The Impact of the Number of Siblings on Men's Adult Earnings: Evidence from Finland, Sweden and the United States

Anders Björklund
Swedish Institute for Social Research, Stockholm University
S - 10691 Stockholm, Sweden
Tel: +46 8 163452, fax: +46 8 154679
anders@sofi.su.se

Markus Jäntti
Department of Economics, Åbo Akademi University
SF - 20500 Turku, Finland
mjantti@abo.fi

September 22 1998

Abstract

We study the relationship between earnings and the number of siblings using representative data from Finland, Sweden and the United States. Regressions that only control for age show that persons with seven or more siblings can expect to earn at least 15 -20 percent less than persons with only a few siblings in Finland and Sweden, and 30-35 percent less in the United States. After controlling for indictors of the parents' budget constraints, there remains a negative siblings effect, at least for the United States. The U.S. results are consistent with a modified version of the hypothesis of a "child-quantity child-quality trade-off".

Acknowledgements

Financial support from The Bank of Sweden Tercentenary Foundation, the Swedish Council for Research in the Humanities, and Swedish Council for Social Research. Early versions were presented at the EALE-conference in Warzaw, September 1994 and at the ESPE-meeting in Lisbon, June 1995. We thank Jan Hoem and his colleagues at the Demography Unit of Stockholm University for useful comments.

1. Introduction

In both the labour economics, psychological and sociological literature, it has for quite a long time been recognised that there is a negative relationship between a person's number of siblings and her performance in the labour market. In their recent article on children's prospects in the Journal of Economic Perspectives, Robert Haveman & Barbara Wolfe (1993) argue that the number of siblings is one of the most important childhood factors that is correlated with well-being in adulthood.

The existence of such a relationship is, in our view, important for at least two reasons. First, from a public policy perspective, if it can be ascertained that the children of such families on average will perform much worse than other children when they enter the labour market as adults, there is an argument for treating families with many children as a target for redistributional measures. Second, from a theoretical point of view, in Gary Becker's influential approach to the economics of the family (see e.g. Becker, 1991) it has been hypothesised that parents face a trade-off between "child quantity and child quality" when making plans about family size and investments in children. The reason for such a trade-off is that parents are restricted by both financial and time budget constraints. Additional children will therefore force the parents to dilute their limited resources on more individuals. Hence, parental investments in a child will, all else equal, be lower the more siblings there are.

These two aspects of the siblings effect on labour market performance might be closely interrelated. Nonetheless, in our views they must be clearly distinguished and they require different types of analyses. In order to describe the distributional profile of policies that are targeted only on the number of children - such as the universal child allowances in Finland and Sweden - it is the unconditional relationship between the number of siblings and the labour market performance as an adult that is relevant and will be used. In a regression framework, there is , however, a good reason to control for age to take the likely correlation between age and earnings into account.[1] The hypothesis of a quality-quantity trade-off, on the other hand, says that parents with given budget constraints face the trade-off. Therefore, it is crucial to control for parents' budget constraints in order to test this hypothesis in an appropriate way.

Our aim with this study is to investigate these two types of relationships between the number of siblings and labour market performance on data from three countries, namely Finland, Sweden and the USA. We confine ourselves to men[2] and to earnings (and hourly

---

[1] Age and the number of siblings are likely to be strongly correlated, because of the long run trends in fertility. Therefore a bias would be introduced if age is not controlled for.

[2] Our limitation to men does not mean that the issues are less interesting for women. On the contrary, we believe that studies of the impact of family background on performance as an adult can provide crucial insights into the mechanisms that cause gender differentials in the labor market. Including women into this study would, however, create methodological problems like sample-selection since many women do not participate in the labour market. Our choice to use data from the 1960s, when even Sweden had low participation rates for women, aggravates this problem.

wages) as the indicator of labour market performance. The advantage with an international comparative perspective on these issues is rather obvious. The first two countries are generally considered as advanced welfare states in the sense that taxes and transfers are high and contribute to a very equal distribution of disposable income in an international perspective (see Anthony Atkinson et al., 1995). Furthermore, most schools are public and free of tuition, a policy that has been motivated by the desire to reduce the impact of parents' economic resources on the opportunity of children. The United States, on the other hand, is the typical "outlier" among Western countries with high inequality of income both before and after taxes and transfers. The school system relies to a much larger extent on private schools with tuition, so there is much more room for parents' economic resources to influence the educational outcomes of children. This in turn implies that the effect of the number of siblings on educational outcomes is potentially much higher.

This international comparative approach is - as far as we know - a novelty with our study. We have instead decided to make some other limitations of our study. For example, we ignore completely the problems of birth order, spacing of children and the sex composition of the siblings. In particular, the impact of the birth order - whether it matters to be the first, the middle or the last child - has been investigated in the previous literature.

We continue the paper with a presentation of our analytical framework in section 2 and a short survey of the previous literature in section 3. We particularly examine whether the previous studies have answered the two issues that motivate our study. The data sets to be used are described in section 4. Our empirical findings are presented in section 5 and a brief summary and discussion of the results concludes the paper.

2. Unconditional and conditional effects - a simple economic framework

Suppose that fertility decisions are intentional and can be understood by a simple model of microeconomic behaviour. The hypothesis of a quality-quantity trade-off suggests that parents with given economic resources face a budget constraint like the one in the lower panel of Figure 1. Earnings on the vertical axis is our proxy for quality. The number of children of the parents equals the number of siblings minus one from the point of view of the child; we have any of these variables on the horizontal axis. In order to identify such a curve, we must assume that heterogeneity of preferences provides observations on different points of the curve as illustrated in the figure.

In most data sets for practical use, there will also be variation in parents budgets constraints. We illustrate two budget constraints in the upper panel of Figure 1, one for a poor family and one for a rich family. Obviously, such differences in economic resources will introduce income effects on fertility. (Again, our discussion follows a micro-economic

Figure 1.

approach.) If children is an inferior good, as illustrated in the figure, there will be another reason for a negative relationship between the number of children (or siblings) and the quality (earnings) of these children. Therefore, the unconditional relationship between the number of children (siblings) and quality (earnings) will consist of two types of effects: the trade-off effect for given economic resources and the income effect. Because the latter can be expected to be negative, we have reasons to believe that the unconditional curve is more negatively sloped than the conditional curve.

3. Previous studies – a short survey

The literature on the relationship between childhood conditions and performance as an adult seems to be enormous, in particular in the United States. We therefore confine ourselves to a short presentation of the studies that emphasise the impact of the number of siblings. Even among such studies our survey will be far from complete.

An early study by R.B. Zajonc & Gregory Markus (1975), which has received much attention, showed a negative siblings effect on the intellectual capacity of 19-year olds in the Netherlands. They used a very large sample size but it is not clear from the paper what control variables that were used.

Frank Stafford (1987) studied the impact on school performance as measured by the ratings of teachers. His sample was quite small (n=77), but it had the great virtue of containing very detailed information about the family during the pre-school years of the children (1975-76) and the performance of the children at school six years later (1981-82). He found negative effects of closely spaced siblings after having controlled for quite many other family characteristics. His controls included family income, mother's education and her predicted wage, as well as various time-use variables like care time, time for helping and teaching, and time for market work.

Eric Hanushek (1992) also examined the impact of the number of siblings on school performance. He used a much larger sample (n around 1700) that originated from the Gary Income Maintenance Experiment and that had been merged with information about the school achievement of the children. The (log of) the number of children in the family had a significant negative effect on all the indicators of achievement that he used. In these regressions a measure of "permanent" income - average income over a five-year period - was included as a control.

Judith Blake (1989) has, in her comprehensive book on family size and achievement, analysed the impact of siblings on both the level of schooling and achievement at school. She used a number of data sources. A general result is that she found negative sibling-effects on both these dimensions of schooling after having controlled for parents' education and occupation. She did not have income among her independent variables.

Daniel Kessler (1991) has done the only U.S. study we are aware of that employs earnings as the dependent variable. He uses a very young sample from the National Longitudinal Survey of Labor Market Experience of Youth (NLSY, so the persons are not older than 28 years. He does not find any family size effects on the level of hourly wages.

Butcher & Case (1994) emphasise the importance of the sibling sex composition, but report results that show strong effects of the number of siblings on educational attainment of both men and women. They use the PSID, the CPS and the NLSW data sets. The negative effect of the number of siblings - that is quadratic so that one sibling is "better" than none - holds also when a number of parental background variables are controlled for.

Turning to Sweden, Bertil Holmlund (1985) used the Level of Living Surveys - as we will do - and examined *inter alia* the determinants of educational attainment as measured by both years of schooling and educational level. He found that the (log of the) number of siblings had strong negative effects on both indicators of schooling in regressions where both parents' education, father's occupation and type of locality during childhood were included as controls. Further, Marja Walldén (1990, 1992) used a large sample of persons born in the Stockholm area in 1953 and found negative sibling effects on both mental capacity and the educational career.

For Finland we do not know of any previous studies on this subject.

4. The data

The Finnish data set originates from a longitudinal data set compiled from the bidecennial censuses in Finland from 1970 to 1990. A household panel was constructed in the following way. First, we drew a random sample of approximately 60 000 individuals. Second, we collected all individuals who lived in the same households as these 60 000 in 1970. Third, for all persons included in stage 2, we collected all individuals who lived in the same household in 1975. This procedure was repeated in 1980, 1985 and 1990, resulting in a sample of approximately 600 000 individuals (see Eriksson and Jäntti, 1997, for details). Thus, for each person who meets our selection criteria, we can go back and study the persons in the household they lived in 1970 and get information on their family members. For Finland we only estimate earnings equations for 1990.

We use two Swedish data sets. The first is the Swedish Level of Living Survey (SLLS, see Erikson and Åberg, 1987). The basic sample in these surveys was representative for the Swedish population in 1968, and has been updated with immigrants and youth to make it representative at the times of the subsequent interviews. The SLLS does not contain information on parents' earnings, but the information on their education and occupation (as reported by the respondents), is quite detailed. We estimate annual and hourly earnings equations from 1968, 1981 and 1990 (1991 for hourly earnings) in order to see if the effects of siblings have changed over time.

The second Swedish data set is a nationally representative sample of over 12 000 pupils in Swedish schools born in 1948. These pupils took part in tests and interviews done in their schools in the spring of 1961. These interviews provide information about family background. The sample has been followed through school and there is also administrative about their educational level and annual earnings in 1993, i.e. at the age of 45.[3] This data set is a research effort undertaken at the Department of Education and Educational Research at the University of Gothenburg (see Härnqvist and Svensson, 1973, for further details).

For the United States, finally, we use the well-known Panel Study of Income Dynamics(PSID), which is representative of the U.S. population. The PSID provides information on the earnings, education and occupation of both parents. For the United States , we estimate equations for 1967, 1979 and 1987. Only household heads are involved in any sample, since only for these do we have all the relevant information.

The siblings variables have been collected in different ways in the three data sets. In the Finnish data set, for each person who meets the sample characteristics, we calculate the number of children living in the household in 1970. This clearly has limitations. Obviously, children who have left the household before 1970 will not be counted as siblings, neither will children born after 1970 be so counted. We are nor able to distinguish between "real" siblings, half-siblings or children of a step-father or step-mother. Thus, the Finnish results should be viewed with some caution. In the Swedish Level of Living Survey the respondents answer the following questions: *"Do you have, or have you had, any brothers or sisters?"* If the answer is yes, the next question is: *"How many? Count even those who have died at an adult age!"*. The PSID employs similar, but not identical, questions to define the number of siblings: *"Did you have any brothers? How many brothers was that? Did you have any sisters? How many sisters was that?"*. The Swedish Gothenburg data collected siblings information by means of a questionnaire to the schools of the pupils.

To define the number of siblings by means of questions like the ones in the Swedish SLLS or the PSID is not without its problems.[4] For example, there is no distinction made between "full" siblings, with whom the respondents have grown up, and "half-siblings" who grew up in another family than the one of the respondent. It is actually unclear how the later cases are treated in questionnaires like the ones used in these surveys. In an attempt to reduce this problem somewhat, we also used samples for Sweden and the United States where we imposed the sample restriction that the respondent has grown up with both

[3] The information about educational level originates form the educational register (*utbildningsregistret*) of Statistics Sweden, and annual earnings from a register based on employers' reports to tax authorities (*ÅRSYS*)

[4] Actually , we are surprised that this data quality issue is so little discussed in the previous studies.

biological parents until the age of 16. In this sample we believe that there are fewer cases of persons who grew up in divorced families with several "half-siblings" as a consequence.

Turning to the earnings variables, the Finnish data stem from registers, collected for tax purposes. For Sweden, the variable annual earnings also comes from such registers, whereas hourly earnings is constructed from interview questions about mode of pay, amount of pay and weekly hours of work. In the PSID both annual and hourly earnings come from interview questions.

Both the Swedish and the U.S. data sets provide measures of retrospectively self-reported measures of "poverty" during childhood that we use. The Swedish question is formulated as follows: *"Did your family have economic difficulties while you were growing up?"*. The question in the PSID is: *"Were your parents poor when you were growing up, pretty well off, or what?"*.

We work with two basic samples. For both samples we exclude those who grew up abroad, the reason being that we want to examine the impact of childhood conditions in the three countries. In the first set of samples, which are confined to Sweden and the United States, we use the broad age limit 20 to 64 years of age. In the second set of samples we employ measures of the income of the fathers for Finland and the United States. This requirement forces us to reduce the age limit of the samples to 29 to 39 years. All age limits refer to the age in the year in which we observe earnings.

Table 1 shows some descriptive statistics of our main samples. The upper panel that is denoted "Broad samples" contains three samples from three different years for Sweden and the United States. It should be noted that the three samples for each country are partly overlapping because of the longitudinal designs of the SLLS and the PSID. Therefore they can not be treated as independent samples. For Sweden we can see that the number of siblings has declined markedly from 1968 to 1981 and again to 1991, whereas we can not find a similar trend for the United States.

For these broad samples we also have a number of other variables at our disposal that we use as indicators of the parents' financial budget constraints. We use the self-reported poverty measures described above. We also use information on both parents' education and on the occupation of the father. Such variables are likely to be rather strong indicators of the parents' budget constraints.[5] In both the SLLS and the PSID the variables that we use are retrospectively reported by the children. The exact definitions of these variables will be presented in an Appendix that will be available at the request from the authors.

The lower panel of Table 1 shows some descriptive statistics of the “narrow samples“. Except for one of the Swedish samples, these samples are much smaller and this follows from the requirement that we want to use direct - and not retrospectively reported -

[5] We have also considered the relevance of information on the occupation of the mother, including whether she was a housewife or not. However, we decided not to use such variables since they are likely to be functions of the number of siblings, i.e. "endogenous".

information on the incomes of the fathers' and estimate the earnings functions for the sons at adult age. For the United States, the father's income is measured as the average earnings over the five years 1967 to 1971. Consequently, we have a measure of the "long-run" income for the father and we believe that such a measure is a better indicator of the budget-constraint during the period of the life when the "investments" in the children were to be done than income in a single year.[6] For Finland the fathers' income is the average of earnings for 1970 and 1975.

5. Empirical analysis

A. Siblings and earnings

We start by looking at the unconditional effects of the number of siblings for Sweden and the United States. The first regression model of Table 2 uses a linear sibling-variable. For all years and both countries the coefficients are by wide margins significantly different from zero. Further, the U.S. coefficients are between 50 and 80 percent higher than the Swedish ones; around .030 for Sweden and .50 for the United States. By and large the coefficients are stable over time for both countries.

Even though many previous studies have employed a linear sibling specification, it can be questioned whether the difference between having one sibling and being the only child is the same as having six instead of five siblings. To allow for such non-linear impacts we have also run regressions with dummy-variables for 1, 2, 3, 4, 5, 6, and 7+ siblings. Regression 2 shows the results for such a specification, again with only age and age squared as controls. The precision for each individual dummy turns out to be rather low, but the explanatory value (the adjusted R-square) raises in four cases of six and is constant in the other two. The results are displayed in Figure 2 and suggest that non-linearities are quite likely. For Sweden the coefficients for one and two siblings are positive for 1990 (although barely significant), and for the United States the same coefficients are positive for 1967. However, the general pattern that the expected earnings are much higher for those with no or a few siblings than for those with many siblings holds firm. The average values of the coefficients for 0-2 siblings and the coefficient for seven or more imply expected earnings advantages for the former group of 21% (1968), 29% (1981), and 36% (1990) for Sweden and 31%, 35% and 51% respectively for United States.[7]

[6] Sample criteria and variables used for the United States and Sweden in our narrow samples are very close to those used in Björklund & Jäntti (1997) and originally by Solon (1992). The economic and econometric arguments in favour of "long-run" measures of the fathers' incomes are discussed in these papers.

[7] We computed these numbers as follows: for Sweden the average coefficient for 0, 1 and 2 siblings was -.049 in 1968 and the coefficient for 7+ siblings was -.239 . The relative advantage for the latter group is: exp(.-049-(-.239))-1, or 21 percent.

Figure 2.

We checked the robustness of these results in two ways. First, we used hourly instead of annual earnings. The results reveal somewhat lower siblings-effects, especially for Sweden. The general pattern, however, is the same and the negative sibling-coefficient is strongly significant. Further, we have imposed the sample restriction that only those who grew up with both biological parents are included. The consequence is somewhat stronger siblings effects. Hence, these two checks do not lead us to question the results.

Next we turn to the results that we obtain when we control for indicators of the parents' budget constraints, i.e. the conditional effects. In regressions 3 to 6 in Table 2 we subsequently add such variables; first, dummies for both parents' educational level, then for the fathers' occupation, then the self reported measure of poverty,[8] and finally the locality of the neighbourhood. The general pattern is that the magnitude of (the negative) sibling-effects declines for both countries. For Sweden the effects are reduced by around half, and for United States even more. So the differences between the United States and Sweden are much smaller after these controls. However, we should note that the indicators of parents' education and father's occupation are more detailed in the U.S. data. Finally, it is important to note that the siblings-coefficients remain significantly different from zero, but not strongly so for Sweden in 1968 and for United States in 1967.

These latter conclusions are not affected when we check the robustness by using hourly earnings and including only those who grew up with both biological parents[9].

Results based on the narrow sample are presented in Table 3. Our main concern here is how the use of fathers' income affects the results in comparison with using only indicators of income like education and occupation. This can be done only for Finland and the United States. We provide Swedish results for the same age groups mainly to show that narrow sample does not give different results than the broad sample for Sweden. We also show results using the Swedish results using Gothenburg data here.

The results suggest that Finland is different from the other two countries. First the unconditional effect is slightly smaller in Finland. Second, after having controlled for indicators of parents’ budget constraints, the coefficient turns positive, even though it is insignificantly different from zero.

There is, however, a disturbing discrepancy between the U.S. results in the two samples, namely that the conditional effects are much smaller in the narrow one. We need to work more to find out the reason for this discrepancy.

---

[8] We are somewhat hesitant to adding this self reported measure of poverty because it can be endogenous with respect to siblings; many siblings during childhood can be the one of the causes of a poor childhood.

[9] We have also experimented by using an interaction term between siblings and age. This can be motivated in several ways, one being that the siblings effects work via schooling and quite often in micro-data one finds that the impact of schooling on earnings grows with age or experience. We found a small tendency to a negative interaction-effect in the equations with hourly earnings. This can explain why Kessler (1991) got insignificant siblings-effects when he used the U.S. NLS-Y data set.

Finally, the results for both Finland and the United States suggest that father's income does the same job as father's education and occupation in reducing the sibling coefficient.

B. Schooling as mediating variable

Having found that the number of siblings really matters for adult men's earnings, we continue by trying to find out why this is so. Previous studies have shown that educational attainment is related to the number of siblings, so it is natural to examine whether schooling is a mediating variable or not. To address this issue we have estimated two equations that can be considered parts of a recursive model. First we add own schooling to the earnings equation so that the indirect effect via schooling is eliminated from the siblings coefficient which then only captures the direct effect of siblings. Second, we estimate equations of educational attainment with siblings as an independent variable. We use the level of educational attainment so an ordered probit is employed as statistical model. For the United States and Sweden we use eight levels of education, whereas we use seven for Finland.

Results from the earnings equations are reported in tables 4 and 5; results from the probit models of educational attainment are reported in tables 6 and 7. In general the results confirm that schooling is a mediating variable, but at most half of the total effect of siblings seems to go via schooling. The probit models strongly confirm previous studies that the number of siblings is strong predictor of educational attainment.

C. Early vs. late effects

The Swedish Gothenburg data allow us to go one more step in investigating how the number of siblings is related to attainment. The data include test scores from three tests undertaken at the age of 13 of the pupils in the sample. This information allow us to examine if the effects of the number of siblings have shown up already at this age, or they show up later when the compulsory school has ended. For the cohort born in 1948, all pupils followed the same compulsory school system until this age and thereafter a choice was made. This choice could involve movement to another place to be able to continue to higher levels of education.

The explore this issue, we first estimated equations for the results at the tests. Then we estimated ordered probit models of educational attainment at the age of 45 conditional upon the results at the tests. The former results are presented in Table 8 and the latter in Table 9. The results show that the number of siblings matter at both levels.

D. Family support and siblings.

The 1991 wave of the Swedish SLLS allow us to examine the mechanisms behind the siblings coefficients in one more way. In this wave, the youngest participants (18-26 years of age) were asked whether they got help from parents or siblings in doing their homework as young teenagers. The results in Table 10 show the number of siblings has a weakly significant negative effect on such support. This finding is consistent with the view that parental time-use is affected by the number of children, and by the siblings from the point of view of the children.

6. Conclusions

Our empirical results can be summarised as follows:

(i) Conditioning only upon age, it is evident that persons with many siblings have on average much lower earnings than those who have no or only a few siblings. This relationship is very strong in all three countries; the expected earnings differentials between those who have no, or a few, siblings and those who have seven were in the range 25-30 percent for Finland and Sweden and 35-40 percent in United States.

(ii) The magnitude of these siblings effects declines markedly for all three countries when control variables that indicate the parents' budget constraints are included. However, the effects do not disappear completely, at least not for the United States.

(iii) We can not conclude with any precision that the siblings-effect is linear, so that getting one or two siblings is disadvantageous (in terms of earnings, of course!) compared to being the only child. Rather, several of our regressions that allow for non-linear effects suggest that having one or two siblings can be better than being the only child.

(iv) Our cross-country comparisons gave unambiguous results about the unconditional relationship between the number of siblings and earnings. In both samples these “effects“ are larger in the U.S. than in Finland and in Sweden. However, when we controlled for indicators of the parents’ budget constraint the results were mixed; in one sample we could not reject the hypothesis of equality between the countries but in another we could. Therefore, we have to be somewhat cautious about the nature of the cross-country differences. However, it seems to be more reasonable to suggest that the fertility pattern in the United States has revealed a stronger negative income elasticity than in Finland and the Sweden than that the trade-offs that parents in the two countries have faced. Of course, it can be questioned whether the basic variables of our study are measured equally enough in

Sweden and the United States in order to be able to infer that there is a marked difference between the countries.

(v) Part of the siblings effect seems to go via educational attainment. Using a Swedish data set, we also found that the number of siblings has a strong impact at test scores already at the age of 13, but the same data set alse revealed that conditional upon the test scores the number of siblings seems to matter for subsequent educational attainment.

What conclusions will these findings lead us to draw about the two basic issues that were the point of departure for our study? First of all, result (i) implies that policies targeted only on the number children can have remarkable distributional precision. A necessary condition for such policies to be successful is, of course, that they do not have other adverse disincentives effects that are so large that the favourable effects are completely counteracted. We think the progressive universal child allowances in Finland and Sweden - that are higher per child from the third child and onwards - are interesting cases to consider from these perspectives.

Second, result (ii) for the United States is consistent with the notion that parents who are subject to a given budget constraint are facing a trade-off between the quality and quantity of children. However, result (iii) forces us to modify this conclusion; the trade-off does not necessarily exist for the first two or three children, but after four or five. It is also a reasonable view that these trade-offs become more pronounced in a society as the American with large income differentials and parents' involvement in the financing of their children's education. However, the evidence from our two samples are inconclusive on this point.

Needless to say, our study leaves a number of unresolved important questions. One concerns the mechanisms that generate the relationships between the number of siblings and earnings. Even though part of the effect goes via schooling, more detailed analysis is needed to find out why the number of siblings seems to matter. More insights into the mechanisms can probably be obtained by studies of the impact of the birth-order and the spacing of children. Separate studies of women would also be useful.

Finally, we want to emphasise the need for better data on childhood conditions. Even such a seemingly unproblematic variable as the number of siblings can be hard to measure accurately with a few questions in a typical survey questionnaire, in particular for those who have grown up in broken families. Most likely the impact of the number of siblings interacts with the type of family where one grows up. Both more frequent use of register information from censuses, and surveys that - like the PSID - follow up the children in the households will improve our knowledge about the relationships between childhood conditions and the well-being during adult life.

References

Atkinson Anthony, Lee Rainwater & Timothy Smeeding (1994), "Income Distribution in OECD Countries: Evidence from the Luxembourg Income Study (LIS)", Paris: OECD.

Becker Gary (1991) *A Treatise of the Family,* Cambridge: Harvard University Press.

Björklund Anders & Markus Jäntti (1997) "Intergenerational Income Mobility in Sweden Compared to the United States", *American Economic Review,* 87(5) 1009-1018.

Blake Judith (1989), *Family Size and Achievement,* Berkeley: University of California Press.

Butcher Kristin & Anne Case (1994) "The Effect of Sibling Sex Composition on Women's Education and Earnings", *The Quarterly Journal of Economics*, Vol. CIX (3), pp. 531-563.

Corcoran Mary, Roger Gordon, Deborah Laren & Gary Solon (1992) "The Association Between Men's Economic Status and Their Family and Community Origins", *Journal of Human Resources*, 27(4), pp. 575-601

Erikson Robert & Rune Åberg (1987) *Welfare in Transition* (Oxford: Clarendon Press.

Eriksson Tor & Markus Jäntti (1997) The Distribution of Earnings in Finland 1970-1990, *European Economic Review*, 41(9) 1763-79.

Hanushek Eric (1992), "The Trade-off between Child Quantity and Quality", *Journal of Political Economy*, vol 100, no.1, pp. 84-117.

Haveman Robert & Barbara Wolfe (1993) "Children's Prospects and Children's Policy", *Journal of Economic Perspectives*, Fall 1993, Vol. 7, no.4, pp. 153-174.

Holmlund Bertil (1985), "Sibling Position and Achievement: the Case of Sweden", *Scandinavian Population Studies*, vol. 7, pp. 100-107.

Härnqvist Kjell & Allan Svensson (1973), “A Swedish Data Bank for Studies of Educational Development“, Sociological Microjournal 7, pp. 35-42.

Kessler Daniel (1991), "Birth Order, Family Size, and Achievement: Family Structure and Wage Determination", *Journal of Labor Economics*, vol 9, no 4, pp. 413-426.

Solon Gary (1992), "Intergenerational income mobility in the United States", *American Economic Review*, vol 82, June 1992, pp. 393-408.

Stafford Frank (1987), "Women's work, Sibling Competition and Children's School Performance", *American Economic Review,* vol. 77, December 1987, pp. 972-980.

Zajonc R.B. & Gregory B. Markus (1975), “Birth Order and Intellectual Development“, *Psychological Review*, Vol. 82, No, 1, 74-88.

Walldén Marja (1990), Sibling Position and Mental Capacity - Reconsdered, Project Metropolitan, Research Report no. 31, Department of Sociology, Stockholm University.

Walldén Marja (1992), Sibling Position and the Educational Career, Project Metropolitan, Research Report no. 37, Department of Sociology, Stockholm University.

Table 1. Sample means and standard deviations (within parenthesis).

Broad samples: 20-64 years of age

	Sweden 1968	Sweden 1981	Sweden 1990	U.S. 1967	U.S. 1979	U.S. 1987
# siblings	3.13 (2.62)	2.42 (2.19)	2.13 (1.84)	4.08 (2.57)	3.94 (2.46)	4.07 (3.09)
Log annual earnings	10.03 (.57)	11.16 (.73)	11.87 (.76)	9.79 (.81)	9.90 (.82)	9.87 (.88)
Log hourly earnings	2.45 (.41)	3.69 (.29)	4.46 (.30)	2.16 (.65)	2.33 (.64)	2.26 (.71)
Age	41.4	40.2	39.9	40.7	36.7	37.8
n	2019 [1634]	1903 [1543]	1877 [1481]	3078	4078	4399

Narrow samples: 29-39 years of age (30-36 for Finland)

	Finland 1990	Sweden 1990	Sweden 1993 (45-year- olds)	U.S. 1987
# siblings	1.64 (1.39)	2.08 (1.65)	2.85 (1.65)	4.13 (2.37)
Log annual earnings	11.29 (1.15)	11.99 (.57)	7.53 (.61)	9.69 (1.02)
Log hourly earnings	-	4.47 (.26)	-	2.17 (0.74)
Age	33.0	35.1	45.0	33.4
n	12812	446 [376]	5226	356

Note: The Swedish samples for which we use hourly earnings only pertain to employed workers and the sample sizes are shown within brackets.

Table 2. Sibling-coefficients from regression models, broad samples. Dependent variable: log of annual earnings. Standard errors within parenthesis, adjusted R-square within brackets.

Controls	Sibling variable	Sweden 1968	Sweden 1981	Sweden 1990	U.S. 1967	U.S. 1979	U.S. 1987
1. Age, age sq.	# sibl.	-.023 [.101] (.005)	-.032 [.130] (.007)	-.029 [.196] (.009)	-.048 [.111] (.006)	-.056 [.172] (.006)	-.055 [.177] (.005)
2. Age, age sq.	1 sibl.	-.072 [.102] (.045)	-.026 [.130] (.053)	.085 [.196] (.055)	.206 [.114] (.074)	.048 [.173] (.067)	.024 [.182] (.086)
	2 sibl.	-.074 (.046)	-.094 (.056)	.048 (.057)	.107 (.074)	.149 (.071)	.0233 (.084)
	3 sibl.	-.150 (.048)	-.061 (.060)	.009 (.062)	.157 (.073)	.037 (.071)	-.036 (.085)
	4 sibl.	-.187 (.053)	-.207 (.073)	.008 (.078)	.023 (.075)	.054 (.074)	-.172 (.088)
	5 sibl.	-.161 (.057)	-.113 (.084)	.050 (.097)	-.044 (.079)	.042 (.078)	-.310 (.094)
	6 sibl.	-.109 (.062)	-.193 (.098)	-.071 (.119)	-.018 (.082)	-.074 (.081)	-.387 (.096)
	7+ sibl.	-.239 (.052)	-.296 (.080)	-.267 (.101)	-.198 (.068)	-.230 (.068)	-.442 (.083)
3. 2 + father's and mother's education	# sibl.	-.019 [.118] (.005)	-.030 [132] (.007)	-.029 [.196] (.009)	-.032 [.143] (.006)	-.032 [.201] (.006)	-.032 [.218] (.005)
4. 3 + father's occupation	# sibl.	-.013 [.137] (.005)	-.021 [.144] (.008)	-.023 [.200] (.009)	n.a.	-.028 [.209] (.007)	-.029 [.225] (.005)
5. 4 + selfrep. Poverty	# sibl.	-.011 [.140] (.005)	-.020 [.144] (.008)	-.019 [.203] (.009)	-.026 [.148] (.007)	-.028 [.209] (.007)	-.025 [.229] (.005)
6. 5 + rural/ urban neigh-bourhood	# sibl.	-.008 [.145] (.005)	-.016 [.147] (.008)	-.019 [.203] (.009)	-.020 [.155] (.007)	-.027 [.209] (.007)	-.025 [.229] (.005)
7. Like 6.	1 sibl.	-.060 [.145] (.044)	-.012 [.146] (.053)	.093 [.203] (.055)	.170 [.158] (.072)	.091 [.209] (.071)	-.043 [.230] (.084)
	2 sibl.	-037 (.046)	-.057 (.056)	.065 (.057)	.091 (.072)	.077 (.070)	-.009 (.082)
	3 sibl.	-.075 (.048)	-.001 (.061)	.037 (.063)	.183 (.072)	-.004 (.070)	-.024 (.083)
	4 sibl.	-.102 (.052)	-.135 (.073)	.049 (.079)	.108 (.074)	.034 (.073)	-.141 (.086)
	5 sibl.	-.060 (.056)	-.039 (.085)	.097 (.098)	.059 (.078)	.063 (.077)	-.204 (.075)
	6 sibl.	.000 (.062)	-.090 (.099)	.015 (.121)	.120 (.081)	-.022 (.080)	-.258 (.094)
	7+ sibl.	-.112 (.053)	-.168 (.083)	-.176 (.104)	-.038 (.067)	-.125 (.069)	-.235 (.083)

Table 3. Sibling-coefficients from various models, narrow samples. Dependent variable: log of annual earnings. Standard errors within parenthesis, adjusted R-square within brackets.

Controls	Sibling var.	Finland 1990	Sweden 1990	Sweden 1993	U.S. 1987
1. Age, age squared	# siblings	-.020 [.001] (.007)	-.028 [.016] (.016)	-.034 [.008] (.005)	-.053 [.042] (.015)
2. Age, age squared	1 sibling	.064 [.002] (.028)	-.064 [.020] (.100)	-.015 [.010] (.058)	-.304 [.030] (.393)
	2 siblings	.067 (.030)	-0.053 (.105)	.002 (.055)	-.299 (.387)
	3 siblings	-.042 (.035)	-.012 (.115)	-.046 (.056)	-.399 (.387)
	4 siblings	-.067 (.047)	-.331 (.135)	-.116 (.058)	-.331 (.392)
	5 siblings	-.088 (.070)	-.280 (.165)	-.076 (.064)	-.363 (.408)
	6 siblings	-.132 (.108)	.048 (.249)	-.233 (.073)	-.673 (.403)
	7+ siblings	-.045 (.175)	-.164 (.188)	-.250 (.069)	-.663 (.389)
3. 2 + father's and mother's educ. + father's occ. + self-rep. pov. + rural/urban neighb.	# siblings	.007 [.044] (.007)	-.027 [.027] (.017)	-.023 [.051] (.005)	-.009 [.183] (.015)
4. Like 3, but father's inc. instead of his educ/occ	# siblings	.007 [.045] (.007)	n.a.	n.a.	-.019 [.138] (.015)
5. 3 + father's income	#siblings	.012 [.051] (.007)	n.a..	n.a.	-.003 [.203] (.015)
6. Like 5, and like 3 for Sweden.	1 sibling	.069 [.047] (.027)	-.049 [.029] (.101)	-.020 [.052] (.057)	-.522 [.199] (.377)
	2 siblings	.097 (.029)	-.069 (.105)	-.002 (.055)	-.571 (.370)
	3 siblings	.056 (.034)	-.023 (.115)	-.043 (.055)	-.637 (.367)
	4 siblings	.080 (.046)	-.319 (.137)	-.094 (.057)	-.638 (.372)
	5 siblings	.052 (.068)	-.255 (.166)	-.033 (.063)	-.459 (.385)
	6 siblings	-.009 (.103)	.074 (.254)	-.179 (.072)	-.563 (.385)
	7+ siblings	.066 (.170)	-.142 (.190)	-.175 (.068)	-.521 (.370)

Table 4. Sibling-coefficients from regression models, broad samples. Dependent variable: log of annual earnings. Standard errors within parenthesis, adjusted R-square within brackets.

Controls	Sibling Var.	Sweden 1968	Sweden 1981	Sweden 1990	U.S. 1967	U.S. 1979	U.S. 1987
1. Like 6 in table 2	# sibl.	-.008 [.145]	-.016 [.147]	-.022 [.203]	-.020 [.155]	-.027 [.209]	-.254 [.229]
		(.005)	(.008)	(.009)	(.007)	(.007)	(.005)
2. 1 + own level of schooling	# sibl.	-.001 [.246]	-.010 [.197]	-.016 [.239]	-.002 [.211]	-.015 [.245]	-.164 [.267]
		(.005)	(.008)	(.009)	(.006)	(.007)	(.005)
3. Like 1.	1 sibl.	-.061	-.012 [.146]	.113	.170 [.158]	.091 [.209]	-.043 [.230]
		(.044)	(.053)	(.054)	(.072)	(.071)	(.084)
	2 sibl.	-.038	-.057	.074	.091	.077	-.009
		(.046)	(.056)	(.057)	(.072)	(.070)	(.082)
	3 sibl.	-.075	-.001	.054	.183	-.004	-.024
		(.048)	(.061)	(.062)	(.072)	(.070)	(.083)
	4 sibl.	-.102	-.135	.050	.108	.034	-.141
		(.053)	(.073)	(.078)	(.074)	(.073)	(.086)
	5 sibl.	-.060	-.039	.103	.059	.063	-.204
		(.056)	(.084)	(.096)	(.079)	(.077)	(.092)
	6 sibl.	.000	-.090	.003	.120	-.022	-.258
		(.062)	(.099)	(.119)	(.081)	(.081)	(.094)
	7+ sibl.	-.112	-.168	-.202	-.038	-.125	-.235
		(.053)	(.083)	(.103)	(.067)	(.069)	(.082)
4. Like 3 + own level of schooling	1 sibl.	-.056 [.247]	-.016 [.196]	.090 [.239]	.146 [.213]	.089 [.245]	-.022 [.268]
		(.041)	(.052)	(.053)	.070	.070)	(.082)
	2 sibl.	.012	-.040	.076	.090	.106	.017
		(.043)	(.054)	(.055)	(.070)	(.068)	(.080)
	3 sibl.	-.041	.020	.057	.183	.051	.020
		(.050)	(.059)	(.061)	(.070)	(.069)	(.081)
	4 sibl.	-.037	-.081	.063	.147	.091	-.070
		(.050)	(.071)	(.076)	(.071)	(.071)	(.084)
	5 sibl.	(-.016)	.007	.115	.095	.127	-.140
		(.053)	(.082)	(.094)	(.075)	(.076)	(.090)
	6 sibl.	.046	-.063	.031	.184	.050	-.179
		(.058)	(.096)	(.116)	(..079)	(.079)	(.092)
	7+ sibl.	-.040	-.123	-.155	.060	-.036	-.134
		(.050)	(.080)	(.100)	(.066)	(.067)	(.081)

Table 5. Sibling-coefficients from various models, narrow samples. Dependent variable: log of annual earnings. Standard errors within parenthesis, adjusted R-square within brackets.

Controls	Sibling var.	Finland 1990	Sweden 1990	Sweden 1993	U.S. 1987
1. Age, age, squared, father's and mother's education, fathers occ., self-reported poverty, rural/urban neighborhood	# siblings	.006 [.043] (.007)	-.027 [.027] (.017)	-.023 [.051] (.005)	-.008 [.183] (.015)
2. 1 + own level of schooling	# sibling	.025 [.089] (.007)	-.024 [.016] (.016)	-.013 [.0147] (.005)	.000 [.208] (.015)
3. Like 1.	1 siblings	.070 [.043] (.027)	-.049 [.029] (.101)	-.020 [.052] (.057)	-.361 [.176] (.378)
	2 siblings	.090 (.029)	-.069 (.105)	-.002 (.055)	-.389 (.370)
	3 siblings	.036 (.034)	-.023 (.115)	-.043 (.055)	-.500 (.369)
	4 siblings	.062 (.046)	-.319 (.137)	-.094 (.057)	-.503 (.375)
	5 siblings	.030 (.069)	-.255 (.166)	-.033 (.062)	-.316 (.387)
	6 siblings	-.059 (.106)	.074 (.254)	-.179 (.072)	-.435 (.389)
	7+ siblings	.048 (.171)	-.142 (.190)	-.175 (.068)	-.400 (.374)
4. Like 3 + own level of schooling	1 sibling	.074 [.087] (.026)	-.054 [.100] (.098)	-.033 [.053] (.054)	-.391 [.206] (.380)
	2 siblings	-119 (.029)	-.061 (.102)	-.018 (.052)	-.445 (.377)
	3siblings	.087 (.034)	-.024 (.111)	-.043 (.052)	-.551 (.375)
	4 sibling	.148 (.045)	-.317 (.133)	-.080 (.054)	-.535 (.381)
	5 siblings	.140 (.067)	-.253 (.161)	-.002 (.059)	-.321 (.395)
	6 siblings	.035 (.104)	.208 (.246)	-.119 (.069)	-.458 (.393)
	7+ siblings	.093 (.167)	-.136 (.183)	-.126 (.065)	-.366 (.380)

Table 6. Ordered probit estimates of the impact of family background on the level of schooling. Broad samples.

Controls	Sibling var.	Sw 1968	Sw 1981	Sw 1990	U.S. 1967	U.S. 1979	U.S. 1987
1. Age, age sq.	# sibl.				-.121 [-7467; .041] (.007)	-.126 [-7036; .038] (.007)	-.093 [-6517; .033] (.006)
2. Age, age sq.	1 sibl.				.128 [-7556; .040] (.075)	.164 [-7171; .038] (.079)	-.080 [-6504; .036] (.093)
	2 sibl.				.019 (.075)	-.008 (.078)	-.121 (.091)
	3 sibl.				-.078 (.074)	-.221 (.078)	-.327 (.092)
	4 sibl.				-.350 (.076)	-.290 (.081)	-.449 (.095)
	5 sibl.				-.369 (.080)	-.444 (.086)	-.637 (.101)
	6 sibl.				-.575 (.083)	-.530 (.089)	-.661 (.104)
	7+ sibl.				-.698 (.069)	-.738 (.075)	-.898 (.090)
3. 1 + father's & mother's educ. , father's occ., self-rep poverty, rural/urban neighb.	# sibl.				-.071 [-7056; .093] (.007)	-.058 [-.6548; .105] (.008)	-.043 [-6106; .094] (.006)
4. Like in 3.	1 sibl.				.064 [-7135; .093] (.076)	-.004 [-6667; .105] (.081)	-.161 [-6104; .095] (.094)
	2 sibl.				-.018 (.075)	-.140 (.079)	-.159 (.092)
	3 sibl.				-.027 (.075)	-.281 (.079)	-.274 (.093)
	4 sibl.				-.206 (.077)	-.270 (.082)	-.373 (.097)
	5 sibl.				-.178 (.081)	-.337 (.087)	-.404 (.102)
	6 sibl.				-.327 (.085)	-.344 (.091)	-.416 (.105)
	7+ sibl.				-.418 (.071)	-.430 (.078)	-.529 (.093)

Table 7. Ordered probit estimates of the impact of family background on the level of schooling. Narrow samples.

Controls	Sibling var.	Finland 1990	Sweden 1990	Sweden 1993	U.S. 1987
1. Age, age sq.	# sibl.	[-18922; .006] -.108 (.007)		[-8946; .006] -.094 (.009)	[-612; .029] -.116 (.020)
2. Age, age sq.	1 sibl.	[-18901; .008] .010 (.026)		[-8936; .007] .037 (.105)	[-606; .038] .355 (.539)
	2 sibl.	-.111 (.028)		.052 (.101)	.593 (.531)
	3 sibl.	-.264 (.033)		-.068 (.102)	.448 (.531)
	4 sibl.	-.466 (.045)		-.136 (.106)	.390 (.537)
	5 sibl.	-.634 (.068)		-.385 (.115)	-.117 (.559)
	6 sibl.	-.499 (.104)		-.576 (.135)	-.115 (.552)
	7+ sibl.	-.368 (.167)		-.570 (.126)	-.453 (.533)
3. 1 + father's & mother's educ., fathers occ., self-rep poverty, rural/urban neighborhood	# sibl.	[-18128; .048] -.090 (.007)		-.069 [-8526; .0529] (.009)	-.053 [-566; .102] (.022)
4. Like in 3.	1 sibl.	-.001 (.026)		.075 [.8519; .0536] (.107)	
	2 sibl.	-.117 (.028)		.077 (.103)	
	3 sibl.	-.228 (.033)		-.034 (.104)	
	4 sibl.	-.383 (.045)		-.052 (.108)	
	5 sibl.	-.547 (.069)		-.259 (.117)	
	6 sibl.	-.385 (.105)		-.423 (.137)	
	7+ sibl.	-.219 (.168)		-.346 (.128)	

Table 8. Regression models of the impact of family background on test scores at the age of 13. Sweden. All persons born in 1948.

Controls	Sibling var.	Test 1 (verbal)	Test 2 (spatial)	Test 3 (logic)
1. No	# siblings	-.68 [.027] (.06)	-.47 [.011] (.06)	-.40 [.007] (.07)
2. No	1 sibling	-.66 (.69)	-.82 [.012] (.76)	1.64 [.011] (.81)
	2 siblings	-.43 (.67)	-.60 (.73)	-1.02 (.78)
	3 siblings	-1.41 (.67)	-.78 (74)	-1.0 (.79)
	4 siblings	-1.82 (.70)	-1.27 (.76)	-1.22 (.82)
	5 siblings	2.61 (.76)	-.3.02 (.83)	-3.08 (.89)
	6 siblings	-4.53 (.88)	-3.31 (.96)	-3.59 (1.03)
	7+ siblings	-5.47 (.82)	-3.74 (.90)	-4.55 (.96)
3. Father's and mother's educ., father's occ., rural/urban neighborhood.	# siblings	-.57 [.101] (.57)	-.36 [.045] (.064)	-.32 [.055] (.068)
4. Like 3.	1 sibling	-.51 (.67)	-.77 (.75)	-1.51 [.058] (.80)
	2 siblings	-.35 (.65)	-.44 (72)	-1.02 (.77)
	3 siblings	-1.28 (.65)	-.58 (73)	-.98 (.78)
	4 siblings	-1.49 (.68)	-.91 (76)	-1.06 (.81)
	5 siblings	-2.06 (.74)	-2.46 (.82)	-2.70 (.88
	6 siblings	-3.75 (.85)	-2.55 (.95)	-3.02 (1.02)
	7+ siblings	-4.48 (.80)	-2.84 (.89)	-3.83 (.95)

Table 9. Ordered probit estimates of siblings coefficients in models of the level of schooling. Sweden.

Controls	Siblings variable	Sweden 1993
1. Father's and mother's educ., father's occ., rural/urban neighborhood.	# siblings	-.069 [-8525.6; .0529] (.009)
2. Like 1.	1 sibling	.075 [-8519.1; .0536] (.107)
	2 siblings	,077 (.103)
	3 siblings	-.034 (.104)
	4 siblings	-.052 (.108)
	5 siblings	-.259 (.117)
	6 siblings	-.423 (.137)
	7+ siblings	-.346 (.128)
3. Like 2 plus test1-test3.	# siblings	-.037 [-8025.4; .109] (.009)
4. Like 3.	1 sibling	.156 [-8020; .109] (.107)
	2 siblings	.131 (.104)
	3 siblings	.049 (.105)
	4 siblings	.049 (.109)
	5 siblings	-.091 (.118)
	6 siblings	-.184 (.138)
	7+ siblings	-.036 (.129)