Proceedings of ICFTBA 2024 Workshop: Human Capital Management in a Post-Covid World: Emerging Trends and Workplace Strategies
DOI: 10.54254/2754-1169/142/2024.LD19010

The Research on Factors Influencing the Fertility Rate -Take the United States as an Example

Yuyu Gao[1,a,*]

[1]*University of California, Irvine, Irvine, 92617, United States*
a. yuyug2@uci.edu
**corresponding author*

Abstract: This study examines the impact of different education levels of parents, income, and housing prices on the number of births in various states in the United States from 2016 to 2021. The data was collected from Kaggle on the number of births and education levels, the historical housing price data from Zillow, and the average income data from Federal Reserve Economics Data (FRED). The number of births was used as the response variable, while average income, housing price, and education level were set as the explanatory variables. The article will apply a multiple linear regression model to analyze the data, and this paper also considers the interaction effect between income and the prices of housing by adding an interaction term in the model. The major findings indicate that housing prices and people with less than a high school education have little or no impact on fertility rates, while higher education levels and income show a strong correlation with fertility.

Keywords: Fertility rate, education level, multiple linear regression, interaction effect, socioeconomic factors.

1. Introduction

Population aging is an unavoidable global trend that could bring social problems such as labor market shortages, pension and social security challenges, and so on [1]. The United Nations estimates that people who are 65 years or older may reach 1.6 billion by 2050 which almost matches the number of children under 12 [2]. The fertility rate is a significant element in balancing the age structure and that is the reason many governments support a higher fertility rate due to the threat of population aging [3]. Furthermore, Li explains that the fertility rate has a positive influence on economic growth, and economic growth also works on the increase in the fertility rate [4]. Therefore, it is necessary to understand the elements contributing to the fertility rate. This paper aims to study different factors that possibly influence the fertility rate, and how some factors interact with each other to apply additional impacts to the fertility rate.

The influence on fertility rate is a complex problem with many factors, and predictions of variation trends in fertility rate are also a popular argument among academics. Several researchers have conducted studies in this field and discovered that the fertility rate corresponds to the price of houses in that place, average income level, and individual's educational background [5-7]. Yi and Zhang used a standard Beckerian model of fertility behavior to test the relationship between house prices and people's fertility decisions in Hong Kong. In addition, they also combined the income effect and the compensated substitution effect with this model to increase the accuracy of the results [5].

DOI: 10.54254/2754-1169/142/2024.LD19010

Moreover, Park analyzed the correlation between housing prices and the fertility rate by bringing in the Shapley decomposition and the forecast error variance decomposition of the panel Vector Autoregression (VAR) model to estimate the effect of the price of the house on the fertility rate [6]. Țarcă et al. pointed out that the average income level obviously impacts the fertility rate fluctuation through a linear regression model in comparison with other socioeconomic factors such as alcohol consumption rate and smoking frequency [8]. Moreover, the impact of educational attainment on the fertility rate has also been examined. Cornett finds that the education level of females could be influenced by the fertility rate within a Demographic Transition Model (DTM) [9]. However, while this model is powerful, it only covers the female data and does not cover the male data. Götmark and Andersson demonstrate that the fertility rate is associated with education in different countries [10]. Nevertheless, previous research could help to prove that these factors (education level, average income, the housing price) could affect the fertility rate, but they are identical studies and did not consider the interaction effect.

In summary, this article focuses on educational attainment, income level, the price of houses, and the interaction between these factors to determine their impact on the fertility rate and build a multi-linear regression model to further study how these factors could impact the fertility rate.

2. Methods

2.1. Data Source

The set of data used in this article is collected from the Kaggle website, birth rates, and related data among 50 states and Washington. DC from 2016 to 2021 sourced from the Centers for Disease Control and Prevention (CDC). The income data comes from the Federal Reserve Economic Data (FRED), and the housing price data is fetched from Zillow, an American tech real-estate marketplace company.

2.2. Variable Selection

The raw dataset included birth data from 2016 to 2021 and parental education levels. In the meantime, since the research purpose of this paper is to examine the relationship between education levels, income, and housing prices on fertility rates, data from the same time period were selected from FRED and Zillow according to time to become my data set, which contains seven variables: state name, state abbreviation, year, number of births, education level, average income, and average housing price. The variables are described as follows in Table 1:

Table 1: List of Variables

Variable	Logogram	Meaning
Under High School	x_1	A dummy variable to check the individual's education level
Under College	x_2	A dummy variable testing the person's education lever
College	x_3	A dummy variable to test the individual's education level
Higher Degree	x_4	A dummy variable to check the individual's education level
Ave housing price	x_5	The average housing price in specific State
Ave income	x_6	The average income in specific state
Number of Births	Y	Total birth numbers in specific group

DOI: 10.54254/2754-1169/142/2024.LD19010

2.3. Method Introduction

In this paper, the proper way is to use multiple linear regression models to compare the power of the two models and the accuracy of the results, including two cases with and without the interaction term. In addition, multiple linear regression models allow the estimation of the effects of multiple independent variables on a dependent variable. This approach helps ensure that the results reflect the true effects of the independent variables on the dependent variable, free from the influence of unobserved, time-invariant characteristics specific to each entity.

3. Results and Discussion

3.1. Model Results

Examining the influence of different education levels, income, and house prices on fertility is the purpose of this study. At the same time, this study is going to use a multivariable linear regression model to estimate the strength and direction of the association between the independent variables and the dependent variables and compare the two cases with and without interaction terms. The general regression model is represented as follows:

$$E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_5 + \beta_5 X_6 + \varepsilon \quad (1)$$

with β_0 represents interception and ε represents the error term.

Table 2. Regression coefficient table

	B	S.E.	T	significance	VIF
Constant	-7174.182	3812.444	-1.88	0.060	
X1	20.164	1914.524	0.01	0.992	1.50
X2	24197.490	1914.524	12.64	0.000	1.50
X3	12195.710	1914.524	6.37	0.000	1.50
X5	0.003	0.008	0.36	0.716	1.81
X6	0.203	0.060	3.38	0.001	1.81

The results for the multiple linear regression analysis are presented in Table 2, where the response variable is the Number of Births, and the explanatory variables are educational levels (under college, college, and higher degree), average housing price, and average income. The coefficient of dummy variables under college and college are 24197.49 and 12195.71, implying that, holding other factors constant, individuals with an education level below college or college level are associated with approximately 24,197 or 12196 more births compared to the reference group in this research individuals with a higher degree. This suggests that lower education levels may be linked to higher fertility rates. Furthermore, both variables have p-values smaller than 0.001, indicating that they have a significant influence on the response variable. Furthermore, the coefficient of average income is 0.2031 with a standard error of 0.06008, and this variable is also statistically significant with a really low p-value of 0.001, indicating that if the average income increases, the number of births also tends to increase. However, variables X_1 and X_5 have high p-values shown in Table 2, this suggests that housing price and the degree level in high school have no or little impact on the number of births, according to the data.

The overall model represents statistical significance, $F(5,1218)=48.73$, $p < 0.0001$, suggesting that the predictors explained a significant portion of the variance in the number of births. The R-square value of 0.1667 presents that approximately 16.71% of the variation in the response variable, the number of births, can be explained by the explanatory variables in the model.

DOI: 10.54254/2754-1169/142/2024.LD19010

From Figure 1 below, the data in the normal distribution P-P plot shows a slightly S-shaped curve, indicating that the residuals generally follow a normal distribution, but the slight deviations at the ends should be investigated further.

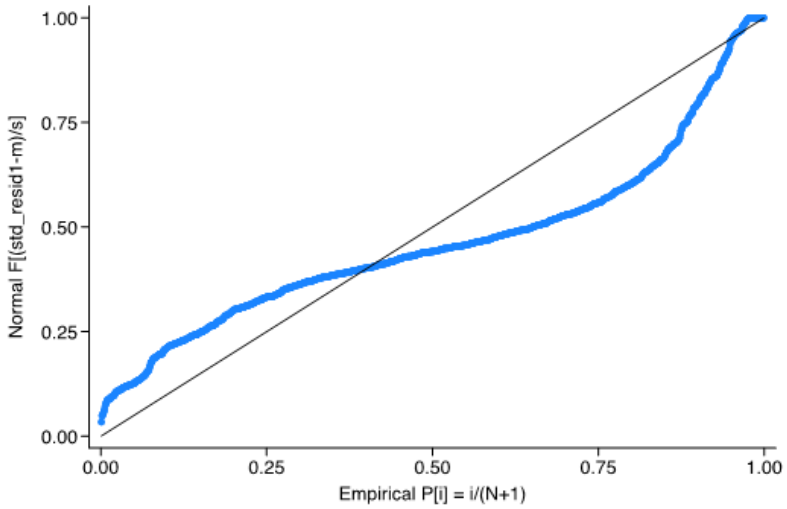

Figure 1: Normalized P-P plots of regression standardized residuals

3.2. Regression with Interaction Terms

Not only can independent variables have an effect on the direct response variable, but the interaction between two independent variables can also have an effect on fertility. In "Declining Fertility in America," Stone said that the article Declining Fertility in America expresses that the majority of women are willing to prioritize buying a home over having children, which reflects the impact of housing on life and decision-making [11]. Therefore, the effect of X_5 on Y may depend on the value of X_6, and the solution is to multiply these two terms together and add the coefficients into the previous regression model:

$$E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_5 + \beta_5 X_6 + \beta_6 X_5 X_6 + \varepsilon \quad (2)$$

With β_0 representing interception, ε represents the error term, and $X_5 X_6$ is the interaction term.

The results of the multivariable linear regression analysis including the interaction term are presented in Table 3 below, where the dependent variable Y means the Number of Births, and the independent variables are educational levels (under college, college, and under high school), average housing price, and average gross income and the interaction term.

Table 3: Regression coefficient table with interaction term

	B	S.E.	T	significance	VIF
Constant	-998.547	8515.380	-0.12	0.907	
X1	20.163	1914.793	0.01	0.992	1.50
X2	24197.490	1914.793	12.64	0.000	1.50
X3	12195.710	1914.793	6.37	0.000	1.50
X5	-0.018	0.027	-0.67	0.504	20.87
X6	0.122	0.116	1.05	0.293	6.77
X5X6	2.6×10^{-7}	3.21×10^{-7}	0.81	0.417	38.04

DOI: 10.54254/2754-1169/142/2024.LD19010

The outcomes of the analysis of regression suggest a statistically significant relationship between certain levels of education and the number of births. Specifically, individuals with "undercollege" education (less than a college degree) have a positive and significant effect on the number of births (β_2 = 24197.49, p < 0.001), with a 95% confidence interval with a range between 20440.83 and 27954.15. This points out that individuals with less than a college education have, on average, a higher number of births compared with other education groups. Similarly, individuals with a college education have a positive, though not statistically significant, coefficient (β_3 = 12195.71, p = 0.500), suggesting a possible relationship with the number of births, although the effect is weaker and more uncertain. Although high school education has a positive coefficient (β_1 = 20.1634), it has no statistical significance (p = 0.992), indicating that there is no strong correlation between the number of births and whether the person has a high school degree in this model.

Among the economic variables, the average housing price and gross income do not show statistically significant relationships with the number of births. The average housing price has a small negative coefficient (β_4 = -0.0181504) but is not significant (p = 0.500), while the average gross income has a positive coefficient (β_5 = 0.1223564), also not significant (p = 0.293). The interaction term (inter) is also not statistically significant (β_6 = 2.60e-07, p = 0.417). The overall R-squared for the model is 0.1671, showing that approximately 16.71% of the variance in the number of births is explained by the predictors in the model. The adjusted R-squared is 0.1630, reflecting a modest fit of the model.

4. Conclusion

In recent years, fertility has been a hot topic in countries with aging populations because labor shortages will become more prominent as the population ages. The purpose of this article is to examine how income, housing prices, and parental education levels affect the fertility rate. Using multiple linear regression model analysis, I concluded that the average income and education level, especially whether the individual gets a college degree or under college degree but higher than a high school degree, strongly correlated with the fertility rate while housing price and whether the individual has a high school degree have little or no impact on the fertility rate. As hypothesized initially, the increasing average income will affect the housing condition and the price of houses and thus stimulate fertility. However, contrary to expectations, the interaction between housing prices and income did not significantly impact fertility rates. This may reflect the complexity of housing dynamics in the U.S., where factors such as state-specific housing demand or urban-rural migration play a role beyond individual income and housing affordability. In addition, in different states, the density of the population is different, with more people living in that state, the housing price would be higher. While housing prices do not have a significant direct effect on fertility rates, income levels, and educational attainment, particularly at the college level, show a strong correlation. The findings suggest that policies targeting fertility rates should focus on improving educational attainment and income levels, particularly for individuals with less than a college education. For example, investment in vocational training programs or higher education access could have a more direct impact on increasing birth rates compared to policies centered on housing affordability.

Furthermore, one of the limitations of this paper is that the income data in the dataset lacks detailed stratification by education level. Additionally, the data does not account for potential cultural or regional differences in fertility decisions. Future research could explore more granular socioeconomic factors, such as disaggregating income levels by both education and geography, and including qualitative data on cultural attitudes toward fertility. All in all, this study underscores the pivotal role of education and income in shaping fertility rates in the U.S. while highlighting that the prices of housing may have a more indirect impact. As fertility rates continue to decline in developed

DOI: 10.54254/2754-1169/142/2024.LD19010

economies, understanding these nuanced relationships is critical for policymakers aiming to address population aging and labor shortages.

References

[1] *Cheng, H., Luo, W., Si, S. et al. (2022) Global trends in total fertility rate and its relation to national wealth, life expectancy and female education. BMC Public Health, 22, 1346.*

[2] *Hu, Q. (2017) Analysis of housing price factors based on the SVAR model. Times Finance.*

[3] *Ronald, L., et al. (2014) Is low fertility really a problem? Population aging, dependency, and consumption. Science, 346, 229-234.*

[4] *Li, Y.J. (2016) The relationship between fertility rate and economic growth in developing countries. Economic Inquiry.*

[5] *Yi, J. and Zhang, J. (2009). The Effect Of House Price On Fertility: Evidence From Hong Kong. Economic Inquiry, 48(3), 635–650.*

[6] *Park, J. (2021) An Empirical Study on the Contribution of Housing Price to Low Fertility. The Journal of the Convergence on Culture Technology, 7(4), 607-612.*

[7] *Hamilton, B.E. (2019) Total fertility rates, by maternal educational attainment and race and Hispanic origin: United States. Healthcare, 70, 5.*

[8] *Țarcă, V., Țarcă, E. and Luca, F.A. (2022) The Impact of the Main Negative Socio-Economic Factors on Female Fertility. Healthcare, 10(4), 734.*

[9] *Cornett, A. (2020) Analyzing the Relationship Between Female Education and Fertility Rate Spring 2020 Edition. BMC Public Health.*

[10] *Götmark, F. and Andersson, M. (2020) Human fertility in relation to education, economy, religion, contraception, and family planning programs. BMC Public Health, 20, 265.*

[11] *Stone, L. (2018) Declining Fertility in America. American Enterprise Institute.*