Education and fertility: an investigation on Italian families

Aldieri, Luigi and Vinci, Concetto Paolo

University of Naples "Parthenope", Italy, University of Salerno

10 June 2011

EDUCATION AND FERTILITY: AN INVESTIGATION ON ITALIAN FAMILIES [1]

by

LUIGI ALDIERI[♦] AND CONCETTO PAOLO VINCI[♣]

ABSTRACT

In this paper we analyse the correlation between the level of education and the number of children in Italy. We select 10,720 Italian families from the 2004-2007 European Union Statistics on Income and Living Conditions (EU-SILC) dataset. Our dependent variable is represented by the number of children ever born to each respondent. Since the number of children ever born is a count variable, Poisson regression is the suitable statistical procedure used to conduct the empirical analysis. First, we estimate the correlation between the female's education and her number of children, and then we use also partner's education to take into account the family dimension. Furthermore, in the context of fertility, zero observations might be due either to the choice not to have children or to impossibility to become a mother. For this reason, we adopt also a more appropriate tool, that is a Zero-Inflated Poisson regression. From the empirical results, we may observe a significant negative correlation between the level of education and the number of children.

KEYWORDS: Fertility, Human Capital, Education

JEL Classification: I21, J13, J24

[1] The authors kindly thank the Department of Economic Studies "Salvatore Vinci" of the University of Naples "Parthenope" and the European Commission, Eurostat, for the availability of EU-SILC (European Union Statistics on Income and Living Conditions) data. Results, conclusions, views or opinions expressed in this paper are only attributable to the authors.

♦ Corresponding author: Dept. of Economic Studies "Salvatore Vinci", University of Naples "Parthenope" (aldieri@uniparthenope.it), Via Medina 40, 80133 Napoli. Tel. +39 081 5474826

♣ Dept. of Economic and Statistical Sciences (DISES), University of Salerno (cpvinci@unisa.it)

1. INTRODUCTION

In the literature, we may observe many works trying to investigate the correlation between education and the number of children. Indeed, Breierova and Duflo (2004) find that women's education increases the age at marriage and decreases the number of children born before the woman turned 15. Aldieri, Barone and Vinci (2006) study the linkages between fertility decisions and human capital of both males and females, proxied by years of schooling. The authors obtain a trade-off between years of schooling and the probability of having a new child. Black, Devereux and Salvanes (2008) find that increases in compulsory schooling reduce the incidence of teenage childbearing in the U.S. and Norway. Monstad, Propper and Salvanes (2008) study the connection between fertility and education. Their results indicate that increasing education leads to postponement of first births away from teenage motherhood and towards women having their first birth in their 20s as well as for a smaller group up to the age of 35-40. However, they do not find that more education results in more women remaining childless or having fewer children. Osilii and Long (2008) pick out a causal effect of education on fertility in Nigeria and this effect is significantly negative. More recently, Kirdar, Tayfur and Koç (2010) analyse the impact of the extension of compulsory schooling in Turkey from 5 to 8 years on the marriage and fertility behaviour of teenage women. They find that the probability of marriage by age 16 is reduced by 44 percent and the probability of giving birth by age 17 falls by 37 percent. The effects of the education policy on the time until the marriage and first birth persist beyond the completion of compulsory schooling. After a woman is married, schooling does not have an impact on the duration until her first-birth. Aldieri, Barone and Vinci (2010) investigate the role of women's education in the transition towards having a second child in Italy. By implementing an event-history model, they find a negative effect of women's education on the transition rate towards a second child. Aldieri and Vinci (2010) find a negative impact of educational level on the number of children in Italy, by considering the grandparents' education as an instrument of parents' education. McCrary and Royer (2011) use age at school entry policies to identify the effects of female education on fertility and infant health. They find that education does not significantly impact fertility. In particular, women born just before and after the school entry date are equally likely to become mothers and give birth at similar age.

Our aim in this paper is to provide further evidence for the correlation between the number of children and educational level both at female level and at family level.

The outline of the paper is as follows: In section 2 we develop a theoretical model about the relation between the number of children and the educational level both at individual level and at family level. Section 3 describes the data and the estimation procedure. Section 4 discusses the empirical results and Section 5 concludes and gives some suggestions for further research.

2. THE MODEL SET UP

2.1 THE BASIC FRAMEWORK

By modifying Kimura and Yasui (2007) where decisions concerning fertility, occupation and education are taken into account and where all of agents live for three different periods: youth, adult age and finally old age; in the first period each subject take delivery of obligatory education and benefits of a predetermined amount of her parents' time; in the second one decisions concerning fertility and labor supply are taken. Moreover, in this period, each agent evaluates the chance of investment in further education in order to supply labor as a more qualified worker. Finally in the old age people consume what saved in adulthood.

The adult subjects are assumed homogeneous, take delivery of utility from children and consumption in the final period, and maximize the following utility function:

$$U_t = \gamma \ln[n_t] + (1-\gamma)\ln[C_{t+1}] \qquad [1]$$

with n_t and C_{t+1} standing respectively for the number of children, and consumption in, while the parameter γ, assumed $0 < \gamma < 1$, measures the relative weight for children.

The budget constraint concerning childcare, work and education supported by workers differs if we distinguish skilled from unskilled subjects; more precisely we have:

$$C_{t+1} = (1+r_{t+1})(h_0 + h_t)[1 - \tau n_{s,t} - \beta h_t] \qquad [2]$$

for high skilled workers, and:

$$C_{t+1} = (1 + r_{t+1})h_0[1 - \tau n_{u,t}] \qquad [3]$$

for the low skilled ones. h_t represents the higher human capital of people who decide to supply labor as skilled workers, h_0 is the compulsory educational level from the first period, $n_{s,t}$ and $n_{u,t}$ are number of children of skilled and unskilled workers, and finally τ and β symbolize respectively the quotas of the wage bill devoted to a child care and to further education.

The maximization processes of eq. [1], with respect to $n_{s,t}$ and h_t, subject to the budget constraint [2] for the high skilled workers, and with respect to $n_{u,t}$, subject to eq. [3] for the low skilled ones, give the following functions[2]:

$$n_{s,t} = \frac{\gamma(1 - \beta h_t)}{\tau} \qquad [4]$$

$$h_{s,t} = \frac{[1 - \tau n_{s,t} - \beta h_0]}{2\beta} \qquad [5]$$

$$n_{u,t}^{*} = \frac{\gamma}{\tau} \qquad [6].$$

By combining eqs. [4] and [5] we obtain the following equilibrium values:

$$n_{s,t}^{*} = \frac{\gamma}{\tau}\left[\frac{1 + \beta h_0}{(2 - \gamma)}\right] \qquad [7]$$

$$h_{s,t}^{*} = \frac{(1 - \gamma - \beta h_0)}{\beta(2 - \gamma)} \qquad [8][3].$$

PROPOSITION 1 *There is a negative relationship between the number of children and education from one hand because high skilled workers have fewer children than low skilled ones, from the other one because in the better educated workers pool the higher educational level the lower number of children.*

[2] The conditions for a maximum are obviously assumed in this and in the following maximization processes

[3] $1 - \gamma > \beta h_0$ is assumed for : $h_{s,t}^{*} > 0$.

PROOF: *By comparing eqs.* [6] *and* [7]*, from which we may easily derive that:* $n_{s,t}^{*} < n_{u,t}^{*}$*, and from inspection of* [4] *and* [5].

3.2 INTRODUCING A FAMILY CONTEXT

We now consider the case in which the decisions unit is the family, in the sense of a couple, instead of the single individual. In this light choices on both the desired number of children and investments in further education are jointly taken within a family.

Three different types of families may be distinguished:

- a high educated family with both the couple's members investing in further education;
- a mixed one with only one investing in further education;
- a low educated family.

The three different constraints are:

$$C_{t+1} = (1+r_{t+1})\left[\left(h_0 + h_t^m\right) + \left(h_0 + h_t^f\right)\right]\left[1 - \tau n_{s,t} - \beta h_t^m - \beta h_t^f\right] \qquad [9]$$

$$C_{t+1} = (1+r_{t+1})\left[\left(h_0 + h_t^f\right) + (h_0)\right]\left[1 - \tau n_{m,t} - \beta h_t^f\right] \qquad [10]^{4}$$

$$C_{t+1} = (1+r_{t+1})2h_0\left[1 - \tau n_{u,t}\right] \qquad [11]$$

where $n_{s,t}$, $n_{m,t}$ and $n_{u,t}$ represent the desired number of children in the high educated, mixed, and low educated families, while h_t^m and h_t^f stand for the higher level of education for males and females.

In the case of a high educated family, assuming that: $\widehat{h}_t = h_t^m = h_t^f$, and maximizing [1] subject to budget constraint [9] with respect to $n_{s,t}$ and h_t we mat easily obtain:

$$n_{s,t} = \frac{\gamma\left(1 - 2\beta\widehat{h}\right)}{\tau} \qquad [12]$$

$$h_{s,t} = \frac{\left[1 - \tau n_{s,t} - 2\beta h_0\right]}{4\beta} \qquad [13]$$

[4] We consider the case with the females more inclined than males to acquire new skills

and as a result the following equilibrium values:

$$n_{s,t}^{*} = \frac{\gamma}{\tau}\left[\frac{(1+2\beta h_0)}{(2-\gamma)}\right] \quad [14]$$

$$\widehat{h}_t^{*} = \frac{(1-\gamma-2\beta h_0)}{2\beta(2-\gamma)} \qquad [15]^{5}.$$

Shifting our attention to the case of a mixed family the maximization process of $[1]$ subject to $[10]$ gives what follows:

$$n_{m,t} = \frac{\gamma(1-\beta h_t)}{\tau} \quad [16]$$

$$h_t^{f} = \frac{[1-\tau n_{m,t}]-2\beta h_0}{2\beta} \qquad [17]$$

from which:

$$n_{m,t}^{*} = \frac{\gamma}{\tau}\left[\frac{(1+2\beta h_0)}{(2-\gamma)}\right] \quad [18]$$

$$h_t^{f*} = \left[\frac{(1-\gamma-2\beta h_0)}{\beta(2-\gamma)}\right] \quad [19]$$

In the final case of a low educated family we easily derive that:

$$n_{u,t}^{*} = \frac{\gamma}{\tau} \qquad [20].$$

[5] $1-\gamma > 2\beta h_0$ is obviously assumed for: $\widehat{h}_t^{*} > 0$

PROPOSITION 2 *Even in the case of the family as decision unit, we can set up what follows:*

There is a negative relationship between the number of children and education from one hand because high skilled and mixed families have fewer children than low skilled ones, from the other one because there is a negative relation between parents' education and number of children in both educated and mixed educated families.

PROOF. *The first part of Proposition 2 is from the comparison of respectively eqs.* [14] *and* [18] *with eq.* [20] . *As far as concerns the second part, it comes from inspections of eqs* [12] *and* [16].

3. DATASET DESCRIPTION AND MODELLING

Our analysis is based on data taken from the European Union – Statistics on Income and Living Conditions (EU-SILC) over the period 2004-2007. The EU-SILC database provides comparable, cross-sectional and longitudinal multi-dimensional data on income, poverty, social exclusion and living conditions in the European Union. The data are gathered by the member states of the European Union and collected by Eurostat. Labour, education and health information is obtained for persons aged 16 and over. Income is mainly collected at personal level, but some components are included in the 'household' section. The selected sample for the analysis of childbirths considers 10,720 Italian families over the period 2004-2007.

Table 1. Definition of used variables.

Number of children ever born	**Dependent variable:** *NC*
	Independent variables:
Lower education	*LowEd*
Higher education	*HighEd*
Higher educated family (higher education for wife and her husband)	*HighHighEd*
Class of income: <10,000*	*H0*
Class of income: 10,000-19,999	*H1*
Class of income: 20,000-29,999	*H2*
Class of income:30,000-39,999	*H3*
Class of income:40,000-49,999	*H4*
Class of income:>50,000	*H5*
Class of age: 25-29	*A1*
Class of age: 30-34	*A2*
Class of age: 35-39	*A3*
Class of age: 40-44	*A4*
Class of age: >45	*A5*
Work=1 if woman works	*Work*
Permanent=1 if woman's work is permanent	*Permanent*
Owner= 1 if woman's family is owner of house	*Owner*
Experience=duration in years of woman's work	*Experience*
Health=1 if woman's health is very good or good	*Health*
Social transfer=Government transfers for children	*ST*
Urbanization=1 if country is densely populated (>50,000 inhabitants per square kilometer)	*UR*

*Yearly family income has been deflated according to GDP deflators (2005 Euro).

All the variables used in the estimated model are collected in the table 1, while table 2 reports the descriptive statistics of our sample.

Table 2. Descriptive statistics

	Mean	SD
NC	1.24	1.056
LowEd	0.86	0.347
HighEd	0.14	0.347
HighHighEd	0.06	0.237
H0	0.26	0.438
H1	0.16	0.367
H2	0.18	0.388
H3	0.16	0.363
H4	0.10	0.300
H5	0.14	0.346
A1	0.03	0.174
A2	0.08	0.265
A3	0.10	0.305
A4	0.12	0.319
A5	0.65	0.477
Work	0.28	0.447
Permanent	0.24	0.428
Owner	0.80	0.400
Experience	19.92	11.665
Health	0.55	0.498
ST	0.39	0.488
UR	0.34	0.473

Our dependent variable is represented by the number of children ever born (*NC*) to each respondent (and to his partner). As can be seen from the table, the mean children ever born for all families is 1.24 with a standard error of 1.056. For our multivariate analysis, we can use the educational levels of the female respondent as a determinant variable and that of her partner in such a manner that we may also identify the high educated families, characterized by a high educational level both for the wife and for her husband. In particular, we use the highest educational level received at the time of the interview. According to Hoem et al. (2001), it would be more useful to use education as a time-variant covariate, but in our case, all respondents have completed their studies before the first child was born. We make a distinction between respondents with a lower education (*LowEd*), with lower education certificate, and higher education (*HighEd*), with a secondary school certificate or a university degree. Results show that most of the people have a lower education, with a mean of 0.86, while the families with both partners high educated indicate a mean of 0.06. We control for age, by using five classes. In terms of other

control variables, the probability of working for females is 0.28, that of having a permanent contract is equal to 24% in our sample, while the mean years of work experience is 19.92. As far as the health conditions are concerned, we consider a dummy which is equal to zero if the woman's conditions are very good or good. In order to control for the economic situation of the households, we take into account six classes of the yearly family income, whether the family is owner of a house and whether the family has received Government transfers for children. Furthermore, in order to control for the geographical variation in fertility tastes and education opportunities in Italy, we include regional dummies in the estimated models. In particular, we consider four regional areas: North-West (reference area), North-East, Center, South. As far as the environmental conditions are concerned, we include a dummy (UR) which is equal to one if the country of residence is densely populated (country has more than 50,000 inhabitants per square kilometer). Finally, we include time dummies in the model to consider a linear trend.

Since the number of children ever born is a count variable, Poisson regression is the statistical procedure to conduct these analysis (Winkelmann and Zimmermann, 1995). The Poisson model is superior to ordinary least squares or other linear models because the distribution of a count variable, such as *NC*, is one that is heavily skewed with a long right tail. The skewed distribution of the *NC* is due to the observed distribution of data with a very low mean, a result which may be attributed to many females desiring few children and few females wanting many children in low fertility countries, such as in Italy. In addition, in the context of fertility, zero observations might be due either to the choice not to have children (i.e. the expected *NC*s are not always 0) or to impossibility to become a mother (i.e. the expected *NC*s are always 0). In order to handle this situation, we also estimate a Zero-Inflated Poisson regression (Winkelmann, 2000).

4. EMPIRICAL RESULTS

By implementing a Poisson regression model we can try to estimate the impact of the females' level of education on the number of children ever born. In Table 3, we estimate four models: a Poisson regression for only female's education (1), a Poisson regression where we consider the correlation between the number of children and high educated families with both partners high educated (2), a Zero-Inflated Poisson regression for only female's education (3) and a Zero-Inflated Poisson regression for high educated families with both partners high educated (4). In particular, for Zero-Inflation Poisson models, we identify two variables which we think may determine the excess zeros: 'health' and 'age'. We may rationalize that the excess zeros in the fertility process may increase if health is not good or if age rises. From the empirical results indicated in Table 3, we can observe a negative correlation between the number of children ever born and educational level, by confirming the results of previous theoretical model. In Poisson model, compared to being low educated females, being high educated ones decreases the expected number of children ever born by 14% ($e^{-0.15}$-1), other aspects being equal. This seems to indicate a substitution effect higher than the income one. This negative effect for Italy confirm that found in Aldieri, Barone, Vinci (2006, 2010) and Aldieri and Vinci (2010). The result is very similar for high educated families with respect to mixed families and low educated families. As far as the control variables are concerned, the yearly family income is positively correlated with the fertility variable by showing an income effect, while experience variable is negatively correlated with the expected number of children, suggesting a problem to reconcile motherhood with a career. Higher is age (A5) lower is the expected number of children, stressing a biological matter to procreate for females starting family later. Having a good health and receiving Government transfers for children affect positively the expected number of children, while living in a country densely populated reduces it. Finally, the Vuong test compares the Zero-Inflated model with an ordinary Poisson regression model. The significant z-test indicates that the Zero-Inflated model is better.

Table 3. Estimation results

	1	2	3	4
Constant	-0.30 (0.058)***	-0.30 (0.058)***	-0.26 (0.063)***	-0.26 (0.063)***
HighEd	-0.15 (0.013)***		-0.13 (0.016)***	
HighHighEd		-0.17 (0.013)***		-0.16 (0.022)***
H1	-0.23 (0.055)***	-0.24 (0.055)***	-0.16 (0.058)***	-0.17 (0.058)***
H2	0.17 (0.052)***	0.16 (0.052)***	0.15 (0.056)***	0.14 (0.056)**
H3	0.41 (0.052)***	0.40 (0.052)***	0.36 (0.056)***	0.35 (0.056)***
H4	0.57 (0.052)***	0.56 (0.052)***	0.50 (0.057)***	0.50 (0.057)***
H5	0.82 (0.052)***	0.81 (0.052)***	0.75 (0.056)***	0.74 (0.056)***
A2	-0.05 (0.026)*	-0.05 (0.026)*	-0.05 (0.030)	-0.05 (0.030)
A3	0.18 (0.023)***	0.17 (0.023)***	0.18 (0.027)***	0.17 (0.027)***
A4	0.25 (0.023)***	0.25 (0.023)***	0.24 (0.027)***	0.25 (0.027)***
A5	-0.08 (0.023)***	-0.07 (0.023)***	-0.05 (0.026)*	-0.04 (0.026)
North-East	0.10 (0.014)***	0.10 (0.014)***	0.10 (0.017)***	0.10 (0.017)***
Center	0.15 (0.014)***	0.15 (0.014)***	0.15 (0.017)***	0.15 (0.017)***
South	0.34 (0.013)***	0.33 (0.013)***	0.32 (0.017)***	0.32 (0.017)***
Work	0.04 (0.021)*	0.03 (0.022)	0.01 (0.026)	0.01 (0.026)
Permanent	-0.01 (0.021)	-0.01 (0.022)	-0.02 (0.027)	-0.01 (0.027)
Owner	-0.02 (0.012)	-0.02 (0.012)	-0.01 (0.015)	-0.01 (0.015)
Experience	-0.01 (0.001)	-0.01 (0.001)	-0.01 (0.001)***	-0.01 (0.001)***
Health	0.13 (0.010)***	0.13 (0.011)***	0.10 (0.013)***	0.10 (0.013)***
ST	0.33 (0.010)***	0.34 (0.010)***	0.34 (0.012)***	0.35 (0.012)***
UR	-0.08 (0.010)***	-0.08 (0.010)***	-0.07 (0.013)***	-0.07 (0.013)***
Pseudo-R^2	0.0749	0.0745		
Vuong-test			Z = 10.34 Pr>Z=0.0000	Z = 10.38 Pr>Z=0.0000

Notes: ***p<0.01, **0.01<p<0.05, *0.05<p<0.10. Robust standard errors are indicated in brackets. Time dummies are included.

5. DISCUSSION AND CONCLUDING REMARKS

The aim of this paper is to investigate from a theoretical and empirical perspective the correlation between the level of education and the number of children both at female decision level and family decision level. For this end, after having described a theoretical where we find a negative link between the level of education and the number of children of the individual and of the family, we use 10,720 Italian families from the 2004-2007 European Union Statistics on Income and Living Conditions (EU-SILC) dataset. Our dependent variable is represented by the number of children ever born to each respondent. Since the number of children ever born is a count variable, Poisson regression is the suitable statistical procedure used to conduct the empirical analysis.

First, we estimate the correlation between the female's education and her number of children, and then we use also partner's education to take into account the family dimension. Furthermore, in the context of fertility, zero observations might be due either to the choice not to have children or to impossibility to become a mother. For this reason, we adopt also a more appropriate tool, that is a Zero-Inflated Poisson regression.

From the empirical results, we can observe a negative correlation between the number of children ever born and educational level, by confirming the results of previous theoretical model. In Poisson model, compared to being low educated females, being high educated ones decreases the expected number of children ever born by 14% ($e^{-0.15}$-1), other aspects being equal. This seems to indicate a substitution effect higher than the income one. The result is very similar for high educated families with respect to mixed families and low educated families. As far as the control variables are concerned, the yearly family income is positively correlated with the fertility variable by showing an income effect, while experience variable is negatively correlated with the expected number of children, suggesting a problem to reconcile motherhood with a career. Higher is age (A5) lower is the expected number of children, stressing a biological matter to procreate for females starting family later. Having a good health and receiving Government transfers for children affect positively the expected number of children, while living in a country densely populated reduces it.

Further investigation is certainty required for a more comprehensive analysis. In particular, it would be interesting to investigate the role of wealth inequality in the human capital accumulation process and in the fertility decision both at individual and at regional perspective, in such a way that it is possible to analyse the relative effects on the local growth rate.

References.

Aldieri L., Barone A. and Vinci C. P. (2006). Human Capital and Fertility decisions in Italy: a Microeconometric Analysis of ECHP Data. *Brussels Economic Review*, Vol. 49 N. 4;

Aldieri L., Barone A. and Vinci C. P. (2010). Education and Second Birth Risks in Italy. *Rivista Internazionale di Scienze Sociali,* forthcoming;

Aldieri L. and Vinci C. P. (2010). An Investigation of the relation between the number of children and education in Italy. MPRA Paper 28534, Library University of Munich, Germany;

Black S. E., Devereux P. J. and Salvanes K. G. (2004). Staying in the Classroom and out of the maternity ward? The effect of compulsory schooling laws on teenage births. *The Economic Journal,* Vol.118 N. 530;

Breierova L. and Duflo E. (2004). The impact of education on fertility and child mortality: do fathers really matter less than mothers? NBER Working Paper 10513;

Hoem J., Prskawetz M. and Neyer G. (2001). Autonomy or conservative adjustment? The effect of public policies and educational attainment on third births in Austria. *Population Studies,* Vol. 55;

Kimura M. and Yasui D. (2007). Occupational choice, education attainment and fertility. *Economics Letters,* Vol. 94;

Kirdar M. G., Tayfur M. D. and Koç Y. (2010). The effect of compulsory school laws on teenage marriage and births in Turkey. Economic Research Forum Working Paper 1035, TERSIAD-Koc, University Economic Research Forum;

McCrary and Royer (2011). The effect of female education on fertility and infant health: Evidence from School Entry Policies using exact date of birth. *American Economic Review,* Vol. 101 N. 1;

Monstad K., Propper C. and Salvanes K. G. (2008). Education and Fertility: Evidence from a Natural Experiment. *Scandinavian Journal of Economics,* Vol. 110, N. 4;

Osilii Una Okonkwo, Long B. T. (2008). Does female schooling reduce fertility? Evidence from Nigeria. *Journal of Development Economics,* Vol. 87;

Winkelmann R. (2000). Econometric Analysis of Count Data. Berlin: Springer-Verlag.

Winkelmann R. and Zimmermann K. F. (1995). Recent developments in count data modelling: Theory and application. *Journal of Economic Surveys,* Vol. 9 N. 1.