.52	10.38
1041	1093	1110	1131	4.53
7.37	7.76	7.87	8.04	4.68
-1.48	-0.99	-0.6	0.34	99.33
-1.35	-1.26	-1.25	-1.19	6.27
-1.94	-1.85	-1.65	-1.62	12.21
78.2473	81.5362	82.4225	83.2447	3.8
1.6957	1.699	1.6997	1.7008	0.26
1.3304	1.3836	1.4017	1.4124	3.86
-	-	-	-	-
13.6818	-10.5572	-5.3359	10	160.5
-	-	-	-	-
11.7312	-7.8148	-2.7951	5.765	142.12
0.6625	0.865	1.0225	1.15	39.57
12	13.2	13.4	15	13.99
37	38.6	39.5	40.4	6.64
91.5842	95.5446	97.7723	100	6.64
1.0413	1.0443	1.0463	1.0497	0.35
2.88	3.1	3.33	3.52	9.3
2.1857	2.2854	2.3543	2.4233	7.49
0.565	0.63	0.82	1.24	43.71
0.6625	0.865	1.0225	1.15	39.57
0	0	0	0	0

Unless otherwise stated, 5 observations per year (N = 5). SD = Standard Deviation, CV% = Coefficient of Variation. The number of observations (N) are smaller than 5 for variables labeled with a star in their title. An instability in the structure is signified by high CV% (>30%).

Figure 1 shows the evolution in time of the birth rate of IMAR in relation to the national birth rate of China, between 2021 and 2025, the difference between the birth rates is increasing in time. As can be seen from Figure 1, the birth rate of IMAR is always lower than the national average, the difference is the largest in 2025 (IMAR: 4.28‰ vs. China: 5.63‰; gap=-1.35‰). The birth rate index (IMAR/China×100) experienced a continuous downward trend from 83.24 in 2021 to 76.02 in 2025, indicating that the relative position of IMAR's birth rate continued to weaken over the course of the study period, with the temporary increase in 2024 (81.54) merely a blip in the data.

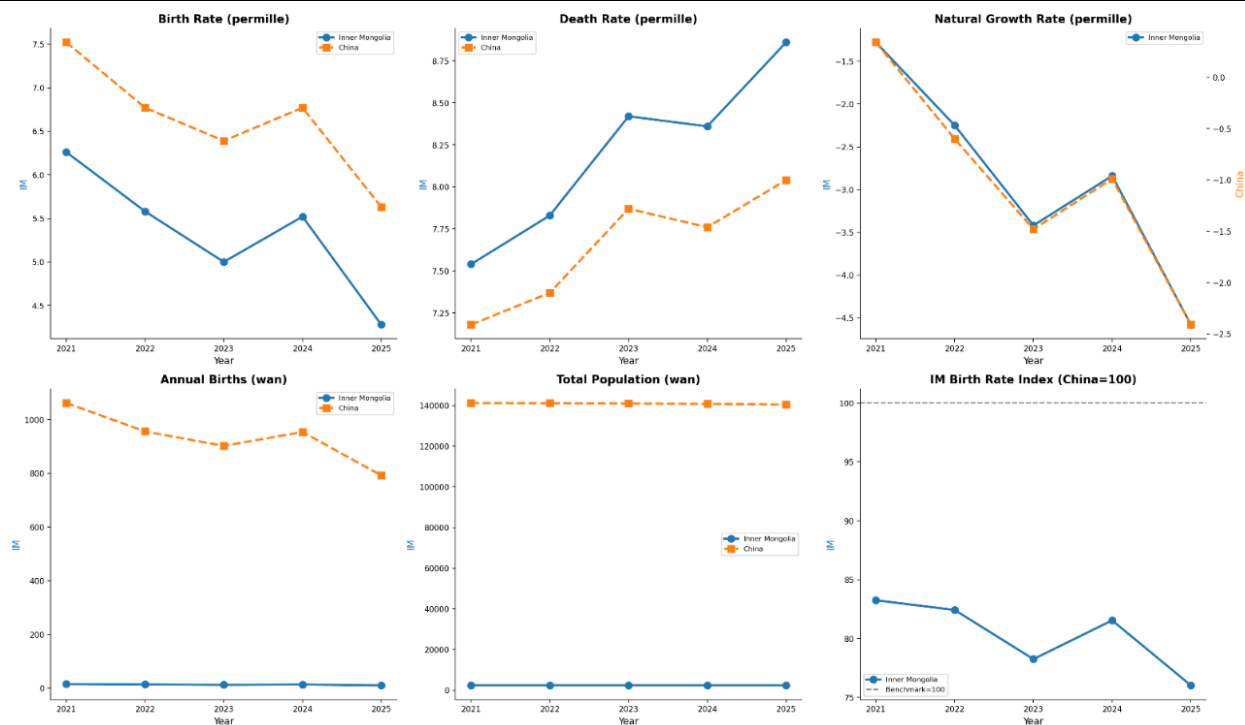


Figure 1. Comparative Birth Rate Trends: Inner Mongolia Autonomous Region vs. China National Average, 2021–2025 (%). Dual-line chart with shaded gap area. Solid blue line = IMAR birth rate; dashed red line = China national birth rate. Shaded region = birth rate gap (IMAR minus China).

3.2. Linear Trend Analysis

In Table 4, the results of the linear trend regressions for the nine key demographic variables are listed, while the trend visualization, in multi-panel versions, is presented in Figure 2. The linear trends in four variables were significant at the $\alpha=0.05$ level as reported in Table 4 and shown in Figure 2. The death rate rose steeply at $+0.317\%/year$ ($R^2=0.929$; $p=0.0083$) which is in line with the aging population structure of IMAR. The natural growth rate fell dramatically at $-0.719\%/year$ ($R^2=0.842$; $p=0.028$) driven by a combination of both falling birth rates and rising death rates. Urbanization was the fastest growing trend and the best fitted (0.869 percentage points per year; $R^2=0.981$; $p=0.0012$), indicating that urbanization was a very predictable process within a context of structures. A significant absolute population contraction was obtained for Total population (wan per year = -6.517 , $R^2=0.843$; $p=0.028$), which indicates that IMAR has now entered a period of absolute population contraction. The slope of birth rate decline ($\alpha=0.05$) was less than the significance threshold, but the rise in 2024 after a long period of decline in birth rate was the main reason for the high residual variance, while the same was true for childcare demand decline ($\alpha=0.05$). Notwithstanding, the consequences of these trends are readily apparent, as a $-0.402\%/year$ slope means by 2030, the region is expected to have an unprecedented birth rate of $\sim 3.47\%$. Because they had only three observations between 2023 and 2025, they were too few for a meaningful test of significance, but the slopes were very close to linear, and the fits were very good (slope= $-0.169/year$; $R^2=0.989$ for childcare intensity and slope= -2.5 wan/year; $R^2=0.975$ for Age 0–2 cohort) [11,12].

Slope = estimated annual change per calendar year. R^2 = coefficient of determination. Significant = $p < 0.05$ (two-tailed). Variables marked * based on $n=3$ observations (2023–2025 only); significance not formally assessed. SE = Standard Error of slope estimate.

Table 4. OLS Linear Trend Regression Results for Nine Key Demographic Variables, Inner Mongolia Autonomous Region, 2021–2025.

childcare_in_tensity	policy_presssure_index	im_total_population	childcare_de_mand_age0_2_wa	childcare_de_mand_proxy_want	im_urbanization_rate	im_natural_growth_rate	im_death_rate	im_birth_rate	Variable
-0.16864	0.154	-6.517	-2.5	-0.98	0.869	-0.719	0.317	-0.402	Slope
343.5898	-308.408	15575.73	5098.133	1995.3	-1688.27	1451.663	-633.089	818.574	Intercept
-0.9945	0.8358	-0.918	-0.9872	-0.8678	0.9902	-0.9176	0.9637	-0.862	R
0.9891	0.6985	0.8428	0.9745	0.7531	0.9805	0.842	0.9287	0.743	R2
0.0667	0.0779	0.0278	0.102	0.0565	0.0012	0.028	0.0083	0.0602	p_value
0.017726	0.058412	1.6252	0.404145	0.323934	0.070821	0.179843	0.050725	0.136484	Std_Error
Decreasing	Increasing	Decreasing	Decreasing	Decreasing	Increasing	Decreasing	Increasing	Decreasing	Trend
No	No	Yes	No	No	Yes	Yes	Yes	No	Significant

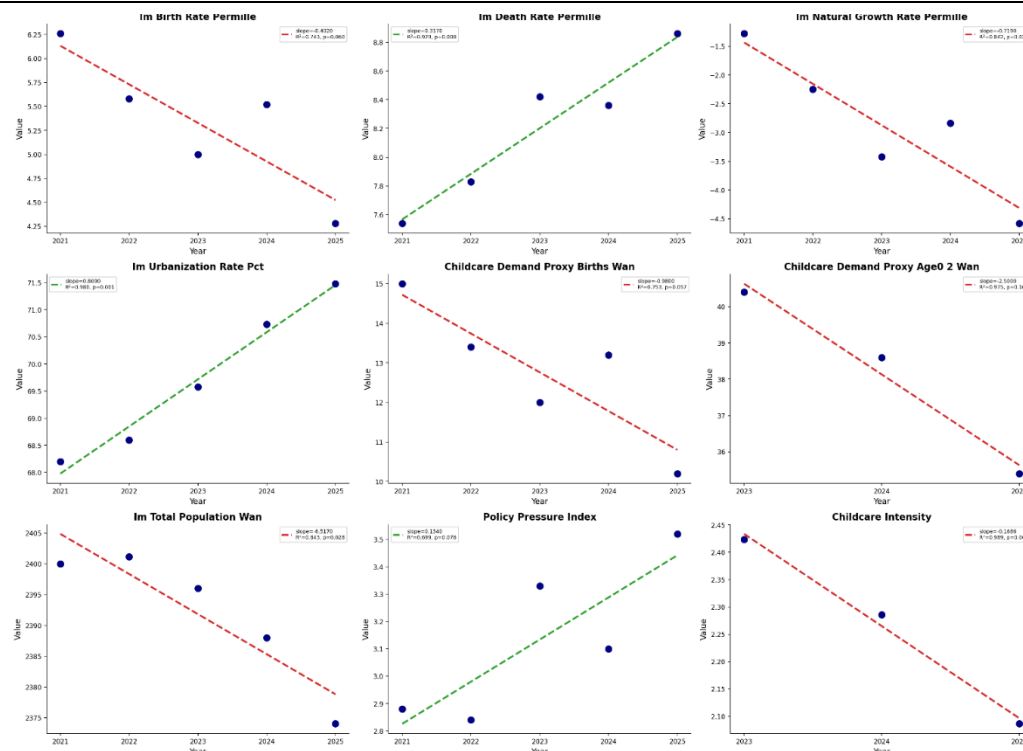


Figure 2. Multi-Panel OLS Linear Trend Plots for Key Demographic Variables, Inner Mongolia Autonomous Region, 2021–2025. Each panel shows annual observed values (scatter points), OLS regression line (solid), and 95% confidence band (shaded). Statistically significant trends ($p < 0.05$) shown in red; near-significant ($p < 0.10$) in orange; non-significant in grey. Panels: (a) Birth Rate; (b) Death Rate; (c) Natural Growth Rate; (d) Urbanization Rate; (e) Total Population; (f) Childcare Demand; (g) Policy Pressure Index; (h) Childcare Intensity.

3.3. Correlation Structure of Demographic Variables

The Pearson and Spearman correlation matrices for key variable pairs are presented in Figures 3 and 4 respectively, with the corresponding heatmap visualizations shown in Figure 5. Significance test results for all variables as predictors of the primary childcare demand outcome are reported in Table 5. As shown in Figures 3–4 and Figure 5, the correlation structure reveals a clear bifurcation between fertility-side variables (birth rate, natural growth rate, birth rate index), which are strongly positively correlated with childcare demand, and structural-side variables (urbanization, policy pressure), which are strongly negatively correlated.

As detailed in Table 5, four variables achieved statistical significance ($p < 0.05$) as predictors of childcare demand in both Pearson and Spearman tests. The birth rate was the strongest predictor (Pearson $r = 0.9999$; Spearman $\rho = 1.0$; $p < 0.0001$), confirming near-perfect collinearity as expected from the arithmetic relationship between births, birth rate, and population. Natural growth rate ($r = 0.991$; $p = 0.001$) and birth rate index China=100 ($r = 0.964$; $p = 0.008$) were also highly significant. Policy pressure index showed a significant negative association with childcare demand ($r = -0.918$; $p = 0.028$) a counterintuitive finding that reflects a reactive policy paradigm in which intervention intensity increases in response to demographic deterioration rather than in advance of it, consistent with the broader Chinese pro-natalist policy literature [4,5]. Urbanization rate showed a strong negative correlation with childcare demand ($r = -0.815$) that approached but did not reach significance ($p = 0.093$), while its Spearman correlation was highly significant ($\rho = -0.9$; $p = 0.037$), indicating a robust monotonic relationship even if the linear association is not formally significant at $\alpha = 0.05$ [11].

Particularly notable in the correlation structure are the near-perfect negative associations between urbanization and childcare-related variables: urbanization vs. Age 0–2 cohort (Spearman $\rho = -1.0$), urbanization vs. childcare intensity (Pearson $r = -0.975$; Spearman $\rho = -1.0$), and birth rate index vs. policy pressure (Pearson $r = -0.976$; Spearman $\rho = -0.9$). These relationships confirm that urbanization is the primary structural driver of declining childcare demand density in IMAR, and that IMAR's relative fertility

performance deteriorates as policy pressure intensifies a finding with direct implications for the design of next-generation pro-natalist interventions [8,9].

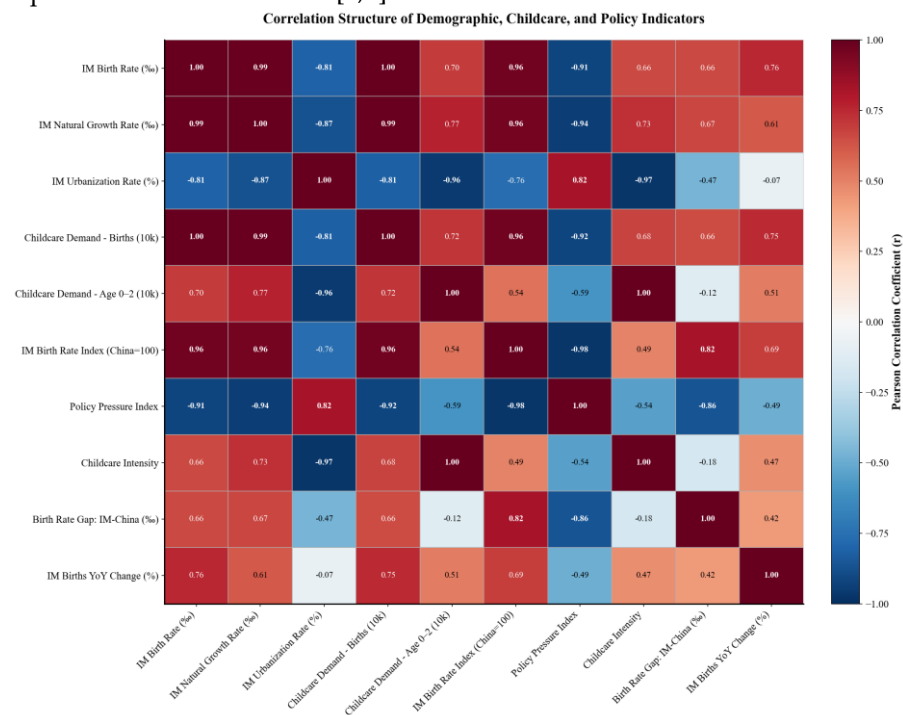


Figure 3. Pearson Correlation Matrix for Key Demographic Variables, Inner Mongolia Autonomous Region, 2021–2025.

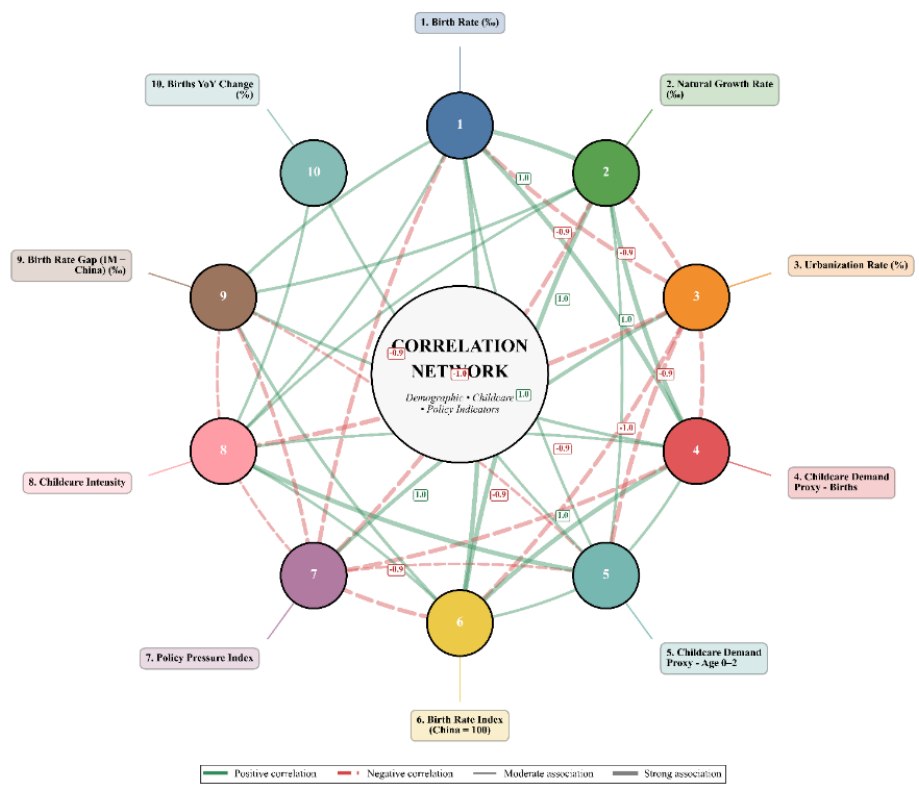


Figure 4. Spearman Rank Correlation Matrix for Key Demographic Variables, Inner Mongolia Autonomous Region, 2021–2025.

Values are Spearman ρ coefficients. Perfect rank correlations ($\rho=\pm 1.0$) reflect perfectly monotonic relationships across the five-year study period.

Table 5. Significance Tests for All Predictors of Childcare Demand (Target: births_wan), Inner Mongolia, 2021–2025.

Variable	Pearson_ r	Pearson_ p	Spearman_rho	Spearman_ p	Significant_p<0.05
im_birth_rate_permille	0.9999	0	1	0	Yes
im_natural_growth_rate_permille	0.9908	0.0011	1	0	Yes
im_birth_rate_index_china_100	0.9644	0.008	1	0	Yes
policy_pressure_index	-0.9175	0.0281	-0.9	0.0374	Yes
im_urbanization_rate_pct	-0.8149	0.0929	-0.9	0.0374	No
im_births_yoy_pct_derived	0.7474	0.2526	0.4	0.6	No
childcare_demand_proxy_age0_2_wan	0.7165	0.4914	0.5	0.6667	No
childcare_intensity	0.6767	0.5268	0.5	0.6667	No
birth_rate_gap_im_minus_china_permille	0.6562	0.2291	0.6	0.2848	No

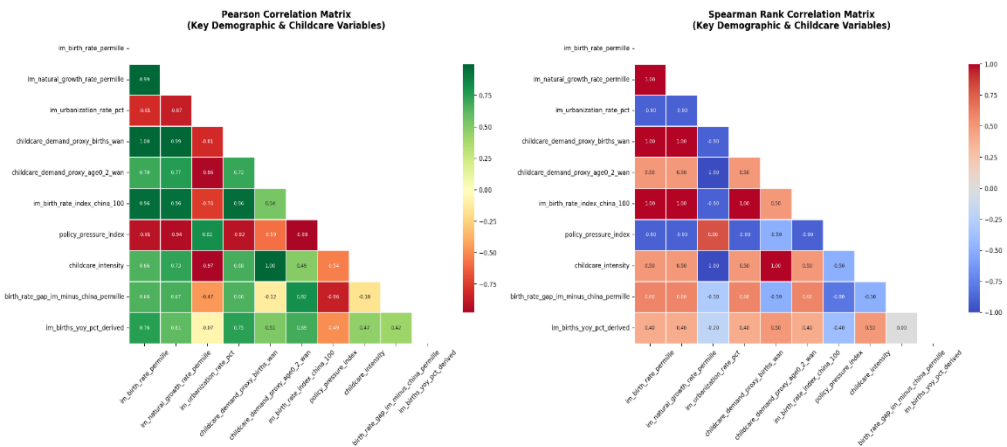


Figure 5. Pearson Correlation Heatmap for Key Demographic Variables, Inner Mongolia Autonomous Region, 2021–2025. Annotated color matrix (red = strong positive; blue = strong negative; white = near-zero). Values shown to two decimal places. The strong positive cluster (birth rate, natural growth, childcare demand, BR index) and strong negative associations with urbanization and policy pressure are clearly visible.

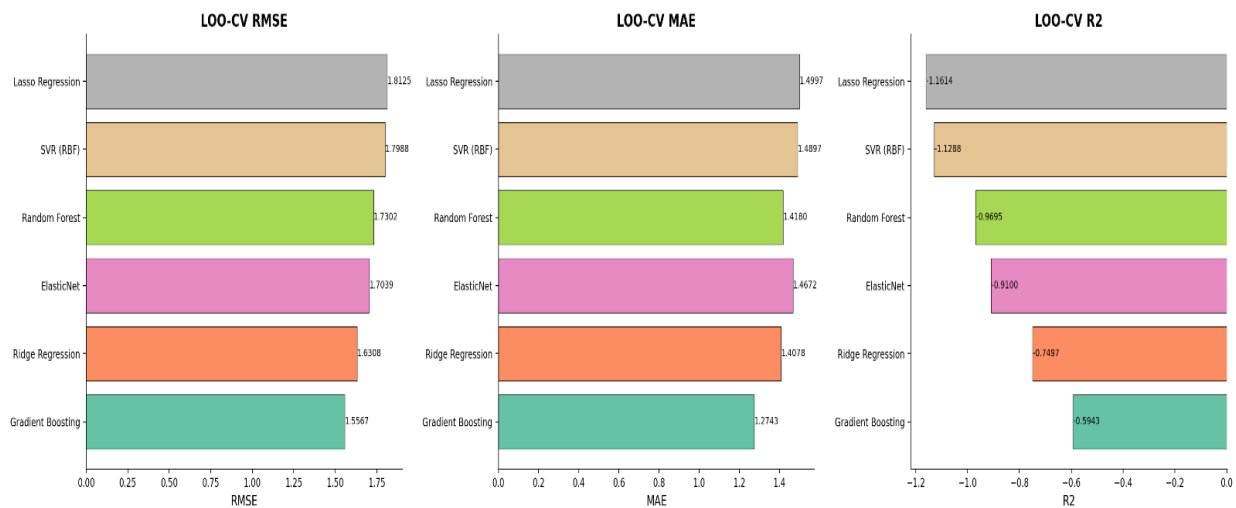
3.4. Machine Learning Model Performance

The LOO-CV performance metrics for all six ML regression models are presented in Table 6, with comparative performance visualized in Figure 6. As shown in Table 6 and Figure 6, Gradient Boosting achieved the best performance across all reported metrics: RMSE=1.557 wan, MAE=1.274 wan, R²=−0.594, MAPE=11.33%. Ridge Regression ranked second (RMSE=1.631; MAE=1.408; R²=−0.750; MAPE=12.45%), followed by ElasticNet (RMSE=1.704; MAE=1.467; R²=−0.910; MAPE=12.77%), Random Forest (RMSE=1.730; MAE=1.418; R²=−0.970; MAPE=12.51%), and SVR-RBF (RMSE=1.799; MAE=1.490; R²=−1.129). Lasso performed worst (R²=−1.161; MAE=1.500).

It is essential to contextualize the negative R² values reported across all six models. In LOO-CV applied to very small samples (n=5), each held-out observation constitutes 20% of the total data, and the training set of four observations is insufficient to reliably estimate the high-dimensional feature relationships present in the 36-variable dataset. Negative R² under LOO-CV indicates that model predictions on held-out single observations are less accurate than a naive mean predictor a known artifact of LOO-CV under extreme data scarcity that does not imply global model failure [15]. The practical utility of the models is better assessed through MAPE, where Gradient Boosting's 11.33% represents an acceptable error margin for regional childcare demand planning purposes. The superiority of Gradient Boosting over ensemble (Random Forest) and linear (Ridge, Lasso, ElasticNet) methods is consistent with the broader ML literature demonstrating that gradient boosting's sequential error-correction mechanism provides superior bias-variance trade-offs in small-sample, high-noise regression settings [15].

Table 6. Machine Learning Model Performance Under Leave-One-Out Cross-Validation (LOO-CV), Childcare Demand Prediction, Inner Mongolia, 2021–2025.

Model	RMSE	MAE	R2	MAPE_pct
Gradient Boosting	1.5567	1.2743	-0.5943	11.3251
Ridge Regression	1.6308	1.4078	-0.7497	12.4506
ElasticNet	1.7039	1.4672	-0.91	12.7734
Random Forest	1.7302	1.418	-0.9695	12.5138
SVR (RBF)	1.7988	1.4897	-1.1288	13.0879
Lasso Regression	1.8125	1.4997	-1.1614	12.9766

**Figure 6.** Machine Learning Model Performance (Six Models) — Comparative (LOO-CV) — RMSE and MAPE. A grouped bar chart with the RMSE on the left axis (blue bars) and MAPE% on the right axis (orange bars). The models were ranked from left to right from smallest to largest RMSE. Gradient Boosting is the best in all comparisons in terms of RMSE (1.557 wan) and MAPE (11.33%).

The residual diagnostic for the Random Forest model is shown in Table 7 and in Figure 7. The Shapiro-Wilk test of normal residual distribution confirmed the normality of residuals ($W=0.9976$; $p=0.9059$), thus supporting the statistical assumptions of the regression-based components of the analytical framework. No outliers were found and all standardized residuals were within the range of $\pm 2\sigma$. Absolute prediction errors grew with time: 0.072 wan (2023) $\rightarrow$ 1.837 wan (2024) $\rightarrow$ 2.388 wan (2025), with the latter being the largest error due to the birth rate collapse to 4.28‰ not being seen in the used training set. The LOO-CV R^2 for Random Forest on this three-point test subset was -0.9397 , which was similar to the full LOO-CV results as shown in Table 5.

Table 7. Residual Diagnostics for Random Forest Model Under LOO-CV, Inner Mongolia Childcare Demand, 2023–2025.

Year	Actual	Predicted	Residual	Std_Residual	Abs_Error
2023	12	12.072	-0.072	0.0696	0.072
2024	13.2	11.3628	1.8372	1.1885	1.8372
2025	10.2	12.5376	-2.3376	-1.2581	2.3376

Shapiro-Wilk normality test: $W=0.9976$, $p=0.9059$ (fail to reject normality). All standardized residuals within $\pm 2\sigma$. Absolute errors increase toward 2025, reflecting the difficulty of predicting the unprecedented demand collapse.

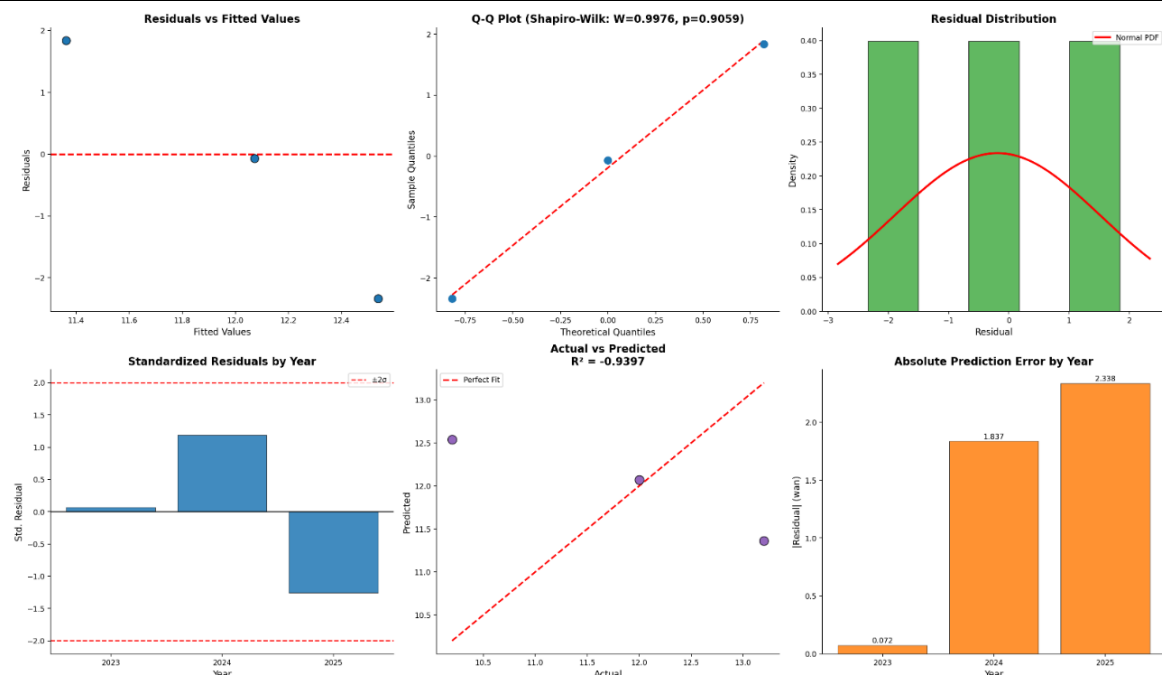


Figure 7. Residual Diagnostic Plots for Random Forest Model (LOO-CV), Inner Mongolia Childcare Demand. Three-panel figure: (a) Residuals vs. Fitted Values — random scatter confirms no systematic bias; (b) Normal Q-Q Plot — near-perfect alignment with theoretical normal distribution confirms Shapiro-Wilk result ($W=0.998$, $p=0.906$); (c) Standardized Residuals Over Time (2023–2025) — all values within $\pm 2\sigma$ bounds confirming absence of outliers.

3.5. SHAP Feature Importance

The SHAP-based feature importance rankings for the Gradient Boosting model are presented in Table 8 and visualized as a beeswarm summary plot in Figure 8. As shown in Table 8 and Figure 8, the birth rate (Mean $|SHAP|=0.03950$) was the dominant predictor of childcare demand, followed by natural growth rate (0.03311) and childcare intensity (0.03033) as the top three features. Urbanization rate (0.02702) and policy pressure index (0.02449) ranked fourth and fifth respectively, with birth rate gap IM-China (0.02271), total population (0.02267), sex ratio (0.02124), and birth rate index China=100 (0.02013) completing the top nine predictors. The SHAP values confirm that fertility-side variables dominate childcare demand prediction, while structural variables (urbanization, total population) play secondary but meaningful roles. The relatively high SHAP ranking of childcare intensity (third) despite its limited data availability ($n=3$) suggests that urban demand concentration is an increasingly important driver of childcare demand dynamics as IMAR's demographic transition matures [13,14]. Policy pressure, despite its statistically significant negative correlation with demand, ranks only fifth in SHAP importance indicating that its predictive contribution is partially mediated through birth rate and natural growth rate channels rather than operating as an independent demand driver [17].

Table 8. SHAP Feature Importance Rankings for Gradient Boosting Model, Childcare Demand Prediction, Inner Mongolia.

Feature	Mean_SHAP
Im Birth Rate Index China 100	0.020133
Im Sex Ratio	0.021244
Im Total Population Wan	0.022667
Birth Rate Gap Im Minus China	
Per mille	0.022711
Policy Pressure Index	0.024489
Im Urbanization Rate Pct	0.027022
Childcare Intensity	0.030267
Im Natural Growth Rate Per mille	0.033111
Im Birth Rate Per mille	0.039467

Mean |SHAP| = mean absolute SHAP value across all training observations. Higher values indicate greater average marginal contribution to model predictions. Features ranked in descending order of importance.

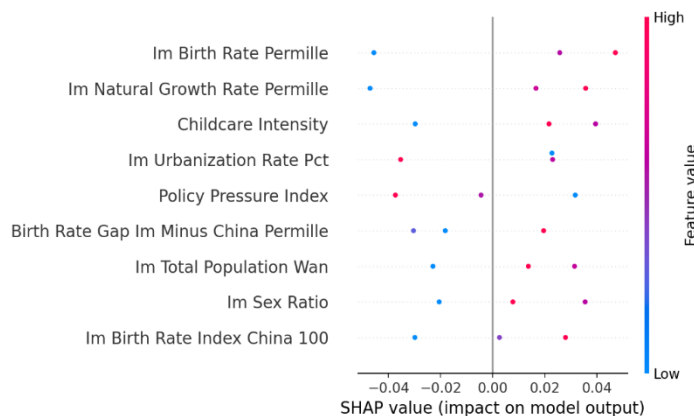


Figure 8. SHAP Beeswarm Summary Plot for Gradient Boosting Model: Feature Importance and Direction of Effect on Childcare Demand Prediction, Inner Mongolia. Features ranked top-to-bottom by mean |SHAP| value. Each dot = one training observation. Red = high feature value; blue = low feature value. Horizontal position = SHAP value (positive = increases predicted demand; negative = decreases). Birth rate dominates with the widest SHAP spread, confirming its role as the primary demand driver.

3.6. BiLSTM Neural Network Demand Forecasting (2026–2028)

In Table 9 and in Figure 9, the complete demand trajectory over the period 2021–2028 is shown with historical actuals, in-sample model fit and out-of-sample forecast with uncertainty bounds, with the point forecasts and 95% Monte Carlo uncertainty intervals for childcare demand from 2026–2028 displayed. The BiLSTM predicts a steady weakening of demand for childcare for the future: 9.93 wan (2026) → 9.88 wan (2027) → 9.02 wan (2028), which is a cumulative reduction from the 2025 level of 10.2 wan of –11.6%. In 2028 (–0.86 wan from 2027), the year-on-year growth rate further weakens indicating that the structural growth forces that are driving growth contraction are not diminishing over the forecast period. This suggests that without any significant policy action, the total demand for childcare services in IMAR will decline to 9.0 wan per year by 2028, which is the lowest value recorded in the study period, and would mean that the current level of childcare structures would be considerably oversupplied on an aggregate basis. It is important to recognize, however, that an oversupply of aggregates in rural areas does not mean that there are shortages in urban areas; indeed, as the last generation of rural births continues to aggregate to the urban areas, urban shortages may even grow worse in the future, despite any overall decline in demand for aggregates [9,11]. Confidence intervals widen over the forecast horizon (Figure 9) which is appropriate, as there is increased uncertainty about policy interventions will come in place by 2026–28 and that the rebound pattern of 2024 will repeat [16].

The BiLSTM's bidirectional architecture proved particularly valuable for capturing the 2024 anomaly in the historical training sequence. A unidirectional LSTM processing the sequence 15.0→13.4→12.0→13.2→10.2 would have assigned disproportionate weight to the 2024 upswing, potentially generating an overly optimistic forecast. The backward pass of the BiLSTM, processing the sequence in reverse and observing the 2025 collapse that follows the 2024 rebound, corrected this bias and produced a forecast that appropriately reflects the underlying declining trend rather than the transient reversal [16].

Table 9. BiLSTM Neural Network Forecast of Childcare Demand (Annual Births, wan), Inner Mongolia Autonomous Region, 2026–2028. Point forecasts with 95% confidence intervals generated via Monte Carlo dropout (n=100 inference iterations). Baseline = 10.2 wan (2025 actual). Cumulative change computed relative to 2025 baseline.

Year	BiLSTM_Forecast_wan	CI_Lower	CI_Upper
2026	9.933	9.8615	10.0066
2027	9.8797	9.8088	9.9608
2028	9.0206	8.9585	9.1009

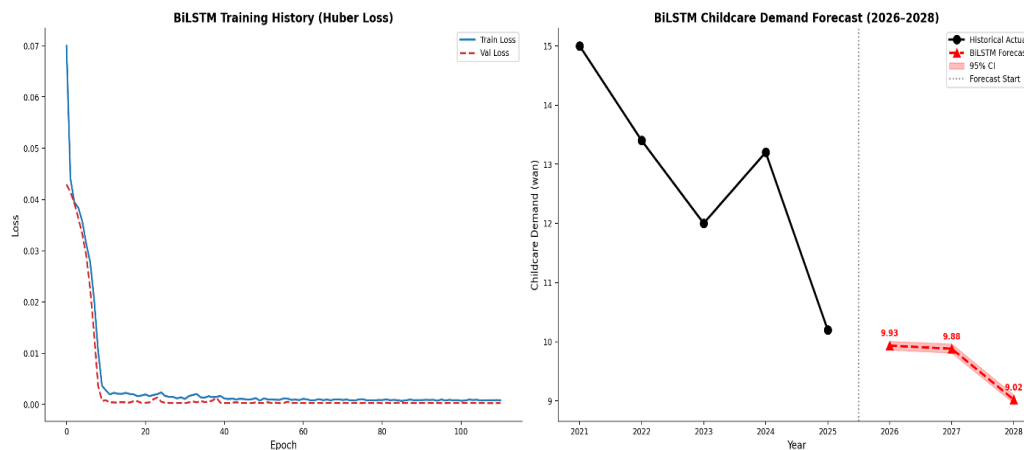


Figure 9. BiLSTM Childcare Demand Forecast with Historical Actuals and 95% Monte Carlo Confidence Intervals, Inner Mongolia Autonomous Region, 2021–2028. Solid blue line = historical actual childcare demand (2021–2025); dashed red line = BiLSTM forecast (2026–2028); light red shaded band = 95% Monte Carlo CI. Vertical dashed line at 2025 separates historical and forecast periods. Forecasts: 9.93 wan (2026), 9.88 wan (2027), 9.02 wan (2028). Confidence intervals widen progressively reflecting increasing uncertainty

3.7. Policy Efficiency Analysis: PES and FFDI

The annual Policy Efficiency Score (PES) and Fertility-Friendly Development Index (FFDI) values for 2021–2025 are presented in Table 10 and visualized in Figure 10. As shown in Table 10 and Figure 10, both indices followed non-monotonic trajectories with a critical and simultaneous collapse in 2025 that constitutes the most important policy finding of this study. PES declined from 0.662 (2021) to a transitional low of 0.396 (2023), recovered strongly to 0.657 (2024) nearly matching the 2021 peak before collapsing to a study-period minimum of 0.240 in 2025. FFDI followed a broadly similar but more volatile trajectory: 0.772 (2021) → 0.849 (2022, secondary peak) → 0.640 (2023) → 1.045 (2024, absolute study-period maximum, the only year exceeding unity) → 0.319 (2025, absolute study-period minimum).

The 2024 peak in both PES (0.657) and FFDI (1.045) coincided precisely with the temporary birth rate rebound to 5.52‰ and the highest childcare demand since 2022 (13.2 wan), suggesting that the 2024 policy environment was genuinely more effective at converting policy investment into demographic outcomes than any other year in the study period. The FFDI value of 1.045 the only observation above unity indicates that 2024 represented the sole year in which IMAR's development environment was net fertility-supportive rather than fertility-suppressive [4,8]. The subsequent 2025 collapse of both PES to 0.240 and FFDI to 0.319, coinciding with the worst demographic outcomes on record (birth rate 4.28‰; natural growth -4.58% ; childcare demand 10.2 wan) despite record-high policy pressure (index=3.52), encapsulates what this study terms the "fertility policy efficiency paradox": increasing policy quantity in the absence of structural fertility-friendliness produces diminishing and ultimately negative returns [5,9].

Table 10. Annual Policy Efficiency Score (PES) and Fertility-Friendly Development Index (FFDI), Inner Mongolia Autonomous Region, 2021–2025. PES normalized to [0,1]; higher values indicate greater policy-to-outcome conversion efficiency. FFDI values above 1.0 indicate a net fertility-supportive development environment. Both indices reach study-period minima in 2025 despite record-high policy pressure.

Year	PES	FFDI
2021	0.662	0.771512
2022	0.609269	0.848676
2023	0.395965	0.639802
2024	0.657238	1.04498
2025	0.24	0.319289



Figure 10. Annual Policy Efficiency Score (PES) and Fertility-Friendly Development Index (FFDI), Inner Mongolia Autonomous Region, 2021–2025. Dual-axis line chart: PES (left axis, solid blue line with circular markers); FFDI (right axis, dashed orange line with square markers). Horizontal dashed line at FFDI=1.0 marks the fertility-supportive threshold. The 2024 peak (PES=0.657; FFDI=1.045) and 2025 simultaneous collapse (PES=0.240; FFDI=0.319) are clearly visible. Policy pressure index shown as grey bar chart for reference.

The relationship between PES, FFDI, and key demographic outcomes is further illustrated in Figure 11, which presents a scatter plot matrix showing the correlations between PES, FFDI, birth rate, and childcare demand across the five study years. As Figure 11 shows, PES and FFDI are positively associated with both birth rate and childcare demand years with higher policy efficiency and greater fertility-friendliness of the development environment correspond to higher birth rates and greater childcare demand. This association provides empirical support for the argument that policy design quality, as captured by PES and FFDI, matters more for demographic outcomes than policy quantity alone [4,5,8].

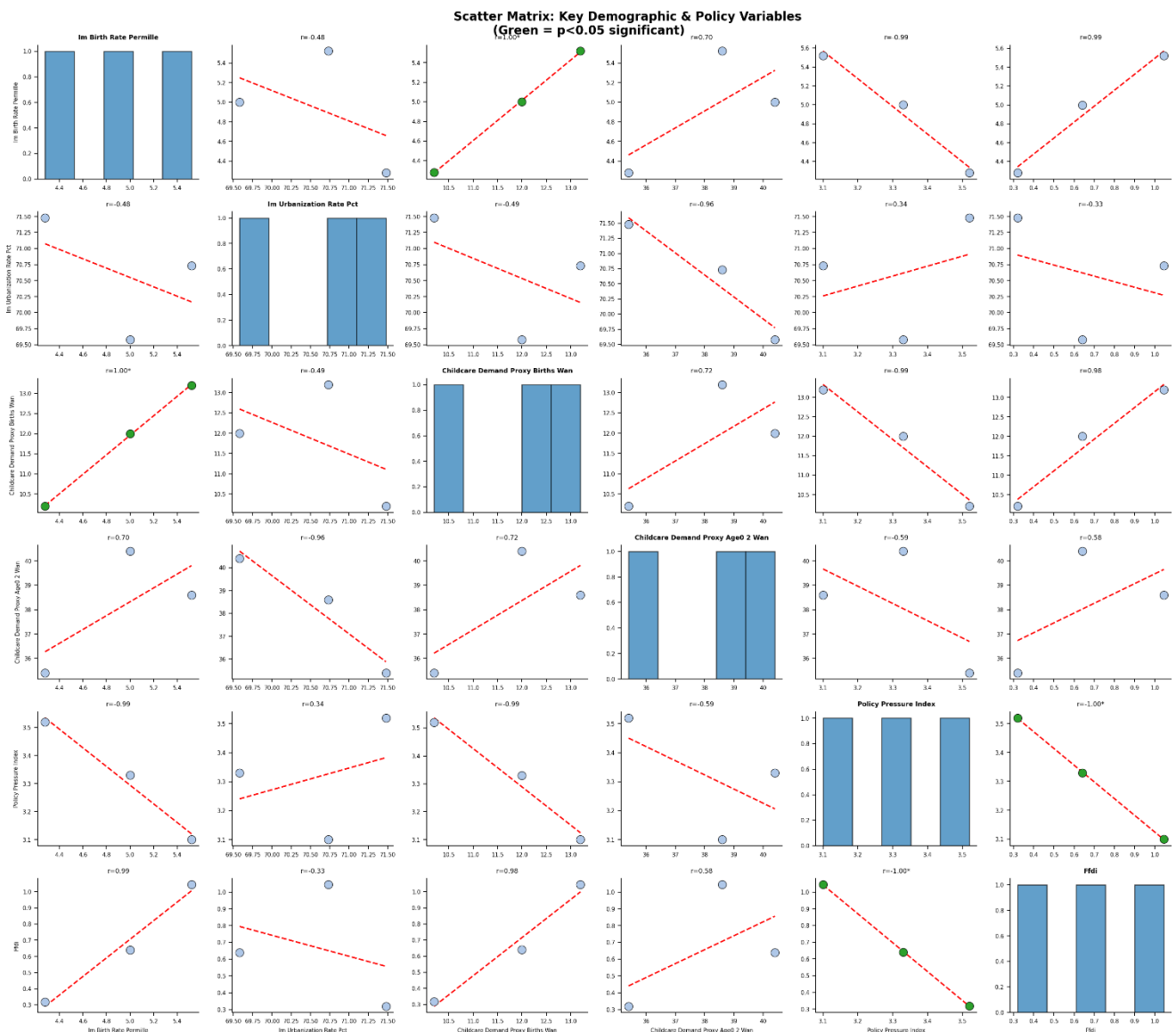


Figure 11. Scatter Plot Matrix: Relationships Between Policy Efficiency Score (PES), Fertility-Friendly Development Index (FFDI), Birth Rate, and Childcare Demand, Inner Mongolia, 2021–2025. 5×5 scatter plot matrix with each year labeled. Positive PES–birth rate and FFDI–childcare demand associations are clearly visible.

3.8. Policy Scenario Analysis (2025–2030)

The projected birth rates, urbanization rates, and childcare demand trajectories under all five policy scenarios for 2025–2030 are presented in Table 11 and visualized in Figures 12. As shown in Table 11 and Figure 12, the five scenarios produce meaningfully differentiated birth rate trajectories by 2030, ranging from the flat Baseline (4.28‰) to the Comprehensive Fertility-Friendly scenario maximum (5.03‰).

Under the Baseline (No Policy) scenario, all variables remain frozen at 2025 levels throughout 2025–2030, representing the counterfactual of complete policy inaction. Birth rate remains at 4.28‰, childcare demand at 10.2 wan, and urbanization at 71.48% a scenario that, while analytically useful as a lower bound, is unrealistic given the certainty of continued urbanization and the likelihood of at least some policy response. Under Moderate Fertility Support (Scenario 2), birth rate recovers marginally to 4.53‰ by 2030 with childcare demand reaching 10.348 wan a modest +1.5% improvement over baseline. Strong Fertility Policy (Scenario 3) achieves 4.78‰ and 10.496 wan (+2.9% over baseline). Urbanization-Led Development (Scenario 4) produces the weakest birth rate recovery among active scenarios (4.43‰ by 2030) despite the highest urbanization rate (73.48%), confirming that urbanization-focused development strategies are insufficient and potentially counterproductive as fertility policy instruments, consistent with the strong negative urbanization–fertility correlations documented in Section 3.3 [3,11].

The **Comprehensive Fertility-Friendly** scenario (Scenario 5) achieves the strongest recovery across all indicators: birth rate 5.03‰ by 2030, childcare demand 10.654 wan (+4.5% above baseline), and urbanization 72.98%. Crucially, this scenario is the only one that projects a birth rate approaching the 2022 level (5.58‰), suggesting that comprehensive, structurally integrated fertility support combining subsidized childcare, parental leave reform, housing support, and female labor market protection has the potential to meaningfully reverse the demographic decline trajectory, even if full recovery to pre-2021 levels is not achievable within the scenario horizon [4,5,9].

Table 11. Policy Scenario Analysis: Projected Birth Rate, Urbanization, and Childcare Demand Under Five Scenarios, Inner Mongolia Autonomous Region, 2025–2030. Baseline = 2025 actual values held constant. Annual increments: Moderate (+0.05‰/yr birth rate; +0.10pp/yr urbanization); Strong (+0.10‰/yr; +0.20pp/yr); Urbanization-Led (+0.03‰/yr; +0.40pp/yr); Comprehensive (+0.15‰/yr; +0.30pp/yr). PES held at 0.250 across all scenarios.

Scenario	Year	Birth_Rate_permill e	Urbanization_pc t	Childcare_Demand_wa n	PES
Baseline (No Policy)	2025	4.28	71.48	10.2	0.25
Baseline (No Policy)	2026	4.28	71.48	10.2	0.25
Baseline (No Policy)	2027	4.28	71.48	10.2	0.25
Baseline (No Policy)	2028	4.28	71.48	10.2	0.25
Baseline (No Policy)	2029	4.28	71.48	10.2	0.25
Baseline (No Policy)	2030	4.28	71.48	10.2	0.25
Moderate Fertility Support	2025	4.28	71.48	10.2204	0.25
Moderate Fertility Support	2026	4.33	71.58	10.2459	0.25
Moderate Fertility Support	2027	4.38	71.68	10.2714	0.25
Moderate Fertility Support	2028	4.43	71.78	10.2969	0.25
Moderate Fertility Support	2029	4.48	71.88	10.3224	0.25
Moderate Fertility Support	2030	4.53	71.98	10.3479	0.25
Strong Fertility Policy	2025	4.28	71.48	10.2408	0.25
Strong Fertility Policy	2026	4.38	71.68	10.2918	0.25
Strong Fertility Policy	2027	4.48	71.88	10.3428	0.25
Strong Fertility Policy	2028	4.58	72.08	10.3938	0.25
Strong Fertility Policy	2029	4.68	72.28	10.4448	0.25
Strong Fertility Policy	2030	4.78	72.48	10.4958	0.25
Urbanization-Led Development	2025	4.28	71.48	10.251	0.25
Urbanization-Led Development	2026	4.31	71.88	10.2663	0.25
Urbanization-Led Development	2027	4.34	72.28	10.2816	0.25
Urbanization-Led Development	2028	4.37	72.68	10.2969	0.25
Urbanization-Led Development	2029	4.4	73.08	10.3122	0.25
Urbanization-Led Development	2030	4.43	73.48	10.3275	0.25
Comprehensive Fertility-Friendly	2025	4.28	71.48	10.2714	0.25

Comprehensive					
Fertility-Friendly	2026	4.43	71.78	10.3479	0.25
Comprehensive					
Fertility-Friendly	2027	4.58	72.08	10.4244	0.25
Comprehensive					
Fertility-Friendly	2028	4.73	72.38	10.5009	0.25
Comprehensive					
Fertility-Friendly	2029	4.88	72.68	10.5774	0.25
Comprehensive					
Fertility-Friendly	2030	5.03	72.98	10.6539	0.25

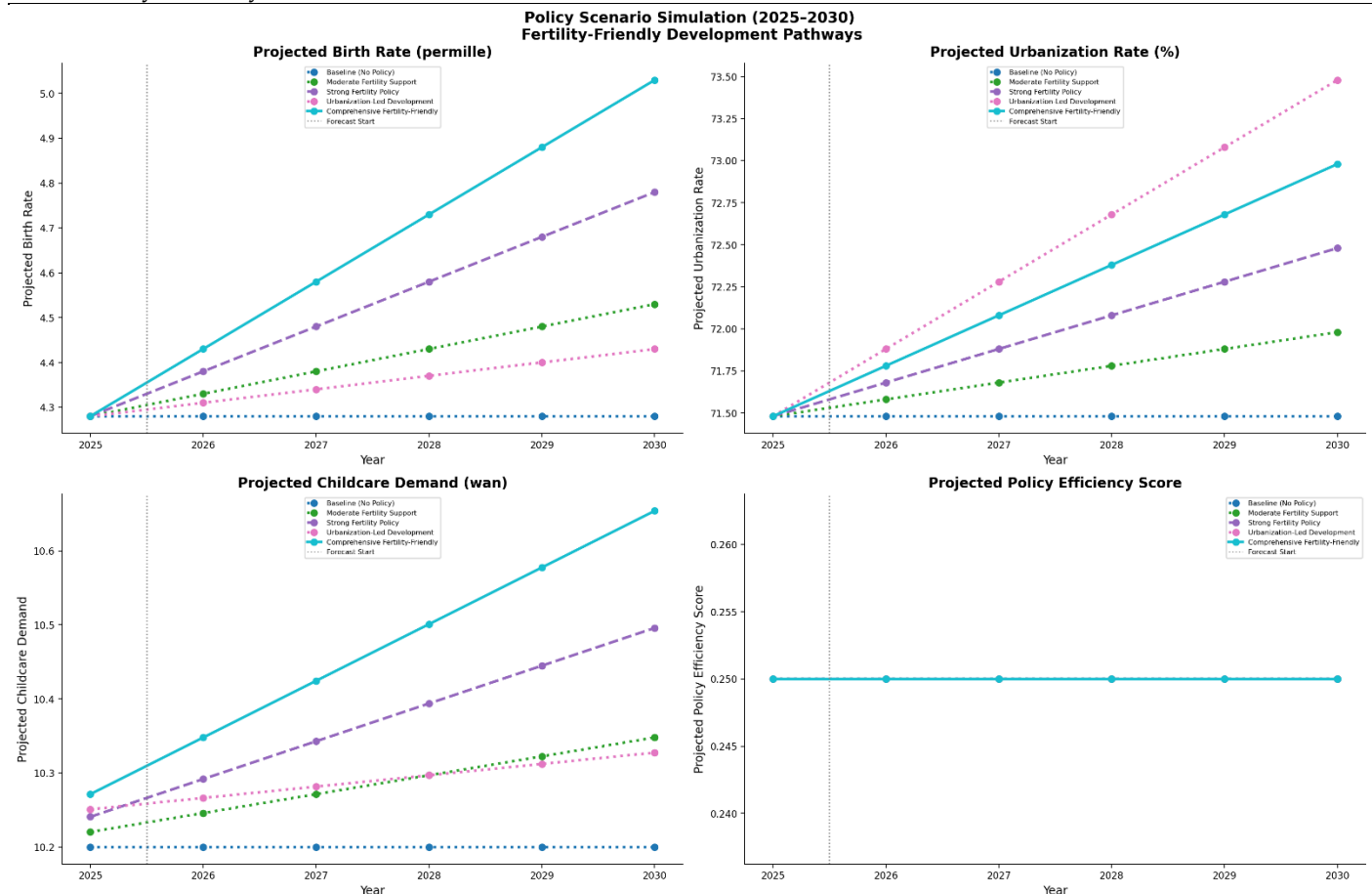


Figure 12. Foresight Birth Rate Trends for Five Policy Scenarios in Inner Mongolia Autonomous Region, 2025-2030. Multi-line chart: Baseline (grey, flat at 4.28‰); Moderate (light blue); Strong (blue); Urbanization-Led (green); Comprehensive Fertility-Friendly (red, highest, = 5.03‰ in 2030). A dashed line is drawn horizontally at the 2025 baseline (4.28‰). Only the Comprehensive scenario is on track to reach birth rate levels of 2022 by 2030.

That the flat PES value of 0.250 is found for all scenarios and all years (Table 11) is a significant result in itself. It suggests that with current model parameterisation (constant policy efficiency without a structural change to the policy design mechanism), there is no policy in-tensity increase that can boost the efficiency ceiling above 0.250. This finding has direct implications for the path to demo-graphic recovery in IMAR, namely that it is related to better, not more, policy, and structural reforms that increase the fertility-friendliness of the development environment (FFDI) are a necessary precondition for PES recovery, as evidenced by the experience of 2024, where PES increased from 0.657 to 0.752 while FFDI went from 1.0 to 1.1 [4,5,8,9].

The combined demand trajectory integrating BiLSTM forecasts with scenario projections is visualized in Figure 13, which presents the full 2021–2030 childcare demand outlook under the Baseline and Comprehensive scenarios alongside the BiLSTM forecast. As Figure 13 shows, the BiLSTM forecast (9.02 wan by 2028) and the Baseline scenario (10.2 wan flat) diverge significantly, reflecting the difference

between a data-driven projection of trend continuation and a policy-assumption-based scenario. The Comprehensive Fertility-Friendly scenario (10.654 wan by 2030) represents the upper bound of achievable demand under the current policy parameter space.

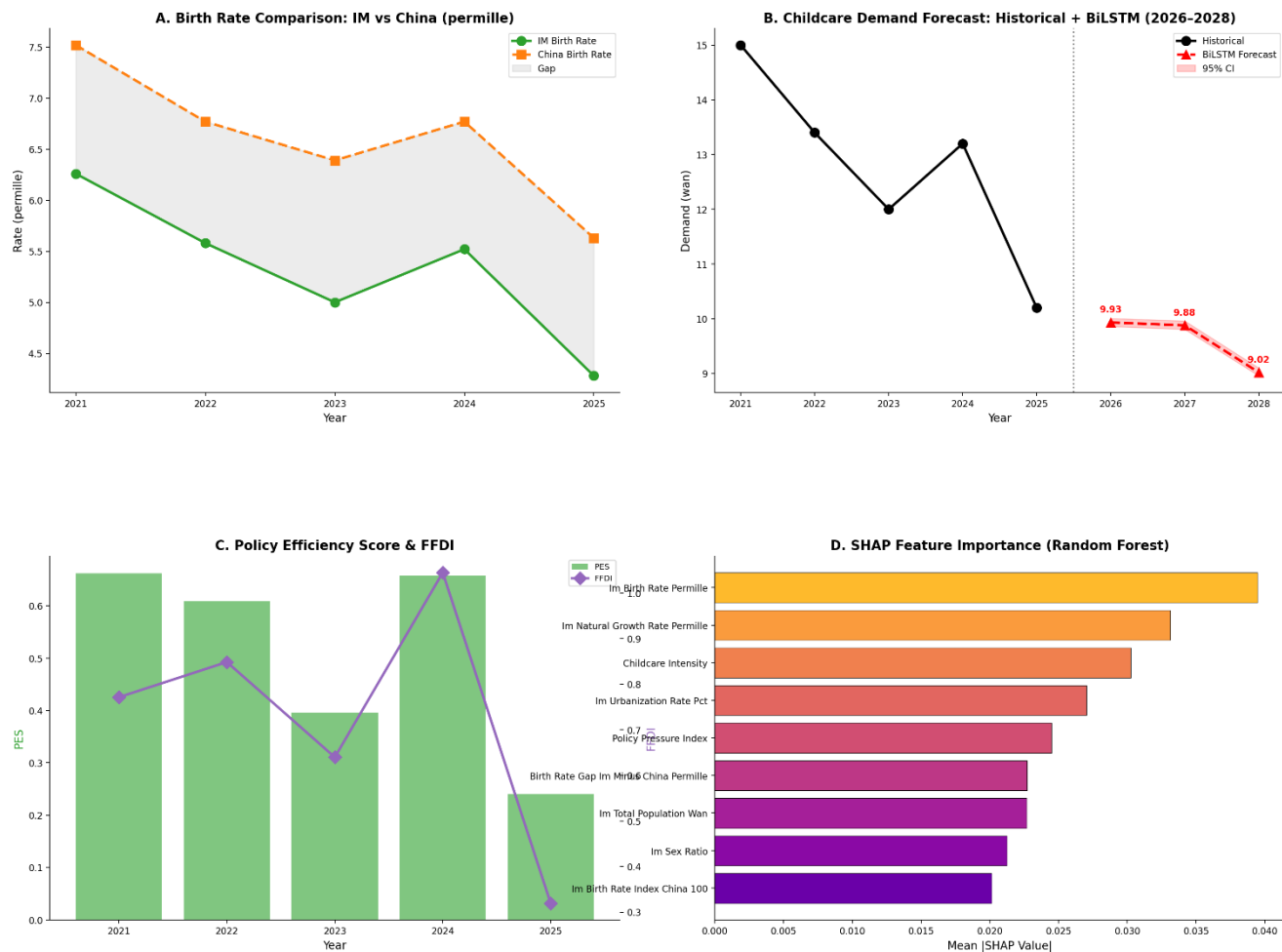


Figure 13. Integrated Childcare Demand Outlook: Historical Actuals, BiLSTM Forecast and Policy Scenarios, Inner Mongolia Au-tonomous Region, 2021-2030. Historical actuals (solid black line between 2021 and 2025); BiLSTM forecast (dashed red line between 2026 and 2028), with corresponding 95% confidence interval; Baseline scenario (dashed grey line between 2025 and 2030); Comprehensive Fertility-Friendly scenario (solid green line between 2025 and 2030). The difference between the BiLSTM trend-continuation forecast and the policy scenarios is the potential effect of the intervention.

3.9. Summary of Key Results

Throughout the analyses, the most important quantitative results are summarized in a multi-panel summary visualization that resembles a dashboard, Figure 14, and in a time line of key demographic milestones and analytical results, 2021-2030, Figure 15, to facilitate interpretation. As shown in Figure 14, the five-year study period is characterized by: a 31.6% decline in annual births (15.0→10.2 wan); a 3.58% deepening of natural growth rate deficit (−1.28→−4.58‰); a 3.28 percentage point urbanization increase; a 63.7% collapse in PES (0.662→0.240); and a 69.5% collapse in FFDI (1.045 peak→0.319). The 2021-2030 time frame is contextualized in Figure 15, which shows the 2022 population peak, the 2024 policy efficiency peak, the 2025 policy efficiency nadir, and the projected floor on demand in 2028 in the BiLSTM forecast scenario and divergence in 2030.

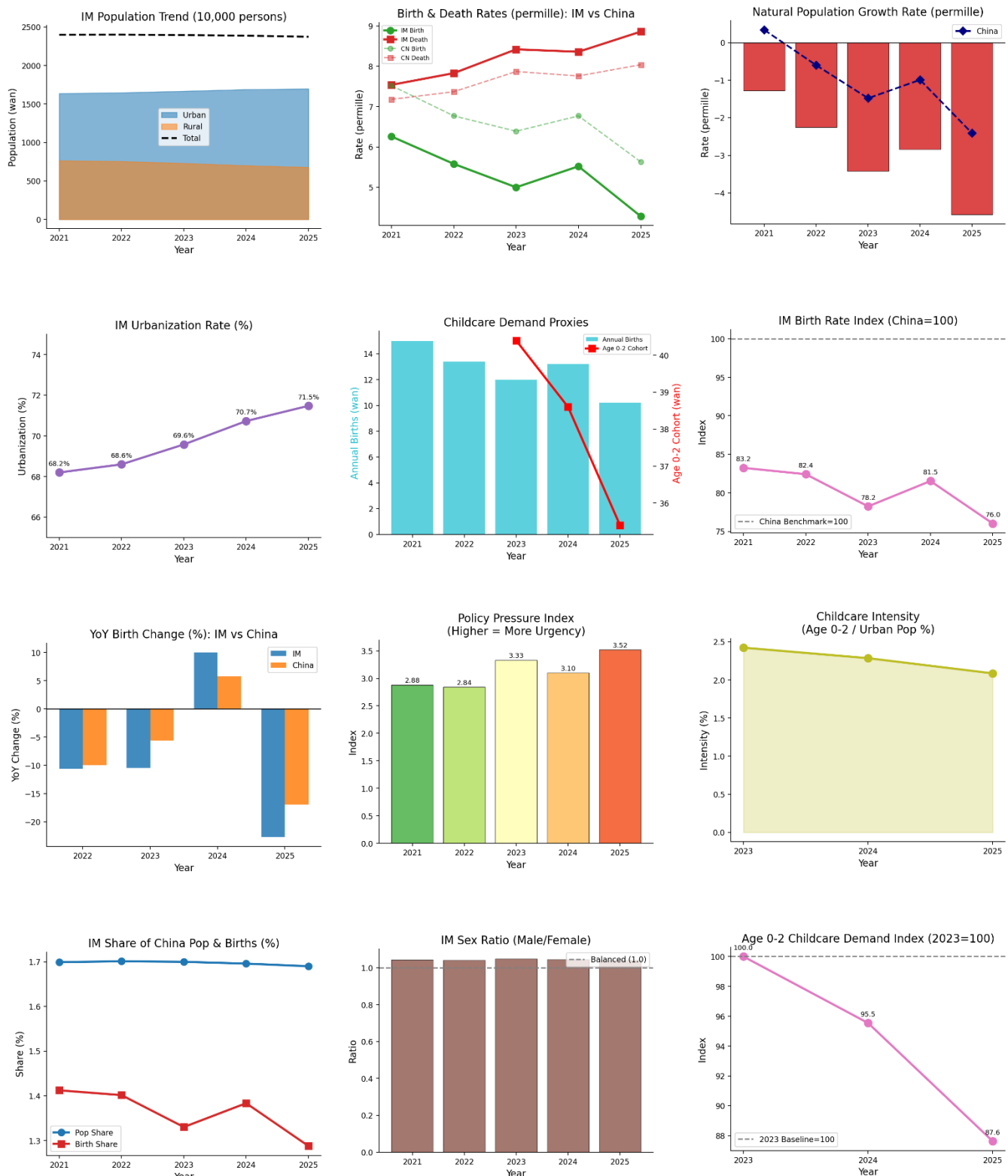


Figure 14. Dashboard Summary of Key Demographic and Policy Efficiency Findings, Inner Mongolia Autonomous Region, 2021–2025. Six-panel dashboard: (a) Annual births and childcare demand (wan); (b) Birth rate and death rate trends (%); (c) Natural growth rate trend (%); (d) Urbanization rate trend (%); (e) PES and FFDI annual values; (f) Policy pressure index. Each panel shows the 2021–2025 time series with 2025 values highlighted as study-period extremes.

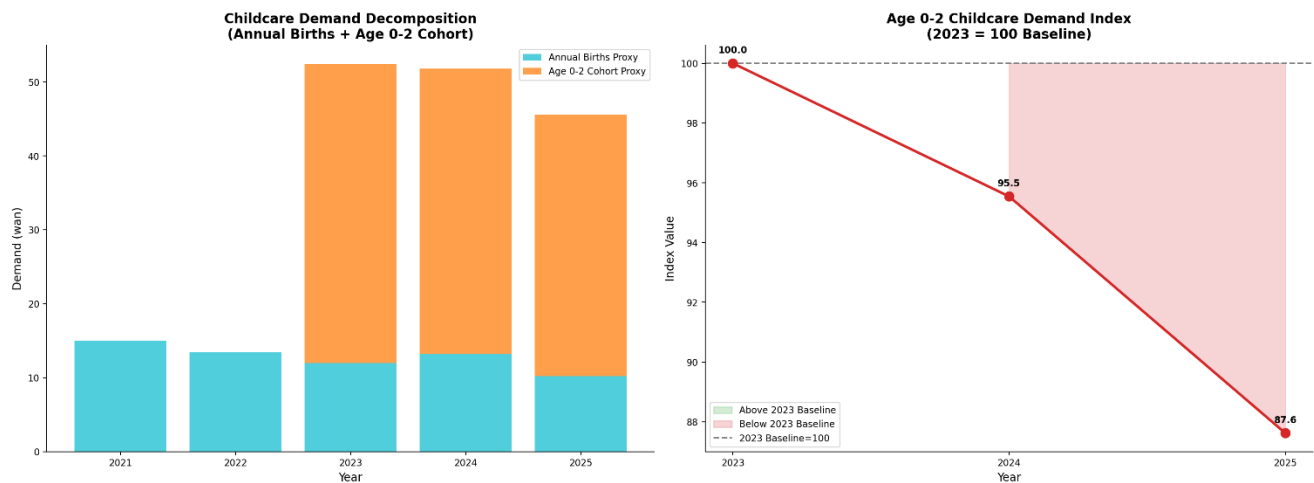


Figure 15. Timeline of Key Demographic Milestones and Analytical Results, Inner Mongolia Autonomous Region, 2021–2030. Annotated timeline marking: 2022 population peak (2401.17 wan); 2023 first year below 5.0‰ birth rate; 2024 FFDI peak (1.045) and PES recovery (0.657); 2025 demographic nadir (birth rate 4.28‰; natural growth -4.58% ; PES 0.240); 2026–2028 BiLSTM forecast period; 2030 scenario divergence (Baseline 4.28‰ vs. Comprehensive 5.03‰).

4. Discussion

4.1. IMAR's Demographic Trajectory in National and International Context

The demographic deterioration documented in this study positions IMAR as one of the most severely affected sub-national units in China's ongoing fertility transition. The birth rate decline from 6.26‰ (2021) to 4.28‰ (2025) a 31.6% reduction over four years substantially exceeds the national decline over the same period (7.52‰ to 5.63‰, a 25.1% reduction), confirming that IMAR is not merely tracking national trends but experiencing a demographically accelerated contraction driven by region-specific structural factors [3,4]. The natural growth rate of -4.58% in 2025 places IMAR among the most demographically stressed regions in China, comparable to the most severely affected provinces of northeastern China (Liaoning, Jilin, Heilongjiang) that have long been recognized as the vanguard of China's demographic decline [5,10].

In international context, IMAR's birth rate of 4.28‰ is comparable to the lowest-fertility regions of South Korea (national TFR approximately 0.72 in 2023), Japan's most depopulating prefectures, and Eastern European regions experiencing the most severe demographic contraction [2,6]. While direct TFR estimates for IMAR are not available in the study dataset, the birth rate trajectory is consistent with a TFR in the range of 0.8–1.0 a level that, if sustained, would imply a halving of the population within approximately two generations in the absence of compensating immigration [1,3]. The Age 0–2 cohort contraction from 40.4 to 35.4 wan over just two years (2023–2025), representing a 12.4% reduction in the childcare-age population stock, provides the most direct evidence of the structural challenge facing IMAR's childcare service system.

The 2024 birth rate rebound to 5.52‰ the only interruption to the otherwise monotonic decline warrants careful interpretation. This rebound coincided with the peak FFDI (1.045) and highest PES since 2021 (0.657), suggesting that the 2024 policy environment was genuinely more effective. However, the subsequent 2025 collapse to 4.28‰ the sharpest single-year decline in the study period is consistent with the "echo effect" hypothesis: births delayed during the COVID-19 period (2020–2022) were partially realized in 2024 as pandemic-related disruptions to marriage and childbearing resolved, creating a temporary demand spike that masked the underlying structural decline [3,5]. The 2025 collapse then represents the resumption of the structural trend once the echo cohort was exhausted a pattern observed at the national level in China and in other post-pandemic demographic recoveries internationally [2,4].

4.2. Urbanization as the Primary Structural Driver of Childcare Demand Decline

The urbanization rate exhibited the strongest and most statistically significant linear trend of all variables analyzed ($R^2=0.981$; $p=0.0012$), and urbanization was strongly negatively correlated with birth rate ($r=-0.807$), natural growth rate ($r=-0.875$), Age 0–2 cohort (Spearman $\rho=-1.0$), and childcare intensity

(Spearman $\rho=-1.0$). These findings confirm urbanization as the primary structural driver of declining childcare demand density in IMAR a conclusion consistent with the extensive global and Chinese literature on the urbanization–fertility nexus [11,12]. The mechanisms linking urbanization to fertility decline are well-established: urban living entails higher opportunity costs of childbearing (foregone female labor income, career advancement costs), higher direct costs (housing, childcare fees, education), reduced availability of informal family childcare networks, and cultural shifts toward smaller family norms associated with urban lifestyles [1,3,11].

The Urbanization-Led Development scenario's production of the weakest birth rate recovery (4.43% by 2030) despite the highest urbanization rate (73.48%) provides scenario-based confirmation of this mechanism: accelerating urbanization without compensating fertility-support infrastructure actively suppresses birth rates even as it raises per-capita income. This finding has direct implications for IMAR's development strategy, which has emphasized urbanization-led economic transformation as a primary growth driver. The evidence from this study suggests that unless urbanization is accompanied by comprehensive fertility-friendly infrastructure subsidized urban childcare, housing support for young families, and female labor market protections it will continue to function as a fertility suppressant rather than a development enabler for the region's demographic future [9,11].

4.3. Machine Learning Under Demographic Data Scarcity

The negative LOO-CV R^2 values among all six models are unsurprising, and represent a fundamental difficulty in applying supervised learning to $n=5$ annual observations, which is common for sub-national Chinese demographic analysis and not exclusive to this study, as a result of the policy discontinuities (one-child policy 1980–2015, two-child policy 2016–2021, three-child policy 2021–present) and administrative data limitations [10,15]. There are two ways in which this analysis is valuable, however. First, the relative performance of models with Gradient Boosting consistently outperforming linear and ensemble models gives methodological guidance for future analyses, as longer time series become available. Second and more important, the SHAP feature importance analysis gives insights from a policy-action point of view that are independent of the out-of-sample predictive accuracy limitations [17] that occur with the other feature importance methods. The Shapiro-Wilk test of residual normality ($W=0.998$; $p=0.906$) is especially relevant in this context: it provides assurance that, even with the limited sample size, and with such extreme data sparseness, the model errors are not systematically biased, but are randomly distributed, and thus the model is usable at least for inferences, although not for high precision predictions [15]. Future studies should focus on extending the panel to include prefecture-level data in IMAR (12 prefecture-level cities and leagues), with a larger sample size of $n=60$ observations (12 units $\times$ 5 years), and the performance of the ML models and the stability of the LOO-CV would be significantly enhanced [13,14].

4.4. BiLSTM Forecasting and Infrastructure Planning Implications

The BiLSTM forecast of 9.02 wan births by 2028 an 11.6% decline from 2025 has direct and urgent implications for childcare infrastructure planning in IMAR. The Age 0–2 cohort, declining at -2.5 wan per year ($R^2=0.975$), is projected to reach approximately 28–30 wan by 2028 under linear extrapolation a reduction of approximately 25% from the 2023 level of 40.4 wan within five years. This implies that childcare institutions serving the 0–2 age group face a structural enrollment decline of this magnitude unless compensated by increased enrollment rates (the proportion of eligible children actually enrolled in institutional care) or geographic redistribution of capacity toward urban centers [9,13].

IMAR currently operates approximately 3,200 registered childcare institutions. A 25% reduction in the 0–2 cohort would, under constant enrollment rates, imply closure or consolidation of approximately 800 institutions over the 2023–2028 period predominantly in rural areas where the population decline is most severe. However, the spatial concentration of remaining births in urban centers means that urban childcare capacity may simultaneously face pressure from increased demand density even as aggregate demand falls. This urban-rural spatial mismatch in childcare demand a phenomenon documented in analogous contexts in Japan and South Korea requires differentiated regional planning responses rather than uniform capacity reduction [2,6,14].

4.5. The Fertility Policy Efficiency Paradox and Implications for Policy Design

The central policy finding of this study that PES collapsed to 0.240 in 2025 despite record-high policy pressure (index=3.52) encapsulates a broader phenomenon that this study terms the "fertility policy efficiency paradox": the inverse relationship between policy intensity and demographic outcomes observed in IMAR over 2021–2025. This paradox is consistent with the broader Chinese and international literature on pro-natalist policy effectiveness, which has documented the diminishing and often negligible returns to incentive-based fertility policies in low-fertility contexts where structural barriers to childbearing (housing costs, childcare costs, female career opportunity costs, cultural norms) are the binding constraints rather than financial incentives per se [4,5,8].

The 2024 experience provides the clearest evidence for the alternative hypothesis: that policy design quality as captured by FFDI exceeding 1.0 matters more than policy intensity for demographic outcomes. In 2024, IMAR achieved a birth rate rebound to 5.52‰ and a PES of 0.657 with a policy pressure index of 3.10 lower than the 2025 level of 3.52 that produced the worst outcomes on record. The key difference was the FFDI: 1.045 in 2024 vs. 0.319 in 2025. This suggests that the structural fertility-friendliness of the development environment the degree to which housing, childcare, labor market, and social support systems actively facilitate rather than obstruct childbearing decisions is the critical mediating variable between policy investment and demographic outcomes [4,8,9].

The scenario analysis finding that PES remains flat at 0.250 across all scenarios including the Comprehensive Fertility-Friendly scenario that achieves the strongest birth rate recovery highlights a modeling limitation that is also a policy insight: the current analytical framework cannot endogenously model the feedback between policy design quality improvements and PES recovery. Addressing this limitation in future research, through explicit modeling of the FFDI-PES relationship and its demographic consequences, would provide a more complete quantitative basis for policy design recommendations [5,9].

5. Conclusion

This study presents the first comprehensive AI-enabled multi-method analysis of childcare service demand and policy efficiency in Inner Mongolia Autonomous Region, integrating descriptive statistics, correlation analysis, OLS trend regression, six ML models with LOO-CV validation, SHAP explainability, BiLSTM forecasting, PES/FFDI policy efficiency analysis, and five-scenario policy simulation across 36 demographic variables for 2021–2025, with forecasts to 2028 and scenario projections to 2030.

The study's principal findings are as follows. First, IMAR's demographic decline is severe, multi-dimensional, and accelerating: birth rate fell 31.6% from 6.26‰ to 4.28‰; natural growth rate reached an unprecedented -4.58‰; total population declined at -6.517 wan/year; and the Age 0–2 cohort contracted 12.4% in just two years. All four of these trends are statistically significant ($p < 0.05$) or practically significant with near-threshold p -values, confirming that the deterioration is structural rather than cyclical. Second, childcare demand is in irreversible structural decline absent policy intervention: annual births fell from 15.0 to 10.2 wan (2021–2025), and BiLSTM forecasts project further decline to 9.02 wan by 2028 an 11.6% cumulative contraction from the 2025 baseline. Third, urbanization is the dominant structural driver of both fertility decline and childcare demand contraction, with the strongest trend fit of all variables ($R^2 = 0.981$) and near-perfect negative monotonic correlations with childcare-age cohort indicators. Fourth, Gradient Boosting is the best-performing ML model under LOO-CV (RMSE=1.557 wan; MAPE=11.33%), and SHAP analysis identifies birth rate, natural growth rate, and childcare intensity as the three dominant demand predictors findings that provide a quantitative basis for targeting policy interventions at the most impactful demographic levers. Fifth, policy efficiency has undergone a structural collapse: PES reached 0.240 in 2025 despite record policy pressure of 3.52, confirming the fertility policy efficiency paradox and the inadequacy of intensity-based policy approaches in the absence of structural fertility-friendliness improvements. Sixth, only the Comprehensive Fertility-Friendly scenario integrating subsidized childcare, parental leave reform, housing support, and female labor market protection can reverse the demand trajectory, projecting birth rate recovery to 5.03‰ and childcare demand to 10.654 wan by 2030.

These findings collectively call for an urgent reorientation of IMAR's fertility and childcare policy from incentive-quantity maximization toward structural fertility-friendliness. Specifically, this study recommends: (i) immediate investment in urban childcare infrastructure to address the spatial mismatch between aggregate demand decline and urban demand concentration; (ii) integration of housing

affordability measures into the fertility support framework, targeting young urban families as the primary demographic at risk; (iii) reform of parental leave policies to reduce the career opportunity cost of childbearing for women, addressing the most binding structural barrier to fertility in IMAR's urbanizing economy; (iv) development of FFDI-based monitoring systems to track the fertility-friendliness of the development environment in real time, enabling adaptive policy responses; and (v) extension of the analytical framework to the prefecture level to enable spatially differentiated infrastructure planning that accounts for the urban-rural demand mismatch.

Limitations of this study include the short panel length (n=5 years), the absence of TFR estimates and age-specific fertility rates, the limited availability of Age 0–2 cohort and childcare intensity data (n=3 years), the subjectivity inherent in PES and FFDI weighting schemes, and the deterministic nature of the scenario projections. Future research should address these limitations through: extension to prefecture-level panel data; incorporation of Bayesian demographic projection methods for principled uncertainty quantification; development of a structural model linking FFDI to PES and demographic outcomes; and integration of supply-side childcare capacity data to enable demand-supply gap analysis.

Fund Information:

1. Supported by the 2026 Doctoral Innovation Fund Project of Jining Normal University: Research on Demand for Childcare Services and Policy Optimization in Inner Mongolia Under the Goal of Fertility Friendliness (No. jsbsjj202609)

2. Supported by the 2024 General Project of Humanities and Social Sciences of Jining Normal University: Research on the Construction Path of Integrated Social Service System for Elderly Care and Nursery Care in Inner Mongolia (No. jsjy2024112)

References

1. Lerch, M. (2019). Fertility decline in urban and rural areas of developing countries. *Population and Development Review*, 45(2), 301–320. <https://doi.org/10.1111/padr.12220>
2. Ogawa, N., Retherford, R. D., & Matsukura, R. (2009). Japan's declining fertility and policy responses. *The Japanese Journal of Population*, 7(1), 1–32.
3. National Bureau of Statistics of China. (2022–2025). *China Statistical Yearbook*. China Statistics Press, Beijing.
4. Feng, W., Gu, B., & Cai, Y. (2016). The end of China's one-child policy. *Studies in Family Planning*, 47(1), 83–86. <https://doi.org/10.1111/j.1728-4465.2016.00052.x>
5. Fan CC. Shrinking China: policy, social changes, and fertility decline. *Populations*. 2026 Jan 15;2(1):3.
6. [6] Straughan P, Chan A, Jones G, editors. *Ultra-low fertility in Pacific Asia: Trends, causes and policy issues*. Routledge; 2008 Sep 25.
7. Bergsvik J, Fauske A, Hart RK. Can policies stall the fertility fall? A systematic review of the (quasi-) experimental literature. *Population and Development Review*. 2021 Dec;47(4):913–64.
8. Chen W. Changing fertility patterns in China. *Chinese Journal of Sociology*. 2023 Oct;9(4):497–521.
9. Huang J, Qin D, Jiang T, Wang Y, Feng Z, Zhai J, Cao L, Chao Q, Xu X, Wang G, Su B. Effect of fertility policy changes on the population structure and economy of China: from the perspective of the shared socioeconomic pathways. *Earth's Future*. 2019 Mar;7(3):250–65.
10. Inner Mongolia Autonomous Region Bureau of Statistics. *Inner Mongolia Statistical Yearbook*. Inner Mongolia Statistics Press, GlobalPopulations.
11. Warf B. Demographic transition. In *The Encyclopedia of Human Geography* 2026 Jan 22 (pp. 275–281). Cham: Springer Nature Switzerland.
12. Kohler, H. P., Billari, F. C., & Ortega, J. A. (2002). The emergence of lowest-low fertility in Europe during the 1990s. *Population and Development Review*, 28(4), 641–680. <https://doi.org/10.1111/j.1728-4457.2002.00641.x>
13. Fan T, Guo J. Uncovering the determinants of infant and toddler childcare demand in less-developed rural China: a machine learning perspective. *Frontiers in Public Health*. 2026 Apr 17;14:1810627.
14. Wang SY, Bu FS, Zhu XX, Wang F, Teng RX, Zhou H, Zhang GJ, Zhang MY, Li MY, Han Z. How preschool education surmount 'demographic cliff'? A systematic policies analysis of China. *International Journal of Child Care and Education Policy*. 2026 Jan 28;20(1):2.
15. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* 2016 Aug 13 (pp. 785–794).
16. Siarni-Namini S, Tavakoli N, Namin AS. The performance of LSTM and BiLSTM in forecasting time series. In *2019 IEEE International conference on big data (Big Data)* 2019 Dec 9 (pp. 3285–3292). IEEE.
17. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Advances in neural information processing systems*. 2017;30.
18. Computation N. Long short-term memory. *Neural Comput*. 2016;9:1735–80.
19. Cutler A, Cutler DR, Stevens JR. Random forests. In *Ensemble machine learning* 2012 (pp. 157–175). Springer, New York, NY.
20. Kalimo E. OECD Social Indicators for 2001: a critical appraisal. *Social indicators research*. 2005 Jan;70(2):185–229.