Swedish Institute for Social Research (SOFI)

Stockholm University

WORKING PAPER 4/2004

PARENTAL SEPARATION AND CHILDREN'S EDUCATIONAL ATTAINMENT: A SIBLINGS ANALYSIS ON SWEDISH REGISTER DATA

by

Anders Björklund and Marianne Sundström

Parental separation and children's educational attainment: A siblings analysis on Swedish register data*

by

Anders Björklund and Marianne Sundström

Swedish Institute for Social Research
Stockholm University
SE-10691 Stockholm, Sweden

Anders.Bjorklund@sofi.su.se Marianne.Sundstrom@sofi.su.se

April 26, 2004

Abstract

This article analyzes whether the commonly found negative relationship between parental separation in childhood and educational outcomes is causal or mainly due to selection. We use data on about 100,000 Swedish full biological siblings, born in 1948-63, and perform cross-section and sibling-difference estimations. Outcomes are measured as educational attainment in 1996. Our cross-section analysis show the expected negative and significant relationship, while the relationship is not significant, though precisely estimated, in the sibling-difference analysis. This finding was robust to the sensitivity tests performed and is consistent with selection, rather than causation, being the explanation for the negative relationship.

JEL Classification: J12, I20, I30

Keywords: divorce, child welfare, siblings estimators

* We thank Donna Ginther, Mikael Lindahl, Oddbjorn Raaum and participants at the ESPE2001-meetings for valuable comments and the Swedish Council for Working Life and Social Research as well as the Nordic Social Science Research Councils for financial support. This paper is a revised version of IZA Discusion Paper 643, entitled "Parental separation and children's educational attainment".

1. Introduction

In the last two decades much research attention has been paid to the relationship between childhood family structure and the educational attainment in adulthood. A common finding of many of these studies is that children who experienced a parental separation or grew up with a single parent incur educational disadvantages compared to those who grew up with both biological parents (see e.g. Jonsson and Gähler 1997; McLanahan and Sandefur 1994; for reviews see Amato and Keith 1991 and Cherlin 1999). While recognizing the presence of selection effects, most previous studies have viewed the relationship as partly causal (e.g. Jonsson and Gähler 1997). Recent work has, however, suggested that the negative association is due to selection rather than causation (see e.g. Ginther and Pollak 2003; Winkelmann 2003; Piketty 2003). Clearly, the policy implications will differ depending on whether the mechanism is seen as selective or mainly causal (cf. Cherlin 1999).

The previous literature offers valid arguments for why both causal effects and selection mechanisms could explain the strong negative cross-sectional relationship between parental separation and educational achievement. Jonsson and Gähler (1997), for example, discuss two main types of causal explanations for the link between family dissolution and children's educational careers. One is called the 'crises model' and stresses the emotional upheaval during the separation process, while the other focuses on the loss of one parent, including his/her income, social and human capital as well as his/her time.

The selection explanation of the cross-sectional relationship recognizes that many of the families that break up were worse off already before the separation. Most studies have attempted to take account of (some) pre-divorce characteristics, such as mother's age at birth, that is, factors that increase marital instability. They typically then find a substantial reduction in the educational disadvantage due to parental separation, but that the disadvantage remains (Ermisch and Francesconi 2001a; Fronstin, Greenberg and Robins 2001; Jonsson and Gähler

1997; McLanahan and Sandefur 1994). That is, some of the factors that increase the risk of parental separation are also associated with lower educational attainment for children, implying that at least part of the negative relationship is due to selection. In addition, intact families may not be the appropriate 'counterfactual' for evaluating the effect of parental separation on children's educational outcomes since the alternative for the couples who separate may not be to live happily together, but rather to live in conflict or to separate later. In fact, work by Gähler (1998) and Amato et al. (1995) suggest that children may be worse off by suffering parental conflict than by a separation.

In order to take differences in family background more efficiently into account and to analyze to what extent the association between parental separation and children's educational attainment is due to causation and selection, we use a sibling-difference approach. We use data on about 60,000 Swedes born in 1951-63 and about 50,000 of their **full** biological siblings born in 1948-63, all of whom have lived with both biological parents for at least some part of their childhood. We estimate both level equations and sibling-difference (fixed-effects) models. The latter models, which are our main contribution to this literature, identify the possible causal effect of separation by comparing the educational attainment of at least one older sibling, who had moved away from home and thus was not directly affected by the separation, with that of at least one younger sibling who was. To focus on separations we exclude persons whose parents died or emigrated before they turned 18.

The outline of the paper is as follows. In Section 2 we present the sibling-difference approach to the analysis of family-structure effects on child outcomes and review previous studies that have used this approach. Section 3 presents the data and the sample. In Section 4 we report our findings and the results of our sensitivity analysis. We end by a concluding discussion.

2. Sibling-difference studies of family structure effects

The advantage with a sibling-difference approach to estimating the impact of family structure on subsequent outcomes is that any omitted variable capturing permanent family characteristics, which are shared by all siblings, cancel out of the estimated equation. As is well known, the presence of such variables will lead to bias in the cross-sectional estimates if they are correlated with the outcome of interest. Nonetheless, the sibling-difference approach is also associated with problems. First of all, the approach identifies the estimated effects from a subset of individuals, namely siblings who differ in their experiences of the childhood event of interest. Thus, only the subset of those who had siblings are used for identification. As a consequence, most survey-based household data sets have only a small number of sibling pairs who have such different experiences, which may lead to imprecise estimates. Our large register-based data set should, therefore, be particularly useful for a siblings-approach.

Further, the sibling-difference estimator is impaired by potential biases. In their careful discussion of this estimator, Ermisch and Francesconi (2001b) mention two situations when this is the case. The first is when family structure – in our case a parental separation – is affected by the 'idiosyncratic endowments' of children. If, for example, the younger child is born with a disability that affects educational attainment and also brings about the separation, the estimated coefficient will not capture the causal separation effect. The second case is when any of the parents develop a problem – their example is that the father becomes addicted to alcohol – which does not directly affect the older sibling but does affect the younger one. If such a non-permanent problem causes the divorce and affects the younger sibling more than the older one, the separation coefficient in the sibling-difference model will overstate the separation effect. The authors emphasize that these two sources of bias also plague the cross-sectional estimator, so they conclude that the underlying assumptions of the

sibling-difference model in general are weaker than those underlying the cross-sectional estimator. These should also be weaker in our study than, for example, in the study by Ermisch and Francesconi (2001b) since we compare full siblings who have all lived with both parents for at least part of their childhood while they also compare half siblings.

The sibling-approach has recently been used in a number of related empirical studies, although we are not aware of any study that focuses solely on the impact of parental separation in the way we do. Ermisch and Francesconi (2001b), Case, Lin and McLanahan (2000) and Ginther and Pollak (2003) study the effect of family structure on children's educational attainment as adults. Using UK data Ermisch and Francesconi (2001b) found that having lived with a single (or divorced) mother as a child significantly lowered (by about 14 percent) the child's probability of achieving A-level qualification according to both the cross-section model and the fixed-effects model. The effect was strongest if the child had lived with a single/divorced mother in early childhood. However, they cannot distinguish the effect of a parental separation in early childhood from that of being born outside of a marriage/cohabitation. Case et al. (2000) compare the educational attainment of women's birth and non-birth children. Controlling for the woman's fixed-effects, they find that children raised by a step, adoptive or foster mother obtain significantly less education on average than do the birth children of the same woman.

Ginther and Pollak (2003) compared the educational outcomes of children from traditional nuclear families to those of half siblings from stable blended families where at least one of the siblings lived with both biological parents. They found that, on average, children from traditional nuclear families did better, but within stable blended families there was no significant difference in educational outcomes between the stepchildren and the biological children of both parents. This finding is consistent with both the selection hypothesis and the disruption hypothesis (i.e., the 'crises model').

3. Data and sample

3.1 Data and variables

Our data set is based on a random sample of the Swedish population born in Sweden between 1951 and 1963.[1] We require that these persons lived in Sweden in 1992. The sample size is 93,333 persons. By means of the population registers at Statistics Sweden we identify the **full** siblings—born in 1948-1963—of these persons. The data allow us to distinguish full siblings from half siblings and siblings who are related through adoptions. In our sibling analysis we use only full siblings while indicators on the presence of half siblings are included in the cross-section analysis (see below). During childhood—defined as 0-17 years of age[2] —we observe the sampled persons and their full siblings in the censuses in 1960, 1965, 1970, 1975 and 1980 (there are no missing observations). We also observe whether their biological parents, other adults, or other children (also aged 0-17 years) live in the same household, but we make no distinction between formally married parents and parents who live in a consensual union. It is, in fact, an advantage of our data that they enable us to observe parental separations, and their links with children's outcomes, also among cohabiting parents.

A separation has occurred when we no longer observe the child living with both parents in the census and if the child is under 18. By contrast, defining *age at separation* as a regressor, we assume that the separation occurs in the middle year between the censuses and measure it as the average of the age in the census where the child was observed with both parents and the age in the censuses in which it was not, provided that it occurred before age 18. If no separation occurs (or if it occurs at age 18 or later) the variable takes the value zero, that is, it is an interaction term. Since the earliest census we have access to is the one from 1960, we cannot observe separations occurring before age nine for those born in 1951 (and

[1] Björklund et al. (2002) present the larger data set from which our sample is drawn.
[2] Limiting childhood to ages below 18 should not be a problem since higher education is free in Sweden and college and university students can receive student loans and allowances. Also, in Sweden the age of legal majority is 18.

not before age eight for those born in 1952 etc.), but we investigate how sensitive our results are to this limitation by estimating the models without the cohorts born before 1959 (see Section 4.3).

We impose a few more restrictions on the final sample that we use in the sibling-difference analysis. First, we require that each sibling lived together with both biological parents in the first census in which he or she was observed. In so doing, we only use observations on individuals who have been exposed to the risk of a parental separation. Second, we require that each sampled person and his/her sibling(s) was observed in at least two censuses, but not necessarily the same ones. Therefore, we do not use siblings born before 1948 as observations. Third, we require that both parents were alive and lived in Sweden at the time of the last census in which the child was observed. By applying this restriction, we focus on separations due to family dissolution rather than separations due to death or emigration of a parent. Fourth, we exclude children who had at least one parent born abroad. These additional restrictions reduce the random sample to about 62,000 and the sibling sample to about 50,000, which results in almost 90,000 observations in the sibling-difference analysis.

We measure outcomes in 1996, so the persons in the random sample and their full siblings are 33-48 years of age when they are observed as adults. We follow much of the previous literature by using educational attainment as our outcome measure. The education information stems from the education register that Statistics Sweden holds for research purposes. At this time, the education register had good coverage of the whole population and contained information about both level and field of education. Furthermore, since education of equal length may imply different levels of earnings and to condense the information on educational achievement into one measure, we construct an index of '*earnings-weighted education*' by estimating a log annual-earnings equation on dummies for all levels and fields

of education plus dummies for age and gender for a sample of the adult population. We define the individual's earnings-weighted education as the coefficient on the dummy variable that indicates the individual's level and field of education, see Björklund and Richardson (2001) for further details. An alternative would have been to impute years of schooling from the register information. But as Card (1999, page 1806) argues, years of schooling is less informative in a country like Sweden with multiple education streams.

In both the cross-section analysis and the sibling-difference analysis we include variables that are likely to influence children's educational outcomes and the risk of parental separation. Thus, in addition to age and gender variables we include indicators of whether, subsequent to any separation, the individual lived with their mother or father or with neither parent. Further, in the cross-section analysis we include dummy variables on whether or not the individual had any older half siblings, and whether these had the same mother or the same father as the individual. Also, we include information on the number of older full siblings. We use only indicators of older full siblings and any older half siblings since their presence is independent of the parent's decision to separate during the childhood of our sampled individual, while the presence of any younger full or half siblings is not. These variables will, of course, cancel out in the sibling-difference analysis. Finally, we include indicators of the age of the parents at the individual's birth. In particular, we include a dummy variable, 'teenage mum', which is equal to one if the mother was 18 years or younger when the individual was born and an indicator, 'old dad', of whether the father was 45 years or older.

3.2 Sample characteristics

We present means and frequencies for the random sample and the sibling sample in Table 1. We see, for example, that the fraction of persons who experienced a parental separation increases over birth cohorts. Further, the mean number of full siblings is higher in the siblings sample than in the random sample, which is as expected since persons who have many

siblings are more likely to be drawn into that sample and persons without siblings are only found in the random sample. Consistently, there is a higher frequency of separations in the random sample and the average number of half siblings is also higher.

Turning to Table 2, we observe some interesting differences between individuals who experienced a parental separation and those who did not. For example, in both samples the persons who experienced a separation during childhood more often had a mother who gave birth before she was 19, they less often had a dad who was 45 years or older when they were born and thus, on average, their parents were younger. Further, they had more half siblings and more older half siblings on both the mother's and the father's side. Finally, individuals who experienced a family dissolution had lower earnings-weighted education, on average.

4. Results

4.1 Cross-section estimates

We start by estimating a set of cross-section equations of the relationship between parental separation during childhood and educational attainment, measured as earnings-weighted education in 1996, for our random sample. In line with previous studies, the results (in Table 3a) show that parental separation indeed has a statistically significant negative association with educational attainment in adulthood. The coefficient of about -.045 corresponds to about one year of schooling and also to approximately one quarter of a standard deviation of the education variable. We then investigate how the separation coefficient is affected when we adjust for characteristics of the family of origin. To that end, we include, first, mother's age at childbirth (a dummy for teenage birth and a linear age term), father's age at childbirth (a dummy for having an older father and a linear age term) and indicators for the presence of any older half-siblings on mother's side and on father's (Model 2). We then see that the separation coefficient is reduced by about one-third, but still remains statistically significant. As

expected, the impact of mother's age at birth is positive, while that of having a teenage mother is negative. Father's age at birth has a negative effect, as does having an older dad. In line with Ginther and Pollak (2003), we find that the presence of any older half siblings affects educational outcomes negatively, but interestingly, more negatively for half siblings on mother's side. Since older half siblings on the mother's side are more likely to live with the sampled individual, a possible interpretation of this result is that domestic conflict and competition over mother's time has a greater impact on educational outcomes than competition over economic resources that arise from any child support paid (or otherwise offered) by the father to the half siblings on his side. Next, we include number of older full siblings (Model 3). (As mentioned, the reason for including only older full siblings and older half siblings is that the number of younger full siblings and half siblings is endogenous with respect to parental separation.) The separation coefficient then turns more negative again. This is because the presence of older full siblings increases marital stability while the presence of older half siblings increases the risk of separation..

In Table 3b we report results for models that are extended with age at separation and an indicator whether the child lived with the father or the mother after separation. By construction, these variables are interacted with separation. We find no significant impact of these additional variables, and the main separation coefficient remains basically unaffected. We also experimented with dummy variables for age at separation but still found no effects.

We also estimated the same models for the sibling sample, see Table 3c. The general impression is that the results are quite similar to those for the random sample in Table 3a. All coefficients have the same sign and the differences in their magnitudes are not large. For example, the main separation effect in Model 1 is -.0449 (std. err. 0.0028) for the random sample and -0.0342 (std. err. 0.0037) for the sibling sample.

4.2 Sibling-difference estimates

The results from our sibling-difference analysis of the association between educational outcomes as adult and experience of parental separation during childhood are presented in Table 4. We see that the separation coefficient is strikingly different from the one in the cross-section analysis. The latter was -.0449 (std. err. 0.0028) in Model 1, implying that the **highest** limit of a 95 % confidence interval is -.0396. The sibling-difference estimate is instead positive but close to zero, .0038, in Model 1. Although the standard error is higher (**.**0068**),** a 95 % confidence interval implies a **lower** limit of -.0076. Even this number implies a quite low causal negative effect of separation.

Age at separation does not enter significantly in the sibling difference models either. The point estimate is negative, though, and thus raises the main separation effect somewhat. But as the mean age at separation is about nine years, the average separation effect does not change much across models.

As to the control variables, we find no effect of which parent the child lived with after divorce (Model 4). A person who had a teenage mother did not incur any educational disadvantage compared to a younger sibling(s) who had not. This suggests that the negative and significant cross-section estimate for teenage mother reflects selection rather than causation. Although we find this result interesting, we are also aware of the fact that many teenage births do not show up in our sample due to our restriction that the child must have been observed with both parents in at least one census. Further, a person who had an older dad had a small, but significant, educational advantage over an older sibling.

4.3 Sensitivity analysis

Are there any other possible explanations than selection for the difference between the separation coefficients in the cross-section analysis and the sibling-difference analysis? In the following we investigate some potential flaws in our analysis. First, is it the case that the

coefficients on separation and age at separation are different for those born in 1959 or later than for the older cohorts (for which we could not observe any parental separation before 1960)? No, the cross-section estimates are highly similar (Table 5) and the sibling-difference estimates (Table 6) point, if anything, to even more positive effects of separation. In neither of these models is age at separation statistically significant.

Second, is the discrepancy between the cross-section and the fixed-effect estimates driven by the separation coefficient being more negative and more significant for persons without siblings? No, this is not the case either. Although the separation coefficient is somewhat more negative in Model 1 (Table 7**)** than for the whole random sample (cf. Table 3a), it becomes almost identical when family background characteristics are taken into account (Model 2)..

Third, is it possible that the difference in results is caused by the siblings being too close in age, so that they are all equally (badly) affected by the family crisis that lead to the separation? We investigate this hypothesis by re-estimating the sibling-difference model only on siblings among whom the age difference was five years or more, but still find no significant effect of separation (Table 8).

A fourth possibility is that the discrepancy between the cross-section and the sibling-difference estimates is due to birth-order effects. This could be the case if, on average, the younger siblings, who are the ones mainly affected by a separation, have better educational outcomes. We test this hypothesis through a re-estimation of the sibling-difference model with indicators for birth-order (oldest, youngest) included. The results show that, in fact, the younger siblings do worse and that the separation coefficient is still small and insignificant (Table 9).

Finally, it is possible that the siblings who identify the separation coefficient in the sibling-difference model, that is, those whose parents separated and among whom at least one

had left home at the time of separation, are deviant in a way that drive the result. In order to investigate this possibility we, first, present means and frequencies for these 2,632 persons from 921 families. The characteristics of these individuals (see Appendix) do not differ much from those of the individuals in the random sample and the siblings sample who experienced a parental separation (cf. Table 1), except that the identifying individuals had somewhat more older full siblings, but fewer half siblings. Next, in order to examine the between-family variation among those who identify the coefficient and to eliminate the within-family variation, we randomly draw **one** individual from each of these 921 families and estimate the same set of cross-section models as above for this particular sample. If these estimates of the separation coefficient were closer to the sibling-difference estimates than to our cross-sectional estimates for the random sample, one could argue that the sibling-difference estimates are driven by the particular sample rather than by selection effects. We find that this is not the case. The estimated separation coefficient for Model 1 in Table 10 (-0.0291) is much closer to our cross-section estimate (-0.0449) for Model 1 in Table 3a than to the sibling-difference estimate of 0.0038 in Table 4. Thus, we conclude that our main finding is mainly – although not completely – driven by selection effects that are controlled for by the sibling approach but not by the cross-sectional analysis. The same conclusion follows from Models 2 and 3 in Table 10.

5. Concluding discussion

This paper has used a siblings approach to analyze whether the commonly found negative association between experience of parental separation during childhood and educational outcomes as adult is causal or mainly due to selection. To that end, we used data on about 100,000 Swedes born in 1948-63 who were full biological siblings and had all lived with both biological parents for some period of childhood and performed both cross-sectional and

sibling-differences analyses. Educational outcomes were measured in 1996. The cross-sectional analysis replicated what most previous researchers have found, namely that persons who experienced a parental separation in childhood incur educational disadvantages compared to those who grew up with both biological parents. Our measure of educational attainment as adult reflects the labor-market return to the highest level and field of education attained and the cross sectional gap associated with a parental separation is equivalent to the average rate of return to one year of schooling in Sweden. When we included some observable correlates of marital instability, such as whether or not the mother gave birth as a teenager and presence of older half siblings, the cross-sectional gap decreased by about one-third. However, when we applied a sibling-difference model that also takes account of unobservable characteristics shared by siblings, we found no impact of parental separation. Thus, an older sibling who lived with both parents during his/her childhood did not have an educational advantage over a younger sibling who experienced a separation in childhood. This finding was robust with respect to a number of sensitivity checks.

Our conclusion that there are no causal separation effects contrasts the recent study by Jonsson and Gähler (1997) that also uses Swedish data. We believe the results diverge because the sibling-difference technique controls more efficiently for family background characteristics than their regression approach, which uses controls on observables such as parental occupation and income. Nonetheless, we note one important difference between the two studies. Whereas they use short-run outcomes, namely Grade-Point Average at age 16 and the probability of continuing from primary to upper secondary school, we use educational attainment in adulthood. Perhaps, the distinction between temporary and permanent effects of parental separation is a useful one for future research on this topic.

REFERENCES

AMATO, P.R. and KEITH, B. (1991). Parental divorce and adult well-being: A meta-analysis. *Journal of Marriage and the Family,* 55, 23-41.

AMATO, P.R., LOOMIS, L. S. and BOOTH, A. (1995). Parental divorce, marital conflict, and off-spring well-being during early adulthood. *Social Forces,* 73, 895-915.

BJÖRKLUND, A. and RICHARDSON, K. (2001). The educational attainment of adopted children born abroad–Swedish evidence. *Paper presented to the ESPE-meetings, Athens, June.*

BJÖRKLUND A., ERIKSSON, T., JÄNTTI M., RAAUM, O. and ÖSTERBACKA E. (2002). Brother correlations in earnings in Denmark, Finland, Norway, and Sweden compared to the United States. *Journal of Population Economics*, 15(4), 757-772.

CARD, D. (1999). The causal effect of education on earnings, in O.C. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics*, vol. 3A, Elsevier: North-Holland.

CASE, A., LIN I. and McLANAHAN, S. (2000). Educational attainment in blended families. *Evolution and Human Behavior,* 22(4), Sept. 2001), 269-289.

CHERLIN, A.J. (1999). Going to the Extremes: Family Structure, Children's Well-Being, and Social Science. *Demography,* 36(4), 421-428.

ERMISCH, J.F. and FRANCESCONI, M. (2001a). Family matters: Impacts of family background on educational attainments. *Economica* 68, 137-156.

ERMISCH, J.F. and FRANCESCONI, M. (2001b). Family structure and children's achievements. *Journal of Population Economics* 14, 249-70.

FRONSTIN, P., GREENBERG, D.H. and ROBINS, P. K. (2001). Parental disruption and the labour market performance of children when they reach adulthood. *Journal of Population Economics* 14, 137-172.

GINTHER, D.K. and POLLAK, R.A. (2003). Does family structure affect children's educational outcomes? *Working Paper* No. 9628, Cambridge, MA.

GÄHLER, M. (1998). Self-reported psychological well-being among adult children of divorce in Sweden. *Acta Sociologica,* 41 (3), 209-225.

JONSSON, J.O. and GÄHLER, M. (1997). Family dissolution, family reconstitution, and children's educational careers: Recent evidence for Sweden. *Demography,* 34, 277-293.

McLANAHAN, S. and SANDEFUR, G. (1994). *Growing up with a single parent*, Cambridge: Harvard University Press.

PIKETTY, T. (2003). The impact of divorce on school performance. Evidence from France, 1968-2002, CEPR Discussion Paper 4146.

SANTOW, G. and BRACHER, M. (1999). Explaining trends in teenage childbearing in Sweden. *Studies in Family Planning*, 30, 169-182.

WINKELMANN, R. (2003). Parental separation and well-being of youths. *IZA Discussion paper No. 894*, Bonn: IZA.

Table 1. Means and frequencies of the random sample and the siblings sample. Standard deviations for earnings-weighted education in parentheses.

	The random sample				The siblings			
	Born 1951-54	**Born 1955-58**	**Born 1959-63**	**All**	**Born 1948-52**	**Born 1953-57**	**Born 1958-63**	**All**
Female	0.491	0.491	0.487	0.489	0.489	0.484	0.486	0.486
Separation[a]	0.033	0.066	0.097	0.068	0.018	0.046	0.072	0.047
Age at sep[b]	10.9	8.6	7.4	8.8	12.1	9.9	8.6	9.4
M.'s age at birth	28.5	28.2	27.6	28.1	26.1	27.1	28.7	27.3
Teenage mum	0.021	0.025	0.032	0.027	0.033	0.024	0.013	0.023
Old dad	0.041	0.042	0.038	0.040	0.038	0.027	0.040	0.028
Mother aft. sep.[b]	0.774	0.836	0.822	0.819	0.866	0.905	0.877	0.886
Father aft. sep.[b]	0.111	0.109	0.134	0.124	0.087	0.078	0.105	0.094
No parent aft sep[b]	0.114	0.055	0.043	0.057	0.047	0.016	0.016	0.020
# of full siblings	1.72	1.73	1.68	1.71	2.76	2.59	2.48	2.60
# of older full sib	0.99	1.12	1.46	1.22	0.77	1.27	2.05	1.39
# of half siblings	0.173	0.259	0.307	0.252	0.156	0.227	0.252	0.214
Any older half siblings mum's	0.059	0.078	0.073	0.070	0.054	0.065	0.061	0.060
Any older half siblings dad's	0.033	0.052	0.061	0.050	0.037	0.050	0.056	0.048
Earnings-weight. Education	0.216 (0.176)	0.215 (0.175)	0.218 (0.177)	0.217 (0.176)	0.201 (0.178)	0.203 (0.171)	0.206 (0.174)	0.203 (0.174)
# of observations	18,302	18,880	24,485	61,667	15,375	17,076	17,726	50,177

Notes: [a] We assume that separations occur in between the two censuses. [b] Conditional on separation.

Table 2. Sample characteristics by experience of parental separation. Standard deviations for earnings-weighted education in parentheses.

	The random sample		**The siblings**	
	No separation	Separation	No separation	Separation
Teenage mum	0.024	0.061	0.022	0.037
Old dad	0.041	0.026	0.028	0.023
Mother's age at birth	28.2	25.7	27.4	26.0
Father's age at birth	31.6	29.1	31.0	29.4
# of full siblings	1.73	1.48	2.62	2.36
# older full siblings	1.22	1.19	1.38	1.54
# of half siblings	0.19	1.16	0.17	1.05
Any older half siblings mum's side	0.067	0.117	0.058	0.102
Any older half siblings dad's side	0.046	0.096	0.046	0.098
Mother after separation	0.0	0.82	0.0	0.89
Father after separation	0.0	0.12	0.0	0.09
Neither parent aft. sep.	0.0	0.06	0.0	0.02
Earnings-weighted education	0.220 (0.178)	0.176 (0.150)	0.205 (0.176)	0.173 (0.148)
# of observations	57,461	4,206	47,833	2,344

Table 3a Cross-section regressions. Dependent variable: earnings-weighted education. The random sample. Standard errors in parentheses. N=61,667.

	Model 1	Model 2	Model 3
Constant	0.4367	0.3762	0.6028
	(0.0856)	(0.0850)	(0.0843)
Age	-0.0103	-0.0099	-0.0201
	(0.0044)	(0.0044)	(0.0044)
Age sq./1000	0.1236	0.1139	0.2254
	(0.0568)	(0.0564)	(0.0558)
Female	-0.0109	-0.0110	-0.0101
	(0.0014)	(0.0014)	(0.0014)
Separation	-0.0449	-0.0349	-0.0350
	(0.0028)	(0.0028)	(0.0028)
Mum's age at birth		0.0036	0.0040
		(0.0002)	(0.0002)
Teenage mum		-0.0327	-0.0295
		(0.0046)	(0.0045)
Dad's age at birth		-0.0011	-0.0005
		(0.0002)	(0.0002)
Old dad (>44 at birth)		-0.0225	-0.0211
		(0.0042)	(0.0041)
Any older half sibs, mum		-0.0508	-0.0590
		(0.0028)	(0.0028)
Any older half sibs, dad		-0.0156	-0.0206
		(0.0033)	(0.0032)
Older full sibs			-0.0245
			(0.0006)
Adj R-sq.	0.005	0.020	0.042

Table 3b Cross-section regressions. Dependent variable: earnings-weighted education. The random sample. Robust standard errors in parentheses. N=61,667.

	Model 1’	Model 2’	Model 3’	Model 4’
Constant	0.4363	0.3754	0.6033	0.6027
	(0.0856)	(0.0850)	(0.0842)	(0.0843)
Age	-0.0103	-0.0099	-0.0201	-0.0201
	(0.0044)	(0.0044)	(0.0044)	(0.0043)
Age sq./1000	0.1234	0.1136	0.2255	0.2252
	(0.0568)	(0.0564)	(0.0558)	(0.0559)
Female	-0.0109	-0.0110	-0.0101	-0.0102
	(0.0014)	(0.0014)	(0.0014)	(0.0013)
Separation	-0.0417	-0.0288	-0.0376	-0.0371
	(0.0090)	(0.0089)	(0.0088)	(0.0089)
Age at separation	-0.00036	-0.0007	0.00030	0.00034
	(0.00096)	(0.0096)	(0.00095)	(0.00096)
Mum’s age at birth		0.0036	0.0040	0.0040
		(0.0002)	(0.0002)	(0.0002)
Teenage mum		-0.0328	-0.0295	-0.0295
		(0.0046)	(0.0045)	(0.0045)
Dad’s age at birth		-0.0011	-0.0005	-0.0005
		(0.0002)	(0.0002)	(0.0002)
Old dad (>44 at birth)		-0.0226	-0.0211	-0.0211
		(0.0042)	(0.0041)	(0.0041)
Any older half sibs, mum		-0.0508	-0.0590	-0.0590
		(0.0028)	(0.0028)	(0.0028)
Any older half sibs, dad		-0.0156	-0.0206	-0.0207
		(0.0033)	(0.0032)	(0.0032)
Older full sibs			-0.0245	-0.0245
			(0.0006)	(0.0007)
With father after separation				-0.0037
				(0.0082)
With neither parent after sep.				0.0074
				(0.0116)
Adj R-sq.	0.005	0.020	0.042	0.042

Table 3c. Cross-section regressions. Dependent variable: earnings-weighted education. The siblings of the persons in the random sample. Robust standard errors in parentheses. N=50,177.

	Model 1	Model 2	Model 3
Constant	0.2748	0.1289	0.3587
	(0.0675)	(0.0675)	(0.0666)
Age	-0.0025	0.0007	-0.0086
	(0.0034)	(0.0033)	(0.0033)
Age sq./1000	0.0225	-0.0009	0.0723
	(0.0413)	(0.0413)	(0.0406)
Female	-0.0081	-0.0078	-0.0078
	(0.0016)	(0.0015)	(0.0015)
Separation	-0.0342	-0.0493	-0.0245
	(0.0037)	(0.0127)	(0.0031)
Mum's age at birth		0.0041	0.0049
		(0.0002)	(0.0002)
Teenage mum		-0.0267	-0.0235
		(0.0042)	(0.0042)
Dad's age at birth		-0.0012	-0.0004
		(0.0002)	(0.0002)
Old dad (> 44 at birth)		-0.0303	-0.0280
		(0.0052)	(0.0052)
Any older half sibs, mum		-0.0566	-0.0591
		(0.0026)	(0.0026)
Any older half sibs, dad		-0.0205	-0.0212
		(0.0033)	(0.0032)
Older full sibs			-0.0263
			(0.0006)
Adj R-sq.	0.002	0.019	0.051

Table 4. Sibling-differences. Dependent variable: earnings-weighted education. Robust standard errors in parentheses. N=87,609. # of groups=35,262

	Model 1	Model 2	Model3	Model 4
Constant	-0.1050	-0.1050	-0.1126	-0.1115
	(0.0766)	(0.0592)	(0.0770)	(0.0770)
Age	0.0135	0.0136	0.0137	0.0136
	(0.0038)	(0.0038)	(0.0038)	(0.0038)
Age sq./1000	-0.1389	-0.1403	-0.1388	-0.1382
	(0.0463)	(0.0464)	(0.0464)	(0.0464)
Female	-0.0073	-0.0073	-0.0073	-0.0073
	(0.0016)	(0.0016)	(0.0016)	(0.0016)
Separation	0.0038	0.0137	0.0143	0.0164
	(0.0068)	(0.0176)	(0.0176)	(0.0151)
Age at separation		-0.0009	-0.0009	-0.0009
		(0.0015)	(0.0015)	(0.0015)
Teenage mum			-0.0029	-0.0028
			(0.0046)	(0.0046)
Old dad (>44 at birth)			0.0134	0.0134
			(0.0062)	(0.0062)
With father after separation				-0.0201
				(0.0155)
With neither parent after sep.				0.0012
				(0.0223)
R-sq. within	0.0036	0.0036	0.0037	0.0038

Note: Only full siblings included. Persons without siblings excluded.

Table 5. Cross-section regressions. Dependent variable: earnings-weighted education. (Standard errors in parentheses.) Persons born in 1959 or later. N=24,485.

	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	-0.6311	-0.5370	-0.4848	-0.7026	-0.7319
	(0.8245)	(0.8314)	(0.8234)	(0.8154)	(0.8154)
Age	0.0515	0.0461	0.0389	0.0540	0.0556
	(0.0472)	(0.0476)	(0.0471)	(0.0467)	(0.0467)
Age sq./1000	-0.7696	-0.6930	-0.5976	-0.8430	-0.8662
	(0.6743)	(0.6799)	(0.6734)	(0.6669)	(0.6669)
Female	-0.0092	-0.0092	-0.0094	-0.0094	-0.0094
	(0.0023)	(0.0023)	(0.0022)	(0.0022)	(0.0022)
Separation	-0.0446	-0.0542	-0.0391	-0.0481	-0.0432
	(0.0038)	(0.0115)	(0.0114)	(0.0113)	(0.0115)
Age at separation		0.00114	0.0008	0.0016	0.0013
		(0.00129)	(0.0013)	(0.0013)	(0.0013)
Mum’s age at birth			0.0044	0.0047	0.0047
			(0.0003)	(0.0003)	(0.0003)
Teenage mum			-0.0269	-0.0250	-0.0249
			(0.0067)	(0.0066)	(0.0066)
Dad’s age at birth			-0.0010	-0.0006	-0.0005
			(0.0003)	(0.0003)	(0.0003)
Old dad (>44 at birth)			-0.0284	-0.0236	-0.0236
			(0.0068)	(0.0068)	(0.0068)
Any older half sibs, mum			-0.0516	-0.0603	-0.0603
			(0.0044)	(0.0043)	(0.0043)
Any older half sibs, dad			-0.0171	-0.0225	-0.0225
			(0.0048)	(0.0047)	(0.0047)
Older full sibs				-0.0217	-0.0216
				(0.0010)	(0.0010)
With father after separation					-0.0055
					(0.0106)
With neither parent after sep.					-0.0436
					(0.0177)
Adj R-sq.	0.006	0.006	0.026	0.044	0.045

Table 6. Sibling-differences. Dependent variable: earnings-weighted education. Persons born in 1959 or later. Robust standard errors in parentheses. N=15,170. # Groups = 7,239

	Model 1	Model 2	Model3	Model 4
Constant	-0.4943	-0.5948	-0.6406	-0.6375
	(1.4378)	(1.0199)	(1.0203)	(1.0206)
Age	0.0362	0.0419	0.0440	0.0439
	(0.0826)	(0.0586)	(0.0586)	(0.0586)
Age sq./1000	-0.4591	-0.5379	-0.5625	-0.5599
	(0.1184)	(0.8401)	(0.8402)	(0.8405)
Female	-0.0034	-0.0034	-0.0033	-0.0033
	(0.0043)	(0.0031)	(0.0031)	(0.0031)
Separation	0.0300	0.0724	0.0712	0.0730
	(0.0219)	(0.0453)	(0.0453)	(0.0459)
Age at separation		-0.0039	-0.0038	-0.0038
		(0.0038)	(0.0038)	(0.0038)
Teenage mum			-0.0105	-0.0105
			(0.0108)	(0.0108)
Old dad (>44 at birth)			0.0387	0.0388
			(0.0193)	(0.0192)
With father after separation				-0.0118
				(0.0426)
With neither parent after sep.				0.0036
				(0.0999)
R-sq. within	0.0028	0.0030	0.0036	0.0036

Note: Only full siblings included. Persons without siblings excluded.

Table 7. Cross-section regressions. Dependent variable: earnings-weighted education. (Standard errors in parentheses.) Persons without full sibling. N=7,705.

	Model 1	Model 2
Constant	0.5389	0.4620
	(0.2496)	(0.2472)
Age	-0.0151	-0.0144
	(0.0128)	(0.0127)
Age sq./1000	0.1951	0.1734
	(0.1640)	(0.1628)
Female	-0.0197	-0.0200
	(0.0041)	(0.0041)
Separation	-0.0521	-0.0362
	(0.0067)	(0.0062)
Mum's age at birth	-	0.0031
		(0.0005)
Teenage mum	-	-0.0312
		(0.0108)
Dad's age at birth	-	0.0000
		(0.0005)
Old dad (>44 at birth)	-	-0.0118
		(0.0113)
Any older half sibs, mum	-	-0.0550
		(0.0051)
Any older half sibs, dad	-	-0.0220
		(0.0071)
Adj R-sq.	0.011	0.034

Note: As these persons have no older full siblings, there is no Model 3 in the table.

Table 8. Sibling-differences. Dependent variable: earnings-weighted education. Siblings born at least five years apart. Robust standard errors in parentheses. N=7,453 # of groups=3,237

	Model 1	Model 2	Model 3	Model 4
Constant	0.4457	0.4427	0.4220	0.4240
	(0.3818)	(0.3818)	(0.3812)	(0.3814)
Age	-0.0137	-0.0137	-0.0130	-0.0131
	(0.0202)	(0.0202)	(0.0202)	(0.0202)
Age sq./1000	0.1722	0.1723	0.1663	0.1680
	(0.2669)	(0.2669)	(0.2665)	(0.2666)
Female	-0.0042	-0.0042	-0.0042	-0.0042
	(0.0049)	(0.0049)	(0.0049)	(0.0049)
Separation	-0.0080	0.0164	0.0183	0.0191
	(0.0242)	(0.0503)	(0.0503)	(0.0525)
Age at separation		-0.0022	-0.0023	-0.0024
		(0.0039)	(0.0039)	(0.0039)
Teenage mum			a)	a)
Old dad (>44 at birth)			0.0117	0.0117
			(0.0142)	(0.0142)
With father after sep.				0.0034
				(0.0508)
With neither parent after sep.				-0.0483
				(0.0437)
R-sq. within	0.0007	0.0008	0.0011	0.0012

Note: Only full siblings included. Persons without siblings excluded. a) Too few cases.

Table 9. Sibling-differences. Dependent variable: earnings-weighted education. Indicators for birth-order included. Robust standard errors in parentheses. N=87,609. # of groups=35,262.

	Model 1	Model 2	Model 3	Model 4
Constant	0.0480	0.0442	0.0394	0.0403
	(0.0843)	(0.0845)	(0.0845)	(0.0845)
Age	0.0081	0.0083	0.0084	0.0083
	(0.0040)	(0.0040)	(0.0040)	(0.0040)
Age sq./1000	-0.1007	-0.1024	-0.1018	-0.1012
	(0.0483)	(0.0484)	(0.0484)	(0.0484)
Oldest	0.0147	0.0147	0.0145	0.0145
	(0.0022)	(0.0022)	(0.0022)	(0.0017)
Youngest	-0.0066	-0.0067	-0.0068	-0.0068
	(0.0022)	(0.0022)	(0.0022)	(0.0022)
Female	-0.0073	-0.0073	-0.0073	-0.0073
	(0.0016)	(0.0012)	(0.0016)	(0.0016)
Separation	0.0040	0.0199	0.0198	0.0218
	(0.0068)	(0.0175)	(0.0175)	(0.0177)
Age at separation		-0.0014	-0.0014	-0.0014
		(0.0015)	(0.0015)	(0.0015)
Teenage mum			-0.0054	-0.0053
			(0.0046)	(0.0046)
Old dad (>44 at birth)			0.0095	0.0095
			(0.0063)	(0.0063)
With father after sep.				-0.0199
				(0.0154)
With neither parent after sep.				0.0019
				(0.0223)
R-sq. within	0.0058	0.0058	0.0059	0.0060

Note: Only full siblings included. Persons without siblings excluded. a) Too few cases.

Table 10. Cross-section regressions. Dependent variable: earnings-weighted education. Standard errors in parentheses. Individuals randomly drawn among families that identify the separation coefficient in the sibling-difference analysis. N=921.

	Model 1	Model 2	Model 3
Constant	-0.5678	-0.4604	-0.1693
	(0.4857)	(0.4856)	(0.4841)
Age	0.0396	0.0336	0.0193
	(0.0240)	(0.0240)	(0.0239)
Age sq./1000	-0.5080	-0.4367	-0.2695
	(0.2944)	(0.2943)	(0.2931)
Female	-0.0115	-0.0110	-0.0086
	(0.0099)	(0.0099)	(0.0098)
Separation	-0.0291	-0.0379	-0.0164
	(0.0117)	(0.0126)	(0.0133)
Mum's age at birth		0.0019	0.0024
		(0.0015)	(0.0015)
Teenage mum		-0.0222	-0.0154
		(0.0276)	(0.0273)
Dad's age at birth		-0.0007	-0.0001
		(0.0012)	(0.0012)
Old dad (>44 at birth)		0.0221	0.0267
		(0.0335)	(0.0331)
Any older half sibs, mum		-0.0269	-0.0317
		(0.0165)	(0.0164)
Any older half sibs, dad		-0.0503	-0.0495
		(0.0217)	(0.0215)
Older full sibs			-0.0215
			(0.0046)
Adj R-sq.	0.008	0.015	0.037

Appendix. Means and frequencies for the individuals in families that identify the separation coefficient in the sibling-difference analysis. Standard deviations for earnings-weighted education in parentheses.

Female	0.489
Separation	0.469
Age at separation	5.11
Teenage mum	0.037
Old dad	0.030
Mother's age at birth	26.3
Father's age at birth	30.1
Mother after separation	0.81
Father after separation	0.15
No parent after separation	0.04
# of full siblings	2.90
# older full siblings	1.55
# of half siblings	0.62
Any older half siblings mum's side	0.110
Any older half siblings dad's side	0.061
Earnings-weighted Education	0.170 (0.150)
# of observations	2,632
# of groups	921