Small Family, Smart Family?
Family Size and the IQ Scores of Young Men

Sandra E. Black
Paul J. Devereux
Kjell G. Salvanes

A B S T R A C T

This paper uses Norwegian data to estimate the effect of family size on IQ scores of men. Instrumental variables (IV) estimates using sex composition as an instrument show no significant negative effect of family size; however, IV estimates using twins imply that family size has a negative effect on IQ scores. Our results suggest that the effect of family size depends on the type of family-size intervention and that there are no important negative effects of expected *increases in family size. However, unexpected shocks to family size resulting from twin births have negative effects on the IQ scores of existing children.*

I. Introduction

Researchers have long reflected on the role of families in influencing the intelligence of their children. Given evidence that intelligence is in part determined by the genetic structure of parents, is there room for post-birth family characteristics to play a role? If yes, how big is this effect? In this paper, we focus on the role of one family characteristic: the effect of family size on IQ test scores.

Family size has long been of interest to researchers, particularly because of the strong empirical regularity that children from larger families tend to have poorer outcomes. There is an extensive theoretical literature on the tradeoff between child quantity and quality within a family that dates back to Becker (1960) and Becker and Lewis (1973). The theory is often cited and is used as the basis for many macro growth models (see Becker and Barro 1988; Doepke 2003). A key element of the model is an interaction between quantity and quality in the budget constraint that leads to rising marginal costs of quality with respect to family size; this generates a tradeoff between quality and quantity.[1] But is this tradeoff real? In particular, is it true that having a larger family has a causal effect on the "quality," or IQ, of the children? Or is it the case that families who choose to have more children are (inherently) different, and the children would have lower IQs regardless of family size?

This paper uses a data set on the male population of Norway to examine the effect of family size on children's IQ, an outcome not previously available in data sets of this size. Importantly, we also address the issue of the endogeneity of family size. Until recent years, the empirical literature on the effects of family size on child outcomes generally relied on ordinary least squares (OLS) estimation and found a negative relationship between family size and child "quality" (usually education), even after controlling for socioeconomic factors.[2] However, few of these findings can be interpreted as causal; family size is endogenously chosen by parents and hence may be related to other, unobservable parental characteristics that affect child outcomes.[3] Also, the absence of information on birth order often means that birth order effects are confounded with family-size effects. We are unaware of any studies of IQ that have attempted to deal with both of these issues.

There is, however, a literature on the causal effects of family size on child educational attainment, starting with Rosenzweig and Wolpin (1980). Recent papers in this literature use arguably exogenous variation in family size induced by the sex composition of children (parents have a preference for variety and so are more likely to have an additional child if the first two are the same sex) and/or the birth of twins (resulting in a family-size increase of two when one was expected) and have generated very little evidence for any quantity-quality tradeoff (Conley and Glauber 2006; Caceres-Delpiano 2006; Angrist et al. 2006). In recent work (Black, Devereux, and Salvanes 2005), we use Norwegian data on cohorts born between 1912 and 1975 and variation in family size induced by the birth of twins to examine the effects of family size on education and earnings. We find little evidence for any family-size effect either in the OLS or IV estimates. In this paper, we can study a more recent set of cohorts—all individuals in our sample were born between 1967 and 1987—and are able to introduce previously unavailable information on IQ and birth endowments including birth weight.

---

1. Rosenzweig and Wolpin (1980) explicitly derive the assumptions under which an exogenous increase in family size should have a negative effect on child quality.
2. See Blake (1989) and the numerous studies cited therein.
3. There is also a literature examining the effect of family size on parental outcomes. See, for example, Bronars and Grogger (1994) and Angrist and Evans (1998).

Given the two types of interventions the instruments represent—one (sex composition) is a planned increase in family size based on parental preferences for variety in the sex composition of their children and the other (twin births) is an unplanned shock to family size resulting in two generally lower birth weight children with zero spacing—it is somewhat surprising that researchers so far have found similar effects of family size on child outcomes using the twins and sex composition instruments (for example, Angrist et al. 2006). In contrast, our estimates differ depending on estimation method. OLS estimates using a rich set of controls suggest that there is no strong relationship between family size and IQ. Likewise, IV estimates using sex composition as an instrument show no significant negative effect of family size (and are precise enough to rule out large negative effects). However, our IV estimates using twins imply that family size has a negative effect on IQ; unexpected shocks to family size resulting from twin births have significant negative effects on the IQ of existing children.

Interestingly, this finding is not consistent with our own earlier work. The question then becomes why we observe a negative effect of increases in family size generated by the birth of twins on IQ when we found little or no effect on education and earnings in our earlier paper. When we examine this issue, we find that this difference is largely due to the different cohorts studied; when education is studied on a more recent cohort, we find family-size effects similar to those of IQ.

Unlike the papers discussed above, we also have information on birth weight for all individuals in our sample. We use this information to address the recent critique of twin studies by Rosenzweig and Zhang (2006), who suggest that much of this literature is biased away from finding negative effects of larger families on child outcomes due to endowment-reinforcing parental investments in children. Importantly, we find that taking account of endowments actually reduces the negative effects of family size that we find using our basic two stage least squares (2SLS) specification, a result that is consistent with compensating rather than reinforcing parental investments.

The paper unfolds as follows. Section II describes the data we use and Section III discusses our empirical strategy. Section IV presents our results and Section V discusses the internal and external validity of our estimates as well as potential mechanisms. Section VI compares the estimates with previous estimates from Norway and Section VII concludes.

II. Data

Our primary data source is the birth records for all Norwegians born in the period 1967 to 1998 obtained from the Medical Birth Registry of Norway. All births, including those born outside of a hospital, are included as long as the gestation period was at least 16 weeks.[4] The birth records contain information on year and month of birth, birth weight, gestational length, age of mother, and a range

---

4. The data also include stillbirths, which constitute approximately 15 per 1,000 births. We exclude these from the sample.

of variables describing infant health at birth. In these data, we are also able to identify twin births. We can measure family size by counting the number of births to each woman. Our sample is composed of families in which the first birth took place during or after 1967 and by 1987. Given that we observe year of birth, we are able to construct family size and indicators for the birth order of each child directly from the births register. In our sample, the distribution of family size is as follows: 10 percent of families have one child, 47 percent have two children, 31 percent have three children, 9 percent have four children, 2 percent have five children, and 1 percent has six or more children.

Using unique personal identifiers, we match these files to the Norwegian Registry Data, a linked administrative data set that covers the population of Norwegians aged 16–74 in the 1986–2002 period and is a collection of different administrative registers such as the education register, family register, and the tax and earnings register. These data are maintained by Statistics Norway and provide information about educational attainment, labor market status, earnings, and a set of demographic variables (age, gender) as well as information on families.[5]

The IQ data are taken from the Norwegian military records from 1984 to 2005. In Norway, military service is compulsory for every able young man. Before entering the service, their medical and psychological suitability is assessed; this occurs for the great majority between their 18th and 20th birthday. IQ at these ages is particularly interesting as it is about the time of entry to the labor market or to higher education.

The IQ measure is a composite score from three speeded IQ tests—arithmetic, word similarities, and figures (see Sundet et al. 2004, 2005; Thrane 1977 for details). The arithmetic test is quite similar to the arithmetic test in the Wechsler Adult Intelligence Scale (WAIS) (Sundet et al. 2005; Cronbach 1964), the word test is similar to the vocabulary test in the WAIS, and the figures test is similar to the Raven Progressive Matrix test (Cronbach 1964). The composite IQ test score is an unweighted mean of the three subtests. The IQ score is reported in Standard Nine (stanine) units, a method of standardizing raw scores into a nine point standard scale with a normal distribution, a mean of five, and a standard deviation of two.[6]

We have IQ scores on about 84 percent of the relevant population of men in Norway.[7] One nice feature of our data is that we can construct family size and both instrumental variables for all persons, including those with missing IQ scores. To examine whether our missing data appears to be systematically related to our variables of interest, we have regressed a dummy variable for whether IQ is missing on family size using both OLS and 2SLS. We find that individuals from larger families are more likely to have missing scores (for example, first-borns from two-child

5. See Møen, Salvanes, and Sørensen (2004) for a description of these data.
6. The correlation between this IQ measure and the WAIS IQ has been found to be 0.73 (Sundet 1988).
7. Eide (2005) examines patterns of missing IQ data for the men in the 1967-87 cohorts. Of those, 1.2 percent died before one year and 0.9 percent died between one year of age and registering with the military at about age 18. About 1 percent of the sample of eligible men had emigrated before age 18, and 1.4 percent of the men were exempted because they were permanently disabled. An additional 6.2 percent are missing for a variety of reasons including foreign citizenship and missing observations. There are also some missing IQ scores for individuals who showed up to the military.

families are 4 percent less likely to have missing scores than first-borns from five-child families), but that this relationship disappears when family size is instrumented with the twin or sex composition instruments.[8] Thus it is unlikely that missing IQ data is causing serious bias to our later estimates.

We exclude families in which no child is born before 1987 as there is no military data available for such families. We use twins to construct our instrumental variables but, as is standard in the literature, we drop twins from our estimating samples. Table 1 presents summary statistics for our sample. We have 388,405 nontwin males in our sample with an average age in 2002 of 25 and an average educational attainment of 12.3 years.

III. Empirical Strategy

A. Twins as an Instrument for Family Size

Rosenzweig and Wolpin (1980) first discuss the idea of using twin births as unplanned and therefore exogenous variation in family size. In their model, parents have an optimal number of children N^*. The birth of twins can vary the actual N from the desired N^* and it is this arguably exogenous variation that is used to estimate the effects of family size on child outcomes. Our general estimation strategy is as follows:

(1) $IQ = \beta_0 + \beta_1 FAMSIZE + X\beta_2 + \varepsilon$

(2) $FAMSIZE = \alpha_0 + \alpha_1 TWIN + X\alpha_2 + \upsilon$

In this case, IQ is the IQ of the child and FAMSIZE is the total number of children in the family. Equation 2 represents the first stage of the two-stage least squares estimation, where Equation 1 is the second stage. The controls in X include indicator variables for cohort, age at test, age of mother at birth, age of mother at first birth, years of education of mother, years of education of father, and log birth weight of the child.[9] We also include controls for the IQ of the father for cases for which this is available.[10]

The TWIN indicator is equal to 1 if the nth birth is a multiple birth, and equal to 0 if the nth birth is a singleton. We restrict the sample to families with at least n births and study the outcomes of children born before the nth birth. In practice, we estimate the specification for values of n between two and four. By restricting the sample to families with at least n births, we make sure that, on average, preferences over family size are the same in the families with twins at the nth birth and those with singleton births. Also, by restricting the sample to children born before birth n, we avoid selection problems that arise because families who choose to have

---

8. The 2SLS estimates are all small, statistically insignificant, and vary in sign.
9. Results are insensitive to the inclusion of birth weight in the regressions.
10. We do this by setting father's IQ to zero for cases in which it is missing and then interacting father's IQ with a dummy variable which is one if we have information on father's IQ and zero otherwise. Both these variables are then included in the regression. Father's IQ is typically unavailable for men whose father was born before 1950.

Table 1
Summary Statistics

Variable	Mean (Standard Deviation)
Age in 2002	25
	(6)
Education	12.3
	(1.9)
IQ (stanines)	5.2
	(1.8)
Mother's education	11.3
	(2.5)
Mother's age in 2002	50.5
	(6.2)
Father's education	11.8
	(2.9)
Father's age in 2002	53.6
	(6.7)

Sample size: N = 388,405 male singletons.
Education is calculated for a subsample of 282,855 men aged at least 21 in 2002.

another child after a twin birth may differ from families who choose to have another child after a singleton birth.

Our methodology avoids two possible pitfalls that can arise when twins are used as instruments. First, by focusing on whether a twin birth occurs for the nth birth, we avoid the problem that families that have more births are more likely to have at least one twin birth. Second, by only using children born before a possible twin birth, we are able to avoid the problem that a twin birth both increases family size and shifts downward the birth order of children born after the twins. For example, if a twin birth occurs at the second birth, the next child born is now the fourth-born child rather than the third. Thus, any estimates using children born subsequent to a twin birth will confound family-size effects with birth order effects.

B. Sex Composition as an Instrument for Family Size

It has been well documented that parents have preferences for variety and so are more likely to have a child at parity n if the previous n-1 children are all of the same sex (Angrist and Evans 1998). As a result, if the sex composition is random, families that randomly had two same-sex children are more likely to have a third, all other things equal, making sex composition a potential instrument. We utilize this instrument for $n = 3$ and $n = 4$, and restrict the sample to families with at least n-1 children.[11]

---

11. In previous work (Black, Devereux, and Salvanes 2005), we chose to use twin births as our instrument

*1. Sample of families with at least two children (*n = 3*)*

When $n = 3$, we are studying families with at least two children and the instrument is whether or not the first two children are the same sex. Given that we only have IQ data on boys, the relevant possibilities for the first two children are BB, BG, and GB, where G is for girl and B is for boy. (Note that we are looking at the outcomes of the first and second children; if the first two children are GG, we will not have IQ information on any of the existing children.). One obvious comparison is to compare the IQ of the first-born children in a sample composed of BB and BG families—here the instrument can be interpreted simply as whether the second child is a boy. A second comparison is to compare the IQ of second-born children in a sample composed of BB and GB families—here the instrument can be interpreted as whether the first child is a boy.

We improve the precision of our estimates by using both comparisons in estimation. To do this, we estimate the following equation using the first two children:

(3) $FAMSIZE = \alpha_0 + \alpha_1 ALLBOYS + \alpha_2 2nd + X\alpha_3 + \upsilon$

Here *ALLBOYS* is an indicator for the first two children being the same-sex, and *2nd* is an indicator for whether the child is the second-born. The controls included in *X* are the same ones included when the twins instruments are used.

It is important to note that this approach is invalid if there are preferences for boys or girls in addition to preferences for variety, as then the probability of a second child (and hence selection into the at least two children sample) will depend on the sex of the first child. The evidence suggests no such sex preferences in Norway.[12]

*2. Sample of families with at least three children (*n = 4*)*

In the case of families with at least three children, our instrument equals one if all of the first three children were boys. We compare the boys in these families to the boys in families where the first two births were either both boys or both girls. In this situation, parents who have a preference for variety either got variety in the third birth (having same sex births prior), at which point they may stop having children, or had another child of the same sex, thereby increasing the probability that they would have another child in order to achieve diversity.

Thus, the only relevant sequences are BBB, BBG, and GGB; no other comparisons are useful because sequences starting with BG or GB are not comparable to those starting with GG or BB. This is because the probability of having a third child differs systematically depending on whether the first two children are the same sex (as is exploited in the $n = 3$ analysis).

---

for family size. However, in the interest of being consistent with the literature, we have included results using sex composition as our instrument as well. We have made significant effort to discuss the assumptions necessary for these both to be valid instruments as well as the appropriate interpretation of the results.

12. There is no statistically significant relationship between the probability of having a second child and the sex of the first-born child despite the enormous sample size of 481,731 families. Note also that we have not included an interaction of second-born with the sex of the first child. We have verified that this interaction term is statistically insignificant in both the first and second stages and the family-size coefficients are not much impacted by its inclusion.

Analogous to the $n = 3$ case, a simple approach is to compare the outcomes of the first two children in BBB families to those in BBG families. Also, assuming no preferences for boys (or girls), an additional comparison is between the outcomes of third-born children in BBB and GGB families.[13] In practice we implement both using families with at least three children whose first two children are the same-sex. The following regression is estimated for the first three children:

(4) $FAMSIZE = \alpha_0 + \alpha_1 ALLBOYS + \alpha_2 2nd + \alpha_3 3rd + X\alpha_4 + \upsilon$

Here *ALLBOYS* is an indicator for the first three children being the same-sex, *2nd* is an indicator for whether the child is second-born, and *3rd* is an indicator for whether the child is third-born.[14]

IV. Family-size Results for IQ

In Table 2, we show the mean IQ and the distribution of IQ scores by family size. There are two clear patterns. First, only children have lower scores than the average child in two-child families. Second, from family sizes of two to five, we see a monotonic relationship that greater family size accompanies lower average scores.

The unconditional relationship between family size and IQ is only suggestive; for example, it could simply represent cohort effects or birth order effects.[15] To better understand the relationship, we run regressions by birth order and regress IQ of children on family-size dummies, presenting estimates from regressions both with and without additional controls to test sensitivity. As described in Section III, these controls include indicator variables for cohort, age at test, age of mother at birth, age of mother at first birth, years of education of mother, years of education of father, father's IQ, and log birth weight of the child.

The results are presented in Table 3. The omitted category in each case is Family with five or more children. The estimates suggest that the highest achieving children come from two- and three-child families, with lower achievement for only children and children from families with four or more.

However, the results are not robust to introducing additional control variables. With the control variables, almost nothing is statistically significant and the estimates are much smaller than before. There still appears to be a penalty to being an only child, but there is little indication that family size matters otherwise. The fact that adding family background variables and log(birth weight) as controls largely elim-

---

13. In fact, in our sample, families are slightly more likely to have a third child if the first two are both boys than if the first two are both girls so completed family size is higher if the first child is a boy (this is the opposite of the finding of Dahl and Moretti 2008, for the United States) While this potentially invalidates the comparison of BBB and GGB families, the effect size is tiny and we have verified that excluding GGB families from the estimation has no meaningful effect on the family-size estimate.

14. As with the $n = 3$ case, we could include interactions between gender of each birth and birth order to fully saturate the model. Doing so has little impact on the estimates.

15. Birth order is highly correlated with family size and has been shown to be strongly related to child outcomes in Norway (Black, Devereux, Salvanes 2005, 2007b; Bjerkedal 2007).

Table 2
Family Size and IQ

Family Size	Average IQ	Mother's Education	Father's Education	N
1	5.2	11.1	11.4	28,083
	(1.8)	(2.5)	(2.8)	
2	5.3	11.3	11.8	170,628
	(1.8)	(2.5)	(2.8)	
3	5.2	11.4	11.9	132,929
	(1.8)	(2.5)	(2.9)	
4	5.1	11.2	11.7	41,805
	(1.8)	(2.6)	(2.9)	
5	4.9	11.0	11.5	10,214
	(1.8)	(2.5)	(3.0)	
6 or more	4.9	10.6	11.2	4,746
	(1.8)	(2.4)	(2.8)	

inates the family-size effect suggests that the family-size results may be driven by unobserved family background characteristics; if unobserved factors impacting IQ have similar correlations with family size as the observed factors, the true causal effect of family size is not very negative.

The IV estimates below take a more direct approach to the issue of causality.

A. Results Using Twins as an Instrument for Family Size

The 2SLS estimates are presented in Table 4. We also include, for comparison, two sets of OLS coefficients estimated on exactly the same samples as the 2SLS estimates; the first set includes controls only for birth order, the second set includes all control variables. As was the case in Table 3, the OLS estimates without controls suggest negative effects of increased family size, but those with controls never provide any evidence for negative family-size effects. Instead they provide very precisely estimated family-size coefficients that are very close to zero.

The first-stage coefficients using the twin instruments are very strong and suggest that a twin birth increases completed family size by about 0.7 to 0.9. These first-stage effects are a little larger than we found in our 2005 paper, presumably because family sizes are smaller for these more recent cohorts.[16] As expected, twins at higher parity have a larger effect on family size, presumably because they are more likely to push families above N^* (the optimal number of children). The t statistics from

---

16. The first-stage coefficients are also much bigger than those of Angrist (2006) using Israeli data. For twins at second birth, they found a first-stage coefficient of only 0.44. This presumably reflects the much larger proportion of large families in the Israeli data.

Table 3
Effect of Family Size on Children's IQ
OLS

	No Controls				Controls			
	First Child	Second Child	Third Child	Fourth Child	First Child	Second Child	Third Child	Fourth Child
One-child family	0.099*				−0.037			
	(0.031)				(0.028)			
Two-child family	0.274*	0.222*			0.034	−0.015		
	(0.029)	(0.031)			(0.026)	(0.028)		
Three-child family	0.256*	0.236*	0.184*		0.052*	0.037	−0.055	
	(0.030)	(0.031)	(0.035)		(0.026)	(0.028)	(0.032)	
Four-child family	0.163*	0.144*	0.093*	0.063	0.051	0.037	−0.028	−0.045
	(0.032)	(0.034)	(0.039)	(0.046)	(0.029)	(0.031)	(0.035)	(0.044)
N	201,789	134,692	42,042	7,879	201,789	134,692	42,042	7,879

* indicates statistical significance at the 5 percent level. Standard errors (in parentheses) allow for correlation of errors within family. Each column represents a separate regression. The regressions with controls include indicators for age, test year, mother's age, mother's age at first birth, mother's education, father's education, father's IQ, and the log of birth weight of the children studied.
Omitted Category is five or more children.

Table 4
Effect of Family Size on IQ of Children
Instrumental Variable Estimation

	OLS (no controls)	OLS (controls)	First Stage	Second Stage	N
Instrument: twin at second birth			0.717*		173,706
(Sample: First child in families with two or more births)			(0.019)		
Number of children in family	−0.052*	0.006		−0.149*	
	(0.005)	(0.005)		(0.052)	
Instrument: twin at third birth			0.839*		135,970
(Sample: first and second children in families with three or more births)			(0.018)		
Number of children in family	−0.082*	−0.002		−0.170*	
	(0.008)	(0.007)		(0.052)	
Instrument: twin at fourth birth			0.903*		44,695
(Sample: first, second, and third children in families with four or more births)			(0.036)		
Number of children in family	−0.046*	0.009		−0.115	
	(0.012)	(0.010)		(0.080)	
Instrument: first two children same sex			0.082*		304,755
(Sample: first and second children in families with two or more births)			(0.003)		
Number of children in family	−0.045*	0.014*		0.065	
	(0.004)	(0.004)		(0.074)	

(continued)

Table 4 (*continued*)

Instrument: first three children same sex (Sample: first, second, and third children in families with three or more births and first two births same-sex)		0.070* (0.005)	94,578
Number of children in family	−0.077* (0.010)	0.008 (0.009)	0.076 (0.171)

* indicates statistical significance at the 5 percent level. Standard errors (in parentheses) allow for correlation of errors within family. N represents number of individuals. The regressions also include indicators for age, test year, mother's age, mother's age at first birth, mother's education, father's education, father's IQ, and log birth weight of the children studied.

the first stage are typically around 40, indicating that there are no concerns about weak instruments in this application.

The 2SLS estimate of the effect on the first child of changes in family size induced by the second birth being a twin birth is −0.15 (0.05). This implies an adverse effect of increased family size on IQ—each additional child reduces IQ of the first child by about 15 percent of a stanine or 8 percent of a standard deviation. The equivalent estimate for families that have at least three births is −0.17 (0.05), and for families that have at least four births it is −0.12 (0.08). Taken together, these three estimates are all more negative than the OLS estimates (the first two are statistically different from OLS) and suggest fairly sizeable negative effects of family size on IQ of about 15 percent of a stanine for each additional child (this is about 1.2 IQ points on a standard IQ scale). This translates into approximately a 1 percent difference in annual earnings as an adult.[17] Unexpected random shocks to family size, induced by the birth of twins, lead to poorer IQ outcomes for existing children.

B. Results Using Same-Sex as an Instrument for Family Size

For the case where the instrument is an indicator of whether the first two children are the same sex, the first-stage estimate on *ALLBOYS,* also reported in Table 4, is 0.081 with a standard error of 0.003. Similarly, Angrist et al.(2006) find a coefficient of 0.073 (0.010) using Israeli data and Conley and Glauber (2006) find that having two same-sex children increases the probability of having a third child by about 0.07 in the United States.

For the case where the instrument is whether the first three children are the same sex, the first-stage estimate on *ALLBOYS* is 0.070 with a standard error of 0.005. While not directly comparable, this is smaller than the Angrist et al. (2006) finding of 0.120 (0.015) using the sample of families with three or more children in Israeli data.

In sharp contrast to the estimates using twins as an instrument for family size, the 2SLS estimates using the sex composition as instruments are both positive. For the sample of families with two or more births the estimate is 0.07 (0.07), and for the sample of families with three or more births the estimate is 0.08 (0.17). While these family-size coefficients are not precisely estimated, the first one does rule out large negative effects of family size on IQ (effects greater than −0.08 in absolute terms, which maps into 0.6 of a point on the standard IQ scale). The contrast with the twins estimates may result from the fact that, in this case, induced increases in family size are expected and result in normally spaced singletons. However, while the estimates are very different in magnitude, the sex composition IV estimates are not statistically significantly different from the family-size estimates using twins – confidence intervals for the twins at second birth and third birth instruments are (−0.251, −0.047) and (−0.272, −0.068) respectively.

---

17. We arrive at this number by regressing log earnings in 2002 on IQ scores for individuals aged at least 30 in 2002. For both cross-sectional and family fixed effects specifications we obtain precisely estimated coefficients on IQ of about 0.06, suggesting an extra stanine in IQ (about half a standard deviation) leads to about 6 percent higher earnings.

V. Internal and External Validity of the Instruments

A. *Internal Validity of Instruments*

1. Randomness

In order for our IV estimates to be consistent, it must be that the instrument is uncorrelated with the error term in Equation 1. One concern is that the occurrence of a twin birth may not be random and may be related to unobservable family background characteristics. By definition, this is untestable, but we do examine whether the probability of twins is related to observed characteristics such as mother's and father's education and father's IQ. To do so, we estimate three linear probability models of the probability of a twin birth at the Nth birth (for $N = 2$, 3, and 4), conditional on having at least N births, using the full set of control variables. Over these three regressions, of the nine coefficients on mother's education, father's education, and father's IQ, none is statistically significant.[18] Given the enormous sample sizes, these results strongly suggest that twinning probabilities are not related to parental characteristics.

A concern might be that, in recent years, fertility treatments have led to an increase in the incidence of multiple births, and these multiple births could then be related to characteristics of the parents. However, this phenomenon did not take off in Norway until the late 1980s so, as can be seen in Figure 1, the probability of a multiple birth is fairly constant over time for the cohorts we are studying (persons born 1967–87).

We use a similar approach to examine the randomness of the sex composition instruments. Once again, we find no statistically significant relationship between any of the family background characteristics and the instrument for either the sample of families with at least two children, or the sample of families with at least three children where the first two children are the same sex.

In addition to examining the relationship between the instrument and the observable characteristics of the family, there are other tests we can conduct. For example, if the birth of twins is truly random, whether the second birth is a twin birth should have no impact on outcomes of the first child that occur before the second birth occurs. We do not have information on early childhood outcomes but we do know birth weight, so we use 2SLS to estimate the "effect" of family size on the log birth weight of existing children, using the full set of control variables used in the later regressions (obviously birth weight is excluded). If the instrument is random, we should see no effect of family size on the birth weight of existing children. For the twins instruments, we get an effect of family size on the birth weight of existing children of 0.013 (0.005) for twins at second birth, 0.007 (0.005) for twins at third birth, and −0.008 (0.008) for twins at fourth birth. Only the first coefficient is statistically significant and it is not very large (a multiple second birth is associated with the first child being about 1 percent heavier at birth). This probably reflects the

---

18. Twin probabilities increase with parity and maternal age at birth (Jacobsen, Pearce, and Rosenbloom 1999; Bronars and Grogger 1994) but this is not a problem as we condition on these variables in the 2SLS estimation.
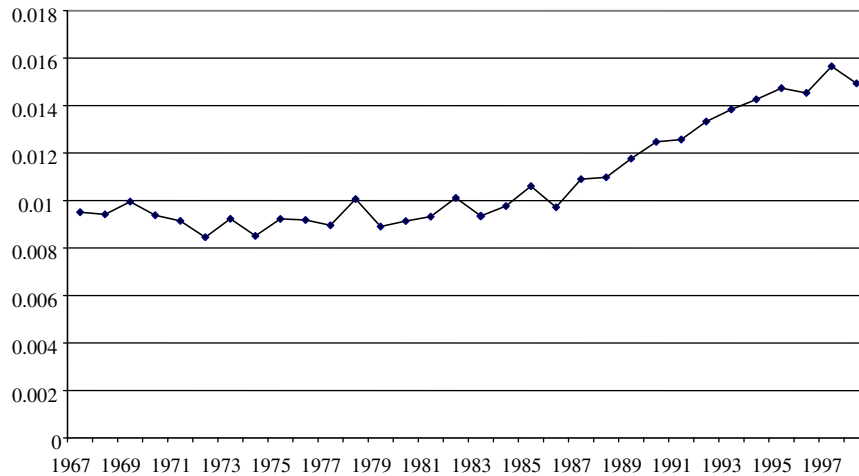
Figure 1
Fraction of Twin Births Out of All Births

fact that larger, healthier women are more likely to have successful twin births and suggests that, in the absence of a control for birth weight, we would have obtained a family-size effect on IQ that was biased toward zero. A value added of our study is that we are able to control for birth weight in the analysis to deal with this possibility.

When we study the effect of the second child being a boy on the log birth weight of the first child, we get an estimate of 0.002 (0.009); the equivalent estimate for the second of our sex composition instruments is 0.036 (0.021). Neither of these is statistically significant. The overall conclusion from this exercise with birth weight is that, with the exception of second-birth twins, there is very little evidence that the instruments are not randomly assigned.

2. Other Threats to Internal Validity

The large first-stage coefficients for the twins instrument $(0.7 - 0.9)$ are reassuring, as they imply that this can be thought of as an experiment with almost perfect compliance—the birth of twins leads to family size increasing by one in almost all cases. This is important, as the major possible biases using twins are that we may be capturing the effects of having closely spaced children, or of having low birth weight children, rather than the effects of having more children. Given that twin births add extra children for almost all families, these possibilities are best seen as threats to the external validity of the twins approach rather than as sources of bias. Later in this section, we consider the external validity issues in more depth.

Because the first-stage coefficients are much lower for the sex composition instruments, the possibility of bias due to direct effects of sex composition on child IQ is greater. Rosenzweig and Wolpin (2000) conjecture that there may be beneficial

effects of having a sibling of the same sex and this would tend to bias away from finding family-size effects using the instrument. The literature is mixed on whether sex composition has direct effects on child outcomes (Dahl and Moretti 2008; Butcher and Case 1994; Conley 2000; Deschenes 2002; all find some evidence of sex-composition effects. Kaestner 1997; Hauser and Kuo 1998; find no evidence of such effects).[19] In our data, we find no evidence for sex composition effects on IQ when we study situations in which the sex composition has no correlation with family size (for example, whether the third child is a boy in families in which the first child is a boy and the second is a girl).

B. External Validity

While it may be reasonable to view both type of instruments as being valid, they constitute very different types of interventions to increase family size. Angrist et al. (2006) discuss the many differences between the twins and sex composition instruments, in particular focusing on how the two instruments affect different parts of the family-size distribution and how the ethnic composition of individuals whose fertility responds to the twins instrument differs from that of individuals who respond to the same-sex instrument. In our case these are not issues, as there is no significant ethnic variation in our sample, and the two instruments affect the family-size distribution in similar ways. Therefore, we concentrate on three different aspects of the twins instrument that do differ from the same-sex one – in the case of twins, the additional family member is smaller (the difference in birth weight between twins and singletons averages about 1,000 grams), there is no space between the births, and the average age of existing children at arrival of the marginal child is younger with twin births.

1. Lower Birth weight of Twins

Recent twins studies (including this one) have estimated the effect of an extra birth at parity N (due to the birth of twins) on the outcomes of children at parities 1 to N-1, for all women who have at least N births. Rosenzweig and Zhang (2006, hence-forth RZ) criticize this approach by arguing that the endowments of twins are lower than those of singletons, and parents may allocate resources differentially across their children on the basis of their endowments. Thus, the effect of adding two twins (with lower birth weights on average) to a family may differ from the effect of adding two singletons.[20]

According to RZ, parents with low birth weight twins are more likely to invest in the heavier first-born child, thereby offsetting the negative effect of increased

---

19. However, the Butcher and Case study found no evidence that sex composition impacted the outcomes of boys.

20. RZ argue that parental investment is likely to reinforce endowment differences and thus bias the literature away from finding negative effects of larger families. In addition, they provide some evidence of reinforcing parental investments using twins data from China. Datar (2005) find evidence that is generally suggestive of reinforcing parental investments among singletons using U.S. data. However, much of the available evidence for developed countries suggests compensatory investment behavior (Behrman, Pollak, and Taubman 1989; Ashenfelter and Rouse 1998; Ermish and Francesconi 2000).

family size on first children. Thus, once we control for birth weight, RZ would predict that the negative effect of increased family size (through twin births) should increase in absolute value. As suggested by RZ, we have tried adding controls for the birth weight of the potential twin birth (using the actual birth weight of a singleton, and the mean birth weight of twins). Interestingly, adding the birth weight controls causes the negative reduced-form effect of twin at second birth on first child IQ to disappear. This is counter to what RZ predict if there were reinforcing investments by parents. Our finding is instead consistent with compensatory investment behavior by parents.

The RZ approach can be thought of as an attempt to create a new experiment in which twins and singletons have the same average birth weight. Unfortunately, in the real world no such experiment is available and their method of conditioning on birth weight may invalidate the exogeneity of twin births. While twinning may be a random event that is orthogonal to family characteristics, birth weight is not. Therefore, conditioning on birth weight creates a correlation between having twins and unobserved family characteristics. For example, families that have twins may be very similar to families that have singletons, but a family that has a 2,000 gram singleton is likely to be different in some ways than a family that has two 2,000 gram twins (for example, the mother may be more likely to have smoked during pregnancy.) Thus, the RZ approach can be seen as an attempt to increase external validity that comes at the cost of reduced internal validity—that is, potential bias.[21] Thus, while our finding here is suggestive that family-size increases through twinning have particularly negative effects on existing children because of compensatory behavior by parents favoring twins, it cannot be conclusive.

2. Birth Spacing

In the case in which we use twins at second birth to study the outcomes of the first child, it may be that the first child is worse off if the second birth is a twin birth because he gets less attention if his siblings are more closely spaced. To examine this, we have looked at nontwin families and estimate the effect of the spacing between later children on earlier children's outcomes. In particular, to simulate the timing of twins, we create a variable for whether the gap between the later children is less than or equal to one year. Also included are the standard set of control variables augmented by a control for family size and a quadratic function of the time gap between children born before the two children whose spacing is being investigated. It is important to note that spacing is not exogenous, so these results should be interpreted with caution.

The estimates are reported in Table 5. The reported coefficients are those relating to the distance in years between the two children immediately after the children listed in Column 1. For example, in the first box, the coefficient of −0.26 suggests

---

21. Recent papers have found that the impacts of birth weight on adult outcomes are generally similar in magnitude in the cross-section to when twin or sibling differences are used in estimation (Black 2007; Oreopoulos 2006). To the extent that twin and sibling differences difference out confounding effects of other factors, this is consistent with the notion that birth weight is not highly correlated with these unobserved factors and hence the bias from the RZ approach may not be large.

Table 5

The Effect of Child Spacing on Previous Children's IQ

	Effect of Gap Less Than or Equal to 1 Year	Effect of Gap Less Than or Equal to 18 months
First child in families of three or more N = 74,996	−0.26* (0.06)	−0.12* (0.02)
First and second child in families of four or more N = 32,772	−0.18 (0.09)	−0.08* (0.03)
First three children in families of five or more N = 9,651	−0.55* (0.17)	−0.19* (0.06)

* indicates statistical significance at the 5 percent level. Standard errors (in parentheses) allow for corre-
lation of errors within family. Each box represents a separate regression. The regressions include indicators
for age, test year, mother's age, mother's age at first birth, mother's education, father's education, father's
IQ, birth order, and the log birth weight of the child studied. Also included is a control for family size
and a quadratic function of the time gap between children born before the two children whose spacing is
being investigated. The reported coefficients are those relating to the distance in years between the two
children immediately after the children listed in column 1. For example, in the first box, the coefficient of
−0.26 suggests that first child IQ is lower by 0.26 if the gap between the second and third child is less
than or equal to one year (in families with at least three children).

that the first child's IQ is lower by 0.26 (about one-quarter of a stanine) if the gap
between the second and third child is less than or equal to one year (in families
with at least three children).

We find that earlier children have lower IQs when later children are more closely
spaced together—children do better if the two children immediately after them are
spaced further apart. If this result can be extrapolated to the case of twins, in which
the space is zero, it implies that the effect of a subsequent case of twins is more
adverse than the effects of having two subsequent singletons.[22] The reduced-form
effect of twins on IQ of earlier born children in Table 4 is about −0.1 for both
twins at second and at third birth. The analogous magnitudes in Table 5 are about
−0.2, suggesting that the 2SLS estimates using twins could easily arise due to
spacing effects and may not be generalizable to a situation in which one adds two
singletons who are spaced apart.

---

22. One caveat here is that, unlike with twins, the spacing of singletons can be chosen by parents and
families that closely space births are different in terms of observables than families that space out births.

3. Age of Existing Children at Arrival of Marginal Birth

A closely related issue is that, in the sample of families with two or more children, generally first-borns are older when a singleton third child is born (7.9 years old on average) than when the third-born child is part of a twin second birth (4.4 years old on average). It is plausible that increased family size has a more negative impact if it occurs earlier in the child's life and this could cause the twins 2SLS estimates to be more negative than OLS or the sex composition 2SLS estimates. There is some evidence that this factor can account for differences in the impacts of the family-size interventions. When one uses the twin at second birth instrument using two different samples (1) where the first child is aged at least four when the potential twin birth occurs and (2) when the first child is aged less than four when the potential twin birth occurs, the 2SLS coefficients on family size are 0.02 (0.07) and −0.26 (0.07) respectively. This is consistent with the idea that the exact timing of the addition of a marginal child may influence the effect on existing children and suggests that some of the differences between the twin and sex composition estimates could be because the extra child arrives "earlier" with twins.[23]

C. Mechanisms through Which Family-size Increases Affect Children

Our results point to an effect of family size that depends on exactly how family size is increased, with evidence of negative effects only from unexpected increases due to twin births. There are many possible reasons why expected increases in family size may not have deleterious consequences on the IQs of existing children. Younger siblings may actually play a beneficial role for children. For example, Zajonc (1976) suggests that older children develop intellectually through having younger children to teach. Also, parents may respond to the expense of an extra child by reducing expenditures on items that are not beneficial to child development such as alcohol or nights out. Additionally, Rosenzweig and Zhang (2006) suggest that there may be an increase in the quantity of household public goods such as internet connectivity when family size increases.

VI. Comparison to Previous Norwegian Estimates of Family-size Effects

In our earlier work using the same methodology but looking at the effect of family size on education and earnings (Black, Devereux, and Salvanes 2005), we found small and generally insignificant effects of family size. The fact that we find significant effects of family size on IQ using the twins instrument in this paper is perplexing, as we would expect an effect on IQ to lead to differences in education and earnings. However, upon further investigation, the differences appear largely due to our use of different cohorts in the two papers. When we do the

---

23. It seems likely that IQ scores are also impacted by investments after the first few years of life as it has been clearly demonstrated that scores on the Armed Forces Qualification Test in the U.S. depend to a large extent on years of schooling of the test taker (for example, by Neal and Johnson 1996).

education analysis on a sample as close as possible to our current sample (individuals aged at least 21 in 2002) and use twin instruments, we find sizeable negative IV effects of family size on the probability of high school graduation (see Table 6).

This discovery led us to consider differences in the effect of family size on IQ over time. In order to increase the number of cohorts in our IQ analysis, we can no longer use the birth records, as that starts in 1967. Using data from the population register (as we did in our 2005 study), we study the effect of family size on IQ for the 1950–84 cohorts.[24]

Using this longer time series data, we have constructed a nine-year moving average (by year of the potential twin birth) of the reduced-form effect of twins at third birth on the IQ scores of first and second-born children.[25] As can be seen in Figure 2, the effect of third-birth twins changed around 1970 from being approximately zero and became quite negative during the 1970s. Clearly, there has been a true change in the effect of twinning as we move from older to more recent cohorts.[26]

As can be seen in Figure 3, there has also been a decline in overall family size over this period. Average family sizes were steady at about 2.7 children for families whose first child was born between 1950 and 1964. However, between 1964 and about 1970, average family size fell precipitously to about 2.4 children and then stabilized at approximately this level for the rest of our sample period. The timing of the family-size change is approximately coincident with the timing of the change in the twins estimates, suggesting these trends may be influenced by a common factor.

One potential factor is changing female labor force participation. In the 1970s, the married women's labor force participation rate increased from about 40 percent to about 60 percent. Thus, changes in the labor force participation of married women in Norway also approximately coincide with the change in the effect of family size using the twins instrument. We are not sure why all of these changes occurred or how they are connected. It is plausible that extra (particularly unexpected) children may be more costly when female labor market participation is higher. Hence, if mothers are working, the effect of twins on existing children may be more negative compared to a situation where mothers were not working.

Another possibility is that falling twin mortality rates over this period may have meant that surviving twin pairs are weaker and hence have a more adverse effect on existing children. However, this does not seem to be an issue as, for the multiple

---

24. In our 2005 study, we did not have access to the birth records. In this new sample, as in our earlier paper, we calculate family size primarily by counting individuals aged 16+ who are present in the population register in 2002. We also obtain information on twin births from the population registry. The twin indicators are unavailable in the population register for all cohorts born after 1984. Given that IQ data are generally available for cohorts starting with those born in 1950, we restrict our data to the 1950-84 cohorts. We exclude individuals not born in Norway from this sample for consistency with the earlier results presented in this paper. We have found our estimates using this methodology are similar but less precisely estimated than those using the information from the birth records. The decline in precision presumably arises because of the greater accuracy of the birth records methodology.

25. We chose twins at third birth because it gives the largest family-size effect in Table 4.

26. Most of the empirical literature has found small effects of family size using either twin or same sex instruments. While it is difficult to be sure, it is possible that our different estimates arise because we use such recent cohorts.

Table 6
Effect of Family Size on Whether Children Finished 12 Years of Schooling Instrumental Variable Estimation

	OLS (No Controls)	OLS (Controls)	First Stage	Second Stage	N
Instrument: twin at second birth (Sample: first child in families with two or more births)			0.726* (0.021)		150706
Number of children in family	−0.030* (0.001)	−0.013* (0.001)		−0.029 (0.015)	
Instrument: twin at third birth (Sample: First and second children in families with three or more births)			0.847* (0.020)		113475
Number of children in family	−0.034* (0.002)	−0.016* (0.002)		−0.039* (0.015)	
Instrument: twin at fourth birth (Sample: first, second, and third children in families with four or more births)			0.900* (0.040)		36927
Number of children in family	−0.016* (0.003)	−0.007 (0.003)		−0.040 (0.025)	
Instrument: first two children same sex (Sample: First two children in families with two or more births)			0.081* (0.003)		260236
Number of children in family	−0.028* (0.001)	−0.012* (0.001)		−0.019 (0.021)	

(*continued*)

Table 6 (*continued*)

Instrument: first three children same sex (Sample: first, second, and third children in families with three or more births and first two births are same sex)		0.069* (0.005)	77410
Number of children in family	−0.028* (0.001)	−0.014* (0.002)	0.023 (0.049)

* indicates statistical significance at the 5 percent level. Standard errors (in parentheses) allow for correlation of errors within family. N represents number of individuals. The regressions also include indicators for age, test year, mother's age, mother's age at first birth, mother's education, father's education, father's IQ, and the log birth weight of the children studied. Children are aged at least 21 in 2002.
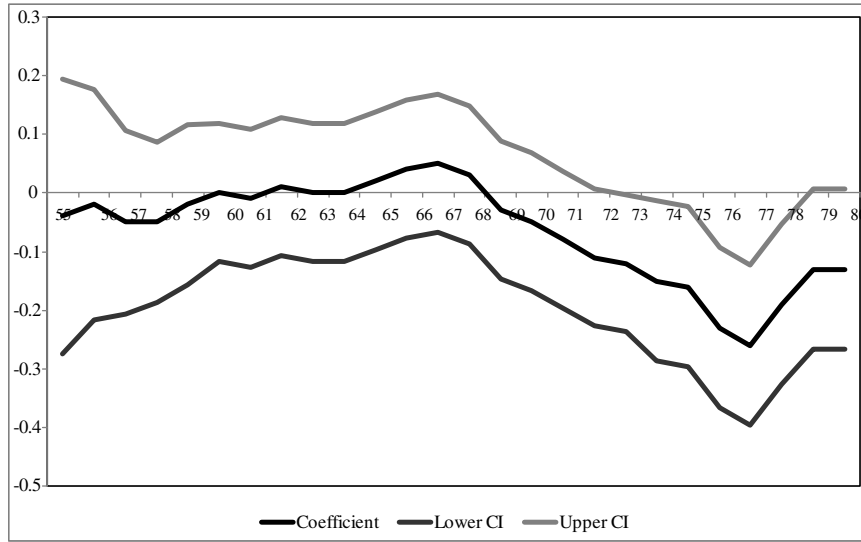
Figure 2
The Reduced-form Effect of Third-Born Twin Births on IQ of First Two Children Nine-Year Moving Average

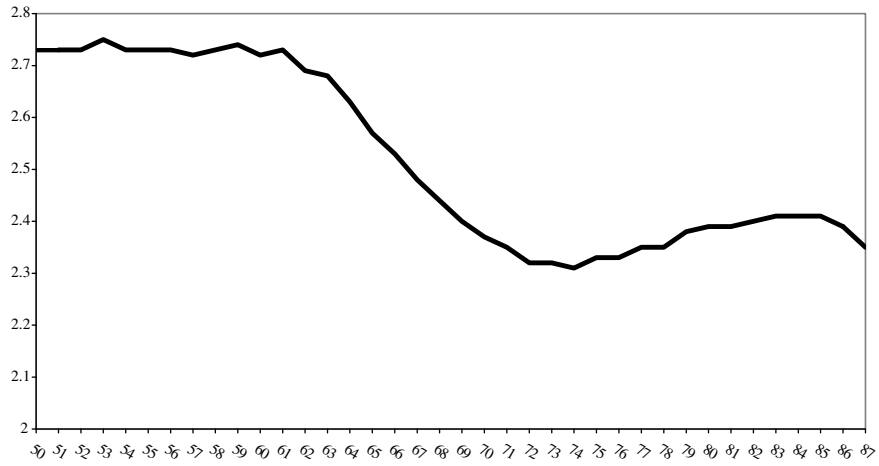The year on the *x*-axis refers to the year of the potential twin birth.

Figure 3
Average Family Size by Year of First Birth (1950–87)

The year on the *x*-axis refers to the year of the mother's first birth.

births in our sample, birth weights generally increased during the decade of the 1970s. The differences between cohorts may also relate to the very different environment concerning contraception and childcare for later cohorts relative to earlier ones. Finally, given our finding that twin births destabilize parental relationships, the increased incidence of family breakdown over time may have led to a greater negative impact of twin births on child outcomes.

VII. Conclusions

Understanding the determinants of a child's cognitive ability is the first step to understanding their longer-run outcomes. In this paper, we examine the role of family size on an individual's IQ, attempting to isolate the causal relationship between the two. Our results point to an impact of family size that depends on exactly how family size is increased. OLS estimates with controls for family background suggest that there is no strong relationship between family size and IQ. Likewise, IV estimates using the same-sex instrument also suggest no negative effect of larger families on IQ. In contrast, our IV estimates using the birth of twins as exogenous variation in family size imply that family size has a negative effect on IQ. However, our investigation of the effects of the zero spacing of twins and their low birth weights suggests that these estimates should not be generalized to other types of variation in family size. Therefore, our conclusion is that there are likely no important negative effect on IQ of *expected* increases in family size that come through normally spaced singleton births.

An intriguing result is that existing children suffer intellectually as a result of the birth of twins and that this effect is only present in recent cohorts. This may reflect that fact that unexpected shocks to family size are becoming increasingly costly as the attachment of both parents to the labor force increases. Because we are able to study an adult outcome (IQ scores around age 18) while using cohorts born as recently as 1987, our quantity/quality estimates are the most "up to date" in the literature. This is potentially important as it is likely that these estimates are more relevant to predicting the impacts of family size on current cohorts than those estimated using much older cohorts that faced very different economic and institutional environments. However, more research will be needed to examine whether this result generalizes to other countries and to children born after our sample period.

References

Ashenfelter, Orley, and Cecilia Rouse. 1998. "Income, Schooling and Ability: Evidence from a New Sample of Identical Twins." *Quarterly Journal of Economics* 113:253–84.

Angrist, Joshua D., and William N. Evans.1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review* 88(3):450–77.

Angrist, Joshua D., Victor Lavy, and Analia Schlosser. 2006. "Multiple Experiments for the Causal Link between the Quantity and Quality of Children." MIT Working Paper 06–26.

Becker, Gary S. 1960. "An Economic Analysis of Fertility." In *Demographic and Economic Change in Developed Countries,* ed. Gary Becker. Princeton, N.J.: Princeton University Press.

Becker, Gary S., and Robert J. Barro. 1988. "A Reformulation of the Economic Theory of Fertility." *Quarterly Journal of Economics* 103(1):1–25.

Becker, Gary S., and H. Gregg Lewis. 1973. "On the Interaction Between the Quantity and Quality of Children." *Journal of Political Economy* 81(2): S279-S288.

Behrman, Jere, Robert Pollak, and Paul Taubman. 1989. "Family Resources, Family Size, and Access to Financing for College Education." *Journal of Political Economy* 97(2):398–419.

Bjerkedal, Tor, Petter Kristensen, Geir A. Skjeret, and John I. Brevik. 2007. "Intelligence Test Scores and Birth Order Among Young Norwegian Men (Conscripts) Analyzed Within and Between Families." *Intelligence*.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. "The More the Merrier? The Effects of Family Size and Birth Order on Children's Education." *Quarterly Journal of Economics* 120:669–700.

———. 2007a. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes." *Quarterly Journal of Economics,* February.

———. 2007b. "Older and Wiser? Birth Order and IQ of Young Men." NBER Working Paper #13237.

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

Bronars, Stephen G., and Jeff Grogger. 1994. "The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment." *American Economic Review* 84(5):1141–56.

Butcher, Kristin F., and Anne Case. 1994. "The Effect of Sibling Sex Composition on Women's Education and Earnings." *Quarterly Journal of Economics* 109(3):531–63.

Cáceres-Delpiano, Julio. 2006. "The Impacts of Family Size on Investment in Child Quality." *Journal of Human Resources* 41(4).

Conley, Dalton, and Rebecca Glauber. 2006. "Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility." *Journal of Human Resources* 41(4).

Conley, Dalton. 2000. "Sibling Sex Composition: Effects on Educational Attainment." *Social Science Research* 29:441–57.

Cronbach, Lee J. 1964. *Essentials of Psychological Testing, 2nd Edition*. London, U.K.: Harper and Row.

Dahl, Gordon, and Enrico Moretti. 2008. "The Demand for Sons." *Review of Economic Studies*. Forthcoming.

Datar, Ashlesha, M. Rebecca Kilburn, and David S. Loughran. 2005. "Health Endowments and Parental Investments in Infancy and Early Childhood." RAND Labor and Population Working Paper No. WR-367.

Deschenes, Olivier. 2002. "Estimating the Effects of Family Background on the Return to Schooling." *Department of Economics, UCSB. Departmental Working Papers*. Paper 10–02. (March 1).

Doepke, Matthias. 2003. "Accounting for Fertility Decline During the Transition to Growth." Working Paper, Los Angeles: University of California at Los Angeles.

Eide, Martha G., Nina Øyen, Rolv Skjærven, Stein Tore Nilsen, Tor Bjerkedal, and Grethe S. Tell. 2005. "Size at Birth and Gestational Age as Predictors of Adult Height and Weight." *Epidemiology* 14:175–81.

Ermish, John, and Marco Francesconi. 2000. "Educational Choice, Families, and Young People's Earnings." *Journal of Human Resources,* 35(1):146–76.

Hauser, Robert M., and Hsiang-Hui Daphne Kuo. 1998. "Does the Gender Composition of Sibships Affect Women's Educational Attainment?" *Journal of Human Resources* 33(3):644–57.

Jacobsen, Joyce P., James Wishart Pearce III, and Joshua L. Rosenbloom. 1999. "The Effects of Childbearing on Married Women's Labor Supply and Earnings: Using Twin Births as a Natural Experiment." *Journal of Human Resources* 34(3):449–74.

Kaestner, Robert. 1997. "Are Brothers Really Better? Sibling Sex Composition and Educational Achievement Revisited." *Journal of Human Resources* 32(2):250–84.

Møen, Jarle, Kjell G. Salvanes and Erik Ø. Sørensen. 2004. "Documentation of the Linked Empoyer-Employee Data Base at the Norwegian School of Economics." The Norwegian School of Economics and Business Administration. Unpublished.

Neal, Derek A., and William R. Johnson. 1996. "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy* 104(5):869–95.

Oreopoulos, Philip, Mark Stabile, Randy Walld, and Leslie Roos. 2006. "Short, Medium, and Long Term Consequences of Poor Infant Health: An Analysis Using Siblings and Twins," NBER Working Paper No 11998.

Rosenzweig Mark R., and Kenneth I. Wolpin. 1980. "Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment." *Econometrica* 48(1):227–40.

———. 2000. "Natural 'Natural Experiments' in Economics." *Journal of Economic Literature* 38(4):827–74.

Rosenzweig Mark R., and Junsen Zhang. 2006. "Do Population Control Policies Induce more Human Capital Investment? Twins, Birthweight, and China's 'One Child' Policy," IZA Working Paper 2082.

Sundet J. M., K. Tambs, P. Magnus and K. Berg. 1988. "On the Question of Secular Trends in the Heritability of IQ Test Scores: A study of Norwegian Twins." *Intelligence* 12:47–59.

Sundet, Jon Martin, Dag G. Barlaug, and Tore M. Torjussen. 2004. "The End of the Flynn Effect? A Study of Secular Trends in Mean Intelligence Test Scores of Norwegian Conscripts During Half a Century." *Intelligence* 32:349–62.

Sundet, Jon Martin, Kristian Tambs, Jennifer R. Harris, Per Magnus, and Tore M. Torjussen. 2005. "Resolving the Genetic and Environmental Sources of the Correlation Between Height and Intelligence: A Study of Nearly 2600 Norwegian Male Twin Pairs." *Twin Research and Human Genetics* 7:1–5.

Thrane, Vidkunn Coucheron. 1977. "Evneprøving av Utskrivingspliktige i Norge 1950–53," Arbeidsrapport nr. 26, INAS.

Zajonc R. B. 1976. "Family Configuration and Intelligence." *Science* 192:227–36.