The More the Sicker? Health, Family Size, and Birth Order

Petter Lundborg[1], Hilda Ralsmark[2], Dan-Olof Rooth[3]

Abstract

Becker's Quantity-Quality model suggests a negative relationship between family size and the health of children. While this relationship has been investigated in recent medical studies, the direction and size of the effect is still unresolved. In this paper, we use a unique dataset on the entire Swedish male birth cohorts between 1965 and 1978 to provide evidence on the effect of an exogenous increase in family size on children's health. Our initial OLS estimates indicate that each additional child is associated with a decrease in height by ¼ cm. The magnitude of this estimate decreases and becomes insignificant when we control for birth order. When we instead estimate the causal effect of family size, using twin births as an instrument for family size, we find a positive effect of family size on height. In addition to this, our fixed effects results also show the existence of negative birth order effects.

JEL classification:
Key words: Quality-Quantity Tradeoff; Family size; Birth Order; Health; Twin Instrument; Fixed Effects.

[1] Lund University, VU University Amsterdam, Tinbergen Institute, IZA, HEP, CED. Email: petter.lundborg@nek.lu.se
[2] Department of Economics and CED, Lund University. Email: hilda.ralsmark@nek.lu.se.
[3] Linnaeus and Lund University, CED, CReAM and IZa. Email: dan-olof.rooth@lnu.se.

1. Introduction

The effect of family size on the outcomes of children has caught the attention of – and been widely debated by – economists, and other social scientists for decades. In an attempt to formalize the relationship between family size and child outcomes in an economic framework, the Quantity-Quality model was introduced in the seminal papers by Becker and Lewis (1973) and Becker and Tomes (1976). The model assumes that the marginal cost of children increases with the quantity of children, and this result in a trade-off between child quantity and child quality.

A relationship between the quantity and quality of children seems to receive support in a number of recent empirical studies in the medical literature, but the health effects are ambiguous. For instance, Strachan (1989) finds a positive Quantity-Quality relationship via a negative correlation between family size and child atopy – the tendency to suffer from one or more allergies, such as eczema. Lawson however, finds that children in larger (3+) families are 2.5 cm shorter than the average height for their age (Sample, 2007). But do such results really reflect a causal effect of family size on health? There might be other, non-causal, mechanisms at play as well. Large families may be inherently different from small families in ways that are difficult, or impossible, to observe. As far as we understand, the mechanisms underlying the relationship between family size and health are poorly understood.

In this paper, we provide evidence on the effect of an exogenous increase in family size on children's health outcomes. To the best of our knowledge, our paper is the first to consider the causal effect of family size on children's health outcomes. We follow a recent literature in economics, which uses twin births as an instrument for family size, in order to estimate the causal effect.

In the medical literature, family size can be interpreted as a measure of the probability of exposure to diseases, and birth order can be thought to relate to the age of exposure. For example, only children can be thought not to be exposed to infections and diseases until they enter school, but later-born children are exposed at an early stage through their older siblings. One famous theory that predicts a positive relationship between family size and health is the hygiene hypothesis, proposed by Strachan (1989).

The model predicts that children in smaller families are less likely to be exposed to diseases spread by interaction with other children. This, in turn, results in a poorer development of the immune system and poorer health. The effect of this might be visible through stunted adult growth or increased susceptibility to diseases as an adult. The effect of birth order, however, is generally thought to depend on the type of infectious agent and how it is related to diseases (Mucci et al., 2004). For example, getting infected with the Epstein-Barr virus during childhood is harmless, leads to an adult immunity, and a reduced risk of Hodgkin's disease (Vianna and Polan, 1978), whereas an infection during adolescence can cause glandular fever (Huen et al., 1993). On the other hand, exposure to the hepatitis B virus at a young age increases the risk of hepatocellular carcinoma (liver cancer) (Hsieh et al., 1992). As a result, it is important that one takes birth order into consideration when one analyses family size effects on health outcomes.

We believe that isolating the causal effect of family size on health is important for a number of reasons. First, family size may be an important input in the child and adolescent health production function. Understanding the determinants of childhood health is important, given the recent evidence on the great importance of childhood health for later life outcomes, such as education and earnings. Case et al. (2005) use data from the 1958 British Birth Cohort Study and find that children that suffer from chronic conditions have lower educational attainment, wages, and employment probabilities compared to other children. Smith (2009) uses data from the Panel Study of Income Dynamics to study the long-term effects of child health on earnings, using sibling fixed effects. The results show that there are significant negative effects of poor health in childhood on adult earnings. Currie et al. (2007) use data on children born in the Canadian province Manitoba between 1979 and 1987, until 2006, to analyze the long run effect of early childhood health on socioeconomic outcomes. They find that health problems in childhood (in particular mental health problems) are significant determinants of adult outcomes. Their results suggest that physical health in early childhood health affects adult outcomes via its effect on future health, and not because of a direct link between health status and cognition (Almond and Currie, 2010). Currie et al.'s (2007) result is particularly interesting for our study since both Canada and Sweden have universal public health insurance. The evidence discussed above suggests that there is a link between child health and adult outcomes. This further highlights the relevance to study the effect of family size on child health outcomes.

A second reason for isolating the causal effect of family size on health is that the results can be used to evaluate the costs and benefits of policies that aim to change fertility levels. Currently, many Western countries are adopting policies to increase fertility levels. For example, in December 2003 the Italian government offered 1.000 euros to couples with one child if they had a second child by the end of 2004 (Kennedy, 2003). Fertility policies might therefore have positive or negative effects on child health, depending on the relationship between family size and health.

In our analysis, we use a unique dataset that combines data on health outcomes from the Swedish Military Enlistment Register with register data from Statistics Sweden (SCB). The data contains information on all males who enlisted for the Swedish Army during the years 1984 and 1997. Since enlistment was mandatory during this period, the data covers more or less the entire population of males that enlisted in these years. In addition, the Multi-generation register allows us to link these individuals to their biological parents and siblings.

The data at hand gives us a number of distinct advantages over previous studies on the relationship between family size and health. First, we are able to exploit twin births as an instrument for family size. Under the plausible assumption that twin births are exogenously given, we are able to isolate the causal effect of family size. Second, the data allows us to consider the effect of family size on height. Height is considered to be one of the best indicators of early life health (Elo and Preston, 1992), and adult height has been widely used as a marker of socioeconomic variations in childhood (Silventoinen, 2003). Moreover, recent research has shown the existence of a substantial height-premium in the labor market, which further motivates our focus on height (Case and Paxson, 2008; Lundborg et al., 2009). Third, measurement errors most likely play a negligible role in our estimates since we are not relying on self-reported height. This stands in contrast to other studies where self-reported health outcomes such as asthma or eczema prevalence in children has been used (Strachan, 1989). Fourth, the data allows us to consider birth order effects with great precision. By running separate regressions by family size, and accounting for family fixed effects, we are able to separate the effect of birth order from that of family size, in the spirit of Black et al. (2005).

Our OLS results show a strong and negative relationship between family size and height, in line with the result in many previous studies. The magnitude of the family size relationship decreases when we control for parental demographic characteristics and socioeconomic background. Similar to Black et al. (2005) we find that family size becomes insignificant when we control for birth order. Our results therefore suggest that the often observed negative relationship between family size and child quality outcomes mainly reflects the influence of omitted variables.

When we instrument family size with twin births, the estimates become larger and change sign, suggesting that there is a significant positive effect of family size on health. Our findings therefore support a Quantity-Quality relationship but, unlike that suggested by the economics literature, we find that it is positive and not negative for health outcomes. Finally, we analyze birth order effects via between- and within-family variation. We find that there are large, significant negative birth order effects on height, and that these increase with birth order.

The outline of the paper is as follows. Section 2 provides a literature review, Section 3 describes the data, Section 4 describes the methodology and results, and Section 5 concludes.

2. Literature review

The relationship between family size and child outcomes has been widely researched by economists as well as by researchers in other fields, such as Sociology and Psychology. An important reason for this interest – As Cunha and Heckman (2007) point out – is that children can neither choose their parents, nor purchase an insurance against poor parents. This creates a market failure, and while traditional arguments of fairness and social justice are applicable, an argument based on economic efficiency that favors investment in disadvantaged children is very appealing. As a result, the effect of family background on child outcomes – and subsequent adult quality of life – is an important research area in Economics.

Much of the past empirical research has shown that family size is negatively associated with child outcomes, and the Quantity-Quality model is seen as providing a good explanation of the mechanisms behind this. This is because supporting evidence of a dilution effect has been found in studies of financial and resource allocation within families. Also, large family size effects have been found in developing countries with scarce parental resources, or in studies on children in poorer households. The key feature in the Quantity-Quality model is that the marginal cost of children with respect to family size increases with child quality, and that the marginal cost of children with respect to quality increase with family size. This creates a trade-off between child quantity and quality (Becker and Lewis, 1973). Quality can be defined in many ways but most empirical papers in Economics focus on education, wages or IQ-scores. Out of these, a large number of studies have found supporting evidence of a trade-off between family size and educational attainment (Blake, 1989; Hanushek, 1992; Joshi and Schultz, 2007; Åslund and Grönqvist 2010). Such results led Haveman and Wolfe (1995) to argue that the number of siblings is one of the most important childhood factors correlated with adult well-being.

Evidence on a Quantity-Quality relationship has also been found in the medical literature, although there is little consensus on the underlying mechanisms, the direction, and the size of the effects. Also, medical research has defined and tested the "sibling effect" in different ways. Alternative definitions are birth order, the number of siblings/family size, the number of older siblings, and the number of brothers (Karmaus and Botezan, 2002). As a result, it is difficult to know which effect that is being tested – family size, birth order, or both. It is clear, however, here are many plausible reasons why birth order can affect health outcomes and confound our results. In the Medical literature, this is related to the age of exposure to diseases. The effect of these diseases, however, depends on the type of infectious agent and how it is related to diseases (Mucci et al., 2004). From a biological perspective, it might be the case that the mother's biological capacity is better the younger she is and the fewer children she has had. Bingley et al. (2000) find a strong association between increasing maternal age at delivery and the risk of child diabetes. This risk is highest in first-born children and decreases with birth order. In the Economics literature, this could be due to the fact that parents invest more time and money in earlier born children. For example, Price (2008) finds that first–born children receive 20-30 more minutes of quality time per day compared to second-born children of the same age. Zajonc (1976) predicts a negative relationship due to the change in the intellectual environment that comes with an increasing family size and rising birth order. For example, first-born children get sole attention and education from their parents, who have adult intelligence, whereas later-born children get attention and learn both from their parents and their older siblings. This results in a difference in the intellectual atmosphere, where the first-born child spends more time in an adult environment, and teaches his younger siblings (this is called the 'teaching effect'), while the young child gets less time with his parents in an adult environment, and does not have the same opportunity to teach his younger siblings. This speaks in favor of being born early into the family. On the other hand, it might be the case that the parents' capacity to rear and raise a child improves with the number of children because they gain experience. In addition, later-born children might have parents with higher income levels and therefore more capacity to allocate resources towards that child. This, in turn, speaks in favor of being born late into the family.

The fact that there are many alternative theories on the effect of birth order on child outcomes is in line with the literature that has looked at birth order effects. The results have been mixed and some support the existence of birth order effects (Belmont and Marolla, 1973; Zajonc, 1976; Zajonc et al., 1979; Sulloway, 1996) and others reject them (Schooler, 1972; Galbraith, 1982; Retherford and Sewell, 1991). This has led to a discussion on methodology and the validity of the research results. Rodgers et al. (2000) argue that the discrepancy is the result of the choice to use either cross-sectional data – giving large birth order effects – or within-family variation – giving small or zero effects. The suggested explanation is that birth order effects are within-family effects, which causes a strong selection bias in the cross-sectional data. The conflicting theoretical predictions make it interesting to investigate the effects of birth order on health more closely and our data allows us to solve many of the problems that arise due to the endogeneity of family size and birth order.

The existence of a positive relationship between family size and health was highlighted by Wickens et al. (1999) who showed an inverse relationship between family size and hay fever, skin prick positivity, and Immunoglobin E (an important class of antibody associated with allergy hypersensitivity). In accordance with this observation, much research has found that a large family size is associated with a lower probability of asthma and allergies (Karmaus and Botezan, 2002) and some cancers (Bevier et al., 2011). Two alternative hypotheses given to explain the sibling effect are the hygiene hypothesis – which suggests that it is the number of infections, vaccinations, use of antibiotics, early day care, and rural versus urban life that affects health – and the in utero programming hypothesis – which suggests that the maternal immune system increases with the number of births and that this is transferred to the child in utero (Ohfuji et al., 2009).

Results that suggest that family size is negatively associated with health outcomes have also been found. For example, Mucchi et al., (2004) use the Swedish Twin Registry data to examine the effect of family size and birth order on tooth loss and periodontal disease. They find that periodontal disease is positively associated with family size and inversely related with birth order. Their findings are consistent with the hypotheses that adult health is related to the probability of exposure in childhood and that an early age of exposure lowers risk. Hatton and Martin (2010) study the effect of family size and birth order in the UK during the 1930s, when poverty was widespread, and find a negative effect of family size on height. They propose two mechanisms to explain the findings – a positive effect of per capita expenditure on food and a negative effect of crowding within the household. The latter effect is related to household cleanliness and hygiene which inhibits an individual's capacity to take up nutrients in the food. It is, however, unclear to what extent the results from a wartime-period are applicable to developed countries today. Horton (1988) finds a similar negative effect of birth order on height-for-age using a multipurpose survey from 1978, collected in the Bicol region in the Philippines. Finally, Lawson and Mace (2008) use the British Avon Longitudinal Study of Parents and Children dataset, and find that children from larger families are shorter (about 2.5 cm) and have a lower growth rate.

The previous evidence on the Quantity-Quality model in terms of health consists of observations of a statistical relationship between family size and different health outcomes. This, however, does not necessarily imply a causal relationship. Instead, family size may be endogenous if larger families are inherently different from smaller families in ways that are difficult to observe. This may create an omitted variable bias in the OLS estimates, which leads to flawed interpretations and policy recommendations. Moreover, since parents choose child quantity and quality, family size cannot be treated as an exogenous variable. In order to address the endogeneity of family size, much of the recent economic research focus on finding an exogenous variation in family size. The most common approach has been to instrument family size with variables such as twin births or sibling-sex composition. The first study that tests for a causal relationship between family size and child outcomes using twin births is Rosenzweig and Wolpin (1980). They find a negative effect of family size on the educational attainment of children but their results are based on a dataset limited by its size: there are only 25 twin pairs in a sample of 1.600 children.

Subsequent studies that have used larger datasets have not found a trade-off to the same extent as Rosenzweig and Wolpin (1980). One of these is Black et al,'s (2005) seminal paper, where data on the entire Norwegian population is used to test for an effect on educational attainment and earnings. They initially find a relationship between family size and child outcomes when they use an OLS estimation but this relationship disappears when they control for birth order or use twins as an instrumental variable. Black et al.'s paper suggests a birth order effect, rather than a family size effect, on child quality. This gives some support to alternative theories such as the Confluence model, which suggests that it is not family size, but rather, the spacing of children and the intellectual atmosphere within the family that matters (Zajonc, 1976).

While almost all studies have looked at the effect of a singleton or twin birth on the earlier born children within the same family, a recent study by Rosenzweig and Zhang (2009) examines the effect on the twins themselves. The argument is that if the quality of the twins is lower compared to their singleton siblings, then their parents might reallocate resources away from the twins, and towards the older singleton siblings. This might offset any Quantity-Quality effect, and make it difficult to use twins as instruments. The authors find a negative relationship between family size and the schooling progress, the expected college enrolment, grades and child health for all children in the family. Some caution with regards to the results should be made, however, since it is difficult to compare twins with singletons (Angrist et al., 2010).

Glick et al. (2007) employs a twin-estimate strategy to test for a causal effect of family size and birth order on height. They find that twins at first birth have a negative effect on the investment in human capital, in particular for children with high birth orders. One weakness to this study, however, is that it instruments family size with twin births at first birth. Since parents that choose to have more children after a twin birth might be different than those that do not, it is generally thought to be preferable to examine the effect on the children born before the twin birth. In that way, one avoids selection bias (Black et al., 2005).

3. Data description

The data used in the empirical analysis is constructed by integrating registers from Statistics Sweden (SCB) and the Swedish National Service Administration. This enables us to link children to their biological parents, and to their siblings, via the Multi-Generation register. The more detailed construction of the dataset is as follows.

The Swedish National Service Administration data contains information on every individual living in Sweden in 1999, and who enlisted in the army between 1984 and 1997.[4] Military enlistment is a two-day procedure and it was mandatory for all Swedish male citizens the year they turned 18 during the period of our analysis. A refusal to enlist resulted in fines, and eventually imprisonment. The only exempted individuals were those 1) with a severe handicap 2) that were institutionalized due to mental disorders 3) that were in imprisonment, or 4) that lived abroad.[5]

The Multi-Generation register has information on the number of children born by the mother up to and including 2008. Since the youngest boys in the data are born in

[4] The individuals had to live in Sweden during 1999 since the enlistment information was initially collected for the 1999 population data.
[5] There is no information available on why a particular individual did not enlist.

1978, it is likely that we observe the complete family size. Each individual is linked to their biological parents and siblings through a unique parental identification number and we only include families in which information on the number of births equals the number of children born by the mother (or where there has been a twin birth and the number of children exceeds the number of births). This allows us to calculate information on birth spacing and family sizes at different ages and to use this to perform a better analysis of the alternative mechanisms behind the birth order effects. Birth order is extracted when calculating birth ranks within families using the full population data in 1999. Hence one could worry that individuals that have migrated or died create a bias in this calculation. However, we expect such bias to be minor since the family size calculated from the Multi-Generation data is almost identical to family size taken from the population data. (The relevant cohorts have a correlation of 0.89.)

Our study population consists of all males born between 1965 and 1978, and that enlisted for the military between 1984 and 1997, and for whom there is full information on relevant variables. In order to avoid any confounding influence of ethnic diversities, we restrict our analysis to native Swedish males, i.e. those born in Sweden to Swedish-born parents.[6] As a result, the population covers around 90% of the native male Swedish population in the relevant cohorts.[7] Finally, we further restrict the data to make the information from different registers coherent. Our final dataset consists of 483,408 observations for our health outcome. This falls to 184,015 when we use variations within families (in-between siblings) to test whether birth order affects health.

In the pursuit of finding a causal family size effect we use a twin strategy for identification. Twins are identified in the registers as having the same mother and being born during the same month and year. Given this information, we construct a twin-instrumental variable that is a dummy variable which equals 1 if the mother has a twin birth at the n^{th} birth (n={2, 3, 4}) and equals 0 otherwise. Similar to Black et al. (2005), we restrict the sample to families with at least n births and study the outcomes of the children born before the n^{th} birth. Hence, estimations are performed for men from

[6] Moreover, non-naitive ethnic groups have a much lower participation rate for enlisting since only about fifty percent (or less) are Swedish citizens, making selective participation an issue for these groups.
[7] Native males born between 1965 and 1978, and for whom we have information on the mother constitute 632,835 individuals.

families with (potential) twin births at the second, third, and fourth birth, respectively. Following the reasoning in the earlier sections, the twins themselves are excluded from the empirical analysis.

Dependent variable

Height (cm): This is measured and reported by employees of the National Service Administration, thus avoiding measurement bias. Since all individuals are 18 at the time of enlistment, one does not need to use height-for-age measurements. In addition, all individuals are likely to have gone through puberty and attained their adult height by the age of 18. This makes it probable that we have data on final adult height outcomes.

Independent variables

In our regressions, we always control for the conscript and mother's year of birth. In addition, we control for age at first birth. This is because the probability of a twin birth increases with age. All estimations also control for year of enlistment which picks up anything specific for the year that the conscript enlisted. We have full information on all parental demographic characteristics, but some information is missing on socioeconomic variables.[8] (When missing it constitutes its own category.) We also include both parents' level of education in 1999, and we choose this year to be sure that the parents have completed their education. In addition to this, we also have earnings data from 1970, 1975, and 1980 (in 1980 prices) that are added up within the families, and then averaged over the three years. We use this to get a notion of permanent family income, and we choose this variable over one plausible alternative – current income – because current income often is a poor proxy for lifetime income (Böhlmark and Lindquist, 2006; Haider and Solon, 2006). Similar to Åslund and Grönkvist (2010), however, we do not condition on parental earnings in our main analysis. Instead, we use it to investigate the potentially heterogeneous effects of family size and to check whether parental characteristics are related to twin births. The same is true for (i) our measure of spacing, which is calculated as the difference in years between the first and the last child born by a mother, as well as

[8] Information on education is missing for 10% of the fathers and 4% of the mothers, while family income is missing for only 2%.

(ii) our two measures of intact families, which in the first case measures whether or not all children in the family has the same father, and in the second case also conditions on that all children lived in the same family household in 1980.

4. Methodology and results

The first part of this section is concerned with the relationship between family size and our health variable; height, when controlling for background factors, and birth order. The second part is concerned with isolating a causal effect of the same outcome variable via an exogenous variation of family size.

Descriptive statistics

Table 1, column 1, shows the distribution of family sizes (defined as the total number of children that the mother has given birth to) for all mothers who gave birth to a boy at least once between 1964 and 1979, and for which the boy enlisted in the army sometime between 1984 and 1997. 57% of the families consist of one or two births, whereas it is uncommon with more than five births (4.7%).

Table 1 also shows the descriptive statistics of the other key variables in the empirical analysis. As can be seen in column 2, the average height is 179.6 cm. With the exception of the one-child family (often found to be disadvantaged in the economics literature) height is negatively associated with family size and this becomes clearer in families with 4+ children. An individual from a two-child family is about 1 cm taller than his 5+ child family counterpart. Many of the other variables; mother's education, mother's age of first birth, income, and the probability of having married parents is also negatively associated with family size.

OLS regression results

We first provide OLS-based estimates of the relationship between family size and height at age 18. In the estimations, we use two different assumptions regarding the effect of family size. First, we assume a linear relationship, where each additional sibling has the same effect on height in the family, irrespective of family size. Second, we allow for a non-linear relationship by adding indicator variables for each family size. In this case, the excluded category is the one-child family.

The estimates from only including family size as the independent variable are reported in Table 2, column A. The first line reports the estimate for the linear specification of family size. The coefficient is significant and negative, and suggests that each additional child in the completed family size is associated with a reduction of height by 0.25 cm. In line 2-5, we instead allow for the more flexible functional form by adding our indicator variables for each family size.[9] The omitted reference category is the one-child family. This gives us a result that was not visible with linear specification of family size, namely that there is an advantage of being from a two-child family compared to a one-child family. This positive relationship is significant but not very large – two-family children are on average 0.16 cm taller than only children. Other than that, the results largely confirm those obtained using a linear specification. For the other family sizes, the coefficients are significant and negative. The main effects are visible in the larger families – Being from a 4-child family and 5+ child family is associated with a fall in average height by roughly 0.5 and 1.1 cm.

Controlling for background characteristics

Most previous literature that has examined the relationship between family size and child quality has used an OLS estimation strategy. There are several reasons why this might not estimate an actual causal relationship. The results can be biased if there are omitted socioeconomic and demographic variables. For example, families with lower socioeconomic status tend to have larger families. As a result, the first OLS regression results are, at best, only indicative of a family size effect. Therefore, we include controls for other factors thought to affect height outcomes. In column B, we control for socioeconomic background, enlistment year and mother's age at first birth. In the linear specifications, the coefficient on the number of children remains significant but falls to about a third of the original estimate (-0.083), leading us to conclude that there was a

[9] We have one category variable for families with five or more (5+) children. This is because very few families have more than five children.

downward bias in the first estimate. We observe the same thing when we include our controls in our regression with family size dummy variables.

Controlling for birth order

As discussed in the introduction and background sections, in the medical literature both family size and birth order are thought to affect health outcomes via their effect on the probability and age of exposure to diseases. In the Economics literature, birth order is also closely related to family size. It is therefore important to include controls for birth order in order for us not confound family size with birth order correlations.

Following this reasoning, we add dummy variables for birth order. The excluded category is the first-born child and the dummies represent the second-born, third-born etcetera. The final dummy variable equals one if the child is the fifth child or greater. The results are given in column C. After including birth order dummies the coefficient on the number of children falls to -0.025 and becomes insignificant. For the non-linear specifications, the difference in the coefficients for the family size dummy variables is not very large. On the other hand, the birth order dummy variables are highly significant and ranges from -0.1 cm for the first born child to -0.3 cm for the 5+ children. This suggests that there is a relationship between birth order and health.

Alternative specifications

Families with the same father

In column D, we re-estimate our model that controls for background characteristics and birth order with a sample that includes children that have the same father. There are several reasons for this. First, we keep the hereditary part of health production (health endowment at birth) the same within each the family. Second, research suggests that children that are raised in nuclear families – two married or cohabiting individuals that are parents to all children in the family – are healthier than children living in a non-nuclear family (Blackwell, 2010).[10] Third, the literature suggests a new partner might

[10] Suggested explanations for this relationship are that a divorce/separation leads to increased stress levels or a change in health-related behavior – such as less physical activity or a poorer diet, or less parental supervision.

divert the mother's attention away from the children and lead to less investment than in a nuclear family structure (Thomson et al., 1994). Fourth, children living with stepparents are more likely to drop out of high school than children living with biological parents (Astone and McLanahan, 1991; McLanahan and Sandefur, 1994). Lastly, our descriptive statistics in Table 1 suggests that larger families are associated with a non-intact family. There is therefore reason to believe that stability in the family structure is positively related to health outcomes for children, and our family size estimates are biased if we do not take this into account. Having the same father can therefore be thought of as having similar health endowments at birth and a stable nuclear family while growing up. Our estimates suggest that the linear relationship between family size and height is similar in both families with and without the same father. The coefficient is small and the variable is insignificant, whereas the non-linear estimates are similar to that in column C.

Children living in the same household

In column E we only include families whose children were all living in the same household in 1980. This is to ensure that all children in the family share (compete for) parental resources. We find a significant but very small (0.05) positive relationship between family size and height. The size of the other estimates is similar to before, but some are now insignificant.

Individuals born in 1965-1972

In column F, we estimate the effects of children born in 1965-1975 (our total sample includes those born between 1973 and 1978). The reason for this is that earlier cohorts might not have had the benefit of birth control, which means that family planning was more difficult. As a result, family size can be thought to have been more exogenous before the availability of birth control, and more endogenous after. Therefore, we test whether or not our health outcomes are driven by the later cohort. We find that the linear effect of family size is negative and significant, but close to zero (-0.06). In our non-linear family size specification, the family size dummies are similar in size, but not as significant as before. Also, our birth order effects are similar to the ones estimated with the full sample.

Results when using different samples

Mother's years of education

Previous research suggests that mother's education is an important determinant of the health of both biological and adopted children. Potential mechanisms behind this relationship are that educated mothers have better knowledge about health care and nutrition, healthier behavior, and can provide safer and more sanitary environments for their children (Behrman and Deolalikar, 1988; Strauss, 1990; Thomas et al., 1990; Desai and Alva, 1998; Glewwe, 1999, Currie and Moretti, 2003). Thus, mother's education might affect the relationship between family size and child health by reducing any negative effects, or increasing any positive effects, of family size on child health.

In addition to this, more educated mothers generally have better health, which leads to better health through genetics (Behrman and Wolfe, 1987; Wolfe and Behrman, 1987). These nurture and nature effects can be interpreted as casual if the mother's education affects her knowledge on health, but as a selective or omitted variable bias if it is a nature effect. Chen and Hongbin (2009) is one of the first papers that separate the nurture and nature effects from each other. They use adopted children from China and find a nurturing effect of mother's education. The effect is similar for both adoptees and biological children. Lastly, education can act as a proxy for financial resources.

We choose to compare samples where the mother has less or more than 12 years of education. This means that we are comparing families with academic or non-academic mothers. The results are given in Table 2.2. We find that the relationship between family size and height is similar for both samples. While our coefficient for number of children in negative and significant for families with non-academic mothers it is also very close to zero (-0.04). The coefficient is close to zero and insignificant for the sample with academic mothers. The remaining coefficients are similar in both samples, with the exception for the 5+child family dummy variable, which is negative for children with non-academic mothers and positive for children with academic mothers. The birth order estimates are similar in both samples.

Family income level

In Economics, the relationship between parental income and child health is called the health gradient (Case et al., 2002; Currie and Stable, 2003; Currie et al., 2008; Propper et al., 2007; Condliffe and Link, 2008; Murasko, 2008; Khanam et al., 2009). The majority of the empirical findings suggest that there is such a thing as a health gradient between parental income and child health, but that it differs between countries and studies, and that there is no agreement on the actual mechanisms behind it (Johnston et al, 2010). The effect of family size on health may differ according to family income levels if families with lower income level face a larger dilution effect with increasing family size than those with higher income level.

In Table 2.2, column D-F, we stratify our sample according to income level. The income level has been calculated by using data on earnings in 1970, 1975, and 1980 (in 1980's prices), which have been added up and averaged over the three years. This gives us a notion of permanent income. In all our samples, the linear family size variable is insignificant and the family size dummies are similar to each sample. This indicates that there is no difference in family size correlations when it comes to income levels. One interesting finding is that the dummy variables for birth order are close to zero and insignificant for the higher family income group. For the lower and middle income group, the birth order variables are negative and similar to each other in size and significance. Our findings indicate that, in general, there is no difference in health gradient when it comes to the relationship between family size and height.

Instrumental variable results

The previous results are based on an OLS estimation strategy, which means that they indicate the presence or absence of correlations between family size and health. This does not, however, imply causality, and the problems associated with OLS regressions on the Quantity-Quality relationship is the bias that arises from endogeneity or simultaneity. Endogeneity occurs if there is a relationship between family size and some unobserved characteristic that also affects the quality outcome. Since it is much easier for couples and individuals to choose if and when to have children after the introduction of birth control (in particular the contraceptive pill), family size is an active choice where factors such as individual preferences and career opportunities play an important role. This creates an omitted variable bias and its direction depends on the correlation between the omitted variables and our sibling size variable and height.

Another cause of bias is simultaneity. This occurs when parents change their view on the optimal family size after they observe the quality of their previous children. For example, if a first-born is of high quality, then the parents might become content with having one child (Behrman and Taubman, 1986). Another potential bias can occur when parents start to have children but realize that they cannot devote as much temporal or financial resources to the last child as they want to (Åslund and Grönqvist, 2010). Black et al. (2005) calls this the 'last child' effect.

One estimation strategy that has been used in the literature on the Quantity-Quality trade-off to solve the problem of endogeneity and simultaneity is twin births. This was first introduced by Rosenzweig and Wolpin (1980). The idea is that parents have an optimal family size that they aspire to have. Thus, when a couple chooses to have a child, they plan for a singleton birth that will contribute to bringing the family closer to the desired size. A twin birth can therefore 'push' the actual family size beyond the optimal size. If twin births can be assumed to be random, then the allocation of suboptimal family sizes is random as well. This randomness is the argument behind using twin births as an exogenous variation of family size.

For twin birth to be a valid instrument it has to be 1) correlated with family size and 2) uncorrelated with the error term ε. A good description of the twin birth IV-approach is given in Black et al. (2005). We repeat it here, slightly adjusted for our research question, for ease of reference. Equation (2) is the first stage of the 2SLS and equation (1) is the second stage.

(1) $\text{HEIGHT} = \beta_0 + \beta_1\text{NUMBER OF CHILDREN} + X\beta_2 + \varepsilon$

(2) $\text{NUMBER OF CHILDREN} = \alpha_0 + \alpha_1\text{TWIN} + X\alpha_2 + \nu$

In this case, HEIGHT is the conscript's height in centimeters; NUMBER OF CHILDREN is total family size and X is a vector of control variables, TWIN is a dummy variable that takes the value 1 if the n^{th} birth is a twin, 0 otherwise.

We use three samples in our analysis. In our first sample, (i), we have first-born children in families with two or more children. In this case, the instrument is whether or not the second birth is a twin or a singleton. In the second sample, (ii), we have first and second born children in families with three or more births. In this case, the instrument is whether or not the third birth is a twin or singleton. In the third and final sample, (iii), we have first, second, and third born children in families with four or more children. In this case, the instrument is whether or not the fourth birth is a twin or singleton.

All our estimates of family size thus concern the effect on children born prior to the twin birth. This is because families that have another child after a twin birth might be different from those that choose not to, and because a twin births complicates the birth order of the children born afterwards. For example, if a twin birth occurs at the 3^{rd} birth the child born after them will be 5^{th} born rather than 4^{th} born. Thus, we only look at children born prior to the twin birth in order not to confuse family size effects with birth order effects. In Table 3.1, we show the OLS, the first stage, second stage IV-estimate, and the reduced form. As can be seen in the first stage column a twin-birth increases family size by between 0.7 and 0.9 children. The effect of twins on final family size is greater the larger the family is before the twin birth. The t-statistics for the first stage are all above 90, suggesting that a twin birth is not a weak instrument for family size.

Having shown that twin births are positively correlated with family size, we next turn to the assumption that the twin birth instrument is exogenous to the error term. If this is not the case, then our IV-estimate will also be biased and inconsistent. In general, this assumption is much more difficult to prove. In our case, it requires that twin births are random. Circa 3% of the annual births in Sweden are twin births, which mean that the overwhelming majority of births are singletons. However, are parents that do get twins different from those that do not? One possible objection to using twin births as an instrument is that the probability of having twins increases with the age of the mother. We therefore control for mother's age and age at first birth in all regressions. Another potential problem is the genetic aspect to twin birth – the propensity that the female body releases two eggs at the time of ovulation instead of one (hyperovaluation) is hereditary. Thus, if you are a twin or have twins in your family history, you are more likely to give birth to twins. However, only women can carry this trait, and non-twin women are unlikely to know whether or not they carry this gene. Therefore, it is likely that twin birth is exogenous.

Our second stage IV-results are given in Table 3.1, column C. In samples (i) and (ii) we find that each additional sibling increases height by 0.49- 0.57 cm (we include the same controls as before, including birth order). We do not find any effect of family size in sample (iii). The effect of a twin birth therefore appears to decrease with parity.

Alternative specifications

Mother's years of education

In Table 3.2, column A, we re-estimate our instrumental variables model for children with mother's that have 12 years of education or less. The argument for this is the same as for the OLS model. The results suggest that for sample (i), the positive effect of family size is only marginally higher for families with non-academic mothers (the difference in the coefficient is 0.05). For sample (ii) and (iii) the effect of family size is insignificant.

Family income level

In column B and C we report estimates for our sample when we stratify it according to whether the family income level is at the lowest or in the middle group. When we do so, the family size effect becomes insignificant in sample (i) for both income groups, whereas the family size effect becomes significant for middle income families in sample (ii). As before, family size is insignificant for sample (iii).

Birth order effects

Even with within-family (fixed) effects, isolating the family size effect from the birth order is difficult. Blake (1989) suggests how to proceed. First, one must control for family size effects. This is because higher born children naturally exist in larger families.

In addition, one should control for cohort effects, since the children are not born in the same year. One must also control for parental background. Thus, in order to obtain credible estimates of birth order one must have many cohorts with same birth orders, and a fertility history of the mother. Given our extensive data, we are able to follow Blake's suggestions in our regression analysis and also compare between- and within family results.

Table 4.1 shows both the between-family variation and sibling fixed effects (within- family variation). Our large sample sizes allow us to run separate regressions by family size, thus mitigating the endogeneity of family size. Each column in Table 4.1 therefore represents a separate regression for each family size. All OLS regressions control for age, mother's age and education, and father's age and education. In the fixed effects model we only include an indicator for age. (When running regressions on 'all families' we also control for family size). The omitted category is the first born child.

In the between-family regression, column A, we can see that birth order is negatively associated with height. When we run separate between-family variations for each family size (column B-E), we find that the effect of being a second child is significant for 2-child families only, and that the effect is small. The results are all negative and significant for the other birth orders. In general, being latter born is associated with being shorter and the negative effect increases with family size. Being a child with birth order 5+ is associated with an average height that is almost 1 cm lower than being first born in a 5+ family.

The sibling fixed effects estimates are also given in Table 4.1 (column F-J). These estimates also suggest that height falls with birth order. The majority of the estimates are more negative than the estimates from the between family model. This suggests that there is a positive bias due to omitted family characteristics in the estimation that uses between family variations. The birth order effects appear to occur at an early stage – our results suggest that second-born children in a two-child family are 0.5 cm shorter than the first-born. As before, there is a monotonic fall in height as birth order increases.

The results in both the OLS between and within family variation suggests that, contrary to what earlier literature has found, there is no first-born disadvantage when it comes to health. On the contrary, first-borns seem to be at an advantage when only looking at birth order effects,

Birth order – sensitivity analysis

We perform two sensitivity analyses to our birth order estimates. First, we check whether or not age where birth order is registered matters. Second, we check whether we are confounding birth order effects with spacing effects.

Age when birth order is registered

Given that parents have more than one child, then family size varies with the age of the child. For example, a child might be an only child for three years, followed by two years as the first born in a two-child family and, finally, the first born in a three-child family at age nine. While the birth order is the same in all of these cases, the family size varies throughout time. We therefore test whether or not the effect of birth order depends on the family size observed at three different periods – age 3, 5 and 9. The results, given in Table 4.2, show that the effect of birth order is independent of when family size is measured. This leads us to believe that the fixed effects results on birth order (Table 4.1) do capture birth order effects.

Spacing and birth order

Not only is family size and birth order of relevance when it comes to analyzing child outcomes. The spacing of children might also matter. For instance, an older child might get less temporal allocation from the parents if a younger child is born within a small time frame. Additionally, fatigue during pregnancy can cause worse parenting (Pettersson-Lidbom and Skogman Thoursie (2009). This can increase the older child's stress levels and cognitive abilities (Center on the Developing Child at Harvard University, 2007). Pettersson-Lidbom and Skogman Thoursie (2009) study the effect of the introduction of a premium in Sweden that was given to parents that had their second child within 24 months after the birth of their first child. They find that this led to a fall in the spacing between the first and second born child by 6 months on average, and that close child spacing (defined as less than two years) and educational outcomes of children are negatively related. The results are similar for first-, second-, third-, fourth-, and fifth-born children.

To test whether or not spacing affects our birth order estimates, we divide our data into two groups depending on whether or not the family had relatively close or long spacing in its family formation. The average spacing is calculated by dividing the total time between the first and last birth by the total number of children. If the average spacing is less than the median, then the family belongs to the below median spacing category, and if the average spacing is more than the median, then the family belongs to the above median spacing category. By doing this, we are able to test whether or not spacing is confounding our birth order estimates. The results are given in Table 4.2, column D. The differences in birth order effects between a family with below median spacing and above median spacing are very small – it is 0.02 for second-born children and 0.06 for third-born children. This suggests that spacing effects do not affect our birth order estimates.

Family size, birth order and height – How large are the effects?

In this paper, we have found significant positive family size effects on height outcomes of children, and significant negative birth order effects. The question is, however, what does an increase in adult height by roughly 0.5 cm per sibling really mean? To answer this, we have calculated the effect of family size on height in terms of standard deviations. This is given in Table 5. The results show the same pattern as before when using height outcomes – the OLS estimate is close to zero and insignificant, whereas the instrument variable results for family size is significant for sample (i) and (ii). The effect of each additional sibling is an increase in height by between 0.075 and 0.09 standard deviations. The birth order effects are all monotonically increasingly negative. Being second-born leads to a reduction in height by 0.06 standard deviations, whereas being a fifth+ child leads to a reduction by -0.17 standard deviations. Lundborg et al. (2009) uses the same data as this study and finds that an additional ten centimeters in height is associated with a raw height premium of six percent per year. Together, these results suggest that family size is related to both adult health and income.

5. Conclusion

In this paper, we examine the effect of family size on health, measured by height at age 18. In order to do so, we first obtain OLS estimates where we control for parental background, cohort effects, and birth order effects. Then, we use twin births as an instrumental variable in order to obtain an exogenous source of variation in family size. Given our unique and detailed dataset we are able to solve many of the limitations that research on the Quantity-Quality trade-off faces.

Initially, our OLS results suggest that family size is negatively correlated with height, but our estimate approaches zero and becomes insignificant when we include background controls, and birth order indicator variables. In general, our non-linear specification of family size gives similar results as our linear specification. When we instead use twin births as an instrument for family size, we find that each additional sibling increases average height in the family by roughly 0.5 cm. This significant effect is found in two of our three samples, and does not appear to vary much depending on mother's educational attainment or family income level.

Both our OLS and IV estimates suggest that birth order matters. In all specifications we find that there is a clear negative and monotonically increasing relationship between birth order and height. We perform two sensitivity checks to see whether or not our birth order effects are confounded by age when family size is measured, and birth spacing, but our results do not change. Our within-family estimates are both more negative and more significant, indicating that birth order estimates on height from between family estimations are upward biased.

Given that height is a good indicator of health status, then our result show that family size has a positive effect on health. Which theory in the literature can explain this positive effect? It does not appear to be a dilution effect of financial and temporal resources, since this would lead to a negative relationship between family size and height. Instead, our findings coincide with the predictions of the hygiene hypothesis, which suggests that children in smaller families have less exposure to diseases, and therefore develop a weaker immune system.

Our results suggest a positive Quantity-Quality trade-off when looking at family size and health outcomes. Interestingly, our results are different from other studies that have used twin instrumental variables and not found a relationship between family size and, for example, educational attainment. A potential explanation for this is that the positive effect of family size on health outcomes dominates the negative effect of family size on alternative quality outcomes. This suggests that Western policies that aim to increase fertility levels might have additional benefits such as better child health.

References

Almond, D. and Currie, J. (2010) Human Capital Development Before Age Five. *NBER Working Paper No. 15827*.

Angrist, J., Lavy, V. and Schlosser, A. (2010) Multiple Experiments for the Causal Link between the Quantity and Quality of Children. *Journal of Labor Economics*, 28, pp. 773-824.

Astone, N. M. and McLanahan, S. (1991) Family Structure, Parental Practices and High School Completion. *American Sociological Review*, 56, pp. 309-320.

Becker, G. S. and Lewis, H. G. (1973) On the Interaction Between the Quantity and Quality of Children. *Journal of Political Economy*, 81, pp. 279–288.

Becker, G. S. and Tomes, N. (1976) Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy*, 84, pp. 143–162.

Behrman, J. and Deolalikar, A. (1988) Health and Nutrition. In Hollis Chenery and T.N. Srinivasan eds. *Handbook of Development Economics, Volume 1*, Amsterdam: North Holland, pp. 631-711.

Behrman, J. and Taubman, P. (1986) Birth Order, Schooling, and Earnings. *Journal of Labor Economics*, 5, pp. 121-45.

Behrman, J. and Wolfe, B. (1987) How does mother's schooling affect the family's health, nutrition, medical care usage, and household sanitation? *Journal of Econometrics*, 36, pp. 185-204.

Belmont, L. and Marolla, F. A. (1973) Birth order, family size, and intelligence. *Science*, 182, pp. 1096-1101.

Bevier, M., Weires, M., Thomsen, H., Sundquist, J. and Hemminki, K. (2011) Influence of family size and birth order on risk of cancer: a population-based study. *BMC Cancer*, 11, pp. not given.

Bingley, P. J., Douek, I. F., Rogers C A. and Gale E A. (2010) Influence of maternal age at delivery and birth order on risk of type 1 diabetes in childhood: prospective population based family study. Bart's-Oxford Family Study Group. *British Medical Journal*, 321, pp. 420-424.

Blake, J. (1989) *Family Size and Achievement*. University of California Press, Berkeley and Los Angeles.

Black, S. E., Devereux, P. J. and Salvanes, K. G. (2005) The More the Merrier? The Effects of Family Size and Birth Order on Children's Education. *Quarterly Journal of Economics*, 120, pp. 669–700.

Blackwell, D. L. (2010) Family structure and children's health in the United States: Findings from the National Health Interview Survey, 2001-2007. *Vital Health Stat 10*, 246, pp. 1-166.

Böhlmark, A. and Lindquist, M. (2006) Life-Cykle Variation in the Association between Current and Lifetime Income: Replication and Extension for Sweden. *Journal of Labor Economic*s, 24, pp. 879-896.

Case, A., Fertig, A. and Paxson, C. (2005) The lasting impact of childhood health and circumstance. *Journal of Health Economics*, 24, pp.356-389.

Case, A., Lubotsky, D. and Paxson C. (2002) Economic status and health in childhood: The origin of the gradient. *American Economic Review*, 92, pp. 1308-1334.

Case, A. and Paxson, C. (2008) Stature and Status: Height, Ability, and Labor Market Outcomes. *Journal of Political Economy*, 116, pp. 499-532.

Center on the Developing Child at Harvard University (2007) A Science-Based Framework for Early Childhood Policy: Using Evidence to Improve Outcomes in Learning, Behavior, and Health for Vulnerable Children. http://www.developingchild.harvard.edu.

Chen, Y. and Hongbin, L. (2009) Mother's education and child health: Is There a nurturing effect? *Journal of Health Economics*, 28, pp. 413-426.

Condliffe, S. and Link, C. (2008) The relationship between economic status and child health: Evidence from the United States. *American Economic Review*, 98, pp. 1605-1618.

Cunha, F. and Heckman, J. (2007) The technology of skill formation. *NBER working paper no. 12840*.

Currie, J., Decker, S. and Lin, W. (2008) Has public health insurance for older children reduced disparities in access to care and health outcomes? *Journal of Health Economics*, 27, pp. 1567-1581.

Currie, J. and Moretti, E. (2003) Mother's Education and the Intergenerational Transmission of Human Capital: Evidence From College Openings. *Quarterly Journal of Economics*, 118, pp 1495-1532.

Currie, J. and Stable, M. (2003) Socioeconomic status and child health: Why is the relationship stronger for older children? *American Economic Review*, 93, pp. 1813-1823.

Currie, J., Stabile, M., Manivong, P. and Roos L. L. (2007) Child Health and Young Adult Outcomes. *NBER Working Papers 14482*.

Desai, S. and Alva, S. (1998) Maternal education and child health: Is there a strong causal relationship? *Demography*, 35, pp. 71-81.

Elo, I. T. and Preston S. H. (1992) Effects of early-life conditions on adult mortality: a review. *Population index*, 58, pp. 186-212.

Galbraith, R. C. (1982) Sibling spacing and intellectual development: A closer look at the confluence models. *Developmental Psychology*, 18, pp. 151-173.

Glewwe, P. (1999) Why Does Mother's Schooling Raise Child Health in Developing Countries? Evidence from Morocco. *Journal of Human Resources*, 34, pp 124-159.

Glick, P. J., Marini, A. and Sahn, E. (2007) Estimating the Consequences of Unintended Fertility for Child Health and Education in Romania: An Analysis using Twins Data. *Oxford Bulletin of Economics and Statistics*, 69, pp. 667-691.

Haider, S. and Solon, G. (2006) Life-Cycle Variation in the Association in the Association between Current and Lifetime Earnings. *American Economic Review*, 96, pp. 169-187.

Hanushek, E. A. (1992) The Trade-off between Child Quantity and Quality. *Journal of Political Economy*, 100, pp. 84–117.

Hatton, T. J. and Martin R. M. (2010) The effects on stature of poverty, family size, and birth order: British children in the 1930s. *Oxford Economic Papers*, 62, pp.157-184.

Haveman, R. and Wolfe, B. (1995) The Determinants of Children's Attainments: A Review of Methods and Findings. *Journal of Economic Literature*, 33, pp. 1829–1878.

Horton, S. (1988) Birth Order and Child Nutritional Status: Evidence from the Philippines. *Economic Development and Cultural Change*, 36, pp. 341-354.

Hsieh, C-C., Tzonou, A., Zavitsanos, X., Kaklamani, E., Lan, S-J. and Trichopoulos (1992) Age at first establishment of chronic hepatitis B virus infection and hepatocellular carcinoma risk: a birth order study. *American Journal of Epidemiology*, 136, pp. 1115-1121.

Huen, D. S., Fox, A., Kumar, P. and Searle, P. F. (1993) Dilated heart failure in transgenic mice expressing the Epstein-Barr virus nuclear antigen-leader protein. *Journal of General Virology*, 74, pp. 1381-1391.

Johnston, D. W., Propper, C., Pudney, S. and Shields, M. (2010) Is there and Income Gradient in Child Health? It depends on whom you ask. *IZA Discussion Papers 4830*.

Joshi, S. and Schultz, P. T. (2007) Family Planning as an Investment in Development: Evaluation of a Program's Consequences in Matlab, Bangladesh. *IZA Discussion paper No 2639*.

Karmaus, W. and Botezan, C. (2002) Does a higher number of siblings protect against the development of allergy and asthma? A review. *J Epidemiol Community Health*, 56, pp. 209-217.

Kennedy, F. (2003) Italy offers families baby-cash. BBC news, [online] 1 December. Available at: http://news.bbc.co.uk/2/hi/europe/3252794.stm [Accessed 16 May, 2011].

Khanam, R., Nghiem, H. S. and Connelly, L. B. (2009) Child health and the income gradient: Evidence from Australia. *Journal of Health Economics*, 28, pp. 805-817.

Lawson, D. W. and Mace, R. (2008) Sibling configuration and childhood growth in contemporary families. *International Journal of Epidemiology*, 37, pp. 1408-21.

Lundborg, P., Nystedt, P. and Rooth, D-O., (2009) The Height Premium in Earnings: The Role of Physical Capacity and Cognitive and Non-Cognitive Skills. *IZA Discussion Paper No. 4266*.

McLanahan, S. and Sandefur, G. (1994) *Growing Up with a Single Parent: What Hurts? What Helps? Cambridge*, MA: Harvard University Press.

Mucci, L. A., Hsieh, C-C, Williams, P. L., Dickman P.W., Björkman, L and Pedersen, N.L (2004) Birth Order, Sibship Size, and Housing Density in Relation to Tooth Loss and Periodontal Disease: A Cohort Study among Swedish Twins. *American Journal of Epidemology*, 159, pp. 499-506.

Murasko, J. E. (2008) An evaluation of the age-profile in the relationship between household income and the health of children in the United States. *Journal of Health Economics*, 27, pp. 1489-1502.

Ohfuji, S., Miyake, Y. Arakawa, M., Tanaka, K. and Sasaki S. (2009) Sibship size and prevalence of allergic disorders in Japan: The Ryukyus Child Health Study. *Pediatric Allergy and Immunolgy*, 20, pp. 377-384.

Pettersson-Lidbom, P. and Skogman Thoursie, P. (2009) Does child spacing affect children's outcomes? Evidence from a Swedish reform. *IFAU working paper 2009:7.*

Price, J. (2008) Parent-Child Quality Time: Does Birth Order Matter? *Journal of Human Resources*, 43, pp. 240-265.

Propper, C., Rigg, J. and Burgess, S. (2007) Child health: evidence on the roles of family income and maternal mental health from a UK birth cohort. *Health Economics*, 16, pp. 1245-1269.

Retherford, R. D. and Sewell, W. H. (1991) Birth order and intelligence: Further tests of the confluence model. *American Sociological Review*, 56, 141-158.

Rodgers, J. L., Harrington, C. H., Van den Oord, E., Rowe, D. C. (2000) Resolving the debate over birth order, family size and intelligence. *American Psychologist*, 55, pp. 599-612.

Rosenzweig, M. R. and Wolpin, K. I. (1980) Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment. *Econometrica*, 48, pp. 227–240.

Rosenzweig, M. R. and Zhang, J. (2009) Do Population Control Policies Induce More Human Capital Investment? Twins, Birthweight, and China's 'One Child' Policy. *Review of Economic Studies*, 76, pp. 1149-1174.

Sample, I. (2007) Study reveals how having an older brother can make you shorter. *The Guardian* [online], 13 September 2007. Available at: http://www.guardian.co.uk/science/2007/sep/13/genetics. [Accessed 11 May 2011].

Schooler, C. (1972) Birth order effects: Not here, not now! *Psychological Bulletin*, 78, pp. 161-175.

Silventoinen, K. (2003) Determinants of variation in adult body height. *Journal of Biosocial Science*, 35, pp. 263-285.

Smith, J. P. (2009) The impact of childhood health on adult labor market outcomes. *The Review of Economics and Statistics*, 91, pp. 478-489.

Strachan, D. P. (1989) Hay fever, hygiene, and household size. *British Medical Journal*, 299, pp. 1259–1260.

Strauss, J. (1990) Households, Communities, and Preschool Child Nutrition Outcomes: Evidence from Rural Cote d'Ivoire. *Economic Development and Cultural Change*, 38, pp. 231-261.

Sulloway, F. J. (1996) *Born to rebel*. New York: Vintage.

Thomas, D., Strauss, J. and Henriques, M-H. (1990) Child survival, height for age and households characteristics in Brazil.*Journal of Development Economics*, 33, pp. 197-234.

Thomson, S. G., Eccles, J. and McLanahan, S. (1994) Family Structure and Child Well-Being: Economic Resources vs. Parental Behaviors. *Social Forces*, 73, pp. 221-41.

Vianna, N. J. and Polan, A. K. (1978) Immunity in Hodgkin's disease: importance of age at exposure. *Annals of Internal Medicine*, 89, pp. 550-556.

Wickens, K., Crane, J., Pearce, N., Beasley, R. (1999) The magnitude of the effects of smaller family sizes on the increase in the prevalence of asthma and hay fever in the United Kingdom and New Zeeland. *Journal of Allergy and Clinical Immunology*, 104, pp. 554-558.

Wolfe, B. and Behrman, J. (1987) Women's schooling and children's health: are the effects robust with adult sibling control for the women's childhood background? *Journal of Health Economics*, 6, pp. 239-254.

Zajonc, R. B. (1976) Family Configuration and Intelligence. *Science*, 192, pp. 227–236.

Zajonc, R. B., Markus H. and Markus G. B. (1979) The birth order puzzle. *Journal of Personality and Social Psychology*, 37, pp. 1325-1341.

Åslund, O. and Grönqvist, H. (2010) Family size and child outcomes: Is there really no trade-off? *Labor Economics*, 17, pp. 130-139.

Appendix

Table 1. Descriptive statistics

	N	Height	Age in 2000	Mother's education	Father's education	Mother's age at first birth	(Log) Family income	Intact Family
					Family size			
1-child family	44,437	179.6	28.8	11.4	11.2	26.5	11.74	1.00
2-child family	247,463	179.7	28.7	11.5	11.5	24.2	11.80	0.93
3-child family	139,916	179.5	29.3	11.4	11.4	22.7	11.75	0.84
4-child family	37,169	179.2	30.0	11.0	11.1	21.8	11.67	0.72
5+ child family	14,423	178.7	30.6	10.3	10.5	20.8	11.50	0.63
					Birth order			
First child	210,670	179.7	29.2	11.6	11.5	24.5	11.73	0.93
Second child	185,194	179.6	28.7	11.4	11.4	23.6	11.80	0.88
Third child	64,238	179.5	29.0	11.1	11.2	22.3	11.78	0.80
Fourth child	16,235	179.1	29.7	10.6	10.7	21.5	11.70	0.72
Fifth+ child	7,071	178.7	30.3	10.0	10.2	20.8	11.52	0.67
All	483,408	179.6	29.0	11.4	11.4	23.7	11.84	0.88

Table 2.1. OLS regression results: Effect of family size on height.

	A	B	C	D	E	F
Family size:						
Number of children	-0.246***	-0.083***	-0.025	-0.028	0.054***	-0.060***
	(0.012)	(0.014)	(0.016)	(0.018)	(0.021)	(0.019)
2-child family	0.163***	0.191***	0.220***	0.246***	0.277***	0.149***
	(0.034)	(0.036)	(0.037)	(0.037)	(0.038)	(0.047)
3-child family	-0.067*	0.119***	0.172***	0.237***	0.275***	0.095*
	(0.037)	(0.040)	(0.042)	(0.044)	(0.048)	(0.052)
4-child family	-0.486***	-0.069	0.024	0.135	0.081	-0.054
	(0.051)	(0.055)	(0.061)	(0.069)	(0.085)	(0.071)
5+ child family	-1.120***	-0.408***	-0.302***	-0.154	-0.066	-0.458***
	(0.073)	(0.079)	(0.103)	(0.130)	(0.186)	(0.112)
Birthorder:						
Second child	-	-	-0.123***	-0.097***	-0.111***	-0.111***
			(0.027)	(0.030)	(0.035)	(0.035)
Third child	-	-	-0.211***	-0.153***	-0.159**	-0.222***
			(0.051)	(0.059)	(0.070)	(0.065)
Fourth child	-	-	-0.322***	-0.216***	-0.060	-0.304***
			(0.087)	(0.103)	(0.133)	(0.106)
Fifth+ child	-	-	-0.315***	-0.248***	-0.187	-0.250***
			(0.141)	(0.176)	(0.282)	(0.164)
R2	0.006	0.014	0.014	0.013	0.014	0.014
No of cases	483,408	483,408	483,408	427,805	359,297	303,875

Notes: Standard errors (in parentheses) allow for correlation of errors within families. Model A include indicators for age and mother's age. Model B also include indicators for mother's education and father's age and education, enlistment year and mother's age at first birth, while Model C then also adds indicators for birth order. Model D through F includes the same set of Xs as Model C but use different samples being analyzed. Model D includes only families in which all children have the same father, Model E further restricts the data in Model D to that all children within the family were also living in the same household in 1980, while Model F includes only those born 1965-1972.

Table 2.2. OLS regression results from different samples.

	C	Mothers edu		Family income		
		≤ 12	>12	Lowest	Middle	Highest
Family size:						
Number of children	-0.025	-0.041**	0.040	-0.017	0.030	0.040
	(0.016)	(0.019)	(0.034)	(0.026)	(0.029)	(0.031)
2-child family	0.220***	0.201***	0.243***	0.246***	0.273***	0.236***
	(0.037)	(0.044)	(0.076)	(0.065)	(0.067)	(0.060)
3-child family	0.172***	0.142***	0.218***	0.204***	0.301***	0.214***
	(0.042)	(0.050)	(0.086)	(0.072)	(0.076)	(0.073)
4-child family	0.024	-0.006	0.120	0.125	0.129	0.173
	(0.061)	(0.072)	(0.131)	(0.098)	(0.109)	(0.118)
5+ child family	-0.302***	-0.364***	0.587**	-0.190	0.019	-0.133***
	(0.103)	(0.115)	(0.279)	(0.146)	(0.192)	(0.241)
Birth order:						
Second child	-0.123***	-0.166***	-0.123***	-0.150***	-0.160***	0.011
	(0.027)	(0.032)	(0.027)	(0.048)	(0.047)	(0.046)
Third child	-0.211***	-0.288***	-0.211***	-0.284***	-0.227***	-0.004
	(0.051)	(0.060)	(0.051)	(0.090)	(0.090)	(0.088)
Fourth child	-0.322***	-0.371***	-0.322***	-0.413***	-0.312**	-0.152
	(0.087)	(0.100)	(0.087)	(0.146)	(0.153)	(0.161)
Fifth+ child	-0.315***	-0.447***	-0.315***	-0.480**	-0.329	0.058
	(0.141)	(0.157)	(0.141)	(0.213)	(0.258)	(0.310)
R2	0.014	0.010	0.007	0.013	0.011	0.009
No of cases	483,408	350,323	112,164	158,854	161,293	163,261

Notes: Standard errors (in parentheses) allow for correlation of errors within families. Model A replicates the estimates for Model C in Table 2.1. The following columns include the same set of Xs as Model C but use different samples being analyzed, dividing the data into mother's education being less than 13 years and above 12 years of schooling and the dividing the data into three classes based on family income (lowest third, middle third and the highest third).

Table 3.1 OLS and IV regression results. Effect of family size on height.

	OLS	First stage	Second stage	Reduced form	N
Instrument: Twin at 2nd birth					
(Sample: First child in families with 2 or more births)	-	0.732*** (0.005)	-		166,233
Number of children in family/Twin	-0.138*** (0.030)	-	0.569*** (0.244)	0.417*** (0.179)	
Instrument: Twin at 3rd birth					
(Sample: First and second children in families with 3 or more births)	-	0.862*** (0.005)	-		103,964
Number of children in family/Twin	-0.132** (0.054)	-	0.485* (0.263)	0.418* (0.227)	
Second child	-0.123** (0.059)		-0.122** (0.059)	-0.123** (0.059)	
Instrument: Twin at 4th birth					
(Sample: First, second and third children in families with 4 or more births)	-	0.865*** (0.009)	-		28,286
Number of children in family/Twin	-0.116 (0.130)	-	0.250 (0.509)	0.216 (0.440)	
Second child	-0.318*** (0.111)	-	-0.316*** (0.111)	-0.317*** (0.111)	
Third child	-0.638*** (0.165)	-	-0.633*** (0.165)	-0.636*** (0.165)	

Notes: Standard errors (in parentheses) allow for correlation of errors within families (OLS-data). All regressions include indicators for age, mother's age and education, father's age and education, enlistment year, age at nth birth, age at first birth, and when applicable, indicators for birth order.

Table 3.2 IV results by mother's education and family income.

	Mothers education <13 years	Family income Lowest	Family income Middle
Instrument: Twin at 2nd birth			
Number of children in family/Twin	0.619**	0.505	0.599
	(0.292)	(0.428)	(0.426)
Instrument: Twin at 3rd birth			
Number of children in family/Twin	0.458	0.000	0.737*
	(0.325)	(0.438)	(0.431)
Second child	-0.148**	-0.275***	-0.051
	(0.070)	(0.099)	(0.103)
Instrument: Twin at 4th birth			
Number of children in family/Twin	0.183	0.653	0.666
	(0.633)	(0.744)	(0.984)
Second child	-0.296**	-0.181	-0.385*
	(0.130)	(0.163)	(0.202)
Third child	-0.587***	-0.518**	-0.722**
	(0.192)	(0.249)	(0.299)

Notes: Standard errors (in parentheses) allow for correlation of errors within families (OLS-data). All OLS regressions include indicators for age, mother's age and education and father's age and education. Fixed effects model only include indicator for age. "All families" also include indicator for family size.

Table 4.1 OLS and sibling fixed effects: Effect of birth order on height

	OLS – between family variation					Sibling fixed effects – within family variation				
	All families	Two-child family	Three-child family	Four-child family	Five+ child family	All families	Two-child family	Three-child family	Four-child family	Five-child family
Height:										
Second child	-0.126*** (0.027)	-0.091** (0.039)	-0.073 (0.046)	-0.165 (0.106)	-0.247 (0.271)	-0.359*** (0.043)	-0.469*** (0.074)	-0.275*** (0.066)	-0.316*** (0.129)	-0.195 (0.277)
Third child	-0.210*** (0.051)	-	-0.189*** (0.076)	-0.276** (0.134)	-0.709*** (0.282)	-0.761*** (0.085)	-	-0.686*** (0.131)	-0.825*** (0.203)	-0.589* (0.316)
Fourth child	-0.320*** (0.087)	-	-	-0.404** (0.181)	-0.930*** (0.315)	-0.941*** (0.140)	-	-	-1.236*** (0.318)	-0.659* (0.395)
Fifth+ child	-0.317** (0.140)	-	-	-	-0.966*** (0.363)	-1.121*** (0.217)	-	-	-	-0.861* (0.524)
No of cases	483,408	247,463	139,916	37,169	14,423	184,015	83,120	72,088	20,641	8,166
Sibling fixed effects	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors (in parentheses) allow for correlation of errors within families (OLS-data). All OLS regressions include indicators for age, mother's age and education and father's age and education. Fixed effects model only include indicator for age. "All families" also include indicator for family size.

Table 4.2 Height and birth order. OLS and sibling fixed effects. Alternative explanations.

	Family size at age 3	Family size at age 5	Family size at age 9	Spacing
Height:				
Second child:	-0.355***	-0.333***	-0.360***	-0.350***
(And below median spacing)	(0.043)	(0.044)	(0.043)	(0.068)
	-	-	-	-0.331***
Above median spacing				(0.048)
Third child:	-0.756***	-0.718***	-0.761***	-0.687***
(And below median spacing)	(0.086)	(0.086)	(0.085)	(0.132)
	-	-	-	-0.746***
Above median spacing				(0.095)
Fourth child:	-0.924***	-0.862***	-0.941***	-0.748***
(And below median spacing)	(0.140)	(0.141)	(0.140)	(0.204)
	-	-	-	-0.989***
Above median spacing				(0.165)
Fifth+ child	-1.048***	-0.937***	-1.121***	-0.861*
	(0.223)	(0.223)	(0.217)	(0.524)
No of cases	184,015	184,015	184,015	184,015
Sibling fixed effects	Yes	Yes	Yes	Yes
No of cases	184,015	184,015	184,015	184,015

Notes: Standard errors (in parentheses) allow for correlation of errors within families (OLS-data). All OLS regressions include indicators for age, mother's age and education and father's age and education. Fixed effects model only include indicator for age. "All families" also include indicator for family size.

Table 5. Family size, birth order and health outcome. Standard deviation results.

	Height (std)
Family size	
Ols – with controls	-0.004
	(0.003)
Twin 1: n ge 2	0.088**
	(0.038)
Twin 2 : n ge 3	0.075*
	(0.041)
Twin 3: n ge 4	0.039
	(0.079)
Birthorder	
Second child	-0.055***
	(0.007)
Third child	-0.117***
	(0.013)
Fourth child	-0.146***
	(0.022)
Fifth+ child	-0.173***
	(0.034)

Notes: Standard errors (in parentheses) allow for correlation of errors within families (OLS-data). All OLS regressions include indicators for age, mother's age and education and father's age and education. Fixed effects model only include indicator for age. "All families" also include indicator for family size.