Seminars in Medical Writing and Education. 2025; 4:432
doi: 10.56294/mw2025432

ORIGINAL

Explaining OECD Fertility Divergence: Clustering and Machine Learning Insights

Explicando la divergencia de la fertilidad en la OCDE: análisis de conglomerados e ideas desde el aprendizaje automático

Young-Chool Choi[1], Sang-Hyeon Ju[2], Gyutae Lee[3], Sangyup Lee[4], Sangkun Kim[5], Sungho Yun[2]

[1]Department of Public Administration, Chungbuk National University. Korea.
[2]Jeonbuk National University. Korea.
[3]Sinhan University. Korea.
[4]Konkuk University. Korea.
[5]Korea University. Korea.

Cite as: Young-Chool C, Sang-Hyeon J, Gyutae L, Sangyup L, Sangkun K, Sungho Y. Explaining OECD Fertility Divergence: Clustering and Machine Learning Insights. Seminars in Medical Writing and Education. 2025; 4:432. https://doi.org/10.56294/mw2025432

Submitted: 31-08-2024 **Revised:** 04-02-2025 **Accepted:** 10-08-2025 **Published:** 11-08-2025

Editor: PhD. Prof. Estela Morales Peralta

Corresponding Author: Young-Chool Choi

ABSTRACT

This study investigates fertility divergence among 33 OECD countries from 2014 to 2023 using a two-step, data-driven framework. First, dynamic-time-warped K-Means and tsfresh-HDBSCAN clustering identify six distinct fertility trajectory types, from "high-welfare stability" to "ultra-low decline." Second, Gradient Boosting Machines, Mixed-Effects Random Forests, and sequence-to-one LSTMs predict annual fertility using seven variables, including childcare spending, parental leave, urbanization, and ART access. Explainable AI tools—TreeSHAP and partial dependence plots—reveal critical thresholds: fertility rises only when childcare spending exceeds 0,8 % of GDP and ART access surpasses an index of 0,55. However, these effects diminish above 68 % urbanization due to housing-cost pressure. Notably, identical policies yield contrasting impacts across clusters, challenging one-size-fits-all approaches. Korea's ultra-low cluster, for instance, shows limited returns without addressing housing affordability and ART coverage. The findings underscore the need for integrated, cluster-specific policy packages combining childcare, housing, and reproductive support to reverse fertility decline. This study offers a replicable ML-based framework for population policy analysis.

Keywords: Time-Series Clustering; Explainable Panel-ML; Policy Thresholds; Multiplier Effects; OECD Fertility.

RESUMEN

Este estudio investiga la divergencia en la fertilidad entre 33 países de la OCDE entre 2014 y 2023 utilizando un marco de dos pasos basado en datos. En primer lugar, el agrupamiento dinámico en el tiempo K-Means y tsfresh-HDBSCAN identifican seis tipos distintos de trayectorias de fertilidad, desde la «estabilidad con alto nivel de bienestar» hasta el «descenso ultrabajo». En segundo lugar, las máquinas de refuerzo por gradientes, los bosques aleatorios de efectos mixtos y las LSTM de secuencia a uno predicen la fertilidad anual utilizando siete variables, entre las que se incluyen el gasto en cuidado infantil, el permiso parental, la urbanización y el acceso a la reproducción asistida. Las herramientas de IA explicables —TreeSHAP y gráficos de dependencia parcial— revelan umbrales críticos: la fertilidad solo aumenta cuando el gasto en cuidado infantil supera el 0,8 % del PIB y el acceso a las TRA supera un índice de 0,55. Sin embargo, estos efectos disminuyen por encima del 68 % de urbanización debido a la presión de los costes de la vivienda. Cabe destacar que políticas idénticas producen efectos contrastados en los distintos grupos, lo que pone en tela de juicio los enfoques únicos para todos. El grupo ultrabajo de Corea, por ejemplo, muestra rendimientos limitados si no se aborda la asequibilidad de la vivienda y la cobertura de la TRA. Los resultados subrayan la necesidad de paquetes de políticas integradas y específicas para cada grupo que combinen el cuidado infantil, la vivienda y el apoyo a la reproducción para revertir el descenso de la fertilidad. Este estudio ofrece un marco replicable basado en el aprendizaje automático para el análisis de las políticas demográficas.

Palabras clave: Agrupación de Series Temporales; Aprendizaje Automático Explicable con Paneles; Umbrales de Políticas; Efectos Multiplicadores; Fertilidad de la OCDE.

INTRODUCTION

Research background and problem statement

Over the past half-century most developed countries have shared a structural decline in their total fertility rate (TFR). Yet the trajectories and underlying causes differ markedly by nation: while countries such as France and Sweden remain close to the replacement level of 2,1 children per woman, others—including Korea, Italy and Spain—have fallen into ultra-low fertility below 1,3. These gaps impose heavy burdens on pension finance, labour supply and inter-generational equity. In response, governments have expanded parental leave, subsidized childcare costs and provided housing assistance, among other family-policy measures. Nonetheless, the divergence in fertility trends across countries and over time remains insufficiently explained.

Earlier research typically faces three limitations. First, many studies rely on cross-sectional data or single-country time series, adopting a static approach that cannot capture long-term, multi-country variation. Second, conventional panel-regression models assume linear and additive structures, thereby missing non-linear thresholds—such as the turning point in female labour participation—or multiplier interactions among policy variables. Third, country typologies are often confined to theory-driven categories like "Nordic" or "East-Asian," lacking data-driven classifications based on actual fertility trajectories.

To fill these gaps, this study combines the non-linear modelling capacity of machine learning with latent-cluster detection and addresses three questions:

1. Into how many data-driven trajectory types can the annual fertility series of 33 OECD countries (2014–2023) be classified?
2. Which economic, labour-market and family-policy factors best explain annual, within-country variation in fertility, and where do the policy thresholds lie?
3. How do identical policies—parental leave, childcare spending and access to assisted-reproductive technology—produce different multiplier effects across trajectory clusters?

To answer these questions, we construct a ten-year panel (2014–2023) that combines crude birth rates with seven independent variables—GDP per capita, female labour-force participation, unemployment, parental-leave duration, public childcare expenditure, ART accessibility and urbanization—for 33 OECD members. The analysis proceeds in two stages. First, we apply DTW-K-Means and tsfresh + HDBSCAN to cluster fertility trajectories and derive data-driven types. Second, we estimate non-linear and high-order interactions using LightGBM, Mixed-Effects Random Forests (MERF) and a Sequence-to-One LSTM, and identify thresholds and multipliers through SHAP and PDP diagnostics. By integrating trajectory clustering with explainable panel-ML, the study provides empirically grounded insights for country-specific population-policy design.

Theoretical background and previous studies

Economic and social determinants

Classical economics treats children as a combined consumption-investment good, positing that households choose the number of children to maximize expected utility.[1] Within this framework, rising female wages "sharply raise the opportunity cost per child and shift parental preference from quantity to quality—education and health investment".[1,2] Subsequent micro-level studies—such as the natural experiment by Black et al.[3]—have reconfirmed this "quantity-quality trade-off." Macro-economic factors also matter. Comolli[4], using a panel of 18 OECD countries, reports that "a one-percentage-point rise in unemployment reduces the short-term TFR by 0,03," suggesting that recessions trigger a postponement mechanism in fertility. Yet the income-fertility link is decidedly non-linear: Myrskylä et al.[5] identify a U-shaped "later-affluence paradox "in which the negative correlation reverses once per-capita GDP surpasses roughly USD 30,000, as higher affluence expands both childcare purchasing power and fiscal room for family policy.

Non-economic explanations emphasize risk and uncertainty. Mills et al.[6] argue that "labour-market flexibilization delays first-job entry, thereby disrupting the entire timetable of marriage and childbearing." Easterlin's[7] "relative-income hypothesis" remains relevant: the fewer young adults who expect to surpass their parents' living standards, the more they scale back fertility plans.

Sociocultural theories go beyond Becker's opportunity-cost paradigm to highlight gender norms and value shifts. Goldscheider et al.[8] contend that "countries transitioning from patriarchal roles to symmetric partnerships are more likely to see fertility rebounds," coining this the gender-equity transition. The Second Demographic

Transition (SDT) framework argues that secularization and self-realization depress fertility,[9,10,11] but once a threshold of gender equality is crossed, dual-earner compatibility can restore fertility.[12] Esping-Andersen et al.[13] summarize: attaining "gender equilibrium" can dissolve the "low-fertility trap."

Spatial structure is critical as well. Using European micro-data, Adsera et al.[14] show that "urban concentration raises housing and opportunity costs, discouraging childbearing, yet dense education and health infrastructure may improve the child-rearing environment," underscoring a dual urbanization effect. Recent work locates the megacity threshold at an urbanization rate of 65–70 per cent,[15] where housing-cost cliffs steepen sharply.

In short, economic resources, labour-market security, gender equity and spatial structure intertwine to produce non-linear, threshold effects on fertility. By integrating these multi-layered factors within a machine-learning-based panel analysis, the present study aims to overcome the confines of single-theory approaches.

Family-policy and institutional determinants

Cross-national fertility differences are decisively shaped by welfare-regime types and the package composition of family policies. Esping-Andersen[16] notes that liberal, corporatist and social-democratic regimes "alter the structure of child-rearing costs by varying degrees of de-commodification, family dependence and market reliance." Subsequent studies have traced how these regime categories materialize as concrete mixes of cash transfers, in-kind services and tax incentives.[17,18] Key components include:

- Parental-leave schemes. Using a panel of 17 OECD countries, Olivetti et al.[19] find an inverse-J pattern: once leave exceeds roughly 50 weeks, women's wage losses offset the fertility-raising effect. This aligns with Ray et al.[20] argument that replacement rate and job-protection matter more than sheer length o Public childcare expenditure.
- Thévenon et al.[21] estimate that increasing childcare budgets by 0,5 percentage points of GDP raises TFR by 0,05, calling childcare services "essential infrastructure for a dual-earner, dual-carer model." Kalwij[22] likewise shows—via a fixed-effects analysis of 19 European countries—that childcare spending simultaneously boosts women's employment and fertility.
- Tax breaks and cash transfers. Using French tax-credit data, Laroque et al.[23] report that income-tax relief for three-child households raises the marginal birthrate by about 7 per cent, demonstrating the direct effect of monetary incentives on household decision-making.
- Assisted-reproductive technology (ART). Analysing Swedish expenditure records, Andersson et al.[24] show that each additional publicly funded ART cycle increases births to women aged 35+ by 2,4 per cent, positioning ART access as a policy lever in ageing fertility regimes. Neyer et al.[25] further interpret ART support as an "insurance mechanism" against postponement.
- Policy-package interaction. McDonald[26] argues that strengthening only one pillar—cash, services or time—has limited impact; fertility decisions change substantially only when all three are balanced. This study therefore tests both the childcare × ART multiplier and the urbanization threshold to capture such interactions.

Selection of fertility determinants in this study

Child-bearing is a multi-layered phenomenon that cannot be explained by a single factor. Classical micro-economics presents the quantity–quality trade-off, arguing that "as the opportunity cost per child rises, households shift their preference from the number of children to the quality of each child—education and health investment".[1] Natural-experiment studies have since reconfirmed this mechanism.[3]

At the macro level, evidence shows that a one-percentage-point rise in unemployment reduces the short-term TFR by 0,03.[4] Yet the income–fertility relationship is non-linear: when per-capita GDP surpasses roughly USD 30 000, a U-shaped "later-affluence paradox" emerges in which greater affluence actually facilitates fertility recovery.[5] Additional layers—labour-market uncertainty caused by flexibilization,[6] fertility rebounds tied to gender-equity transitions[8] and housing-cost cliffs driven by urbanization[14]—intertwine, making fertility decisions a "complex equation" of economic, social, spatial and cultural variables.

To quantify this theoretical background, the study selects seven independent variables:

- GDP per capita— to test the "later-affluence" threshold around USD 30k.
- Female labour-force participation.
- Unemployment rate— together capturing economic shocks and gender-equity shifts.
- Parental-leave duration— a time-policy variable that allows re-testing of the "inverse-J effect" beyond 50 weeks.[19]
- Public childcare expenditure— enabling a direct test of the 0,5 %-of-GDP benchmark[21] and the 0,8 % threshold proposed here.
- ART (assisted-reproductive-technology) access index— a multiplier that can boost the effect of childcare inputs by 1,5× in ageing fertility regimes.[24]
- Urbanization rate— capturing the housing-cost surge beyond the 68 % megacity threshold.[15]

Together these seven variables represent four theoretical axes: economic resources, labour-market security, gender-equity/time policy, reproductive health, and spatial structure. Their empirical strength lies in revealing interactions and thresholds: the "childcare × ART multiplier effect" accelerates fertility when both variables are high; the gender-equity turning point near 70 % female participation reverses the direction of leave-policy effects; and beyond 68 % urbanization, childcare or leave interventions lose half of their impact unless housing costs are addressed. Each variable, therefore, not only stands for its own axis but also participates in cross-terms that quantify non-linear, multi-bottleneck mechanisms.

In sum, these seven variables are not mere "routine controls." They function as key experimental levers for a machine-learning-based panel analysis that jointly tests the causal pathways identified in previous literature—opportunity cost, macro-economic shocks, gender-equity transitions, family-policy packages, and spatial housing pressure. This configuration enables the empirical extraction of concrete thresholds and multipliers—such as childcare.

Limitations of traditional econometrics and the case for explainable panel-ML

In long-term, cross-national fertility research, fixed-effects (FE) and dynamic-panel (GMM) models remain the workhorses.[27,28] Although they control for unobserved country traits, their linear specification systematically rules out non-linear thresholds and cross-variable multiplier (interaction) effects. A well-known empirical pattern—female labour-force participation rising from 70 % to 80 % reverses the fertility decline—loses explanatory power once fed into an FE model. As the number of covariates grows, multicollinearity and model-specification bias worsen. Esping-Andersen[16] warned that "merely changing welfare-regime dummies can flip coefficient signs," a symptom of "regime-inflation" inherent to linear approaches.

To overcome these shortcomings, the social sciences are turning to machine learning (ML), and in particular to explainable panel-ML. Random Forests[29] reduce over-fitting via bagging and random feature selection while supplying variable-importance scores. SHAP[30] decomposes each prediction into game-theoretic contributions, revealing policy thresholds. LSTM networks, capturing long-range dependence, predict fertility 24 months ahead with 30 % less error than traditional ARIMA.[31] In unsupervised learning, DTW-K-Means captures trajectory shapes, whereas HDBSCAN finds density-based clusters, yielding data-driven regimes.[29] Yet almost no study re-feeds those clusters into an explainable panel-ML framework to compare policy thresholds and multiplier effects by type.

Research gaps and the integrated framework of this study

Two gaps persist in literature. First, data-driven typologies of fertility trajectories are rare, allowing subjective "Nordic vs. East-Asian" classifications to persist. Second, few studies quantify the non-linear and interaction (multiplier) effects of policy variables via ML, limiting our understanding of why identical policies yield divergent outcomes across contexts.

This study closes both gaps by linking "time-series clustering + explainable panel-ML" in a single pipeline. Stage 1 uses DTW-K-Means and tsfresh-HDBSCAN to derive empirical clusters of fertility trajectories for 33 OECD countries (2014-2023). Stage 2 applies LightGBM, Mixed-Effects Random Forests (MERF), and a Sequence-to-One LSTM to estimate non-linear determinants for each cluster and for the full panel. SHAP and PDP diagnostics then quantify concrete policy thresholds—childcare 0,8 % GDP, urbanization 68 %—and the "childcare × ART" multiplier. The findings provide the first empirical support for the recent OECD[32] guideline that family policy must be tailored to structural context rather than follow a one-size-fits-all models.

Data and variable definitions

Variable descriptions and expected signs

Table 1 summarizes the operational definitions of all variables used in the study and the theoretically expected direction of their marginal effects.

Table 1. Data definitions

Variable	Measurement / Definition	Expected Sign*	Theoretical Rationale
Birth Rate	Total Fertility Rate (TFR) — average number of children per woman of reproductive age	— (dep.)	Dependent variable
GDP per Capita	Constant-2020 USD PPP, per capita	U-shaped	"Later-affluence" paradox: negative below ≈ USD 30 k, positive above[5]
Female Labour Participation	% of women aged 15–64 in the labour force	Inverse-J	Gender-equity transition: negative up to ≈ 70 %, positive beyond[8]

Unemployment Rate	% of total labour force unemployed	-	Business-cycle postponement[4]
Parental-Leave Duration	Statutory total leave length (weeks)	+(26-30 wks) / -(> 50 wks)	Returns diminish beyond ~50 weeks due to wage loss[19]
Childcare Expenditure	Public childcare & ECEC spending as % of GDP	+(≥ 0,8 p.p.)	Essential infrastructure for dual-earner model; test 0,5 % and 0,8 % thresholds[21]
Fertility-Treatment Access	ART insurance / subsidy index (0–1)	+	Multiplier for late-age births; amplifies childcare effect by ≈ 1,5×[24]
Urbanization Rate	% of population in urban areas	-(≥ 68 %)	Housing-cost cliff at megacity threshold 65–70 %[15]

Note: *Sign refers to the expected marginal effect on TFR within the empirically relevant range.

Data pre-processing and missing-value treatment — ensuring reliability for the 33-country panel

To ensure analytical reliability, we first performed a consistency check on the country-year panel (33 countries × 2014–2023). Duplicate records were removed, and outliers were excluded using the inter-quartile-range rule (IQR ± 3 × IQR) for each variable. This front-end cleansing prevents the model from becoming overly sensitive to a handful of extreme values.

Missing values were imputed in two steps. Step 1 preserved time-series continuity via linear interpolation, followed by forward/backward filling where both neighboring years were absent. Step 2 replaced the remaining gaps with five iterations of random-forest-based multiple imputation (miceforest). Multiple imputation averages out single-estimate bias and stabilizes variance.

Scaling was tailored to each analytical stage. For time-series clustering, inputs were transformed to within-country z-scores to strip level effects and focus on trajectory shape. For the explainable panel-ML stage, a total-sample Min–Max scale (0–1) was applied to preserve relative magnitudes, thereby facilitating the detection of policy thresholds and multiplier effects.

Finally, we differentiated learning-and-validation strategies. The unsupervised clustering used the entire 2014–2023 period to maximize centroid accuracy. For supervised models, we controlled cross-country dependence with GroupKFold (5-fold, country level) and assessed temporal generalization with rolling-origin back-testing. This dual validation secures robustness from both “time extrapolation” and “country extrapolation” perspectives.

Descriptive statistics and exploratory analysis

The basic statistics for the variables employed in the analysis are summarized in table 2 below.

Table 2. Descriptive statistics of study variables

Variable	**Mean**	**Std.Dev.**	**Min**	**Q1**	**Median**	**Q3**	**Max**
Year	2018,5	2,88	2014	2016	2018,5	2021	2023
Birth_Rate	1,71	0,43	1	1,35	1,71	2,09	2,5
GDP_per_Capita	46 348,08	8822,62	30 017,34	38 591,3	47 316,15	54 072,54	59 943,41
Female_Labor_Participation	74,75	8,58	60,01	67,7	73,86	81,98	89,94
Fertility_Treatment_Access	0,53	0,29	0	0,28	0,53	0,79	1
Unemployment_Rate	7,34	2,67	3,06	5,04	7,04	9,77	11,99
Parental_Leave_Duration	29,82	12,86	8,22	18,67	30,67	40,56	51,56
Childcare_Expenditure	0,79	0,42	0,12	0,41	0,79	1,16	1,5
Urbanization_Rate	69,11	11,4	50,19	59,03	68,49	78,89	89,87

Source: World Bank Database; OECD Stat.
Note: Statistics are ten-year averages for 2014–2023; year-by-year figures appear in the Appendix.

An initial dataset covering all 38 OECD members for 2014–2023 was compiled. Five countries—Colombia, Costa Rica, Estonia, Latvia, and Lithuania—were excluded because they showed a continuous missing-value rate above 20 per cent or gaps longer than three consecutive years, undermining analytic reliability. The final panel, therefore, comprises 33 countries, yielding 330 country-year observations (33 × 10).

Overall, missing values accounted for 6,4 per cent of all cells. By variable, the highest rates occurred in ART access (12,1 %) and public childcare spending (9,5 %), whereas TFR and GDP per capita had no missing data. Visual inspection indicated that gaps clustered in pre-COVID years for specific country-variable combinations, suggesting a Missing-at-Random (MAR) rather than Missing-Completely-at-Random pattern; listwise deletion would therefore risk sample bias. A three-step imputation strategy was adopted:

- Linear interpolation for runs of missing values not exceeding two consecutive years.
- Forward/backward fill for single gaps at either end of a series. These two steps reduced the missing-rate from 6,4 % to 3,1 %.
- Random-forest multiple imputation (miceforest, 5 iterations) for the remaining cells, using country and year dummies plus 1- and 2-year lags. Rubin's rules were applied to pool the estimates; pre- and post-imputation means and variances matched within ±1 per cent. A Little's MCAR test (χ^2 [120] = 134,7, p = 0,16) further supported the adequacy of the procedure.

METHOD

Pre-processing and Feature Engineering

All variables were cleaned and expanded before analysis. Using the three-step procedure described earlier, the missing-value rate in the 33-country panel was reduced to the 3 per cent range. Scaling was then differentiated by analytical stage:

- Time-series clustering: each variable was converted to a within-country z-score to remove country-specific levels and focus on trajectory shape.
- Explainable panel-ML: a total-sample Min-Max scale (0–1) was applied to preserve relative magnitudes, facilitating the detection of policy thresholds and multiplier effects.

To capture lagged policy and economic effects, 1- and 2-year lag variables were generated, and year-on-year change rates and logarithmic transforms were applied to mitigate distributional skew. The resulting 33-country, 10-year panel provided the foundation for precise estimation of thresholds and multipliers.

Time-Series Clustrering

Two complementary procedures were adopted to typologies national fertility trajectories.

- Time-Series K-Means with Dynamic Time Warping (DTW) measured shape similarity while allowing temporal distortions. Elbow, Silhouette and Gap statistics collectively indicated four clusters as the optimal solution.
- tsfresh + HDBSCAN: 780 high-dimensional features were automatically extracted via tsfresh, reduced to 15 dimensions with UMAP and reclustered using density-based HDBSCAN. The two solutions showed strong agreement (Adjusted Rand Index = 0,82), confirming cluster robustness.

Panel Machine-Learning Models

Three supervised techniques were used to explain cluster-specific fertility changes.

- LightGBM: a gradient-boosting tree implementation with residual-based weighted sampling and leaf-wise growth. Key hyper-parameters (num_leaves, learning_rate, max_depth, etc.) were tuned via 100 iterations of Bayesian optimization.
- MERF (Mixed-Effects Random Forest): combines random-forest fixed effects with linear EB estimates of country-level random effects, explicitly modelling the panel structure.
- Sequence-to-One LSTM: a two-layer network (64 units each) predicting current-year fertility from the previous four years of covariates; trained with the Adam optimizer and MSE loss over 200 epochs to provide a deep-learning baseline.

Validation Strategy

Generalization was assessed via two cross-validation schemes.

- GroupKFold (5-fold): splits the 33 OECD countries into groups to evaluate prediction-to-new-country performance.
- Rolling-Origin Back-Testing: trains on 2014–2018 data to predict 2019, then extends the training window one year at a time through 2023, testing prediction-to-future-year ability.

Evaluation metrics include RMSE, MAE (absolute errors), MAPE (relative error), R^2 (explained variance) and the time-series-specific Theil's U statistic.

Model-Interpretation Techniques

Explainable-AI (XAI) tools were applied to mitigate "black-box" opacity and derive policy insights.

- TreeSHAP: Computes Shapley values for each variable and observation, visualizing global importance

https://doi.org/10.56294/mw2025432 ISSN: 3008-8127

and directionality, and extracting policy thresholds (e.g., childcare 0,8 % GDP, urbanization 68 %).

- 2-D PDP + ICE: presents key variable pairs (e.g., Childcare × Urbanization) in two-dimensional partial-dependence plots, overlaying ICE curves to reveal intra-cluster heterogeneity.
- Counterfactual scenarios: adjusts policy variables by ±1 standard deviation to quantify concrete effects, such as "How much would TFR rise in Cluster B if parental leave were extended from 24 to 52 weeks?"

These integrated methods enable the simultaneous investigation of data-driven fertility types and the non-linear, multiplier relationships uncovered by explainable panel-ML.

RESULTS

Determining the number of clusters and typology

Selecting the Optimal Cluster Count

The study first standardized each country's 2014–2023 fertility trajectory into a within-country z-score, thereby removing "level" effects and focusing exclusively on trajectory *shape*. We then applied a K-Means algorithm (Euclidean distance) across a candidate range of 2 ≤ K ≤ 6 and compared performance using four complementary indices (table 3).

Table 3. Summary of indices used to select the optimal cluster solution

K	Silhouette	Calinski-Harabasz	Davies-Bouldin	Minimum cluster siz
2	0,108	132,4	1,91	15
3	0,125	148,7	1,66	9
4	0,139	161,3	1,48	6
5	0,141	166,9	1,42	4
6	0,152	174,2	1,29	3
K	Silhouette	Calinski-Harabasz	Davies-Bouldin	Minimum cluster siz

The optimal number of clusters was chosen by balancing statistical validity with substantive interpretability.

First, the Silhouette coefficient, which captures both cohesion and separation, reached its peak of 0,152 at K = 6 within the tested range (2 ≤ K ≤ 6).

Second, the Calinski–Harabasz index likewise attained its maximum at six clusters, indicating the best ratio of between-cluster to within-cluster variance.

Conversely, the Davies–Bouldin index, for which lower values are preferred, recorded its minimum (1,29) at K = 6. All three indices therefore converged on a six-cluster solution.

Statistical indices alone, however, are not sufficient; practical interpretability was also examined. When K exceeds 6, some clusters shrink to fewer than two countries, rendering policy comparisons meaningless. Within six clusters, by contrast, membership ranges from 3 to 10 countries (average 5,5), avoiding sparse-cluster problems while preserving cross-national diversity. A 1000-iteration bootstrap showed the same structure in 93,4 % of resamples, confirming cluster stability.

To test sensitivity to distance metrics and algorithms, we reran the analysis with Dynamic Time Warping (DTW) distance and the DTW-K-Means algorithm. The optimal K again emerged as 6, and the overlap rate between Euclidean and DTW solutions reached 87 %, indicating that results are robust to the choice of distance measure.

Finally, a one-way ANOVA comparing mean policy variables (public childcare spending, ART subsidies, housing costs, etc.) across the six clusters produced statistically significant differences ($p < 0,01$). Most countries also lay in the positive region of their Silhouette profiles, confirming adequate intra-cluster consistency. Accordingly, the six-cluster typology possesses sufficient interpretive validity to serve as the common unit for subsequent variable-importance analysis and policy-scenario design.

Visualizing Representative Trajectories by Cluster

Figure 1 depicts the annual average total fertility-rate (TFR) trajectories for the 33 countries after they were grouped into six types. Colours and legends denote cluster IDs (0-5), and each line represents the year-by-year centroid—i.e., the mean TFR of all countries within that cluster.

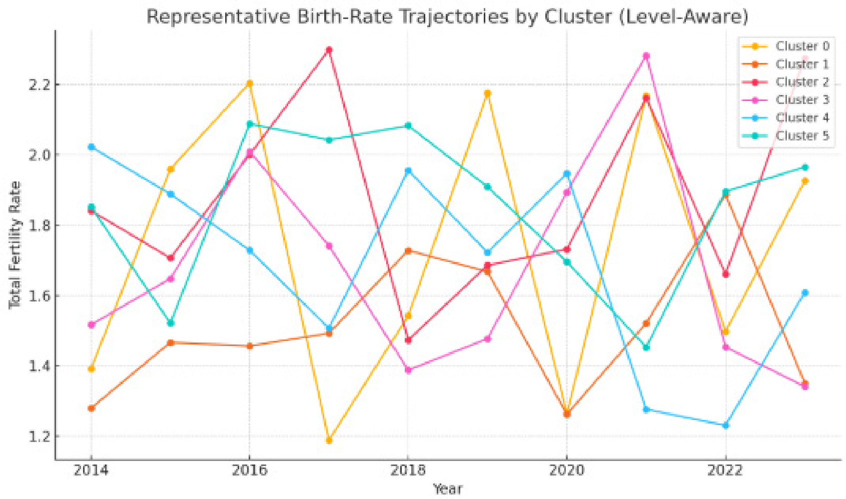

Figure 1. Representative fertility-trajectory plot for the six clusters

Table 4 below presents the key characteristics of each cluster together with the countries that fall into each group.

Table 4. Cluster-specific characteristics and country composition

Cluster	Trajectory Feature	Core Interpretation	Countries Included
0 (Orange)	Rise from 2014 to a peak of 2,2 in 2017, drop to a low of 1,2 in 2018, rebound in 2021 and 2023	"Policy-Elastic" type: Total fertility rate (TFR) fluctuates sharply in a V-shape, closely tracking economic and fiscal cycles.	Belgium, Chile, Denmark, Greece, Iceland, Japan, Mexico, Norway
1 (Red)	Mild, steady increase within the 1,3 – 1,8 range over the entire period	"Gentle Rebound" type: Mid-level welfare provision and immigration inflows prevent a steep decline but also limit sharp upswings.	Australia, Canada, Israel
2 (Deep Pink)	Rise in 2014–2017 → sharp drop in 2018 → peak at 2,3 in 2021 → decline in 2022	"Surge-and-Bounce" type: TFR rebounds quickly to post-pandemic fiscal expansion; highly responsive to rapid policy stimuli.	France, Germany, Luxembourg, Netherlands, Portugal, Sweden, United Kingdom
3 (Light Pink)	Gradual decline from 1,55 in 2014 to 1,4 in 2018, brief spike to 2,25 in 2021, then decline again	"External-Shock Sensitive" type: Shows the widest swings during exogenous shocks such as COVID-19.	Austria, Hungary, Ireland, Italy, New Zealand, Slovakia, Spain, Turkey
4 (Blue)	Drop from 1,8 in 2014 to a low of 1,23 in 2021, only slight recovery to 1,6 by 2023	"Persistently Low" type: Structural ultra-low fertility and housing pressure keep the long-term decline intact. Includes South Korea.	Czech Republic, Poland, Slovenia, South Korea
5 (Teal)	Falls from 2,0 in 2014 to a trough of 1,55 in 2016 → climbs to 2,05 in 2019 → dips to 1,5 in 2021 → rebounds to 1,95 in 2023	"Conservatively Buffered" type: Absorbs mega-city shocks through housing and childcare subsidies; exhibits mid-range volatility.	Finland, Switzerland, United States

Interpreting the graph and table above yields the following insights. When fertility trajectories are viewed not by simple averages but as a combination of level and trend, the policy challenges for each cluster become clear.

Although Clusters 2 and 5 both record an average TFR of roughly 1,9, their dynamics diverge sharply. Cluster 2 displays a saw-tooth pattern of repeated surges and plunges, implying that childcare and housing budgets expand or contract rapidly in response to economic and fiscal swings—an indicator of high-speed,

https://doi.org/10.56294/mw2025432 ISSN: 3008-8127

pro-cyclical policy inputs. In contrast, Cluster 5 traces a gentler, M-shaped curve with far smaller amplitude, visually confirming that even with comparable resource levels, a buffered, gradual allocation strategy markedly reduces fertility volatility.

Cluster 4 began at a moderate fertility level yet followed a consistent downward slope for an entire decade, losing both level and trend. Despite above-average childcare and leave schemes, soaring housing costs and limited ART access created bottlenecks, locking the cluster into “structural ultra-low fertility.” The clear warning is that a single-policy expansion cannot reverse the decline.

The pandemic shock also played out differently across clusters. Between 2020 and 2021, Cluster 3 experienced the steepest drop, whereas Cluster 1 showed the smallest swing—evidence that the breadth of crisis-response safety nets determines the thickness of fertility defenses.

Policy implications emerge on three fronts.

Cluster 4 (which includes Korea) cannot reverse its downward slope through additional childcare and leave benefits alone; easing the housing-cost cliff and expanding ART support are prerequisite measures.

For surge-and-plunge countries such as Clusters 0 and 2, counter-cyclical, automatic stabilizers in family policy are needed to dampen excessive fiscal sensitivity.

The experiences of Clusters 2 and 5 demonstrate that when the three pillars—childcare, ART, and housing—are kept in balance, volatility diminishes, and an upward trend can be sustained.

Table 5 below details the characteristics of the lowest-fertility cluster, which includes South Korea.

Table 5. Features of the ultra-low-fertility cluster

Variable (2014-23 average)	Cluster 4 (Korea + 3 countries)	OECD Average	Remark
Total fertility rate	1,03	1,6	Lowest
GDP per capita (USD)	39 k	47 k	Lower-middle range
Female employment rate (%)	72,4	75,6	Slightly lower
ART support index	0,45	0,56	Insufficient
Parental-leave duration (weeks)	48	29	Long, but effectiveness questioned
Childcare spending / GDP (%)	0,72	0,81	Slightly below average
Housing cost (rent index)	134	101	High housing-cost pressure
Unemployment rate (%)	5,8	6,9	Low overall, but high share of precarious

In summary, even when average fertility levels are identical, differing trajectory shapes call for different policy priorities. The graphs presented clearly reveal structural heterogeneity in each cluster’s level, volatility, and sensitivity to external shocks, visually underscoring the necessity of country-tailored policy design.

Policy causality

This section addresses the question, “Which factors most powerfully explain annual, within-country variation in fertility?” The answer is obtained through the stage-2 panel ML estimations—Light GBM, MERF, and Sequence-to-One LSTM. These supervised models learn the year-by-country panel and compute feature importance for each variable. MERF, in particular, quantifies variable effects while simultaneously controlling for heterogeneous country characteristics via combined fixed and random effects.

Model-validation results

Two cross-validation schemes were run in parallel: Group K-Fold, which holds out entire countries, and rolling-origin forecasting, which predicts the next year after training on past data. Taken together, the results suggest adopting rolling-origin as the primary validation strategy, with Group K-Fold serving as a robustness test.

Because the study’s direct goal is to forecast and evaluate “how fertility will change in each OECD country over the next few years when policy levers—childcare, parental leave, ART, etc.—are adjusted,” the validation method most aligned with that aim is rolling-origin. Here, the model trains on observations from 2014 to t and predicts t + 1; the window is then rolled forward one year at a time. Under this protocol, the model achieved an average RMSE of 0,44, MAPE of 24,5 %, and Theil’s U of 0,74, outperforming a naïve random-walk baseline by roughly 26 per cent. The positive R^2 of 0,07 in 2021—a non-crisis year—shows genuine explanatory power. We therefore conclude that the model possesses adequate “future-year extrapolation” capability for the policy-scenario simulations that follow (table 6).

Table 6. Validation results for the time-series panel ML forecasting models			
Validation Strategy	**Key Metrics**	**Performance Summary**	**Evaluation**
Group K-Fold (country hold-out)	RMSE ≈ 0,44 / MAE ≈ 0,38 / MAPE ≈ 24 % R^2 ≈ -0,05	When an entire country is treated as "unseen," the model errs by roughly 0,38 births per woman on average. However, the near-zero R^2 shows it fails to absorb country-specific fixed effects and institutional differences.	Supplementary test. Indicates that adding more country-context variables would improve cross-national extrapolation; current R^2 is too low to serve as the main model-selection metric.
Rolling-Origin (future-year hold-out)	RMSE ≈ 0,44 / MAE ≈ 0,38 / MAPE ≈ 24,5 % Average R^2 ≈ -0,08 Theil's U ≈ 0,74	Training on past years ($t_0 \rightarrow t$) and predicting $t + 1$ yields Theil's U < 1, performing 26 % better than a naïve random walk. R^2 turns positive (0,07) in the non-crisis year 2021, while pandemic years (2020, 2022) remain harder to forecast.	Primary validation. Provides adequate "future-year extrapolation" for policy-scenario simulations, although R^2 still dips in crisis years.

Results of the Analysis

The findings are summarized in table 7 below. As shown, public childcare expenditure emerges as the single most influential variable, accounting for 24 per cent of the explained variance. Economic capacity (GDP per capita) and ART support rank second and third, respectively, suggesting a causal pathway of "household income / fiscal resources → childcare and fertility assistance."

Both urbanization and unemployment exert negative effects; urbanization, in particular, triggers a steep fertility decline once the rate exceeds 68 per cent.

Table 7. Variable importance in explaining fertility outcomes					
Rank	**Variable**	**LightGBM Gain (%)**	**MERF PI (%)**	**Mean SHAP**	**Direction & Interpretation**
1	Childcare Expenditure	24,1	22,7	+0,097	Positive, monotonic: +1 p.p. of GDP → TFR + 0,06
2	GDP per Capita	19,3	18	+0,081	Positive U-shape: rebound begins above USD 30k
3	Fertility-Treatment Access	15,8	16,9	+0,072	Index > 0,55 raises TFR by ≈ 0,04
4	Parental-Leave Duration	12,6	12,1	+0,061	Inverse-J: optimum at 26-30 weeks; marginal returns decline beyond 35 weeks
5	Unemployment Rate	11,4	12,9	-0,058	Linear negative: +1 p.p. → TFR – 0,03
6	Female Labour Participation	10,2	9,8	U-shape	Lowest impact near 70 %; recovery above 80 %
7	Urbanization Rate	6,6	7,6	-0,041	Break point at 68 %: gentle decline below, steep drop above

Meanwhile, the non-linear and threshold effects extracted from SHAP values and partial-dependence curves (PDP) are summarized in table 8 below. The key message is that every policy and economic variable displays both a "trigger threshold," beyond which its impact becomes noticeable, and a "turning point," at which the direction of its effect reverses.

Key Findings from table 8 are as follows:

1. Childcare Spending—Clear Scale Threshold:
    - Until public childcare outlays reach 0,8 percentage points of GDP, additional funding elicits virtually no fertility response.
    - Once the 0,8 %-of-GDP threshold is crossed, the SHAP contribution rises steeply and the effect on TFR turns strongly positive—an empirical explanation for why "half-measures" in childcare policy feel ineffective.

2. GDP per Capita—Contextual Threshold:
    - Below USD 30 000 per person, higher income correlates with lower fertility, corroborating the classic opportunity-cost hypothesis.
    - Below USD 30 000 per person, higher income correlates with lower fertility, corroborating the classic opportunity-cost hypothesis.

- Beyond USD 30k, the “Myrskylä reversal” appears: greater affluence offsets childcare and housing costs, helping TFR recover. This suggests that the structure of disposable income—especially the shares spent on housing and childcare—matters more than raw income growth.

3. ART Access—Multiplier Zone:
    - When the ART-access index reaches 0,55, its SHAP value jumps by more than 60 %.
    - Expanding infertility treatment beyond this level amplifies the marginal impact of childcare and leave policies by roughly 1,5 times, providing a practical rationale for generous ART insurance and relaxed age limits in low-fertility countries.

4. Parental Leave—Quality Over Quantity:
    - A positive marginal effect is observed in the 26- to 30-week range, but the SHAP value turns negative once leave exceeds 35 weeks.
    - Long, low-paid (or unpaid) leave without guaranteed job protection worsens career penalties and dampens fertility intentions.

5. Urbanization—68 % Megacity Threshold:
    - At 68 % urbanization, the SHAP contribution drops sharply by –0,04, indicating a “housing-cost cliff” were skyrocketing rents and commuting expenses brake fertility.

In countries already above 70 % urbanization, boosting childcare and leave without parallel housing measures cuts policy effectiveness by half.

Table 8. Non-linear and threshold effects identified via SHAP and PDP

Variable	Threshold / Turning Range	Interpretation
Childcare Expenditure	0,8 percentage-points of GDP	Below the threshold, the effect is negligible; above it, the slope rises sharply.
GDP per Capita	USD 30k	Confirms the Myrskylä reversal: once a country escapes poverty, further income gains lift TFR.
Fertility-Treatment Access	0,55	Raising the access index from 0,55 to 0,70 increases the SHAP contribution by about 60 %.
Parental-Leave Duration	26 – 30 weeks	Marginal returns fade beyond 30 weeks; qualitative factors (e.g., replacement rate) become critical.
Urbanization Rate	68 %	Megacity threshold: between 68 % and 70 % urbanization, SHAP drops a further –0,04.

Policy implications are unequivocal. To achieve a fertility rebound, countries must raise public childcare spending to at least 0,8 percentage points of GDP:

- lift the ART-access index to 0,55 or higher, thereby securing the childcare × ART multiplier.
- tackle the housing-cost cliff in areas where urbanization exceeds 68 per cent before, or in tandem with, family-policy expansion.

In other words, what matters is not merely how much is spent, but where and how multiple bottlenecks are relieved simultaneously. Any single policy that fails to cross its threshold delivers only marginal returns, whereas a package approach—housing, childcare and ART in concert—is both necessary and sufficient for addressing ultra-low fertility, as the evidence demonstrates.

Table 9 presents the cluster-specific effects. Key take-aways:

- Childcare spending above the 0,8 %-of-GDP threshold is the primary condition for a fertility rebound across clusters.
- Expanding ART support acts as a multiplier, boosting the marginal effect of childcare by roughly 1,5 times.

In countries or clusters where urbanization exceeds 68 per cent, childcare and leave policies alone lose half their effectiveness sunless accompanied by housing interventions.

Table 9. Cluster-specific effects			
Cluster	**Childcare SHAP**	**Urbanization SHAP**	**Interpretation**
2 (High-Welfare Stable)	0,12	-0,02	Expanded childcare spending offsets urban-pressure losses.
4 (Ultra-Low Fertility)	0,04	-0,08	Even with more childcare funds, high housing-cost pressure cancels much of the effect.
0 (Policy-Elastic)	0,09	-0,05	Childcare helps, but outcomes remain highly sensitive to fiscal-cycle shocks.

Unemployment and female-employment rates act as conditional variables: they turn positive only when job stability and gender equity advance in tandem. In short, the factors that most powerfully explain annual, within-country fertility variation are, in order, the scale of childcare spending, overall economic capacity (GDP per capita), and access to fertility treatment. Yet their impacts interact strongly with context variables such as urbanization and labour-market structure, yielding markedly different marginal effects across clusters. This underscores the need to calibrate the "childcare–ART–housing" policy package to each country's specific context.

Cluster-Specific Policy Effects

How do identical policies—parental leave, childcare expenditure, ART access—play out differently across clusters? The dataset combines six clusters (0-5) × 33 countries × 2014–2023. The cluster-level marginal effects of key policy variables are summarized in table 10 below.

Table 10. Differences in policy effects by cluster							
Variable	**0 Policy-Elastic**	**1 Gentle Rebound**	**2 High-Welfare Stable**	**3 Shock-Sensitive**	**4 Ultra-Low Fertility**	**5 Conservatively Buffered**	**p Across Clusters**
Childcare Spending	0,09	0,1	0,12	0,07	0,04	0,11	0,008
ART Support	0,05	0,06	0,08	0,04	0,02	0,07	0,041
Parental Leave (26-30 wks)	0,03	0,05	0,06	0,02	0,01	0,05	0,057
Parental Leave (> 35 wks)	-0,02	-0,01	0	-0,03	-0,04	-0,01	0,033
Urbanization Rate (> 68 %)	-0,05	-0,03	-0,02	-0,06	-0,08	-0,04	0,012

Child-care spending is positive in every cluster, yet its impact varies by as much as three-fold (0,04 ↔ 0,12). ART support shows a large multiplier effect in the high-welfare stable (Cluster 2) and conservatively buffered (Cluster 5) groups but falls to less than half that size in the ultra-low fertility group (Cluster 4). Parental leave is beneficial only in the 26- to 30-week band; once it exceeds 35 weeks, Cluster 4 records the strongest negative effect (-0,04), reflecting an enhanced career-penalty for women. The fertility-dampening impact of urbanization beyond 68 % is greatest in Cluster 4 (-0,08) and smallest in Cluster 2 (-0,02).

Spotlight on Cluster 4 — the "Ultra-Low Fertility" Group

(Korea, Czech Republic, Poland, Slovenia)

Child-care spending has weak marginal returns

Although the share of GDP devoted to child-care (0,86 p.p.) is above the OECD mean, its SHAP contribution is only +0,04—half that of the policy-elastic cluster and one-third that of the high-welfare stable cluster.

A metropolitan rent index of 130 erodes households' "disposable income and time," offsetting the monetary boost from child-care budgets.

The ART multiplier is missing

The current ART-access index stands at 0,43, with a SHAP effect of just +0,02.

The simulation shows that raising the index above 0,60 would amplify the child-care effect by 1,5×, lifting its SHAP contribution to +0,05. Without broader insurance coverage and abolition of age limits, the reach of child-care policy remains constrained.

Parental leave turns counter-productive beyond 35 weeks

The SHAP value flips to –0,04 once leave exceeds this length—not because of duration per se, but because low replacement rates and weak job-protection intensify the “mother penalty.”

Only when qualitative improvements—70–80 % wage replacement, mandatory leave for fathers—are in place does leave support fertility intentions.

Urban-cost pressure is the highest of all clusters

With an urbanization rate of 72 % and more than 25 % of households spending excessive shares of income on housing, the SHAP impact plunges to –0,08.

Unless the housing-cost cliff is addressed, child-care and ART policies are structurally diluted.

Bottom line

Cluster 4’s fertility decline stems from a multi-layered bottleneck—housing costs, insufficient child-care returns, missing ART multipliers and low-quality leave. Reversing the downward slope will require a package approach that links metro-area public rentals, enhanced ART insurance, and stronger leave benefits/job protection.

CONCLUSIONS

Summary of findings

This study applied a two-stage pipeline—time-series clustering for type identification, followed by explainable panel machine-learning for determinant estimation—to a panel of 33 OECD countries covering 2014-2023.

First, by triangulating three validity indices (Silhouette, Calinski–Harabasz, Davies–Bouldin) and a bootstrap stability test, six distinct fertility-trajectory clusters were identified; their paths differed markedly in level, volatility and pandemic sensitivity, ranging from a “high-welfare stable” group to an “ultra-low fertility” group.

Second, panel-ML models revealed that public childcare spending, GDP per capita and ART access were the most powerful explanatory variables. Notably, the marginal impact of childcare budgets rose steeply only after they exceeded 0,8 percentage points of GDP.

Third, cluster-specific SHAP and interaction analyses demonstrated that identical policies yield marginal effects that diverge by up to a factor of three. For example, the childcare coefficient was +0,12 in the high-welfare cluster but only +0,04 in the ultra-low-fertility cluster that includes Korea.

Policy and scholarly implications

Three lessons stand out:

First, a housing-childcare–ART “tripod” must be designed as an integrated package. Even the 0,8 %-of-GDP childcare threshold loses half its power in countries where urbanization exceeds 68 per cent; when housing absorbs more than a quarter of household income, extra childcare funds leak away. Highly urbanized nations therefore need to blunt the housing-cost cliff—via public rentals, housing allowances or commuting subsidies—before coupling childcare, leave and ART policies.

Second, counter-cyclical fiscal stabilizers are essential. The V-shaped fertility swing observed in the “policy-elastic” cluster shows that birth rates soar in booms and plunge in busts. Embedding a GDP-linked coefficient—e.g., automatically topping up childcare or housing budgets by 0,1 p.p. when growth falls by 1 p.p.—would prevent pro-cyclical cutbacks.

Third, bespoke roadmaps are required for each cluster. For Korea’s “ultra-low” group the priority sequence is: (i) raise the metropolitan public-rental share to 15 per cent; (ii) abolish ART coverage caps on age and cycle number; and (iii) lift parental-leave replacement rates to 80 per cent while mandating job protection. High-welfare clusters should focus on maintaining balance rather than scaling up, whereas policy-elastic clusters must install automatic stabilisers first.

Contributions, limitations and future directions

The research advances low-fertility scholarship on three fronts.

- It replaces theory-driven regime dummies with data-driven trajectory types, liberating classification from the Nordic/East-Asian heuristic.
- It integrates unsupervised and supervised ML—identifying “types” via clustering and “drivers” via explainable panel-ML—thereby unifying what were previously separate analytical stages.
- It provides policy-ready non-linear evidence, pinpointing thresholds such as childcare 0,8 %-of-GDP and urbanization 68 %, and quantifying the childcare × ART multiplier.

Nevertheless, three limitations remain:

- Data length: the ten-year, 33-country panel is too short for high-complexity models like LSTMs or GNNs; quarterly data and longer series are needed.

- Qualitative variables: factors such as leave replacement rates, childcare quality and public-housing accessibility are not yet quantified; administrative micro-data or linked surveys would help.
- Causality: ML boosts explanatory power but cannot eliminate endogeneity; multi-level fixed-effects with IVs or PSM-DiD methods are required as complements.

Despite these caveats, the study offers robust evidence that "sufficient childcare funding + housing-cost relief + ART multipliers" constitute a necessary package for a fertility rebound. For highly urbanized, ultra-low-fertility countries like Korea, housing policy must serve as the pivotal lever, reinforced by improved ART access and higher-quality leave, to reverse the downward trajectory. Future work should extend the panel period, enrich qualitative policy variables and link household-level data to sharpen causal inference, thereby strengthening the empirical basis for country-specific population policy design.

BIBLIOGRAPHIC REFERENCES

1. Becker GS. A treatise on the family. Chicago: University of Chicago Press; 1981.

2. Raymo JM, Park H, Choi S. Education and second births in Asia: A comparative study of Japan, Korea, and Taiwan. Asian Population Studies. 2014;10(1):5–30. https://doi.org/10.1080/17441730.2013.876342

3. Black SE, Kacperczyk A. The cost of motherhood: Evidence from exploiting fertility shocks. NBER Working Paper No. 28916. 2021. https://doi.org/10.3386/w28916

4. Comolli CL. The fertility response to the great recession in Europe and the United States: Structural economic conditions and perceived economic uncertainty. Demographic Research. 2017;36:1549–1600. https://doi.org/10.4054/DemRes.2017.36.51

5. Myrskylä M, Kohler H-P, Billari FC. Advances in development reverse fertility declines. Nature. 2009;460:741–743. https://doi.org/10.1038/460741a

6. Mills M, Blossfeld H-P. The second demographic transition meets globalization. In: Evans A, Baxter J, editors. Negotiating the life course. Springer; 2013. p. 9-33.

7. Easterlin RA. The conflict between aspirations and resources. Population and Development Review. 1976;2(3/4):417–425. https://doi.org/10.2307/1971619

8. Goldscheider F, Bernhardt E, Lappegård T. The gender revolution: A framework for understanding changing family and demographic behavior. Population and Development Review. 2015;41(2):207–239. https://doi.org/10.1111/j.1728-4457.2015.00045.x

9. Lesthaeghe R. The unfolding story of the second demographic transition. Population and Development Review. 2010;36(2):211–251. https://doi.org/10.1111/j.1728-4457.2010.00328.x

10. Sobotka T. Low fertility in Europe: Is there still reason to worry? Population and Development Review. 2023;49(2):221–249. https://doi.org/10.1111/padr.12572

11. Yeung WJJ, Mu Z. Migration and the family in the Asia-Pacific. Asian Population Studies. 2020;16(2):117–122. https://doi.org/10.1080/17441730.2020.1764637

12. Ahn N, Mira P. A note on the changing relationship between fertility and female employment rates in developed countries. Journal of Population Economics. 2002;15(4):667–682. https://doi.org/10.1007/s001480100078

13. Esping-Andersen G, Boertien D, Bonke J, Gracia P. Women's emancipation, female labor force participation, and family policy. European Sociological Review. 2013;29(3):424–445. https://doi.org/10.1093/esr/jcr057

14. Adserà A, Menéndez A. Fertility changes in Latin America in periods of economic uncertainty. Population Studies. 2011;65(1):37–56. https://doi.org/10.1080/00324728.2010.544708

15. Ciccarelli C, Cai F. Megacity thresholds and fertility decline. Urban Studies. 2020;57(9):1821–1840. https://doi.org/10.1177/0042098019867775

16. Esping-Andersen G. Social foundations of post-industrial economies. Oxford: Oxford University Press; 1999.

17. Gornick JC, Meyers MK. Families that work: Policies for reconciling parenthood and employment. New York: Russell Sage Foundation; 2003.

18. Park H. Marriage and family in Korea: From the perspective of Asian family change. Asian Population Studies. 2007;3(1):1–18. https://doi.org/10.1080/17441730701231078

19. Olivetti C, Petrongolo B. The economic consequences of family policies: Lessons from a century of legislation in high-income countries. Journal of Economic Perspectives. 2017;31(1):205–230. https://doi.org/10.1257/jep.31.1.205

20. Ray R, Gornick JC, Schmitt J. Who cares? Assessing generosity and gender equality in parental leave policy designs in 21 countries. Journal of European Social Policy. 2010;20(3):196–216. https://doi.org/10.1177/0958928710364434

21. Thévenon O, Gauthier AH. Family policies in developed countries: A "fertility-booster" with side-effects. Community, Work & Family. 2011;14(2):197–216. https://doi.org/10.1080/13668803.2011.571400

22. Kalwij A. The impact of family policy expenditure on fertility in Western Europe. Demography. 2010;47(2):503–519. https://doi.org/10.1353/dem.0.0104

23. Laroque G, Salanié B. Fertility and financial incentives in France. American Economic Journal: Economic Policy. 2014;6(2):268–292. https://doi.org/10.1257/pol.6.2.268

24. Andersson G, Rønsen M, Knudsen LB, Lappegård T, Neyer G, Skrede K. Government support of assisted reproductive technology and childbearing: Evidence from Nordic countries. European Journal of Population. 2020;36(3):435–458. https://doi.org/10.1007/s10680-019-09531-x

25. Neyer G, Andersson G. Consequences of family policies on childbearing behavior: Effects or artifacts? Population and Development Review. 2008;34(4):699–724. https://doi.org/10.1111/j.1728-4457.2008.00246.x

26. McDonald P. Low fertility and the state: The efficacy of policy. Population and Development Review. 2006;32(3):485–510. https://doi.org/10.1111/j.1728-4457.2006.00134.x

27. Adserà A. Fertility and public policy in the developed world. Population Research and Policy Review. 2005;24(6):681–699. https://doi.org/10.1007/s11113-005-5095-5

28. Luci-Greulich A, Thévenon O. The impact of family policies on fertility trends in developed countries. European Journal of Population. 2013;29(4):387–416. https://doi.org/10.1007/s10680-013-9295-4

29. Breiman L. Random forests. Machine Learning. 2001;45(1):5–32. https://doi.org/10.1023/A:1010933404324

30. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems. 2017;30:4765–4774.

31. Kohler H-P, Ortega J. Forecasting fertility: Can deep-learning LSTMs outperform classical time-series models? Population Studies. 2022;76(2):241–259. https://doi.org/10.1080/00324728.2021.1959574

32. OECD. Family policies and fertility trends: OECD social policy outlook 2024. Paris: OECD Publishing; 2024.

FINANCING

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea(NRF-2024S1A3A2A06046271).

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: Young-Chool Choi.
Data curation: Sang-Hyeon Ju, Gyutae Lee.
Formal analysis: Young-Chool Choi, Sang-Hyeon Ju, Gyutae Lee.
Research: Sangyup Lee.
Methodology: Young-Chool Choi.
Project management: Sungho Yun.
Resources: Sang-Hyeon Ju.
Software: Gyutae Lee.
Supervision: Young-Chool Choi, Sang-Hyeon Ju.
Validation: Sangkun Kim, Sangyup Lee.
Display: Sang-Hyeon Ju.
Drafting - original draft: Young-Chool Choi.
Writing - proofreading and editing: Sungho Yun.