Carleton College Department of Economics

Working Paper Series

Be Fruitful and Multiply: Religious Denomination as Instrumental Variable in Estimates of the Family Size Effect

by Nathan D. Grawe

No. 2005-05

Department of Economics
Carleton College
One North College Street
Northfield, MN 55057
Telephone: (507) 646-4109
Facsimile Number: (507) 646-4044

November 2005

Thanks to Scott Drewianka, John Heywood, and seminar participants at the University of Wisconsin-Milwaukee for helpful comments. The National Child Development Study data used in this study are maintained by the Data Archive at the University of Essex and are Crown Copyright. Portions of this research were completed with funding assistance from Carleton College. All remaining errors are my own. This paper represents the views of the author and does not necessarily reflect the opinion of Carleton College.

Abstract
Endogeneity resulting from fertility choice inevitably complicates ordinary least squares estimates of the effect of family size on child achievement. Previous studies have addressed this concern by instrumenting with twinning or older-sibling sex composition. These instruments have been criticized for the large sample size required by the former and the possible endogeneity of both. This paper proposes religious denomination (Catholicism, in particular) as an alternative instrument. Using data from the British National Child Development Study, I estimate the effect of family size in cognitive test scores, emotional development measures, and earnings. While ordinary least squares estimates show a quality-quantity trade-off in all three measures of achievement, instrumental variables estimates indicate either zero or positive effects of additional siblings. By contrast, birth order does not consistently affect child achievement.

I. Introduction

Estimation of the family size effect (or quality-quantity trade-off)—the relationship between number of children and subsequent child achievement—is inherently complicated by the endogenous fertility choice which is the foundation of economic models of fertility. A portion of this endogeneity is easily addressed. For instance, if mothers with limited market opportunity choose to have many children due to the relatively low cost of this decision, then intergenerational transmission of market opportunities creates a negative relationship between family size and child performance. Controls for mother's education mitigate the biases caused by observable human capital. Nevertheless, unmeasured components of mother's market ability or other sources of endogeneity may still bias ordinary least squares estimates.

Several existing studies address endogeneity concerns with instrumental variables. To date, the chosen instruments include twinning propensity and the sex composition of early children. The former approach, first proposed by Rosenzweig and Wolpin (1980), requires large sample sizes due to the low frequency of twinning in the general population. For example, Rosenzweig and Wolpin find a quality-quantity trade-off in education using an Indian sample including 25 families with twins. Using a sample of nearly 1.5 million observations (and roughly 20,000 twins) from the Norwegian census, Black et al. (2005) successfully address sample size concerns and report a family size effect which disappears when birth order controls are included.

Twinning is a problematic instrument in two ways. First, large data sets are not available in many countries, especially countries with lower incomes. In addition to limiting empirical work, this challenge hinders application to theory. For instance, theory suggests that credit constraints can evidence themselves in a negative family size effect. (See Becker and Lewis 1973 and Willis 1973.) Because testing this hypothesis requires estimates from many countries with varying degrees of credit market access, the issue is not well-suited to using twinning as an instrument. Moreover, a wide literature on child spacing (and the sense of most parents!) suggests that having twins is distinctly more stressful than typical additions to family size. (See

Black et al. 2005 for evidence consistent with this hypothesis.) If true, this effect produces negative bias in family size estimates.

The second instrument found in existing studies exploits cultural tastes over sex composition. In Western economies, parents whose first two children are of mixed sex are less likely to have a third child than parents with two same-sex children; in Eastern countries parents often continue having children until a boy is born. Black et al. (2005), Conley (2004), Goux and Maurin (forthcoming), Lee (2004), and Qian (2005) apply this instrument the question of family size effects on educational outcomes and find either no effect or a positive effect. However, as noted by Black et al., a substantial literature suggests a direct impact of sex composition on child performance—a result which would compromise the instrument. (For example, see Butcher and Case 1994 and Conley 2000.)

This paper introduces a third possible instrument for family size: religious denomination. Due to church teaching concerning birth control and fertility (most notably in Pope Paul VI's 1968 encyclical letter *Humanae Vitae*), Catholics have traditionally achieved greater family size than other Christian denominations. This instrument is applied to data drawn from the National Child Development Study (NCDS), a panel covering more than 40 years. While the data available for previous studies has limited the examination of family size effects to a narrow time frame and/or to one measure of child achievement, the length and breadth of the NCDS panel permits exploration of family size effects across the life cycle and in a variety of child achievement measures. Moreover, the standardized test scores contained in the NCDS allow for a subtler study of cognitive attainment than that permitted by blunt measures such as years of education, enrollment, or the probability of being held back in school.

In the following section of the paper, the data are described. Section III then explores the validity of the religious denomination instrument in more detail. Section IV presents results which are qualitatively consistent with those reported in other studies using instrumental variables. Specifically, the quality-quantity trade-off observed in OLS estimates for cognitive

performance, emotional development, and adult earnings disappear or are reversed in IV estimates. Unlike existing studies, the magnitude of the positive family size effects found using religious denomination as an instrument are often very large—many times the size of the negative effect estimated by OLS—and in several cases statistically significant. In contrast to the Norwegian results document by Black et al. (2005), the British family size effect is not sensitive to the inclusion of birth order controls. Section V then examines estimates of birth order effects, reporting no consistent evidence of a "pecking order" in Britain. The final section concludes.

II. Data

The National Child Development Study (NCDS) data used in this study are part of a longitudinal panel following all children born in Great Britain during the first full week of March 1958. One advantage of the NCDS is the length and breadth of the survey which permits multiple measures of child achievement across the life cycle. The survey reports standardized math and reading test scores at ages seven, 11, and 16. In addition, standardized test assessments of susceptibility to social adjustment disorders (the Bristol Social-Adjustment Guide total syndrome score or BSAG) were also collected in the first two waves. (A higher value on this emotional development measure indicates that the child is susceptible to a larger number of social syndromes. Thus a low score represents greater "performance".) To facilitate the interpretation of subsequent results, these cognitive and emotional scores are normalized within the full NCDS population to be mean zero with standard deviation of one. In addition, the survey contains earnings data for ages 33 and 41.

While the survey does not report family size directly, when the study subjects were ages seven, 11, and 16 their parents reported the number of older and younger siblings living in the household. The sum of these accurately captures household family size so long as the NCDS subject has no sibling more than 11 years older or 16 years younger than she. In all regressions, family size is defined as the number of children in the family at the time of achievement measurement; completed family size is used only in regressions for age-16 reading and math and

adult earnings.[1] Table 1 presents the distribution of family size at each age within the NCDS data used in this study.[2]

[Table 1 goes here]

Several controls are added to focus attention on the family size effect. First, as noted above mother's human capital may influence both the achievement of the child (through intergenerational ability transmission) and family size (through the opportunity cost of fertility). All regressions include a variable measuring years of education for the mother.[3] Second, as noted by Kessler (1991) and others, birth order is naturally correlated with family size. Black et al. (2005) show that, at least in the case of Norway, this theoretical concern is practically important as the inclusion of birth order controls eliminates family size effects in that study. To avoid attributing birth order effects to family size, I include dummies for birth orders one through six and "seven or more". In addition, a dummy variable for "last born" captures the effects of "relative birth order," distinguishing between being, say, third of three as opposed to third of five children. The coefficient on this last variable may capture the gain for last-born children predicted in models which emphasize the greater access to parent time these children have after older siblings achieve school age (Hanushek 1992). In addition, all regressions contain a dummy variable for whether the individual is a twin or triplet.

Religious affiliation is recorded in the fourth wave of the NCDS at age 23. A religion dummy variable divides the data between Catholics and Protestant/non-religious respondents.[4] Table 2 shows the distribution of religious denominations reported in the NCDS along with the

[1] Results are not sensitive to replacing contemporaneous with ultimate family size in the age-seven and age-11 regressions.
[2] Kolmogorov-Smirnov tests find no evidence that the family size distributions for the population at large differ from those after imposing the sample selection rules used here.
[3] Assuming assortative matching in the marriage market following Becker (1973), father's years of education may also capture some element of mother's ability. Inclusion of this additional control does not meaningfully alter any presented results.
[4] The exclusion of those affiliated with non-Christian religions eliminates one possible source of endogeneity in the religion variable. Non-Christian religious affiliation is highly correlated with race. Thus, racial discrimination may create a direct connection between earnings and religion in these observations. The analysis was repeated excluding all non-religious respondents. The results generally showed the same patterns reported below.

average family size within each religious group. The data confirm that Catholics tend to have larger families (3.97 children, on average, compared to 3.32 for the total Christian/non-religious population).

[Table 2 goes here]

For the purpose at hand, the religious denomination data in the NCDS contains two types of measurement errors. First, it is the religious denomination of the subject's parents which affects the number of siblings. An "inter-generational" measurement error results when an NCDS subject's age-23 denomination differs from her parents' affiliation. Second, it is the parents' denomination during fecundity which influences fertility. An "intra-generational" measurement error occurs if the affiliation of NCDS parents changed over the course of their life so that they raised their children in a different denomination than that which influenced their own fertility choices.

Cases of inter- and intra-generational affiliation switching present two potential problems in the context of this study. First, if affiliation changes are endogenous, IV estimates are inconsistent. This issue is explored in more depth in the following section. However, even if affiliation switches are random with respect to child achievement, adding noise to the instrument might weaken it. Of course, a weak instrument, while consistent, may suffer from considerable bias. While the next section shows that religious denomination is a strong instrument for family size, it may interest readers to know a bit about the frequencies of these inter- and intra-generational measurement errors.

It is not possible to estimate the degree of intergenerational religion switching in the NCDS, but Tomes (1984) studies the issue in a population of working US men ages 25-64. Among Catholics and Protestant/non-religious, he finds that fewer than 10 percent report an affiliation different from that of their upbringing. Given that the NCDS measure of religious denomination is taken at age 23, it seems likely that intergenerational measurement error in the present sample is much smaller than Tomes' estimate based on a much older sample. The NCDS

does provide information on intragenerational affiliation changes—though only in the NCDS subject generation. Table 3 shows that among Christians and non-religious respondents, over 96 percent of reported the same affiliation at the ages of 23 and 41.[5] Given reduced allegiance to religious denominations over time, it seems reasonable to assume that affiliation switching among parents of the NCDS subjects was no greater than that among the subjects themselves. As such, four percent probably represents an upper bound estimate of the degree of intragenerational affiliation switching. In total, the available religious denomination measure appears to be accurate enough for effective use as an instrument.

[Table 3 goes here]

Table 4 presents summary statistics for all variables included in the regression analysis.

[Table 4 goes here]

Section III: The validity of religious denomination as an instrument for family size

The validity of the estimation approach adopted here rests on the two assumptions made of any instrument: significant covariation with the independent variable and exogeneity of the instrument. The former issue is readily set aside. Given only one instrument with one possibly endogenous independent variable, the standard rule of thumb requires a first-stage regression F statistic of roughly 10 or greater. (More precisely, a value of 8.96 ensures that the actual size of a 5% test is no less than 15%; a value of 16.38 ensures the actual size is no less than 10%. See Stock et al. 2002 and Stock and Yogo 2003.) In the regressions reported in the next section, all of the first-stage regression F statistics easily exceed this standard. For the cognitive achievement regressions, the statistics range between 138.97 and 140.99. The F statistics in the first-stage regressions of the emotional achievement (age-seven and age-11) regressions are 130.35 and 135.94. In the earnings regressions the first-stage F statistics are smaller (44.80 and 35.69 for men at ages 33 and 41 and 37.00 and 37.41 for women at ages 33 and 41), but still well beyond

[5] It should be noted, however, that the switching rate is much higher among Catholics. Of those who report Catholic affiliation at age 23, around 20% report Protestant/non-religious affiliation at age 41.

the critical values cited in Stock and Yogo (2003). In sum, there is no doubt that religious denomination and family size share enough common variation to set aside concerns of weak instrument bias.

This focuses the question entirely on the issue of instrument exogeneity. Clearly, affiliation is chosen by individuals and so is, in the broad sense, endogenous. This is not a problem, however, unless the choice is related to our measures of educational, emotional, and market achievement. Broadly speaking, there are two ways in which affiliating with the Catholic Church might be correlated with child outcomes. First, success/failure may increase the propensity for an individual to claim membership in the Church (achievement begets denomination). Second, through discrimination and/or differential work ethics, religious denomination may determine educational attainment, emotional development, and earnings (denomination begets achievement).

Results presented above in Table 3 and in Tomes (1984) suggest that religious denomination is largely a matter of upbringing. Even if affiliation-switching were common, instrumental variables (IV) estimates are consistent so long as the decision to change is uncorrelated with achievement. Tomes (1984) examines the relationship between religious affiliation and earnings in the US and finds that, among Catholics and Protestants, those who switch are not unlike those who maintain a constant affiliation.[6] In the NCDS we can imitate this work, testing whether the average achievements of those who switch affiliations between the ages of 23 and 41 differ from those of the population as a whole.[7] Table 5 reports the mean levels of achievement for life-long Catholics and Protestants/non-religious with those who left each respective affiliation. Columns three and six report test statistics for equal means between the groups that switched as compared with their stable counterparts. The results for earnings match

[6] Tomes does find that those leaving the Jewish faith earn less than the average individual raised in Judaism.
[7] As with the main analysis, the sample includes Christians and non-religious respondents.

the finding in Tomes (1984): religion switching is not significantly correlated with earnings.[8] However, some early childhood achievements are statistically different among those who change affiliations. Two measures (reading at ages seven and 11) are higher among those leaving the Catholic Church than among life-long Catholics. An additional five measures (reading at age seven, math at ages 11 and 16, and both social adjustment scores) show lower performance among those joining the Catholic Church than life-long Protestants/non-religious. Because Table 5 examines changes between age 23 and age 41, they are only suggestive of correlations with switches leading up to the age-23 religion measure which is the instrument in this study. But taken at face value, all seven significant correlations suggest that Catholic affiliation may be correlated with lower performance. If this is the case, then IV estimates of family size effects based on religious affiliation include a negative bias.

[Table 5 goes here]

Setting aside questions of affiliation changes, religious association may yet be endogenous due to religious discrimination or religious difference in attitudes toward work and scholastic achievement. There are reasons to suspect both issues may be relevant. First, anti-Catholic discrimination is institutionalized in British law. The Act of Settlement of 1701 prohibits Catholics from the throne (and so limits the marital choice of royals as well). And while no law prohibits a Catholic from serving as Prime Minister, the Catholic Relief Act of 1939 does prohibit a Catholic from advising the Sovereign on bishop appointments in the Church of England—a responsibility of the Prime Minister. Senior Church of England bishops gain seats in the House of Lords, a body which has experienced significant reduction in legislative power yet retains a powerfully symbolic authority to veto House of Commons bills. Popular expression of

[8] These results are robust to the inclusion of controls for job tenure, ethnicity, marital status, union/staff association membership, manager and supervisor position, size of firm, type of organization (private firm, nationalized industry, local authority, etc.), job security, and prospects for promotion. (The race/ethnicity controls capture effects of being black or of Asian decent. In particular, no control for Irish ethnicity is included since this may capture discrimination based on religion.) Similarly, the results do not change when the sample is limited to full-time workers.

these institutions is seen in the recent controversy surrounding Prime Minister Blair's possible conversion to Catholicism. The propensity for such institutional discrimination to spill over into the labor market is widely discussed in the context of Ireland, but is undoubtedly a concern even in Great Britain. To the extent that Catholics have suffered under discrimination which reduced their earning opportunities, IV estimates of family size effects are biased toward finding a negative impact of additional siblings.

Even absent discrimination, Catholic attitudes toward education and work may differ from those of Protestants. Of course, this hypothesis has a long lineage which dates back to Weber's *The Protestant Ethic and the Spirit of Capitalism*. If, as Weber argues, Calvinist roots drive the modern Protestant to greater educational and economic achievement in an effort to demonstrate divine selection, then religious denomination directly effects earnings and so cannot instrument for family size. In particular, the IV estimates would once more include a downward bias which predisposes us to find a negative family size effect. [9]

In sum, the fact that very few individuals change religious denomination suggests it is largely determined by upbringing. There is some evidence that the very small number of individuals who change identification away from Protestant/non-religious (Catholic) affiliation are more likely to be low (high) childhood achievers. This along with anti-Catholic discrimination and a "Protestant work ethic" would negatively bias IV estimates. As a result, the IV estimates in the next section should be viewed as lower bounds.

Section IV: The effect of family size

The family size effect estimates using ordinary least squares (OLS) and IV estimation are reported in Table 6. [10] To examine the role of birth order in creating the family size effect, the

[9] Gruber (2005) finds that religious participation in general increases earnings. In itself, it is not clear how this finding impacts the denominational instrument used here.

[10] To check whether the impact of different family structures drives these results, the analysis was repeated using only those individuals whose parents were both in the household at ages seven, 11, and 16. Roughly 25 percent of the sample is eliminated by this restriction. The results are robust to this change in sample

analysis is done both excluding (columns 1 and 3) and including (columns 2 and 4) controls for birth order. Column 1 shows that when birth order controls are excluded OLS predicts large, negative impacts of additional siblings. On average, each sibling reduces cognitive test scores by roughly 1/10th of a standard deviation and increases susceptibility to social syndromes by more than 1/20th of a standard deviation. In adulthood, these losses contribute to an earnings reduction of roughly 2.5% (except for young women who experience and even larger loss of more than 4%).

[Table 6 goes here]

Unlike the results reported by Kessler (1991) and Black et al. (2005) for the US and Norway respectively, the family size estimates in the NCDS are robust to the inclusion of birth order controls. In fact, the estimates reported in column 2 show an even larger negative effect of siblings in 8 of the 12 regressions. And 9 of the12 regressions remain significant at the 1% level with one more each significant at the 5% and 10% levels. The only evidence suggesting that the inclusion of birth order controls explains away the family size effect is found in the age-41 earnings of men. This fact underscores the value of observing a diverse set of achievement measures drawn from across the life cycle.

With or without birth order controls, the IV estimates consistently find a family size effect that is less negative than OLS—so much so that in 9 of the 12 cases with birth order controls the estimated effect is actually positive including all four earnings regressions. (The number of positive estimates is reduced to 6 when birth order controls are excluded.) In five cases, the difference between OLS and IV estimates are large enough to reject equality using the test proposed in Wu (1973). (This is true both with and without birth order controls.) Importantly, both earnings regressions for men show statistically significant differences between OLS and IV estimates. Looking at the IV estimates without reference to the OLS results, when

---

selection. The results are also qualitatively unchanged when the birth order dummies are replaced by a first-born control.

birth order controls are included there are three statistically significant family size coefficients. All are positive. Again, both earnings regressions for men fall in this category.

Not only are the family size effects for male earnings statistically significant, they are practically large as well. The addition of one sibling predicts an earnings increase of 21.7% at age 33 and 57.6% at age 41 (when calculated at the mean level of earnings for both age groups).[11] These differences correspond to increases of 0.45 and 0.68 standard deviations within the population as a whole. Remembering that potential sources of endogeneity in religious denomination discussed in the previous section produce a downward bias in the IV estimates, these large, positive effects can be viewed as lower bounds on the true effect. While existing theory has emphasized the rival forces of substitution between siblings, these results suggest a need to consider sibling complementarity.

Section V: The effect of birth order

As noted above, the addition of birth order controls does not eliminate or even mitigate estimated family size effects. We next turn to the question of birth order effects in their own right. Tables 7 and 8 present the birth order coefficients estimated by OLS and IV respectively. Unlike the consistent results for family size, no regular birth order patterns are apparent. A few regressions warrant brief comment.

In Table 7 the only regressions showing systematic birth order effects are those for the age-seven achievements. All three show the same pattern: higher birth order children do better than the first born (and, if point estimates are taken at face value, better than lower birth order siblings) with the exception of the last child who suffers from a significant "last-born" deficit. Other achievements (age-16 cognitive and emotional achievement and age-41 earnings for women) show part, though not all, of this pattern.

[Table 7 goes here]

[11] When the six birth-order dummies are replaced with a dummy for first-borns, the results are qualitatively similar, but less dramatic (and more believable) in magnitude. At age-33, each additional sibling produces a 12.7% gain; at age-41 the gain is 29.7% per sibling.

The IV estimates in Table 8 are far more consistent. All achievements except math at ages 11 and 16 show a negative effect of birth order with a "last-born" premium—a mirror image of the effect found in several of the OLS regressions. This pattern across birth order is statistically significant in age-seven math and both male earnings regressions. Given the paucity of statistically significant coefficients it would not be prudent to make strong claims on these results. But it bears noting that the pattern in the IV results roughly conforms with the theoretical prediction in Hanushek (1992) based on children competing for parent time. (That said, it is equally compatible with sociological models of favoritism toward first and last-borns.)

[Table 8 goes here]

Section VI: Conclusion

The predictions drawn from the economics of the family are inherently difficult to test empirically due to the endogeneity in fertility which is the basis for the very same models. This paper explores the family size effect in Britain using a novel instrument: Catholic affiliation. Extensive panel data drawn from the National Child Development Study (NCDS) provide measures of cognitive, emotional, and market achievement across the life cycle. Whereas ordinary least squares estimates consistently show a negative effect of additional siblings, the instrumental variables estimates show a large, positive effect. These results are generally consistent with recent work done in Norway by Black et al. (2005) using twinning as an instrument. However, unlike Black et al., in the NCDS birth order does not consistently matter in itself and the inclusion of birth order controls does not notably alter estimates of the family size effect.

References

Becker, Gary S. "A Theory of Marriage: Part I," *Journal of Political Economy*, Jul./Aug. 1973, 81(4): 813-846.

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

Black, Sandra E.; Paul J. Devereux; and Kjell G. Salvanes. "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education," *Quarterly Journal of Economics*, May 2005, 120(2): 669-700.

Butcher, Kristin F. and Anne Case. "The Effect of Sibling Sex Composition on Women's Education and Earnings," *Quarterly Journal of Economics*, Aug. 1994, 109(3): 531-563.

Conley, Dalton. "Sibling Sex Composition: Effects on Educational Attainment," *Social Science Research*, Sept. 2000, 29(3): 441-457.

Conley, Dalton. "Parental Education Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility," http://homepages.nyu.edu/~dc66/pdf/sibship_size.pdf, 2004.

Goux, Dominique and Eric Maurin. "The Impact of Overcrowded Housing on Children's Performance at School," *Journal of Public Economics*, forthcoming.

Gruber, Jonathan. "Religious Market Structure, Religious Participation, and Outcomes: Is Religion Good for You?" NBER Working Paper 11377, May 2005.

Hanushek, Eric A. "The Trade-off between Child Quantity and Quality," *Journal of Political Economy*, Feb. 1992, 100(1): 84-117.

Kessler, Daniel. "Birth Order, Family Size, and Achievement: Family Structure and Wage Determination," *Journal of Labor Economics*, Oct. 1991, 9(4): 413-426.

Lee, Jungmin, "Sibling Size and Investment in Children's Education: An Asian Instrument," IZA Discussion Paper No. 1323, http://ideas.repec.org/p/iza/izadps/dp1323.html, Sept. 2004.

Qian, Nancy. "Quantity-Quality: The Positive Effect of Family Size on School Enrollment in China," http://www.econ.brown.edu/fac/Nancy_Qian/Papers/qq_pap&tab_1017.pdf, 2005.

Rosenzweig, Mark R. and Kenneth I. Wolpin. "Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment," *Econometrica*, Jan. 1980, 48(1): 227-240.

Stock, James H.; Jonathan H. Wright; and Motohiro Yogo. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business and Economic Statistics*, Oct. 2002, 20(4): 518-529.

Stock, James H. and Motohiro Yogo. "Testing for Weak Instruments in Linear IV Regression." Unpublished working paper, http://ksghome.harvard.edu/~JStock/pdf/rfa_6.pdf, 2003.

Tomes, Nigel. "The Effects of Religion and Denomination on Earnings and the Returns to Human Capital," *Journal of Human Resources*, Autumn 1984, 19(4): 472-288.

Willis, Robert J. "A New Approach to the Economic Theory of Fertility Behavior," *Journal of Political Economy*, Mar.-Apr. 1973, 81(2, part 2): S14-S64.

Wu, De-Min. "Alternative Tests of Independence between Stochastic Regressors and Disturbances," *Econometrica*, Jul. 1973, 41(4): 733-750.

Table 1: Family size distribution

	Family size at age…					
	seven		eleven		sixteen and older	
	(count)	(percent)	(count)	(percent)	(count)	(percent)
1	683	8.4	656	7.5	658	7.3
2	2903	35.6	2791	31.7	2735	30.2
3	2141	26.3	2291	26.0	2347	25.9
4	1252	15.4	1483	16.9	1577	17.4
5	554	6.8	734	8.3	792	8.7
6	302	3.7	408	4.6	463	5.1
7	146	1.8	198	2.3	225	2.5
8	94	1.1	124	1.4	144	1.6
9	43	0.5	73	0.8	70	0.8
10	21	0.3	29	0.3	32	0.4
11	10	0.1	9	0.1	14	0.2
12	2	0.0	2	0.0	5	0.1
13	2	0.0	2	0.0	2	0.0

Note: These distributions are not for the full NCDS population. Sample selection criteria: observation must include data on mother's education, show Christian or no religious denomination, and contain at least one of the child achievement measures.

Table 2: Average family size by religious denomination

	Frequency	Mean family size
No religion	3729	3.39
Non-denominational	333	2.95
Catholic	994	3.97
Church of England	3172	3.14
United Reformed Church	39	2.57
Baptist	78	3.11
Methodist	214	2.98
Other Christian	508	3.16

Note: These distributions are not for the full NCDS population. Sample selection criteria: observation must include data on mother's education and contain at least one child achievement observation.

Table 3: Age-23 and age-41 religious denomination in the NCDS sample

		Age-23 affiliation	
		Protestant/non-religious	Catholic
Age-41 affiliation	Protestant/non-religious	8,117 (86.4%) [98.7%]	254 (2.7%) [21.7 %]
	Catholic	106 (1.1%) [1.3%]	915 (9.7 %) [78.3 %]

Note: Percent in parentheses reports the fraction of the entire sample. Percent in brackets reports the fraction conditional on age-23 affiliation.

Table 4: Summary statistics for regression variables

	Mean	Standard deviation
Age-33 log earnings		
Men	9.38	0.44
Women	8.57	0.77
Age-41 log earnings		
Men	9.72	0.67
Women	8.92	0.82
Age-7 reading score	0.07	0.95
Age-11 reading score	0.07	0.98
Age-16 reading score	0.09	0.94
Age-7 math score	0.03	0.98
Age-11 math score	0.08	0.99
Age-16 math score	0.07	0.99
Age-7 social adjustment score	-0.06	0.96
Age-11 social adjustment score	-0.08	0.95
Ultimate family size	3.32	1.70
Birth order=2 dummy	0.34	0.48
Birth order=3 dummy	0.14	0.35
Birth order=4 dummy	0.06	0.23
Birth order=5 dummy	0.02	0.14
Birth order=6 dummy	0.01	0.09
Birth order=7 or higher dummy	0.00	0.05
Last-born dummy	0.26	0.44
Twin dummy	0.02	0.15
Mother’s education (years)	9.97	1.62
Catholic denomination dummy	0.11	0.31

Table 5: Mean achievement and changes in religious denomination between ages 23 and 41

	Mean (standard error) of achievement for those who…					
Achievement measure	Remain Catholic (1)	Changed to Protestant / non-religious (2)	|t statistic| for H0: (1)=(2)	Remain Protestant / non-religious (3)	Changed to Catholic (4)	|t statistic| for H0: (3)=(4)
Cognition measures						
Reading, age 7	0.05 (0.03)	0.23 (0.09)	1.80*	0.11 (0.01)	-0.06 (0.07)	2.64***
Math, age 7	0.13 (0.04)	0.24 (0.11)	0.96	0.06 (0.01)	-0.03 (0.06)	1.36
Reading, age 11	0.03 (0.03)	0.26 (0.12)	1.77*	0.11 (0.01)	0.05 (0.07)	0.86
Math, age 11	0.00 (0.03)	0.17 (0.11)	1.42	0.13 (0.01)	-0.10 (0.06)	3.55***
Reading, age 16	0.02 (0.04)	0.18 (0.11)	1.35	0.12 (0.01)	0.08 (0.06)	0.68
Math, age 16	-0.05 (0.04)	0.04 (0.12)	0.77	0.12 (0.01)	-0.13 (0.06)	3.91***
Emotional measures						
BSAG, age 7	-0.08 (0.03)	-0.14 (0.10)	0.55	-0.11 (0.01)	0.07 (0.07)	2.72***
BSAG, age 16	-0.10 (0.03)	-0.03 (0.11)	0.59	-0.12 (0.01)	0.06 (0.08)	2.24**
Earnings measures						
Men, age 33	9.37 (0.03)	9.43 (0.06)	0.84	9.39 (0.01)	9.44 (0.06)	0.92
Men, age 41	9.73 (0.04)	9.70 (0.11)	0.27	9.70 (0.13)	9.77 (0.06)	1.06
Women, age 33	8.63 (0.05)	8.67 (0.17)	0.21	8.57 (0.02)	8.44 (0.09)	1.38
Women, age 41	9.01 (0.05)	9.11 (0.15)	0.62	8.91 (0.80)	8.93 (0.09)	0.18

Note: The BSAG (Bristol Social-Adjustment Guide total syndrome score) is a standardized test which assesses susceptibility to social adjustment disorders. The greater the score, the more susceptible the child is to social syndromes.
** Statistically significant at 10% level*
*** Statistically significant at 5% level*
**** Statistically significant at 1% level*

Table 6: The effects of family size on various measures of achievement

	Ordinary least squares estimates		Instrumental variables estimates		t statistics for H0: $b_{IV}=b_{OLS}$		Sample size
Achievement measure	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cognition measures							
Reading, age 7	-0.125***	-0.165***	-0.029	0.218	1.84*	1.83*	7965
	(0.007)	(0.017)	(0.054)	(0.223)			
Math, age 7	-0.048***	-0.099***	0.100*	0.523**	2.67***	2.79***	7938
	(0.007)	(0.017)	(0.057)	(0.247)			
Reading, age 11	-0.149***	-0.155***	-0.087*	0.026	1.02	1.29	7524
	(0.006)	(0.012)	(0.048)	(0.144)			
Math, age 11	-0.116***	-0.123***	-0.142***	-0.196	-0.56	-0.51	7521
	(0.006)	(0.011)	(0.049)	(0.144)			
Reading, age 16	-0.161***	-0.166***	-0.100**	-0.003	1.15	1.34	6732
	(0.007)	(0.013)	(0.048)	(0.133)			
Math, age 16	-0.117***	-0.116***	-0.145***	-0.197	-1.13	-0.62	6706
	(0.006)	(0.011)	(0.049)	(0.133)			
Emotional measures							
BSAG, age 7	0.067***	0.118***	-0.036	-0.308	-1.90*	-1.91*	7941
	(0.008)	(0.018)	(0.056)	(0.238)			
BSAG, age 16	0.072***	0.096***	0.019	-0.061	-1.10	-1.10	7523
	(0.007)	(0.013)	(0.051)	(0.150)			
Earnings measures							
Men, age 33	-0.027***	-0.026***	0.063	0.196*	1.93*	2.19**	2204
	(0.006)	(0.009)	(0.041)	(0.112)			
Men, age 41	-0.023***	-0.011	0.146**	0.455**	1.97**	2.42**	2216
	(0.008)	(0.014)	(0.069)	(0.223)			
Women, age 33	-0.043***	-0.027*	0.020	0.094	1.39	0.74	1944
	(0.010)	(0.016)	(0.092)	(0.194)			
Women, age 41	-0.023**	-0.037**	0.073	0.194	1.28	1.07	2320
	(0.011)	(0.017)	(0.089)	(0.229)			
Birth order controls?	No	Yes	No	Yes	No	Yes	

Notes: (1) All regressions include controls for being a multiple (twin or triplet) and for years of mother's education.
(continued on next page)

Notes (cont.)

(2) To avoid differences due solely to different samples, all estimates for a given achievement measure use a common sample which requires observations of birth order, mother's years of education, status as a multiple, and religious denomination. The results are not sensitive to this requirement, however.

(3) The BSAG (Bristol Social-Adjustment Guide total syndrome score) is a standardized test which assesses susceptibility to social adjustment disorders. The greater the score, the more susceptible the child is to social syndromes.

(4) The t statistic for H0: $b_{IV}=b_{OLS}$ is found by the method proposed in Wu (1973) and is asymptotically identical to a Hausman test for equality between ordinary least squares and instrumental variables estimates.

(5) Standard errors, in parentheses, are corrected for heteroskedasticity.

** Statistically significant at 10% level*

*** Statistically significant at 5% level*

**** Statistically significant at 1% level*

Table 7: The effects of birth order on various measures of achievement: ordinary least squares estimates

	Birth position						Last-born	
Achievement measure	2	3	4	5	6	7+	dummy	Sample size
Cognition measures								
Reading, age 7	0.069**	0.133***	0.111	0.136	0.161	0.355**	-0.128***	7965
	(0.030)	(0.048)	(0.070)	(0.101)	(0.138)	(0.173)	(0.034)	
Math, age 7	0.150***	0.163***	0.209***	0.252***	0.417***	0.441***	-0.089**	7938
	(0.032)	(0.048)	(0.069)	(0.093)	(0.127)	(0.165)	(0.035)	
Reading, age 11	-0.035	-0.012	0.016	0.035	0.021	0.358**	-0.011	7524
	(0.027)	(0.041)	(0.061)	(0.092)	(0.141)	(0.172)	(0.032)	
Math, age 11	0.032	-0.003	0.010	0.029	0.088	0.351*	-0.027	7521
	(0.028)	(0.041)	(0.059)	(0.090)	(0.128)	(0.192)	(0.033)	
Reading, age 16	-0.024	-0.017	-0.047	-0.012	-0.043	0.105	-0.063*	6732
	(0.027)	(0.044)	(0.069)	(0.110)	(0.187)	(0.246)	(0.033)	
Math, age 16	-0.005	-0.072	-0.059	-0.112	0.048	0.232	-0.047	6706
	(0.029)	(0.043)	(0.062)	(0.091)	(0.120)	(0.196)	(0.035)	
Emotional measures								
BSAG, age 7	-0.070**	-0.163***	-0.168**	-0.278***	-0.243*	-0.273	0.138***	7941
	(0.031)	(0.048)	(0.074)	(0.097)	(0.137)	(0.185)	(0.035)	
BSAG, age 16	-0.053**	-0.053	-0.142**	-0.087	-0.208	0.029	0.095***	7523
	(0.027)	(0.044)	(0.065)	(0.110)	(0.168)	(0.232)	(0.033)	
Earnings measures								
Men, age 33	0.018	-0.007	0.027	-0.078	0.073	0.044	0.002	2204
	(0.023)	(0.032)	(0.051)	(0.085)	(0.129)	(0.150)	(0.028)	
Men, age 41	-0.013	-0.071	-0.022	-0.292**	0.051	-0.160	0.009	2216
	(0.033)	(0.055)	(0.076)	(0.119)	(0.145)	(0.172)	(0.043)	
Women, age 33	0.030	-0.029	-0.041	-0.267	0.437	0.435	0.080	1944
	(0.045)	(0.064)	(0.095)	(0.141)	(0.351)	(0.324)	(0.052)	
Women, age 41	-0.006	0.042	-0.008	-0.044	0.402**	0.729***	-0.059	2320
	(0.042)	(0.062)	(0.096)	(0.125)	(0.172)	(0.216)	(0.050)	

Notes: (1) All regressions include controls for family size, being a multiple (twin or triplet), and for years of mother's education.
(continued on next page)

Notes (cont.)

(2) The BSAG (Bristol Social-Adjustment Guide total syndrome score) is a standardized test which assesses susceptibility to social adjustment disorders. The greater the score, the more susceptible the child is to social syndromes.

(3) Standard errors, in parentheses, are corrected for heteroskedasticity.

** Statistically significant at 10% level*

*** Statistically significant at 5% level*

**** Statistically significant at 1% level*

Table 8: The effects of birth order on various measures of achievement: instrumental variables estimates

	Birth position						Last-born	
Achievement measure	2	3	4	5	6	7+	dummy	Sample size
Cognition measures								
Reading, age 7	-0.322	-0.635	-1.072	-1.459	-1.831	-2.318	0.477	7965
	(0.229)	(0.450)	(0.691)	(0.932)	(1.163)	(1.560)	(0.353)	
Math, age 7	-0.482*	-1.080*	-1.703*	-2.331*	-2.809*	-3.884*	0.891**	7938
	(0.252)	(0.495)	(0.762)	(1.029)	(1.288)	(1.726)	(0.390)	
Reading, age 11	-0.210	-0.380	-0.560	-0.790	-0.990	-0.911	0.317	7524
	(0.142)	(0.295)	(0.462)	(0.661)	(0.818)	(1.025)	(0.262)	
Math, age 11	0.102	0.144	0.241	0.361	0.494	0.862	-0.159	7521
	(0.142)	(0.295)	(0.461)	(0.664)	(0.814)	(1.025)	(0.262)	
Reading, age 16	-0.176	-0.349	-0.567	-0.742	-0.948	-1.057	0.244	6732
	(0.126)	(0.273)	(0.428)	(0.605)	(0.759)	(0.979)	(0.251)	
Math, age 16	0.071	0.094	0.201	0.253	0.500	0.812	-0.200	6706
	(0.127)	(0.275)	(0.428)	(0.601)	(0.746)	(0.964)	(0.252)	
Emotional measures								
BSAG, age 7	0.365	0.690	1.143	1.494	1.966	2.691	-0.533	7941
	(0.244)	(0.480)	(0.735)	(0.993)	(1.242)	(1.660)	(0.376)	
BSAG, age 16	0.099	0.265	0.356	0.629	0.669	1.129	-0.189	7523
	(0.149)	(0.309)	(0.483)	(0.691)	(0.863)	(1.081)	(0.275)	
Earnings measures								
Men, age 33	-0.186*	-0.442**	-0.703*	-1.047**	-1.234*	-1.593*	0.404**	2204
	(0.104)	(0.219)	(0.364)	(0.498)	(0.667)	(0.849)	(0.202)	
Men, age 41	-0.415**	-0.952**	-1.544**	-2.316**	-2.793**	-3.629**	0.852**	2216
	(0.193)	(0.422)	(0.723)	(0.980)	(1.387)	(1.662)	(0.403)	
Women, age 33	-0.087	-0.281	-0.415	-0.790	-0.191	-0.473	0.315	1944
	(0.193)	(0.405)	(0.605)	(0.852)	(1.067)	(1.522)	(0.383)	
Women, age 41	-0.228	-0.421	-0.723	-1.059	-0.767	-0.732	0.375	2320
	(0.221)	(0.455)	(0.698)	(1.012)	(1.166)	(1.463)	(0.429)	

Notes: (1) All regressions include controls for family size, being a multiple (twin or triplet), and for years of mother's education.
(continued on next page)

Notes (cont.)

(2) The BSAG (Bristol Social-Adjustment Guide total syndrome score) is a standardized test which assesses susceptibility to social adjustment disorders. The greater the score, the more susceptible the child is to social syndromes.

(3) Standard errors, in parentheses, are corrected for heteroskedasticity.

** Statistically significant at 10% level*

*** Statistically significant at 5% level*

**** Statistically significant at 1% level*