THE ANALYSIS OF THE EFFECT OF WOMEN’S EMPLOYMENT RATE ON FERTILITY ACROSS STATES IN UNITED STATES

Author Ai Liu, Zhanchu Quan, Ruling Wang

Abstract

This paper utilizes the approach of cross-sectional data to analyze the effect of women's employment rate on fertility across states in the United States. In recent years, there is a decline in women's fertility rates in the United States. Considering the economic consequences of a declining fertility rate, this paper aims to specifically analyze and quantify the causes of the low fertility rate. We will primarily look into women's employment rate while also taking into consideration other factors such as household income, related children of the householder, personal health care expenditures, and women's educational level.

I. Introduction

The fertility rate in the United States has been fairly steady for the previous three decades before the Great Recession. The fertility rate normally varied within a relatively small range with fewer kids being born during hard times and a rebound in births during periods of higher economic prosperity. However, the fertility rate in the United States has dropped sharply and shown no signs of recovery since the Great Recession of 2007 (Kearney et al., 2022). This broad national tendency has the potential to have far-reaching consequences on Americans' future: with a diminishing fertility rate, future generations will undoubtedly be burdened by a rapidly aging society and population decrease, through which many socioeconomic benefits would be hampered and national development could eventually be downcasted. However, the United States is not unparallel in this regard, since it has been observed to become a globalized trend, and the entire world appears ill-prepared for this fertility catastrophe. For instance, developed states in the global north such as Italy have witnessed a rapid drop in the fertility rate from 4 children per household in the 1960s to an average of 1.2 children in 2019, as highlighted by PBS. Such a drastic decline has caused significant social and economic repercussions, like the closing of schools and hospitals which further raises the cost of child-bearing, and the shrinking labor force contributes to economic recess (Livesay, 2019). Not to mention Japan, which has been historically bothered by a depressed fertility rate. Although its fertility rate has stopped declining in recent years, there is little effect on all efforts the Japanese government has done to boost the fertility rate. A record low of 1.26 hypothetical lifetime birth per woman was observed in 2005, and it increased to 1.3 in 2021- even when the pandemic effect is considered, the ratio hasn't been higher than 1.5 in more over three decades (Reidy, 2022). The Atlantic attributes this low rate to the Japanese-characterized culture that emphasizes the importance of males getting a regular job and acting as the breadwinner in marriage, while it is relatively hard to scout for "regular employment" since the Heisei Depression after the Japanese asset price bubble in the early 90s. On the other side, such a culture that emphasizes men's roles as breadwinners has established a key prerequisite for marriage and childbearing. Since irregular workers are regarded as less desirable marriage partners, and premarital childbearing is uncommon in the country's traditional eastern culture, with a significant rise in the number of irregular workers after the 90s, Japan's fertility issue has been further intensified (Semuels, 2021). This low fertility rate has led to extreme population aging in Japanese society and it's aging faster than any other nation in the world. As estimated by the National Institute of Population and Social Security Research (2017), approximately 36.19 million people in Japan were over the age of 65 in 2020 and this number is projected to peak at 39.35 million by 2042. By 2065, the elderly will account for no less than 38.4% of its total population.

Similarly, Japan's situation has implications for the United States, where temporary jobs become more common as many labor opportunities are slipped overseas. Such harsh economic prospect may deteriorate the overall fertility rate in the United States as its record lows keep dwindling year over year - from a historic low in 2018 with 59.0 births per 1,000 women aged 15 to 44 to a new low of 55.8 in 2020 (Livingston, 2019; Chappell, 2021). While the decline in fertility rate may be attributed to various socioeconomic factors such as economic environment, shifting societal and gender standards, household income, and female educational and work opportunities, the purpose of this study is to test the hypothesis that there is a significant association between the female employment rate and fertility rate across different states in the United States on both the national and regional levels. In this study, we will carefully examine how this relationship reacts among different states where women and their circumstances of conception and childbearing are of different characteristics. For this purpose, cross-sectional data analysis is applied. We opted to look into this relationship because there are few existing research that utilizes cross-sectional data to examine the relationship between female employment and fertility rates in different states in the United States. Our cross-state analysis presents a better understanding of these empirical associations at the national level, which is essential for comprehending the mechanisms underlying fertility change as a social phenomenon. Therefore, our hypothesis states that female labor force participation rates have a strong association with the fertility rate, so as more or fewer females participate in work, the fertility rate could be affected either way. The sets of contributing elements to the states' diverged results are usefully revealed by our cross-state analysis, providing a unique perspective on the current issue.

II. Literature Review

Jaba et al. (2016) conducted a study assessing the effect of female employment rate on the variation of total fertility rate over the period 2002-2012 and confirmed the relationship between the two. Taking into account the analysis of the employment rate of women aged 25-54 years, the research's result proves that women's participation in the labor market has a significant positive effect on the total fertility rate in the EU.

The study specifies that the welfare state model adopted and the specific labor market characteristics, such as political regimes and geographical aspects, among different EU countries result in different relationships between the two. In contrast to the other groups, the Anglo-Saxon social model group of countries exhibits a negative relationship between the employment rate of young women and the overall fertility rate. The social policies providing substantial help may, therefore, result in the negative equilibrium relationship between young women's employment rate and fertility rate in the countries in the group, rendering young women not motivated to enter or re-enter the labor market, not wishing to reduce their state-funded child care facilities. Neither the employment rate of young women nor female part-time work significantly affects fertility rates for the Continental social model group. Furthermore, for the group of nations using the Anglo-Saxon social model, GDP has a highly significant and advantageous effect on fertility rates. On the other hand, GDP harms fertility for the nations that belong to the Continental Social Model.

Media et al. (2020) verified how the fertility of Riau Province was affected by several social aspects. Mantra (2003) proposed that changes in demographic transition have been affected by various factors, including demographic and non-demographic ones. As fertility is a crucial component of demographic change, Media et al. (2020) examined such effect by using time series and cross-section data acquired between the years 2010 and 2017 and incorporating three measurements, including per capita GRDP, the number of working women, and poverty. Riau Province's GRDP is dominated by the agricultural, forestry, and fishing sectors, a typical traditional society with low production, education, and income levels, signifying children are commonly viewed as investments in production factors. Therefore, per capita GRDP shows a strong positive impact on fertility in Riau Province given that childbearing is more of a rational form of economic choice, illustrating the substitution effect and income effect (Todaro, 2000). Similarly, the number of working women (age 15 and over) exhibits such an effect, since high participation in work denotes the low education and income level of this society's very pre-modernization stage - typically linked with lower first marriage age for women and thus increase fertility. Yet Media et al. (2020) concluded no discernible connection between poverty and fertility. Fertility changes brought on by industrialization, which limited childbirth, will have little of an impact on the poor who is mostly found to live in rural areas (Chrisaniani, 2005) and have a strong propensity to uphold their culture and traditions (Baudin, 2010).

Siegel (2017)'s study is one of the few recent ones that examines the impact of factors such as women's relative wages on fertility in the United States and highlights the gender wage gap. The entire study was inspired by the idea that while increases in women's wages would theoretically reduce fertility, the fertility rates in the United States have remained fairly stable over the past 40 years against a backdrop of sustained increases in women's wages. On the other hand, a simultaneous trend of women spending less time on household chores has also been observed. This article differs from previous literature in that it aims to examine and quantify the endogenous response of men and women to the allocation of work hours at home and in the workplace and its impact on fertility. Although Siegel (2017) did not employ cross-section analysis to investigate the effect of wage increases on fertility, we may apply some of his models to explain compiled data in our study as both studies require domestic data from the United States. Siegel (2017) deems the previously idealistic complete division of labor within a family unrealistic, arguing the impact of the female wage increase on fertility is determined by the combination of substitution and income effect: when female wages rise, the cost of home production rises due to the substitution effect, lowering the fertility rate. On the other hand, as female wages increases, men's free time becomes available for household chores, cutting the cost of childbirth for women and keeping fertility rates from falling, indicating the income effect. Siegel (2017) establishes a log-linear relationship that computationally describes the effect of wage increase on fertility when the income effect is constrained and validates the result using calibration data from 1965 to 2005.

According to Siegel (2017)'s data up till 2005 and his derived model, it was predicted that the fertility rate will most likely be U-shape under the influence of income and the substitution effect. With the more recent data acquired, we will be able to further test his model in the period after 2005. If the data is consistent with Siegel (2017)'s model, we will be able to employ his model to explain data in our own cross-section analysis. Otherwise, we would have to identify other factors that affect recent fertility and reflect them in our variables selection.

III. Data

This study examines the employment and fertility of women across states. Given the purpose of the study, a sample size of 50 states, excluding one federal district (Washington D.C.), and Puerto Rico due to lack of data on fertility rate, was required. These states were chosen because they have the most reliable data that is consistently available through the official database. It can be challenging to find reliable data that is consistent and accurate for women's employment analysis as numerous studies have found that various factors and influences affect women's employment in the US. For example, Media et al. (2020) illustrated how fertility shows a strong positive association with per capita GRDP and the number of working women, while Jaba et al. (2016) indicate the employment rate of young women and fertility rate is negatively correlated for the Anglo-Saxon social model group. Albeit Caldwell (1980) contends that the impact of mass schooling is the main factor triggering the fertility decline in the West and relative wages and household specialization could also play a part (Siegel, 2017). Among those many societal influences that are believed to have impacts on women's employment status, we choose four: household income, percentage with related children, personal health care expenditures, as well as level of education, hoping to show a such intricate social effect on women's employment, with the number of females over age 16 in labor force being our explanatory variable and the fertility rate per 1000 females over the age of 15 being the dependent variable.

Table 1 – Variable Descriptions

Variable Descriptions	
fert[a]	General fertility rate per 1,000 women aged 15–44 (‰)
thouslab[b]	Number of females over age 16 in labor force per 1000 of females (‰)
medinc[b]	Median household income (dollars)
percchild[b]	Percentage with related children of household under 18 years old (%)
percchealexp[c]	Personal health care expenditures by state of provider (million of dollar)
percedu[b]	Percentage of female over 25 years with a bachelor's degree or higher (%)

Note. Data are from CDC, User Guide to the 2019 Natality Public Use File[a], U.S. Census Bureau, 2019 American Community Survey 1-Year Estimates[b], Centers for Medicare & Medicaid Services (2019). *Health Expenditures by State of Provider* .[c]

Main sources for independent variables are listed above including 2019 selected economic characteristics according to ACS 1-year estimates data profiles, and education attainment data compiled from the same data profile. The year of 2019 in used because it is the most recent data available from before the covid pandemic, which may cause some timing difference in its effect on some of the variables we consider. Our primary dependent variable fertility rate is from the CDC WONDER database for the year of 2019.

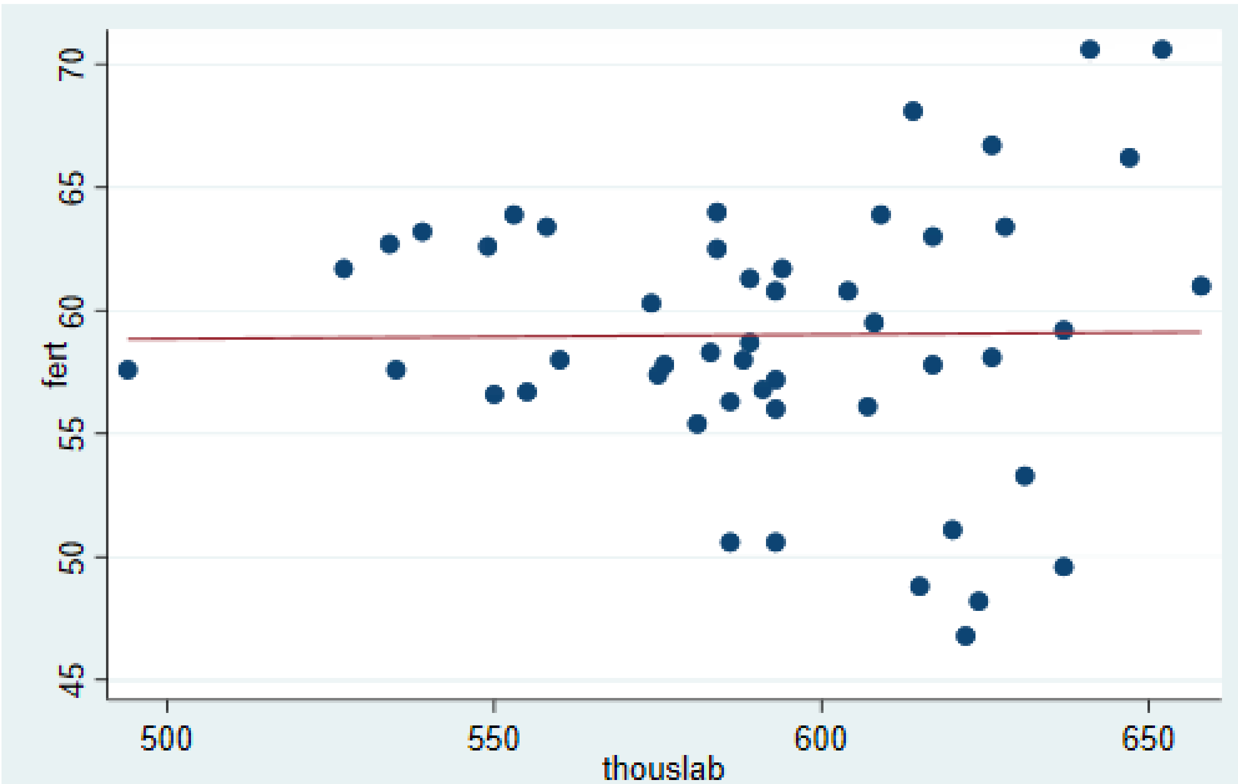
The table below shows the descriptive statistics for each variable.

Table 2 – Variable Descriptive Statistics

Summary Statistics				
Variable	Mean	Standard Deviation	Minimum	Maximum
fert	59.01	5.43	46.8	70.6
thouslab	592.92	35.47	494	658
medinc	64976.22	10604.02	45792	86738

percchild	13.46	3.76	7.1	23.8
percchealexp	63196.78	72749.58	4546	385480
percedu	33.30	5.29	22.3	45.7

Figure 1 – Scatterplot of fert vs. thouslab



Since *medinc* is in dollars and the mean value is also around 65000, in future studies we will consider the value of log(*medinc*) so that the coefficients will show more accurately the effect of medinc on fertility. Similarly, *percchealexp* has a unit in million of dollar and relatively large deviation, we would also consider to take the log level in further studies. The scatter plot shows that our simple regression model may not be a good fit for fertility. In addition, we do observe some outliers in the data, with the leftmost point being West Virginia, which has a low labor force participation rate and average fertility rate. We tried to build a model that excluded West Virginia, but the simple regression model and the fitting line did not change much.

There are a few things to look at to make sure the aforementioned models meets Gauss Markov Assumptions before running regression analysis on the data gathered.

1. Linear in Parameters:

First, the parameters we are estimating are linear given our linear regression equations as below.

Model 1:

$$fert = \beta_0 + \beta_1(thouslab) + u$$

Model 2:

$$fert = \beta_0 + \beta_1(thouslab) + \beta_2(medinc) + \beta_3(percchild) + \beta_4(percchealexp) + \beta_5(percedu) + u$$

Model 3:

$$fert = \beta_0 + \beta_1(thouslab) + \beta_2 log(medinc) + \beta_3(percchild) + \beta_4(percedu) + u$$

Model 4:

$$fert = \beta_0 + \beta_1(thouslab) + \beta_5(percedu) + u$$

2. Random Sampling

Secondly, all data used in this study are from random sampling from populations across the world. Since the data was gathered from credible sources (CMS, US Census, and CDC), this check is made. Since the analysis of this paper is about American states, the sampling done here is among the states. The sample size is 50 states. Thus, each data point from the samples therefore follows the population equation.

3. Non-Collinearity

For our simple regression model, the third assumption to be made is the values of the explanatory variables are not all the same. For our multiple regression model, the third assumption to be made is that the regressors are not perfectly correlated with each other. In our sample (and therefore in the population), none of the independent variables is constant. We test the correlation among our variables in STATA and got the results as following. As shown in the table, there're relatively strong correlation between *thouslab* and *percchild* and between *medinc* and *peredu*, which may cause bias to our regression models. We will also test the joint significance of these two pairs of variables in the Extensions section.

Table 3 – Multicollinearity Test

	thouslab	*medinc*	*percchild*	*perchealexp*	*percdeu*
thouslab	1.0000				
medinc	0.6724	1.0000			
percchild	-0.8423	-0.7844	1.0000		
perchealexp	-0.1026	0.2125	0.0481	1.0000	
peredu	0.7262	0.8375	-0.7391	0.1400	1.0000

4. Exogeneity:

The fourth check to make is exogeneity- the expected value of the error term, u, is zero, meaning the value of the explanatory variable (variables) must contain no information about the mean of the unobserved factors. Fertility rate has been studied in many researches and there are many factors that could affect fertility rate and the coefficients. Although it is still uncertain if we have factors affecting fertility rate left unconsidered, we assume there's no exogeneity in our regression models here. Also, we have tried our best to include possible independent variables in our multiple regression models to minimalize things end up in the error, where the zero conditional mean assumption is much more likely to hold, and exogeneity is minimalized.

5. Homoskedasticity:

The last assumption is for homoscedasticity- the variance of the residual, or error term, in a regression model is constant and not affected by any change in any explanatory variable holding other variable constant. In our models, it is likely that the value of the variance of the residuals is constant for each state's fertility rate, meaning that the explanatory variable's value are unrelated to that of the unobserved factors.

IV. Results

1. Simple Linear Regression model (Model 1):

$$\hat{fert} = 57.992 + 0.172(thouslab)$$

2. Multi Linear Regression model (Model 2):

$$\hat{fert} = -20.262 + 0.1603(thouslab) + 0.000267(medinc) + 0.7851(percchild)$$

$$+ 7.23\cdot10^{-8}(percchealexp) - 1.3110(percedu)$$

3. Modified Multi Linear Regression model (Model 3, 4):

$$fert = \beta_0 + \beta_1(thouslab) + \beta_2 log(medinc) + \beta_3(percchild) + \beta_4(percedu) + u$$

$$fert = \beta_0 + \beta_1(thouslab) + \beta_5(percedu) + u$$

Combine results from the simple model and the multiple regression models in one table (See Table 19.3)

3. Results Interpretation

We observe that perchealexp appears to have an extremely small coefficient, which may indicate that this perchealexp variable may be overspecified, as it also appears to be insignificant at the 10% level, we consider dropping perchealexp in Model 3. Moreover, as medinc has a unit of the dollar with a mean of 64976.22, therefore log(medinc) is considered instead of medinc in Model 3. In Model 4, we decided to drop all variables that are relatively insignificant, which are medinc (or log(medinc) in model 3), percchild, and perchealexp. We will further test our decision with F-test in the robustness test section.

Table 4 – Regression Models Summary

Dependent Variable: Growth Rate					
		Model 1	Model 2	Model 3	Model 4
Independent Variables	*thouslab*	0.0017 (0.0221)	0.1603*** (0.0293)	0.1572*** (0.0288)	0.1172*** (0.0227)
	medinc		0.0003** (0.0001)		
	log(medinc)			17.04** (7.15)	
	percchild		0.7851** (0.3115)	0.7995** (0.3161)	
	perchealexp		7.23e-8		

			(7.93e-6)		
	percedu		-1.3110*** (0.1975)	-1.2800*** (0.1964)	-1.066*** (0.152)
intercept		57.99 (13.12)	-20.26 (20.65)	-190.95** (82.21)	25.04** (10.40)
R-square		0.0001	0.60	0.58	0.51
Standard errors indicated in parentheses * Coefficient significant at 10%, ** at 5%, *** at 1%					

According to our regression result through STATA, we observe that our simple linear regression model appears not to be not a good fit as it has an R-square value of 0.0001, which suggest that the simple model are able to explain 0.1% of the change in fertility rate.

After other variables such as *medinc*, *percchild*, *perchealexp*, and *percedu* are added to Model 2, we have an R-square value of 0.60, suggesting that this model is a better fit for our explanatory variable. Almost all variables except perchealexp and the intercept β_0 are significant at 5% level. Our main independent variable thouslab appear to have a slightly positive coefficient with significant level of 1%, it is different from our hypothesis which assumes a negative relationship with all other variables held constant.

The results for model 3 are similar compared to model 2, but with a 5% significant intercept. The variable thouslab still has a slight positive coefficient at the 99% confidence level, contrary to our hypothesis. In model 4 we drop more variables and the coefficient of thouslab drops further to 0.1172 but still seems to be significant at 1% level. The only other independent variable *percedu* has a negative coefficient with the same significance level of 1%. Although the R-square decreases from 0.60 to 0.51 compared to model 2, the change is still relatively small and we will further explain the effect of removing the three variables from model 2 by F-test in the extended section.

V. Extension

F-Test:

Based on our estimation results earlier, *medinc* and *percchild* are only significant at 95% confidence level but not at 99%, while *perchealexp* is not significant at 90% confidence level. Thus, we want to test the joint significance of these three variables.

We considered MLR1 as our unrestricted model. (See Appendix MLR1 for STATA regression)

$$fert = \beta_0 + \beta_1(thouslab) + \beta_2(medinc) + \beta_3(percchild) + \beta_4(percchealexp) + \beta_5(percedu) + u$$

And we drop *medinc*, *percchild,* and *perchealexp* in our restricted model. (See Appendix MLR3)

$$fert = \beta_0 + \beta_1(thouslab) + \beta_5(percedu) + u$$

From STATA we got the Sum of Squared Residuals (SSR) for our unrestricted model is 584.279603, and 706.198469 for our restricted model. We considered the following hypotheses:

$$H_0 : \beta_2 = \beta_3 = \beta_4 = 0$$
$$H_1 : \text{null hypothesis is false}$$

We then got our F-statistics as:

$$F = \frac{(706.198469-584.279603)/3}{584.279603/44} = 3.06042408261 \approx 3.060$$

An F-value of 3.060 is smaller than the critical value of 3.15, thus we conclude at 95% confidence level that *medinc*, *percchild,* and *perchealexp* have no effect on *fert* after $thouslab$ and $percedu$ have been controlled for and therefore can be considered to be excluded from the model.

Besides, we detect relatively strong collinearity between *thouslab* and *percchild* and between *medinc* and *peredu*. We decide to do an F-test in order to find the joint-significance of these two groups of variables to see if they are jointly insignificant. The two F-test below would further the pragmatic economic applications of our study.

We have one of the restricted models not including *thouslab* or *percchild* as below.

$$fert = \beta_0 + \beta_2(medinc) + \beta_4(percchealexp) + \beta_5(percedu) + u$$

From STATA (see MLR2.1 in appendix for regression model) we got the Sum of Squared Residuals (SSR) for our unrestricted model is 584.279603, and 997.992057 for our restricted model. We considered the following hypotheses:

$$H_0 : \beta_1 = \beta_3 = 0$$
$$H_1 : \textit{null hypothesis is false}$$

We then got our F-statistics as:

$$F = \frac{(997.992057-584.279603)/2}{584.279603/44} = 15.5776000758 \approx 15.578$$

An F-value of 15.578 is larger than the critical value of 3.15, which means the *thouslab* and *percchild* are jointly significant at 95% confidence level.

For our second F-Test, we have our restricted model not including *medinc* and *peredu.*

$$fert = \beta_0 + \beta_1(thouslab) + \beta_3(percchild) + \beta_4(percchealexp) + u$$

From STATA (see appendix MLR1.2 for regression) we got the Sum of Squared Residuals (SSR) for our unrestricted model is 584.279603, and 1208.11124 for our restricted model. We considered the following hypotheses:

$$H_0 : \beta_2 = \beta_5 = 0$$
$$H_1 : \textit{null hypothesis is false}$$

We then got our F-statistics as:

$$F = \frac{(1208.11124-584.279603)/2}{584.279603/44} = 9.11998640085 \approx 9.120$$

An F-value of 9.120 is larger than the critical value of 3.15, which means the *thouslab* and *percchild* are jointly related at 95% confidence level.

Thus we don't consider removing the two pairs of variables with highest correlations, *thouslab* and *percchild* and *medinc* and *peredu,* in our model, although their high correlation might impose inaccuracy to some extent in our model.

Dummy

We note that a state's law on abortion can also affect the fertility rate, and to take this into account, we decided to add a dummy variable for whether a state bans abortion completely. As a restriction on abortion would prevent a woman from terminating her pregnancy, we expect a state that fully bans abortion holding all other variables constant may have a higher fertility rate compared to states where abortion is allowed. Information on abortion laws was collected from the New York Times, most recently updated on November 14th of 2022. Specifically for our binary dummy variable abortion, we define abortion = 0 in states where abortion is fully banned and abortion = 1 in all other cases, including those states with legal or only partially legal abortion laws. With this new dummy variable, we come up with our model 5 (see appendix MLR4 for STATA regression):

$$fert = \beta_0 + \delta_0(abortion) + \beta_1(thouslab) + \beta_2(medinc) + \beta_3(percchild) + \beta_4(percchealexp) + \beta_5(percedu) + u$$

Table 5 – Regression Models Summary

Dependent Variable: Growth Rate			
		Model 2	Model 5
Independent Variables	*abortion*		-1.9909 (1.4288)
	thouslab	0.1603*** (0.0293)	0.1577*** (0.0290)
	medinc	0.0003** (0.0001)	0.0002** (0.0001)
	percchild	0.7851** (0.3115)	0.7256** (0.3112)
	perchealexp	7.23e-8 (7.93e-6)	3.96e-8 (7.85e-6)
	percedu	-1.3110*** (0.1975)	-1.2513*** (0.2001)
intercept		-20.26 (20.65)	-19.2727 (20.4474)
R-square		0.60	0.61

Standard errors indicated in parentheses * Coefficient significant at 10%, ** at 5%, *** at 1%

According to our result above, we may conclude that there is a negative association of our dummy variable *abortion* on fertility rate. However with a t-value of -1.39, p-value of 0.171 and a 95% confidence interval including 0, we may also conclude that abotion is insignificant at 95% level. It is possible that since we only differentiate between fully abortion ban and all other conditions, in states where abortion is partially banned, we may need to consider multiple dummy variables to include all cases in our future research.

VI. Conclusion

As shown in the linear regression model, Model 1, there is a slight positive linear relationship between the labor force participation rate of females over the age of 16, *thouslab*, and the fertility rate. In this model, the main and only independent variable being tested against the main dependent variable indicating the fertility rate is women's employment, and a 0.172 coefficient is returned for the dependent variable. According to the Robustness test of Model 1, our null hypothesis that female employment does not affect fertility rate fails to be rejected at any significance level as the t-value being 0.08, is unacceptable near zero. As visible in the STATA output, this model produces an R^2 value of 0.0001 that is surprisingly low, meaning the *thouslab* variable explains at the most of 0.1% of the variability in their fertility rate around its mean. Given this low value, it is not a stretch at all to conclude that *thouslab* plays an insignificant role in determining fertility - that being said, the *thouslab* variable remains far from the best explanatory variable of fertility rate.

Four extra variables are corporated in our multiple regression models to assist us in better grasping the ceteris paribus effect. Our first multiple regression Model 2 has a greater coefficient for the *thouslab* variable and shows a clearer positive trend. We discovered that one of the newly included control variables, *perchealexp*, does not appear to be worthy of further examination due to its low coefficient, which implies negligible effects on fertility rates and is statistically insignificant. As a result, the perchealexp variable is determined to be removed. Another modification was that in Model 3, we used the logarithm value of *medinc* instead of the variable units in US dollars, which resulted in an excessively large coefficient. By taking the logarithm value of it, we effectively lower the coefficient without changing the original relationship. In fact, after removing *perchealexp* and logarizing *medinc*, Model 3's R^2 value declined only little from that of Model 2, from 0.60 to 0.58, showing that the differences between the two are minor and that using *log(medinc)* is acceptable. We went a step further by removing *log(medinc)* and *percchild*, which have too low t-scores and are only significant at 95% confidence interval but not at 99%. In addition, our F-test revealed that *medinc*, *percchild*, and *perchealexp* together have no statistical significance on fertility rate at 95% confidence level. With a final t-score of 5.84, our null hypothesis is successfully rejected in Model 4 at both the 95% and 99% confidence levels. The reasonably high R2 value of 0.51 indicates that making the additional modifications was the proper decision. For our dummy variables, we considered the state laws and legislation on abortion. However, its t-value from STATA output indicates it is insignificant.

Despite our best efforts to demonstrate the ceteris paribus effect between female employment and fertility rate, our study is still susceptible to bias because some factors are omitted. Factors omitted, for example, women's emotional status may have a temporary impact on childbearing, and such factors are practically hard to qualify and quantify in reality. In addition to the omitted variables, two sets of variables have strong collinearity: medinc and perdedu as well as thouslab and perchild. Our F-test results demonstrate that they are jointly significant, therefore we opt not to remove them whilst high-collinearity variables are likely to bring biases into the model. Some efforts could be made in the future, such as increasing the sample size by considering data from individual counties rather than one entire state, as well as increasing the number of variables to hopefully account for other factors such as regional cultural and religious differences and local public policy on abortion.

References

Chappell, B. (2021, May 5). *U.S. Birthrate Fell By 4% In 2020, Hitting Another Record Low*. NPR. Retrieved November 18, 2022, from https://www.npr.org/2021/05/05/993817146/u-s-birth-rate-fell-by-4-in-2020-hitting-another-record-low

Jaba, E., Chirianu, I., Balan, C., Robu,, I.-B., & Roman, M. D. (2016, March). The analysis of the effect of women's participation in the labor market on fertility in European union countries using welfare state models. *Economic Computation and Economic Cybernetics Studies and Research*, *50*(1), 69-84. ResearchGate. Retrieved November 18, 2022, from https://www.researchgate.net/publication/296994918_The_analysis_of_the_effect_of_women's_participation_in_the_labor_market_on_fertility_in_European_union_countries_using_welfare_state_models

Kearney, M. S., & Levine, P. B. (2022, August). *The Causes and Consequences of Declining US Fertility*. Aspen Economic Strategy Group. Retrieved November 18, 2022, from https://www.economicstrategygroup.org/wp-content/uploads/2022/08/Kearney_Levine_081222.pdf

Livesay, C. (2019, November 25). *In Italy, rising anxiety over falling birth rates*. PBS. Retrieved November 18, 2022, from https://www.pbs.org/newshour/show/in-italy-rising-anxiety-over-falling-birth-rates

Livingston, G. (2019, May 22). *Is U.S. fertility at an all-time low? Two of three measures point to yes*. Pew Research Center. Retrieved November 18, 2022, from https://www.pewresearch.org/fact-tank/2019/05/22/u-s-fertility-rate-explained/

Media, R., Putro, T. S., & Rosyetti, R. (2020). The analysis of factors affecting fertility in Riau Province in 2010-2017. *Indonesian Journal of Economics, Social, and Humanities*, *2*(1), 39-50. https://doi.org/10.31258/ijesh.2.1.39-50

Population Projections for Japan (2017): 2016 to 2065. (2017). National Institute of Population and Social Security Research. Retrieved November 18, 2022, from https://www.ipss.go.jp/pp-zenkoku/e/zenkoku_e2017/pp_zenkoku2017e.asp

Semuels, A. (2017, July 20). *The Mystery of Why Japanese People Are Having So Few Babies*. The Atlantic. Retrieved November 18, 2022, from https://www.theatlantic.com/business/archive/2017/07/japan-mystery-low-birth-rate/534291/

Siegel, C. (2017, March). Female relative wages, household specialization and fertility. *Review of Economic Dynamics*, *24*, 152-174. https://doi.org/10.1016/j.red.2017.01.010

The New York Times. (2022, May 24). *Tracking the states where abortion is now banned*. The New York

Times. Retrieved November 18, 2022, from
https://www.nytimes.com/interactive/2022/us/abortion-laws-roe-v-wade.html

Appendix

States used in research:

Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Hawaii, Idaho, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, Utah, Vermont, Virginia, Washington, West Virginia, Wisconsin, Wyoming.

Stata Output:

MLR1

```
. regress fert thouslab medinc percchild perchealexp percedu

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(5, 44)        =     12.94
       Model |  858.825397         5  171.765079   Prob > F        =    0.0000
    Residual |  584.279603        44  13.2790819   R-squared       =    0.5951
-------------+----------------------------------   Adj R-squared   =    0.5491
       Total |    1443.105        49  29.4511224   Root MSE        =     3.644

------------------------------------------------------------------------------
        fert | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
    thouslab |   16.02945   2.928166     5.47   0.000     10.12811    21.93078
      medinc |   .0002666   .0001071     2.49   0.017     .0000508    .0004824
   percchild |   78.50697   31.15338     2.52   0.015     15.72145    141.2925
 perchealexp |   7.23e-08   7.93e-06     0.01   0.993    -.0000159    .0000161
     percedu |  -131.1037   19.75384    -6.64   0.000    -170.9149   -91.29245
       _cons |  -20.26154   20.65255    -0.98   0.332    -61.88402    21.36093
------------------------------------------------------------------------------
```

MLR1.1

```
. regress fert medinc perchealexp percedu

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =      6.84
       Model |  445.112943         3  148.370981   Prob > F        =    0.0007
    Residual |  997.992057        46  21.6954795   R-squared       =    0.3084
-------------+----------------------------------   Adj R-squared   =    0.2633
       Total |    1443.105        49  29.4511224   Root MSE        =    4.6578

------------------------------------------------------------------------------
        fert | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      medinc |   .0002489   .0001167     2.13   0.038     .0000141    .0004837
 perchealexp |  -9.68e-06   9.38e-06    -1.03   0.308    -.0000286    9.21e-06
     percedu |  -89.49271   23.08162    -3.88   0.000    -135.9536   -43.03181
       _cons |   73.25228   4.371163    16.76   0.000     64.45358    82.05097
------------------------------------------------------------------------------
```

MLR1.2

```
. regress fert thouslab percchild perchealexp

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(3, 46)        =      2.98
       Model |  234.993758         3  78.3312528   Prob > F        =    0.0409
    Residual |  1208.11124        46  26.2632879   R-squared       =    0.1628
-------------+----------------------------------   Adj R-squared   =    0.1082
       Total |    1443.105        49  29.4511224   Root MSE        =    5.1248

------------------------------------------------------------------------------
        fert | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
    thouslab |   8.975402   3.855234     2.33   0.024     1.215218    16.73559
   percchild |   100.6107   36.17315     2.78   0.008     27.79794    173.4235
 perchealexp |  -9.09e-06   .0000101    -0.90   0.375    -.0000295    .0000113
       _cons |  -7.170703   27.17549    -0.26   0.793    -61.87214    47.53073
------------------------------------------------------------------------------
```

MLR3

```
. regress fert thouslab percedu

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(2, 47)        =     24.52
       Model |  736.906531         2  368.453266   Prob > F        =    0.0000
    Residual |  706.198469        47  15.0254993   R-squared       =    0.5106
-------------+----------------------------------   Adj R-squared   =    0.4898
       Total |    1443.105        49  29.4511224   Root MSE        =    3.8763

------------------------------------------------------------------------------
        fert | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
    thouslab |   11.71895   2.270819     5.16   0.000     7.150651    16.28725
     percedu |  -106.6385   15.22916    -7.00   0.000    -137.2756    -76.0014
       _cons |   25.03876   10.39843     2.41   0.020     4.119826     45.9577
------------------------------------------------------------------------------
```

MLR4

```
. regress fert abortion thouslab medinc percchild perchealexp percedu

      Source |       SS           df       MS      Number of obs   =        50
-------------+----------------------------------   F(6, 43)        =     11.33
       Model |  884.066381         6  147.344397   Prob > F        =    0.0000
    Residual |  559.038619        43  13.0008981   R-squared       =    0.6126
-------------+----------------------------------   Adj R-squared   =    0.5586
       Total |    1443.105        49  29.4511224   Root MSE        =    3.6057

------------------------------------------------------------------------------
        fert | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
    abortion |  -1.990904   1.428842    -1.39   0.171    -4.872438    .8906302
    thouslab |   .1577101   .0290326     5.43   0.000     .0991602      .21626
      medinc |   .0002787   .0001063     2.62   0.012     .0000644    .0004931
   percchild |   .7255837   .3111957     2.33   0.024     .0979977     1.35317
 perchealexp |   3.96e-08   7.85e-06     0.01   0.996    -.0000158    .0000159
     percedu |  -1.251261    .200111    -6.25   0.000    -1.654823   -.8476984
       _cons |  -19.27266    20.4474    -0.94   0.351    -60.50878    21.96345
------------------------------------------------------------------------------
```