Fertility Decline and Economic Consequences: A Statisical Analysis of Armenia's Demographic Challenges

Margarita Hambaryan

American University of Armenia

CS-108 Statistics, Section A

Dr. Ashot Abrahamyan

May 10, 2025

CONTENT

Introduction

Context & Motivation

Armenia faces a critical demographic challenge: its birth rate has plummeted to just 1.7 births per woman as of 2022, far below the replacement level of 2.1 required to sustain population size(Craig, 1994). This decline, combined with a rapidly aging population, poses serious risks to the country's long-term economic trajectory. In particular, a smaller future workforce may reduce overall economic output, strain public resources, and limit the country's growth potential.

This study explores the long-term economic consequences of Armenia's falling fertility rate. We aim to evaluate whether declining birth rates are likely to result in weaker GDP growth in the decades ahead. To do this, we analyze historical data and use statistical modeling to assess and forecast this relationship.

Specific Research Question/Hypotheses

Hypothesis 1: A decline in Armenia's fertility rate leads to lower long-term GDP growth, as fewer births reduce the future working-age population.
Hypothesis 2: There is an association between fertility rate categories and total labor force categories in Armenia.

Brief Outline Of Data Source

We use publicly available macroeconomic and demographic data sourced from the World Bank. These datasets include annual observations ranging from 1960 to 2023, depending on the variable. We use a 20-year lag to link fertility rates with corresponding future GDP outcomes.

Data Description

Source and Collection Method

We chose 5 datasets for this analysis:
Fertility Rate of Armenia,
GDP growth of Armenia,
Labor Force Participation Rate of Armenia,
Age Dependency Ratio of Armenia,
Total Labor Force(in absolute numbers) of Armenia.

All datasets used in this project were obtained from the World Bank World Development Indicators database, which provides standardized, high-quality data for cross-country comparisons. The variables were downloaded in CSV format and reshaped in R using a standardized preprocessing script.

Key Variables and Units

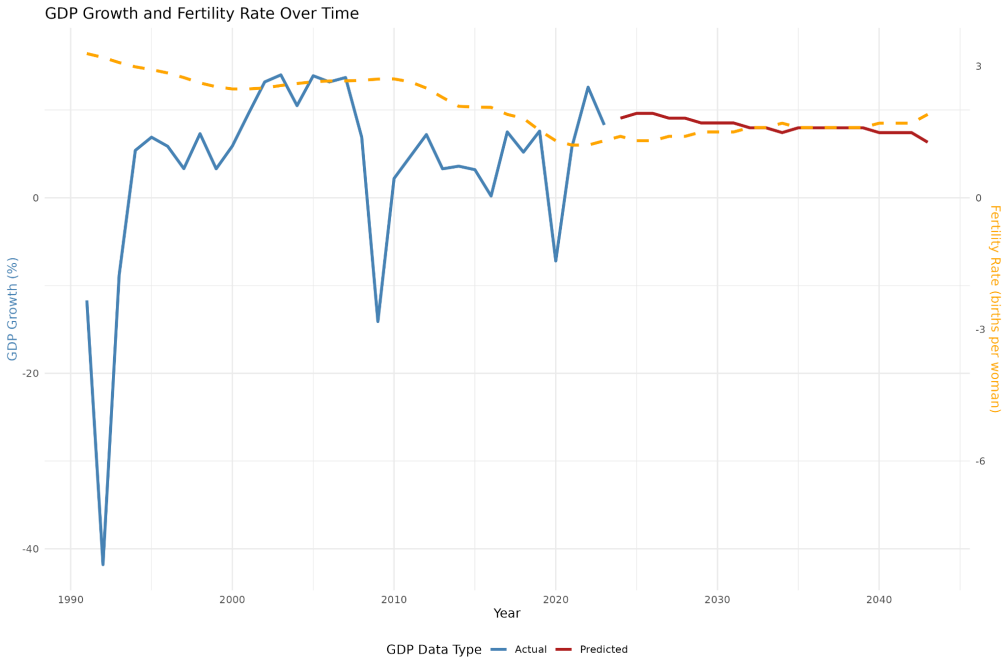
Variable Name	Description	Unit	Type
GDP_Growth	Annual real GDP growth rate.	Percent(%)	Numeric
Fertility_Rate	Average births per woman.	Births per woman.	Numeric
Fertility_Rate_Lag20	Fertility rate lagged by 20 years.	Births per woman.	Numeric
Labor_Force	Total labor force.	Number of people.	Numeric
LFPR	Labor Force Participation Rate	Percent(%)	Numeric
Dependency_Ratio	Age dependency ratio (% of dependents to workers)	Percent%)	Numeric
Fertility_Rate_Category	Categorical version of fertility rate (Low/Medium/High)	Category	Categorical
Labor_Force_Category	Categorical version of fertility rate (Low/Medium/High)	Category	Categorical

Exploratory Data Analysis

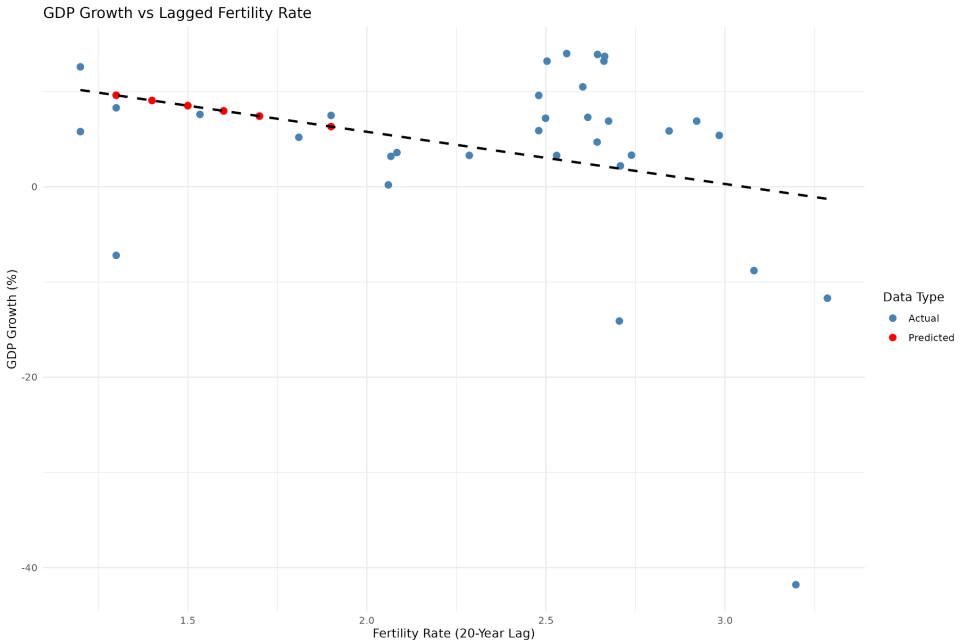
Summary Statistics

Variable	Mean	Standard Deviation	Min	Max
GDP Growth (%)	5.76	9.58	-43.89	29.14
Fertility Rate	1.80	0.32	1.40	2.60
Age Dependency Ratio	50.61	6.43	40.87	65.46

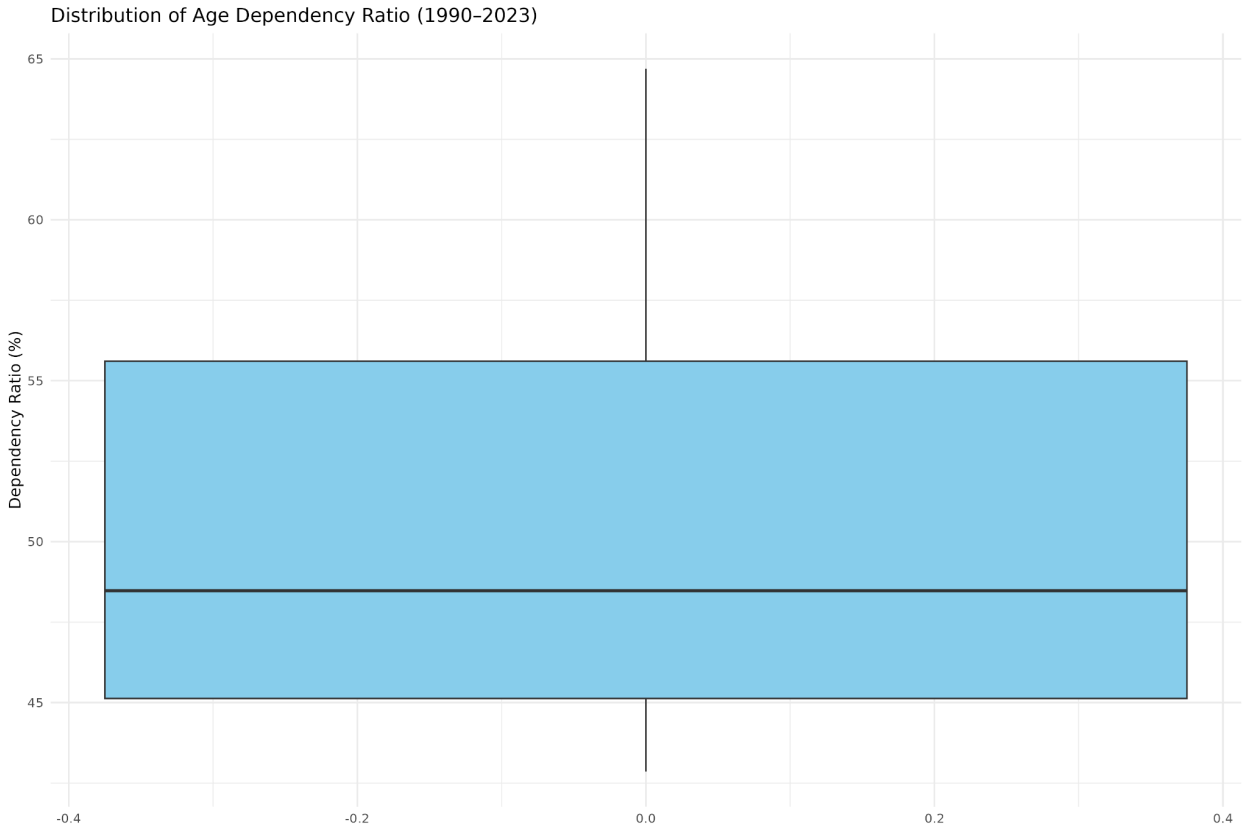
Vizualizations



This dual-axis line plot illustrates how Armenia's **GDP growth** (left axis) and **fertility rate** (right axis) evolved from 1990 through projected years. We observe a **general decline in fertility rate** over time. GDP growth is **more volatile**, but stabilizes in recent years and projections. This visual helps connect demographic shifts (fertility decline) with economic trends.



This scatterplot models GDP growth as a function of the fertility rate with a 20-year lag. Key insights: there is a negative relationship, suggesting that lower past fertility may reduce current GDP growth. Predicted values are highlighted in red, and a regression line shows the overall downward trend. This is important to show, since the graph and actuall model differ and may result in dissusions.



This boxplot displays the distribution of Armenia's age dependency ratio.The median is around 48%. The interquartile range is tight, indicating moderate variability. A few years may exhibit mild outliers, but overall, dependency rates remain within a consistent band.

Missing Values and Outliers

Missing Data: No missing values were found in the merged dataset used for analysis. All observations were retained for years with complete GDP, fertility, and demographic indicators.

Outliers: GDP growth shows major outliers in the early 1990s, likely due to post-Soviet transition and economic collapse. These have been retained, as they reflect important historical context. Age dependency ratio and fertility rate are relatively stable.

Why These Patterns Matter

The negative relationship between lagged fertility rates(we used 20 years) and GDP growth suggests a critical demographic-economic link. If fertility remains below replacement levels, it may result in smaller future workforces, higher dependency burdens, slower GDP growth.

In contrast, stable or increasing fertility (in balance with productivity growth) may help sustain long-term economic performance. This insight directly addresses the research question by highlighting how current demographic choices impact future economic outcomes, reinforcing the importance of forward-looking population and labor policies.

Methodology

To examine the economic impact of demographic trends, we estimated a simple linear regression model where GDP growth is the response variable, and fertility rate with a 20-year lag is the main predictor:

$$\text{GDP\_growth}_t = \beta_0 + \beta_1 \cdot \text{Fertility}_{t-20} + \epsilon_t$$

This model is appropriate given the research hypothesis: that past fertility affects current labor supply and thus economic growth. A linear model was chosen due to its interpretability and suitability for exploratory analysis, as recommended for initial hypothesis testing (Wasserman, 2004).

To further investigate whether fertility rates and overall labor force levels are statistically associated, a Pearson's chi-square test of independence was applied(Wasserman, 2004).

Software and Packages

All analysis was conducted using R. The following packages were used: tidyverse, ggplot2, lubridate, broom, scales, stargazer.

Results

The linear regression analysis examined the relationship between fertility rates (lagged by 20 years) and subsequent GDP growth. The model estimated that for each additional child per woman, GDP growth decreased by approximately 5.48 percentage points, holding other factors constant ($\beta = -5.482$, $SE = 3.220$). This effect size, though substantial, was only marginally statistically significant ($p = 0.0987$), indicating a potential but not definitive negative association. However, due to the relatively high p-value, there is insufficient evidence to reject the null hypothesis at the conventional 5% significance level, suggesting that the observed relationship may be due to chance.

The intercept was estimated at 16.745 (*SE* = 7.897), which is statistically significant at the 5% level (p = 0.0421), suggesting that in the hypothetical case of a zero fertility rate (lagged), GDP growth would average around 16.75%.

Model diagnostics showed an R-squared value of 0.086, meaning approximately 8.6% of the variability in GDP growth was explained by lagged fertility rates. The overall model fit was marginally significant, with an F-statistic of 2.899 on 1 and 31 degrees of freedom and a corresponding p-value of 0.0987. The residual standard error was 10.42, indicating a moderate level of unexplained variance.

For the second hypothesis, to investigate the potential association between fertility rates and the total labor force, a chi-square test of independence was conducted. Fertility rate and labor force were each divided into three categories (Low, Medium, High) based on tertiles. The contingency table was analyzed using Pearson's chi-square test. The results yielded a chi-square statistic of $\chi^2(4, N = 33) = 9.483$, with a p-value of 0.0501. Since the p-value slightly exceeds the 0.05 threshold, the result is not statistically significant at the 5% level. Therefore, we fail to reject the null hypothesis, suggesting that there is insufficient evidence to conclude a statistically significant association between fertility rate categories and labor force categories in Armenia.

Discussion

This analysis explored whether fertility rates influence economic growth two decades later. The results could suggest a negative relationship—countries with higher fertility rates 20 years ago tend to experience slower GDP growth today(even if it failed). While the association could be only marginally statistically significant, the effect size was economically meaningful, indicating that demographic dynamics may substantially shape long-term economic performance.

However, this analysis has important limitations. First, the model relies on a simple bivariate regression, which omits potentially confounding variables such as education, health, political stability, and global economic integration. These omitted variables could bias the estimated relationship. Second, the relatively low R-squared indicates that fertility explains only a small portion of the variation in GDP growth, suggesting the influence of many other factors.

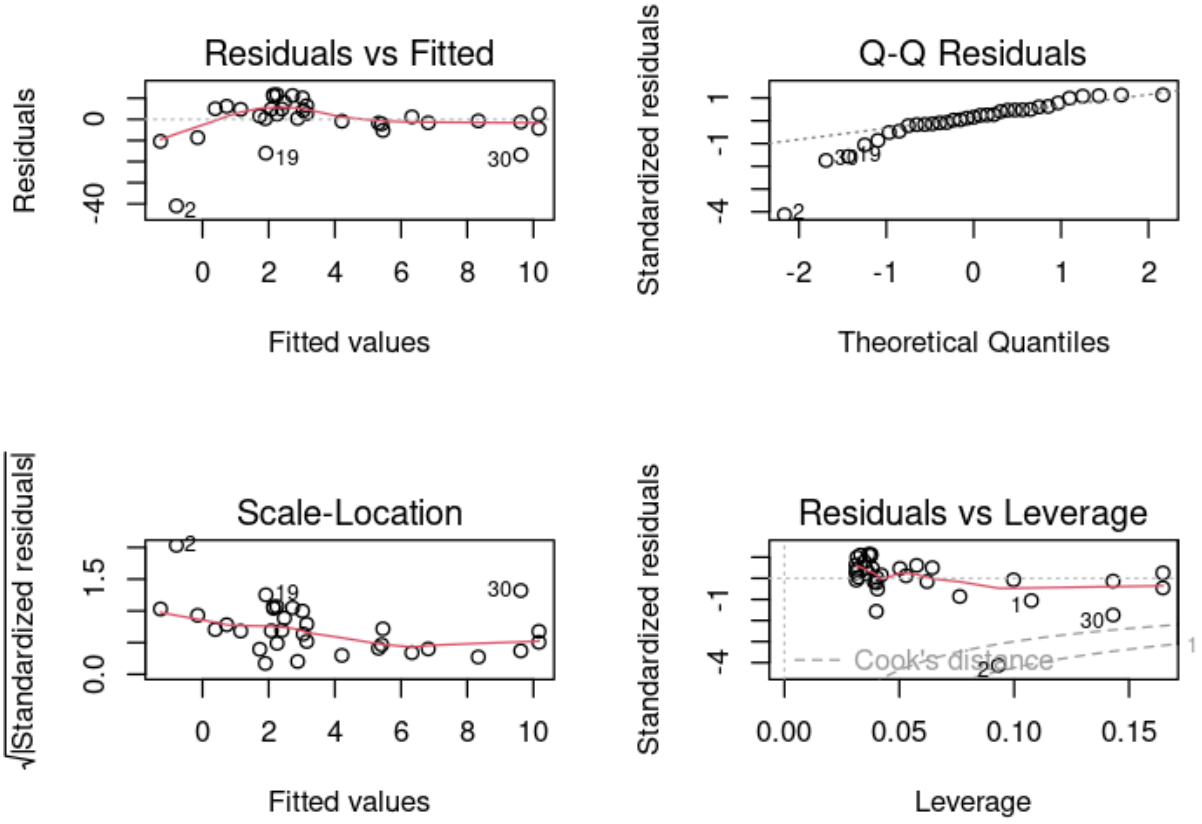
Speaking about the limitations of the second testing, a key limitation in this analysis is that several of the expected cell counts were below 5, which violates one of the assumptions of the chi-square test. This is indicated by the warning that the chi-square approximation may be incorrect(we also saw this while executing the script). Such small expected frequencies can lead to inaccurate p-values, weakening the reliability of the test. A larger sample size maybe could provide more robust results.

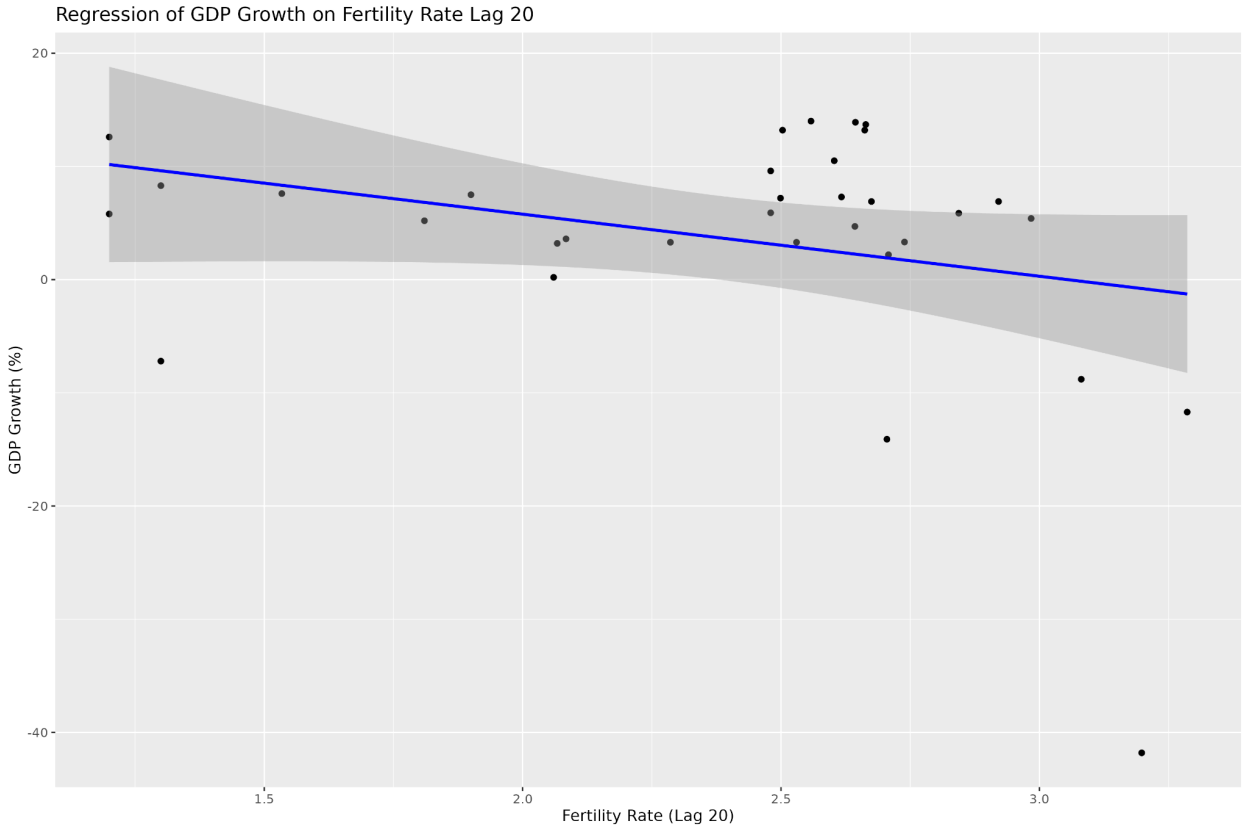
Conclusion

The linear regression analysis revealed a potential negative relationship between fertility rates lagged by 20 years and GDP growth; however, the association was only marginally significant, and the evidence was insufficient to confidently reject the null hypothesis at the 5% level. Similarly, the chi-square test

examining the association between categorized fertility rates and labor force levels produced a p-value slightly above the 0.05 threshold, indicating no statistically significant relationship. Thus, across both models, the evidence did not support rejecting the null hypotheses, underscoring the need for further investigation with larger datasets and alternative analytical methods to draw more definitive conclusions.

References & Appendix





Craig, J. (1994). Replacement level fertility and future population growth. *Population Trends*, (75), 20–22. https://pubmed.ncbi.nlm.nih.gov/7834459/

Wasserman, L. (2004). *All of Statistics: A Concise Course in Statistical Inference*. Springer Science & Business Media.

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Wickham, H. et al. (2019). *Welcome to the tidyverse*. *Journal of Open Source Software*, 4(43), 1686.

World Bank. (n.d.). *Age dependency ratio (% of working-age population)*. World Development Indicators. https://data.worldbank.org/indicator/SP.POP.DPND

World Bank. (n.d.). *Fertility rate, total (births per woman)*. World Development Indicators. https://data.worldbank.org/indicator/SP.DYN.TFRT.IN

World Bank. (n.d.). *GDP growth (annual %) – Armenia*. World Development Indicators. https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=AM

World Bank. (n.d.). *Labor force participation rate, total (% of total population ages 15+) (modeled ILO estimate) – Armenia*. World Development Indicators. https://data.worldbank.org/indicator/SL.TLF.ACTI.ZS?locations=AM

World Bank. (n.d.). *Labor force, total – Armenia*. World Development Indicators. https://data.worldbank.org/indicator/SL.TLF.TOTL.IN?locations=AM