The negative educational gradients in Romanian fertility

Cornelia Muresan
Jan M. Hoem (hoem@demogr.mpg.de)
The negative educational gradients in Romanian fertility
by Cornelia Muresan and Jan M. Hoem

Abstract
In Western countries, rates of second and third births typically increase with educational attainment, a feature that usually disappears if unobserved heterogeneity is brought into the event-history analysis. By contrast, second and third birth rates decline as one moves across groups with increasing education in a country like Romania, and the decline becomes greater if one adds unobserved heterogeneity in the analysis. The present paper demonstrates this and shows that the feature is retained in the presence of control variables like age at first birth, period effects, and others.

1. Introduction
In a tradition going back to the seminal book by Gary Becker (1981), economists typically hypothesize that women’s fertility will decrease as their educational attainment increases because childbearing is dominated by the opportunity cost. A long string of demographic contributions have shown the opposite pattern for countries in the West, namely that in the populations studied in recent years, more highly educated women have had higher fertility at parities above 0 than women with less education (Kravdal 2001, Hoem et al. 2001, Kreyenfeld 2002, Oláh 2003, Kreyenfeld and Zabel 2005, Köppen 2006). Conversely a negative educational gradient has been found (as predicted by the Beckerman theory) for some countries in Eastern Europe. (For Albania see Gjonca et al. 2008, Fig. 4; for the Ukraine see Perelli-Harris 2008, Figure 5; for an investigation with particularly clear results, see Koytcheva 2006, who studied patterns for first and second births in Bulgaria in her Chapter 6.) In the present paper we use event-history analysis with a number of control variables to show that negative educational gradients in fertility are present for Romania for all birth orders 1 through 3. We suggest that Romanian educational gradients may be negative precisely for the reasons given in the Beckerian theory (i.e., opportunity cost considerations and quantity-quality trade-off), for we have little reason to suspect a strong income effect, particularly not during the socialist period when education did not play any important role in wage differentials. We speculate that the rapid increase in income dispersion during the subsequent market-oriented period produced a growth in the return to education which was likely to increase the opportunity costs of childbearing particularly strongly for better educated women. We suggest that in addition women with higher educational attainment must have been more successful in avoiding the coercive measures of the pronatalist policies in the socialist period and that they benefitted more when more childbearing-friendly conditions appeared. The lower childbearing intensities for women with higher levels of education appear to be persistent during the whole period we investigate. There does not seem to be much room in the Romanian data for the explanations often given for the positive educational gradients at positive parities in Western data, namely that the more highly educated that pass into motherhood is a select group of particularly childbearing-prone women, something which should show up in a disappearance of the positive educational gradient at parities above 0 when unobserved heterogeneity is brought into the picture. In Romania, the negative educational gradient is strengthened when we control for unobserved personality.
characteristics. This suggests that other selection processes are at work than in Western countries.

In the present paper we first display negative educational gradients in parity-progression rates for cohorts in the Romanian census of 2002. Such rates have the advantage that they reflect the structure of women’s final parity at the end of childbearing, on the other hand they do not reflect explicitly the childbearing dynamics that lead up to the end product and so may be confounded by changing age patterns and developing birth intervals. These issues are kept under control in hazard regression, which is the method that we use for most of our analysis. The bulk of our paper is concerned with the pattern of educational attainment in relative risks of childbearing of birth orders 1, 2, and 3, which is our main focus of interest. In order not to detract attention from this main topic, we relegate our results concerning other covariates to Appendix 2.

2. Trends and patterns in Romanian cohort fertility

From the Romanian census of 2002 we have computed parity progression ratios (PPRs) for selected five-year birth cohorts of women, organized by parity (for parities 0, 1, and 2) and by educational attainment, measured at the end of their childbearing period and at ages where education is normally completed. The group denoted “mid ed” have a high-school diploma or a certificate of vocational qualifications, those denoted “hi ed” have a university-level diploma or more, and the “no ed” group has neither of these qualifications. For more details, see our Appendix 1. Our main results appear in Figure 1, in which a striking feature is that at each of the given parities, the PPR is lower the higher the women’s educational attainment is. The diagram also suggests that at each step in the childbearing progression, fertility has declined over the twenty-some cohorts involved.

Figure 1. Parity progression ratio (PPR), trend across birth cohorts, for each parity and educational level

3. Period fertility: Data and method

In a closer inspection of Romanian fertility trends and patterns we have subjected the data of the first round of the national Generations and Gender Survey (GGS),
collected at the end of 2005, to some scrutiny. The sample consists of 11,986 respondents (5,977 men and 6,009 women) aged from 18 to 79 years at the time of interview, but our interest is focused on the 6,002 women who had proper childbearing and educational histories.

In the GGS, respondents were asked to report the date (month and year) of each event. Consequently, our timing estimates have the precision of one month. We have imputed the middle of the month as the exact time of each birth, our dependent variable. Mothers are counted as exposed to risk only after 9 months from the previous birth. We right-censored observations after 15 years, or when the woman reached age 40, or if she had not had another birth by the date of the interview.

Our key explanatory variable is educational attainment. Unfortunately, we do not really have complete educational histories, as only the highest educational level at the time of interview and the year and month of completion are reported in the first-round GGS. We cannot use the final educational level as a time-constant covariate, because the results would be strongly biased for several reasons, as discussed, for example, by Kravdal (2004) or Hoem and Kreyenfeld (2006) in their papers about anticipatory analysis. We have chosen the same solution as the latter authors, namely to impute a current educational level which is time-varying and whose value changes when the respondent is deemed to complete her final educational level. Beside the three categories of low, middle, and high educational attainment which we explained in Chapter 2 above, we have a separate category for women who are regarded as still under education in the current month. Details can be found in our Appendix 1.

For first births we apply a hazard regression model with an intensity \( h_i(t) \), which for the \( i \)th woman has the form

\[
\ln h_i(t) = y_i(t) + \sum_k \beta_{ki} x_{ik}(t) + z_{i0} (t + 1950).
\]

Here \( t \) is the process time, defined as her age (counted from age 12). The log-baseline intensity \( y_i(t) \) is represented as a linear spline function of \( t \). Furthermore, \( x_{ik}(t) \) is an

---

1 The Romanian Generations and Gender Survey (GGS) was conducted in the framework of the international Generations and Gender Programme (GGP) with the financial support of the United Nation Fund for Population Activities (UNFPA) and the Max Planck Institute for Demographic Research (MPI/DR). More details about the program can be found on the website of the Population Activities Unit of the United Nations Economic Commission for Europe (UNECE PAU), http://www.unece.org/pau/ggp, the coordinator of the whole project, and also on the website of MPI/DR, http://www.demogr.mpg.de.

2 We eliminated seven women: one for having an improper childbearing history, five for not having the information needed to impute educational histories, and one for having a first birth recorded before age 12. Next, For the analysis of second-birth intensities, we retained 4,801 mothers with one child, and for the analysis of third-birth intensities we retained 3,036 mothers of two children. We eliminated from analysis mothers of twins at a previous birth and mothers with an age interval of less than 9 months between their previous and analyzed births. We also eliminated from the higher-order births analysis mothers aged more than 40 years at a previous birth, mothers closer than 9 months to age 40, and mothers whose children were less than 9 months old at the time of interview, as they were not at risk of a supplementary birth during our period of observation.

3 Instead of considering the time to each birth as our dependent variable and the exposure to risk since 9 months after the previous birth, we might have considered the time to each conception (that ended in a live birth) as our dependent variable and the time at the last previous birth as the start of exposure to risk. Both strategies are equally good.

4 An investigation by Zabel (2007) suggests that this solution should be pretty reliable.
indicator of whether she has reached educational level \( k \) at process time \( t \) (including whether she is deemed to be under education at that process time) and \( \exp(\beta) \) is the relative risk of occurrence of a first birth for educational category \( k \) (with \( \beta \neq 0 \) for \( k = 1 \), say). Finally, a second piecewise linear spline \( z_{a2} \) is supposed to catch the effect of calendar time (counted since the beginning of the year 1950). The argument \( c_{i0} \) is the calendar time at which we take individual \( i \) to start to be exposed to the risk of a first birth, i.e., the calendar month in which she turns 12. Thus \( c_{i0} + t - 1950 \) is the calendar month in which the respondent is \( 12 + t \) years old, counted from the beginning of 1950.

For second births we use a corresponding specification

\[
\ln h_2(t) = y_2(t) + \sum_k \beta_k x_{a2}(t) + z_{c2}(c_{i1} + t - 1950) + z_{a2}(a_1 - 18),
\]

where \( z_{a2} \) is another linear spline used to pick up the effect of the age \( a_1 \) at first birth (counted from age 18) and the other items are quite similar to those of the intensity of a first birth. Process time \( t \) is now months since first birth, and \( c_{i1} \) is the calendar month of the first birth. Note that the educational-level binaries vary by birth order because process time \( t \) differs from one birth order to the next.

We extend this specification to third births by using a parallel third-birth intensity of the form

\[
\ln h_3(t) = y_3(t) + \sum_k \beta_k x_{a3}(t) + z_{c3}(c_{i2} + t - 1950) + z_{a3}(a_1 - 18),
\]

where \( c_{i2} \) is the calendar month of the second birth. Note that the spline \( z_{a3} \) is supposed to pick up the effect of the woman’s age at first (not second) birth on the third-birth intensity. We suppose that what is important is the age at entry into motherhood, not the age at last previous birth.

These specifications of \( \{ h_{i\ell}(t); \ell = 1, 2, 3 \} \) are supposed to reflect the impacts of the observed covariates on the intensities when we do not account for unobserved covariates, i.e., when we assume that all female respondents with the same values of \( \{ x_{ii}(t) \} \), \( c_{i\ell} \), \( a_1 \), and so on have the same childbirthing intensities, as specified. Alternatively we may try to account for unobserved differences in the individual characteristics of the respondents. As usual among users of the software aML, we do this by adding an unobserved-heterogeneity factor \( U_{i\ell} \) to the above formula for \( \ln h_{i\ell} \) (for each \( \ell \)) and by assuming that the triple \( \{ U_{i1}, U_{i2}, U_{i3} \} \) is tri-normally distributed with a zero mean \( (EU_{i\ell} = 0) \) for each \( \ell \), some set of variances \( \text{var} U_{i\ell} = \sigma_{i\ell}^2 \), and correlations \( \rho_{\ell\ell'} = \text{correlation}(U_{i\ell}, U_{i\ell'}) \) for \( \ell \neq \ell' \). (The triples are taken to be independent across individuals \( i \).) We would expect each correlation to be positive, for we would assume that a woman who is unusually childbearing-prone, say, for the values that she has on the observed control variables, will be so across all birth orders. Note that we can specify that two of the unobserved-heterogeneity factors are identical by specifying the corresponding correlation to be identically 1 (for all individuals). For other correlations the \( \{ \rho_{\ell\ell'} \} \) are estimated along with the \( \{ \sigma_{i\ell}^2 \} \) when the model is fitted to the data.

We need to give some further attention to the specification of the calendar-time splines \( \{ z_{a\ell} \} \). The defining items are the location of their nodes, which we select to reflect
major political and institutional changes in Romanian society. The fall of the communist regime at the end of 1989 is an important node for our calendar, but there are also other breaks caused by major changes in social policy in recent Romanian history. For instance there have been important changes in abortion legislation, and we want them to be reflected in our calendar splines. Abortion was legalized in 1957 and was widely used to limit family size. In 1967 abortion was suddenly banned, which resulted in a doubling of the number of births in the next three years. Abortion was again legalized in 1989 on the eve of the old socio-political regime. There have also been other changes in family policy. The first forms of financial support for children were introduced in 1956, but the massively pronatalist Romanian policy was only implemented from 1967. The demographic policies of the “golden age” of Ceausescu’s regime lasted between 1967 and 1989. The emerging industry needed an increasing work force, including working women, and the Romanian family policy was set and developed in those years with this in view. Incentives and coercive measures together resulted in more births of all orders as compared to the periods before 1967 or after 1989. Moreover, higher-order birth risks were very sensitive to periodic re-enforcements of pronatalist policies in June 1973 and March 1984. The birth rate declined every time the corps-control weakened and it rose again when measures against illegal abortion measures were strengthened (Mureşan 1996, 2008). After 23 years of the authoritarian communist regime, 1990 marked the start of freedom and of a re-definition of family policies. One year of paid childcare leave for working mothers was introduced in 1990 and this was extended to two years in 1997. Paternity leave was introduced in 1999, childcare benefits for “insured” mothers in 2003, and other incentives have been introduced repeatedly (Mureşan et al. 2008), targeted to increasing the low level of fertility.


For the baseline splines we located the nodes simply so they fit the data suitably after some experimentation, and we used the same practical criterion for the nodes of the splines that represent the effect of age at first birth. For details, see the tabulation in Appendix 2.

4. Main finding: The negative educational gradient in parity-specific fertility

Our main results about the effect of educational attainment on fertility in the Romanian data are highlighted in Table 1. The columns for our analysis without unobserved heterogeneity show that the educational gradient is negative for all the birth orders we consider in our study, and most strongly so for third births. The effects do not really change radically when we introduce heterogeneity, except that the negative gradient becomes steeper at each birth order. It is as if the effect of unobserved factors is revealed more strongly when we account for them explicitly (without specifying what they represent in terms of personality characteristics). We do not see much trace of the selectivity of entry into motherhood for the more highly educated that is such a prominent trait in current explanations of the positive educational gradient in countries in the West.
Table 1. Relative risk of childbearing by educational status, First, second, and third births

<table>
<thead>
<tr>
<th></th>
<th>Without unobserved heterogeneity</th>
<th>With unobserved heterogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First birth</td>
<td>Second birth</td>
</tr>
<tr>
<td>Lo ed</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mid ed</td>
<td>0.75</td>
<td>0.75</td>
</tr>
<tr>
<td>Hi ed</td>
<td>0.67</td>
<td>0.65</td>
</tr>
<tr>
<td>In ed</td>
<td>0.28</td>
<td>0.63</td>
</tr>
</tbody>
</table>

To get this far we have experimented some with the specification of the distribution of the heterogeneity triple \(\{U_{i1}, U_{i2}, U_{i3}\}\), essentially by working on their correlation matrix. In our preferred specification we find a positive but negligible correlation between \(U_{i1}\) and \(U_{i2}\) (we get \(\hat{\rho}_{12} = 0.15\)) and no determinable correlation between \(U_{i3}\) and the two other heterogeneity factors; in fact we cannot get convergence of the iterative estimation process that aML uses if we try to introduce a \(\rho_{13}\) or a \(\rho_{23}\). The operational consequence is that we take both of these correlations to be zero. Thus, we cannot verify our hypothesis that unobserved characteristics that would make a woman particularly childbearing-prone, would manifest themselves across all parities. We are surprised by this finding and can only suggest that one keeps an eye on this feature in future fertility analyses in Romania and other East European countries.

Another striking feature of Table 1 is the high fertility of mothers who are under education, both for the transition to the second and to the third birth. It is as if being under education is not much of an obstacle to further childbearing for women who already have entered motherhood. We suspect that a finding of this nature may be connected to a weakness in the determination of periods in which a respondent is under education, a weakness that our data share with all data from the first round of the Gender and Generation surveys. It would have been better to collect proper educational histories than to impute them, as we have been forced to do. If this is the case, then we will find out when data from the second-wave GGS surveys become available, for they are supposed to pay more attention to individual educational histories. Meanwhile, we offer the following observations in favor of the possibility that the Romanian educational system really has made continued childbearing easier than it would have been otherwise.  

\[5\] Romania may actually share this feature with other countries. In fact, using imputations similar to ours for the time-varying educational-status covariate, Köppen (2006: 317-318) finds that enrolled West German women only have a 23% to 33% lower second-birth risk than women who have completed their education. She also finds that French mothers enrolled in education only have a 14% lower second-birth risk than highly educated women with a completed education and an even higher risk (24%) compared to women with a medium level of education. For Bulgaria, where the enrolment variable construction benefitted from more detailed information about all interruptions in people’s educational careers, Koytcheva (2006, p. 208) found a lower birth risk (by 42%) for women in education compared to those with a completed education (Koytcheva 2006: 208). But when the enrolment variable was constructed with assumptions similar to ours, she found the enrolled mothers to have only a 32% lower second-birth risk than women who had completed education (Koytcheva 2006: 193).
In Romania, mothers enrolled in education most likely are either in a part-time or a distance-learning type of enrolment, i.e. they attend high-school or vocational training in the evening and/or study through correspondence courses, at weekends, or work toward a tertiary education through distance learning. All of these forms of alternative learning were available to a certain extent in socialist times, and they are even more so in the continuously developing educational environment of a democratic Romania.

The socialist regime had the ambition to provide working people with access to education, and it organized evening education at the level of vocational and high schools for them. In such schools the rhythm of teaching was less demanding and the requirements more relaxed, which was compensated for by a duration of study that was longer than in the usual full-time education. Young parents, forced by the need to leave daytime school early in order to support their families, still had the opportunity to continue their studies in the evening. Less available because of the scarcity of openings was part-time attendance at a university; however, it did exist in the form of correspondence courses. Such possibilities may have facilitated the combination of enrolment and continued childbearing, which may make the high fertility levels in the last row of Table 1 more plausible.

5. Possible explanations of the negative gradients

The negative educational gradient in recent Romanian fertility can possibly be explained by two aspects of the Beckerian theory (beside the specific context of Romania). First, more highly educated working mothers may have risked missing wage increases and may have lost skills during maternity leave and childrearing leave. Second, better educated women may have had a stronger preference for higher-quality children instead of for a higher number of children.

Both of these aspects are sides of rational choice theory, but we believe that in Romania the second aspect is the more relevant. More highly educated women have probably wanted to have children who have a high-level human capital and who are well integrated into the future of a modern society, and therefore their mothers are likely to have invested much more in their children’s education than mothers with a lower educational level. (Note that the latter usually live in rural areas or have grown up with “rural” norms and values.) The more educated women are more likely to have higher ambitions for their children and therefore can be expected to concentrate more energy, time, and money in a lower number of children.

It is more costly to rear a child adequately for a more highly educated woman, not only because of her higher ambition but also because that responsibility is considered more a duty of the mother than of the father in the Romanian society. Even if at the workplace there is considerable gender equality, in families household chores and childrearing are rather regarded as a matter for the woman.6 We extend this theory by arguing that highly educated and therefore generally more ambitious women are likely to feel a harder burden of childrearing, and this may lead them to limit their number of children more strongly than others.

---

6 Hârăguș (2005) has found evidence of a considerable gender imbalance family work in Romania. Incoherence between the levels of gender equity in individually-oriented institutions and family-oriented institutions leaves women facing a difficult choice between work and family, and MacDonald (2000) has argued that low fertility is correlated with the degree of such incoherence.
Our findings concerning the fertility effect of educational attainment is based on data for a long period which covers both socialist and post-socialist regimes. The Beckerian “opportunity cost” hypothesis should work well in a society were wage differentials are large and are connected to educational levels. This was not the case in socialist Romania, and people had an income that did not much reflect their educational level. The introduction of market-economy rules brought more returns to education, and correspondingly higher-opportunity-cost considerations began to lead to a strong role in lowering the birth risk of more educated women, as we have already indicated at the beginning of this paper.

When even more highly educated women have their first birth early, they have little reason to feel any time squeeze in their progression to subsequent births and are likely to have them at a rather leisurely pace. This feature in itself may produce rates of second and third births that are lower for the more highly educated, something which will be reflected in negative educational gradients of childbearing in the manner that we have picked up in our intensity regressions.7

Appendix 1: Construction of the variable “current educational attainment” and considerations concerning exposure time

Given the lack of complete educational histories, what we can do to impute a current educational level from our GGS data is first to separate for each individual the time spent in education from the time after education has been completed, and then to distinguish those who have finished education by educational level attained. Only women who have completed at least lower-secondary education (which is compulsory in Romania) have reported the date they achieved it, so for women without education or with primary education only we have had to impute the date they ended taking education. We set this date as the respondent’s 12th birthday, which we take to be the age when a woman becomes exposed to the risk of giving birth. Every respondent is considered to be enrolled in education all the time until the date of her declared completion of final education. Such a definition of educational enrolment is strongly anticipatory, i.e. it assumes that the respondent did not take breaks between completing different educational levels. Until the late 1990s, when alternatives to daytime higher education started to emerge with some regularity, there were limited opportunities to return to the educational system and thus we can have confidence in our imputation for such periods. By contrast, strong anticipatory bias could affect our findings post 2000, since increasingly more students, especially women, including mothers, have returned to the educational system in order to gain a first degree diploma as is required by the new market-economy professions. Without making further assumptions we cannot distinguish between women currently enrolled in high school, vocational school, or undergraduate studies. Those who were under education at the time of interview are considered to be enrolled full time. After completion, we have assigned the reported educational level to the respondent.

For the sake of simplicity we have grouped the educational levels into just three categories, low, medium, and high, as follows:

- “Low level of education” means no academic qualifications. Respondents who did not report that they had graduated from high school, but who had attended

7 We thank Gerda Neyer for this insight.
primary school, middle school, and high-school (first cycle only, giving a maximum of 10 grades, ISCED levels 0, 1, and 2) are in this category. They have no more than compulsory education.

- “Medium level of education” means high school or a vocational qualification. High school graduates with a “diploma de bacalaureat” and graduates with a lower or an advanced vocational education (ISCED levels 3 and 4) are in this category.
- “High level of education” means at least a university degree. All holders of an academic degree are in this category, as are graduates of short-cycle and long-cycle university programs, and of more advanced tertiary education (ISCED levels 5 and 6).

In total we have four educational statuses, three of them categorized as with a completed education and one additional category which amalgamates the statuses of being currently enrolled no matter at which level of education. We operate with no transitions between education levels, except from the status “in education” into a “low” or a “medium” or a “high” educational level. Transition to motherhood or to a higher parity is possible from each of the four educational statuses without any change in the educational level.

**Appendix 2: Other features of our period analysis**

a) Baseline intensities for first, second, and third births

For completeness we include diagrams for the baseline hazards for each of the three birth orders that we study. See Figures 2A, 2B, and 2C. In each case we include the baseline both for the model without unobserved heterogeneity and for the model with the heterogeneity factors as specified in Section 3. These baselines display the effect of process time in each case. To us they do not contain any surprises.

**Figure 2A. First-birth baseline, by age of mother**
b) Period effects

Figure 3 contains the period effects for each birth order. For second and third births we see clear effects of the variations in public family policy, and in all curves we see a general effect of institutional and economic changes, in particular the fertility decline after the end of communism, as described at the end of Chapter 3 above. The prohibition of abortion in 1967 appears as a clear jump in the period curves for second and third births in that year. We are surprised that there is not really a corresponding jump in the curve for first births. As Mureșan (2008) has shown based on other data, this suggests that women did not use abortion as a means of fertility control before their first birth in any way as extensively as did women who had already entered motherhood.

The curves in Figure 3 reflect the findings in models with no unobserved heterogeneity. The corresponding diagrams for models where unobserved heterogeneity has been included are largely similar. To save space we do not display them here.
c) The effect of age at first birth on the intensities of second and third birth

The impact of age at first birth on second- and third-birth risks appears in Figure 4. The figure shows trivially that the second- ad third-birth intensities decline uniformly as age at first birth increases. These curves come from our models without heterogeneity elements. The corresponding curves from models with heterogeneity factors are not much different and are not displayed here. We believe, however, that the models we have used up to this point are too simple to pick up subtler effects of age at first births on later fertility. Entering motherhood at age 22, say, must have a different social meaning\(^8\) for women with much education (few of whom become mothers at such a tender age) than it has for women with very little education (most of whom actually have their first child before age 22). Conversely, very few women with little education get a first child in their late 20s while this is quite a normal age at entering motherhood for the more highly educated. We plan to pay attention to the interaction between educational attainment and age at first birth in future work on second- and third-birth risks. We expect unobserved heterogeneity to play a more prominent role in that analysis than it has done in the present paper, because educational differentials in attitudes and values would surely be reflected in different ages at first birth.

\(^{8}\) See B. Hoem (1996).
Figure 4. Second- and third-birth intensities, regressor splines by age at first birth

References


