AN ECONOMETRIC ANALYSIS OF THE RELATIONSHIP BETWEEN FEMALE PARTICIPATION IN THIRD LEVEL EDUCATION AND THE BIRTH RATE

Anne Talbot

Senior Sophister

Studies have shown that women who have a higher education tend to choose to have less children. Anne Talbot tests this theory with a well-considered econometric analysis of the relationship between female participation in third level education and the birth rate, using other interesting variables such ashealth and unemployment.

Introduction

High educational attainment levels are an imperative indicator of the standards of living of a country, but the further development of a country relies on its future generations to improve these standards. In a famous quote, US President Herbert Hoover stated that "children are our most valuable natural resource".

In this project I have completed a regression analysis of the relationship between female participation in third level education and the birth rate. Other independent variables include mothers' average age at first birth, infant mortality rate, the percentage of GDP spent on health by the government, legality of abortion, unemployment rate of women and the marriage rate. This analysis is based on panel data over the time period 2000-2007 for the EU-15 countries[1].

Women's role in society is no longer to be the bread maker, but often the bread winner. Births outside of marriage are more common than ever before, people's attitudes towards women's roles in society, towards education and towards having children have all changed.

Observation and analysis shows that the birth rate has declined drastically in the recent past; in Ireland alone the birth rate has shrunk from 4.1 in 1970 to 1.95 in 2000 (Central Statistics Office). Coinciding with this decrease has been an increase in educational attainments, more notably those of women. As education participation at third and higher levels continues to soar, what effect does this impose on future generations, or lack thereof? Intuition suggests that there is a direct and indirect relationship present. By constructing an econometric analysis of birth rates I aim to identify, analyse and describe this relationship.

Literature Review

There are few studies analysing the effect of education on the child birth rate, and I was unable to find specific studies for third level education. There are however countless studies on the child birth rate and the influences on it. Studies used cross-sectional data, time series data, and panel data. In almost all studies of the child birth rate, it has been proven that a higher level of schooling leads to a lower birth rate and many examples are shown by contrasting the Western World with the Developing World where education is not as widely available. I am more interested in the effects of education on the birth rate in EU-15 countries. According to a study by the National Centre for Health Statistics, Centres for Disease Control and Prevention in America (NCHS, 1997):

"A woman's educational level is the best predictor of how many children she will have". Many studies are done in the United States, and a paper by Mathews and Ventura's (1994) implies that racial background is an extremely influencing factor on the number of children women had. However, in the US, where third level education is estimated to cost $22,000 per child per year[4], many Hispanic Americans and African Americans (the groups with the highest birth rates) were in lower paid jobs than whites (the group with the lowest birth rate), which would imply that fewer Hispanics attended college due to the higher costs and so it is hard to distinguish results.

Studies are also often done at a national level and aim to analyse the effect of migration on the birth rate. Countries such as France and Ireland have seen huge net immigration figures in the early 2000s, whereas many British have moved permanently to Spain and other southern countries. In 2005 alone, Eurostat reported figures of net international immigration to the EU of close to 1.8 million. In Ireland the immigration of women of child bearing age in the boom years lead to an increase in the birth rate.

I aim to show whether tertiary education has or does not have an impact on the birth rate that is worth recognising or disregarding in its own right.

Empirical Approach

$$Birth = \beta_0 + \beta_1 educ + \beta_2 mar + \beta_3 age + \beta_4 health + \beta_5 unemp + \beta_6 infmort + D_1 abort + u$$

Dependent Variable:

Y *(Birth)*: The dependent variable Y represents the live birth rate represented by the total live births in a country for a given year. *A live birth* is defined as "the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of pregnancy, which after such separation breathes or shows any other evidence of life such as beating of the heart, pulsation of the umbilical cord or definite movement of voluntary muscles, whether or not the umbilical cord has been cut or the placenta is attached" (UNECE). The data for this variable was obtained from the UNECE

(United Nations Economic Commission of Europe) and is assumed to be adequate and accurate for the sample years.

Independent Variables X_i:

X_1 *(educ)*: Third level education of women. The first explanatory variable is the percentage of women in third level education, expressed as a percentage of total persons in third level education. This data was obtained from the UNECE Statistical Division Database, so is assumed to be adequate and accurate. I would expect there to be a negative relationship between my dependent variable, the birth rate, and my first independent variable.

X_2 *(mar)*: Marriage. The second independent variable is the number of new marriages in a country per year. Previous studies have shown that women who are married are more likely to have children than those who aren't and so it is an important variable to include. I would expect a strong positive relationship between the number of births and number of marriages. This data was obtained from UNECE Statistical Division Database and is assumed to be adequate and accurate.
A plot of birth, education and marriage rate showed similar trends, implying that they change with each other.

X_3 *(age)*: Age. The third independent variable is the mean age of women at the time of birth of their first child. I expect that age will have a negative effect on the child birth rate since older women are less likely to have multiple children ,due to the natural cycle of life. This data was obtained from the UNECE Statistical Division Database and the CSO, thus is assumed to be adequate and accurate.

X_4 *(health)*: Health. The fourth independent variable represents the percentage of GDP spent on the healthcare system by the government. Since it is reasonable to assume that more money spent on the healthcare system improves the system, I expect that this variable will positively affect my dependent variable. This data was obtained from the Eurostat statistical database and is assumed to be adequate and accurate.

X_5 *(unemp)*: Unemployment. The fifth independent variable represents the number of women unemployed as a percentage of the labour force. The unemployment rate is defined as "the share (in per cent) of the unemployed in the labour force" (UNECE). The data I have collected is from UNECE and is assumed accurate and adequate. I expect unemployment rate to positively affect the birth rate.

X_6 *(infmort)*: Infant mortality Rate. The sixth independent variable represents the infant mortality rate, defined as "the number of deaths of infants under one year of age per 1000 live births in a given year" (UNECE). I expect a high infant mortality rate to negatively affect the birth rate. This data was obtained from the Eurostat statistical database and is assumed to be adequate and accurate. A plot of the Infant mortality rate showed it to be normally distributed.

X_7 *(abort)*: Abortion. I included the seventh variable as a dummy variable since abortion is not legal in all EU-15 countries. I assigned a value of one if abortion was legal in the country for the

given year and a value of zero if abortion was illegal in the country for the given year. I expect the legality of abortion to decrease the birth rate, this is reasonable to assume based on the purpose of abortion. Table one provides the breakdown of abortion as a dummy variable and the number of observations where it is legal (1) and illegal (0).

Table 1

	Overall		Between		Within
abort	Freq	Percent	Freq	Percent	Percent
0	19	15.83	3	20	79.17
1	101	84.17	13	86.67	97.12
Total	120	100	16	106.67	93.75

U: Error term. The Error term represents the unobservable or uncountable variables such as natural ability to have children, or demands for contraception. This must be included as my econometric analysis, based on my independent variables, may be incomplete. This term is assumed to have a mean of zero and to be normally distributed.

Approach

"Panel data (also known as longitudinal or cross-sectional time-series data) is a dataset in which the behaviour of entities are observed across time" (Data Statistical Services).

This project's entity is countries. With panel data there are various techniques used to test the data. I must use the Hausman test to determine whether fixed or random effects describe our data more appropriately. If the data is found to be fixed then we must test for heteroscedasticity and the least square dummy variable model (LSDV) provides a good means to understand fixed effects. If the data is found to be random then the "variation across entities is assumed to be random and uncorrelated with the independent variables included in the model" (Data Statistical Services). We must use the Breusch and Pagan Lagrangian multiplier test. The Pasaran CD test is used to test whether the residuals are correlated across entities since cross-sectional dependence can lead to bias in tests results. Autocorrelation is hard to test for with panel data and so I will not test for autocorrelation in my model.

Description of data set

I collected my data from three main sources. The first being UNECE, the second being Eurostat and the third being CSO Ireland. These data sources provided the most comprehensive and up to date data for my desired countries. I decided to base my model on the EU-15 countries since I feel that they are comparable in terms of living standards, employment levels and GDP per capita

levels. They also have free or relatively cheap third level education fees, along with a fully affiliated ECTS university credit system. My data is base on the time period 2000-2007. This is due to the amount of comparable data available for these years. Initially I had planned to have a wider time period however data availability restrictions impeded this.

I chose my variables carefully and based on resources available. I included the marriage rate as marriage is often associated with families and children. I included the average age of a woman at first birth because a woman's ability to have children is restricted by age so the older a woman is the more likely it is she will have fewer children. I decided to include the percentage of GDP spent on health by the government as in all of my chosen countries, child birth and post natal care are free and so the quality of such service would be reflected in the expenditure, which would impact on the birth rate. I included the unemployment rate of women because women who are long-term employed are often career driven and having children may not be a priority. I included the infant mortality rate since a high mortality rate could be a deterrent to having children. I decided to include abortion since, if abortion is legal and women chose to have abortions, the birth rate decreases.

Table 2 contains the variables, their means, standard deviations, minimum and maximum values. Bir, ed, marr and ump are the logs of birth, educ, mar and unemp respectively. I took the logs of these variables because they varied greatly between countries and over time.

Table 2

Variable	Obs	Mean	Std. Dev.	Min	Max
year	120	2003.5	2.300895	2000	2007
country	120	8	4.338609	1	15
birth	120	274385.2	276267	5303	830288
educ	91	498664.5	446513.2	91158	1352421
mar	108	90533.03	95781.9	1472	305637
age	89	28.26292	0.864237	26.4	30
health	118	8.918644	1.213454	5.8	11.1
unemp	120	7.145	2.893367	1.9	18.4
abort	120	0.841667	0.366584	0	1
infmort	117	4.122222	0.790473	1.8	6.2
bir	120	11.88175	1.293377	8.576028	13.62953
ed	91	12.70731	0.908923	11.42035	14.11741
marr	108	10.73635	1.341818	7.294377	12.63015
ump	120	1.87893	.435091	.6418539	2.912351

Empirical Results

After taking the log of birth, education, unemployment and marriage (which was conducted to improve the functional form) I ran the general multiple regression model in Stata. Since my model consists of panel data I had to ascertain whether I had fixed or random effects in my model; this was done by using a Hausman test and my results showed I had fixed effects. Table 3 contains the results of my fixed effects model.

Table 3

Log Birth			
Log Education	-4.81e-07**	Observations	68
	(2.33-e-07)		
Log Marriage	0.0321***	Adjusted R-squared	0.62
	(0.071)		
Age	-0.001	Sigma_u	0.915
	(0.023)		
Health	0.0226	Sigma_e	0.033
	(0.014)		
Log Unemployment	-0.049**	RHO (fraction of	0.999
	(0.02)	variance due to u_i)	
Abortion	-0.025		
	(0.032)		
Infant Mortality	-0.014		
	(0.015)		
Constant	8.59***		
	(0.915)		
F test (all u_i=0)	F(12,48) = 74.54	Prob>F=0.0000	

Robust standard errors in parenthesis. *significant at 10%; **significant at 5%; ***significant at 1%.

We can see that education, marriage, unemployment and the intercept are significant at the 5% significance level while marriage and the intercept are significant at the 1% significance level. At the 15% significance level health would be significant also. Contrasting to this is age, which has a p-value of almost one and so is highly insignificant in my model. Looking at the signs of the coefficients, we see that education, age, unemployment, abortion and infant mortality negatively affect the birth rate, which concurs with my expectations. The relatively large coefficient of marriage implies that marriage has the greatest influence on the birth rate.

The R-squared value is a goodness of fit test for the model. The overall R-squared value of 0.62 implies that 62% of the variation in the birth rate is explained by this model.

The F-test is "used to test the overall significance of the regression" (Data Statistical Services). Since the p-value for the F-test is <0.05, I can reject the null hypothesis that all of the Beta coefficients are zero and conclude that at least one of my independent variables has a significant effect on the birth rate.

The rho value of 0.99866553, also called the intra-class correlation, represents the "variance not explained by differences across entities" (Data Statistical Services).

Table 4 shows the contrasting coefficients when different techniques are used. Here Dummy represents the regression line where we are controlling for the unobserved heterogeneity. Fixed and Dummy are very similar, reinforcing the decision to use the fixed effects model.

Table 4

Variable	Fixed	Random	Dummy
Education	-0.127	0.255	-0.128
Marriage	0.332	0.578	0.341
Age	0.0004	-0.0191	-0.005
Health	0.0172	0.024	0.024
Employment	-0.009	-0.006	
Abortion	-0.027	-0.006	-0.025
Infant Mortality	-0.0116	-0.01	-0.152
Unemployment			-0.055
Country 2			.42120689
Country 3			-.13495557
Country 4			-.10887974
Country 5			1.9388043
Country 6			1.7276863
Country 7			.27509024
Country 8			-.11282649
Country 9			0
Country 10			0
Country 11			.80203807
Country 12			.26888844
Country 13			1.4287954
Country 14			.29237861
Country 15			1.8710752
Constant	9.8470606	2.8464006	9.3335514

Since my model is a fixed effects model I can ignore the Breusch and Pagan Lagrangian multiplier test. The Pasaran CD (cross-sectional dependence) test is used to test whether the residuals are correlated across countries. After running the test we see that there is no cross-sectional dependence. To test for heterscedasticity, we use the function *xtest3*. We obtain a p-value of 0.000 and since the null hypothesis states we have homoscedasticity we reject the null and conclude that our model is heterscedastic, which must be controlled for using the 'robust' option. Tables 2, 3 and 4 used the robust option.

Extensions

Since my model does not account for all variables that affect the birth rate, such as contraception use, extensions to my model would be possible. The data available was limited and I was unable to find data on a large time scale. Also, contraception data was not available to me. Such data may be available to certain organisations, or may be available in the future and so further analysis of my topic would be appropriate. My model looks at the EU-15 countries. Perhaps extensions to this set could be to include all EU countries, the US or Japan to account for a cultural impact on the birth rate.

Conclusion

With regards to the original data set, the main conclusion is that the most important factor affecting the birth rate is the marriage rate. Due to the reasoning for marriage and the social ideas of marriage, this makes sound and logical sense. Other studies found marriage to have a high impact on the birth rate also, both in developed and less developed countries.

Looking in particular at education, which was my variable of interest, it does, as expected, have a negative effect on the birth rate. The main reasons for this negative effect is due to women who complete third level education having more awareness of contraception and how to use it properly – this may not come from the education received in lectures, but as a result of college life and the availability of information on such matters. A direct and indirect relationship exists.

In various other papers on birth rates, it is often concluded that, while third level education does not always reduce the first-birth rates, "the effect on higher order birth rates was found to be significant" (Demographic-Research), implying that education does reduce the birth rate, but with regards to the number of children women have, not the number of women who have children.

The purpose of this project was to analyse the effect female participation had on the child birth rate and to provide suggestions for this effect, and so, I concur with John Dewey, who famously said "Education is a social process. Education is growth. Education is, not a preparation for life; education is life itself". Not only do we obtain a third level education, we learn valuable life lessons and general knowledge that improve our lives and it is the collective result of female participation in third level education that negatively impacts on the birth rate.

References

1: The EU-15 countries are Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxemburg, Netherlands, Portugal, Spain, Sweden, United Kingdom.

2: $22,000 based on the average fees for one academic year in a college in the USA (America.gov).

America.gov, [Online]. Available at: www.america.gov

Baum, C. F., 2006. An Introduction to Modern Econometrics Using Stata. 1st ed.: Stata Press.

Data and Statistical Services, Princeton University, [Online]. Available at: http://dss.princeton.edu

Demographic-Research, [Online]. Available at: www.demographic-research.org/Volumes/Vol3/3/html/4.html

Gujarati, D., 2004. *Basic Econometrics*. 5th ed.: McGraw-Hill.

Mathews, T.J. and Ventura S. J., 1994. 'Birth and Fertility Rates by Educational Attainment: United States, 1994'. National Centre for Health Statistics, Monthly Vital Statistics Report, 45(10).

UNECE (United Nations Economic Commission for Europe), [Online]. Available at: http://w3.unece.org/pxweb/Dialog/Concepts.pdf

Wooldridge, J. M.,2009. *Introductory Econometrics, A Modern Approach*. 4th ed.: Thomson.

Data Sources

America.gov, [Online]. Available at: www.america.gov

Central Statistics Office (CSO), [Online]. Available at: www.cso.ie

Data and Statistical Services, Princeton University,[Online]. Available at: http://dss.princeton.edu

Demographic-Research, [Online]. Available at: www.demographic-research.org/Volumes/Vol3/3/html/4.html

Europa, [Online]. Available at: www.europa.eu

Eurostat, [Online]. Available at: www.epp.eurostat.ec.europa.eu

Stata, [Online]. Available at: http://www.stata.com/meeting/1nasug/drukker.pdf

UNECE (United Nations Economic Commission for Europe), [Online]. Available at: www.unece.org