Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size

By JOSHUA D. ANGRIST AND WILLIAM N. EVANS *

Research on the labor-supply consequences of childbearing is complicated by the endogeneity of fertility. This study uses parental preferences for a mixed sibling-sex composition to construct instrumental variables (IV) estimates of the effect of childbearing on labor supply. IV estimates for women are significant but smaller than ordinary least-squares estimates. The IV are also smaller for more educated women and show no impact of family size on husbands' labor supply. A comparison of estimates using sibling-sex composition and twins instruments implies that the impact of a third child disappears when the child reaches age 13. (*JEL* J13, J22)

An understanding of the relationship between fertility and labor supply is important for a number of theoretical and practical reasons. First, economists and demographers have developed a variety of models linking the family and the labor market. Empirical studies of childbearing and labor supply are sometimes seen as tests of these models (e.g., Reuben Gronau, 1973; Mark R. Rosenzweig and Kenneth I. Wolpin, 1980b; T. Paul Schultz, 1990). Second, the link between fertility and labor supply might partly explain the postwar increase in women's labor-force participation rates if having fewer children causes an increase in labor-force attachment (Mary T. Coleman and John Pencavel, 1993). Evidence for this thesis includes Claudia Goldin's (1995) study, which shows that few women in the 1940's and 1950's birth cohorts were able to combine childbearing with strong labor-force attachment. Other researchers have also drawn a link between fertility-induced withdrawals from the labor force and lower wages of women (e.g., Gronau, 1988; Sanders Korenman and David Neumark, 1992). So perhaps childbearing keeps women from developing their careers.

Any success in disentangling the causal mechanisms linking fertility and labor supply should shed light on other substantive issues as well. For example, reductions in female labor supply could increase the total time parents devote to child care, making at least some children better off (see, e.g., Frank P. Stafford, 1987; Francine Blau and Adam J. Grossberg, 1992). Some theories of family behavior also suggest that changes in wives' earnings affect marital stability (Becker et al., 1977; Becker, 1985).

Not surprisingly, given the wide and longstanding interest in the connection between childbearing and labor supply, hundreds of empirical studies report estimates of this relationship. The vast majority of these studies find a negative correlation between fertility (or family size) and female labor supply.[1] As

* Angrist: Department of Economics, Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02139; Evans: Department of Economics, University of Maryland, College Park, MD 20742. The authors thank seminar participants at the Council of Economic Advisors, Brown University, the University of Maryland, MIT, the University of Michigan, Northwestern University, Princeton University, the University of Virginia, and the July 1996 NBER Labor Studies meetings, as well as Susan Athey, Peter Diamond, Jon Gruber, Judy Hellerstein, Jim Poterba, David Weil, and three anonymous referees for helpful discussions or comments. John Johnson and Amanda Honeycutt provided excellent research assistance. Special thanks go to Duncan Thomas who stimulated our interest in the subject of parental sex preferences. The authors bear sole responsibility for the content of this paper.

[1] There is less work on the effects of children on husbands' labor supply. See Pencavel (1986 Table 1.17) for a few estimates, which suggest a positive association between fathers' labor supply and the number of children. The relationship between husbands' and wives' labor sup-

noted in two recent literature surveys, however, the interpretation of these correlations remains unclear. In his assessment of the "economics of the family," Robert J. Willis (1987 p. 74) writes, "... it has proven difficult to find enough well-measured exogenous variables to permit cause and effect relationships to be extracted from correlations among factors such as the delay of marriage, decline of childbearing, growth of divorce, and increased female labor force participation ... ." Martin Browning (1992 p. 1435) expresses similar views: "... although we have a number of robust correlations, there are very few credible inferences that can be drawn from them."[2]

Skepticism regarding the causal interpretation of associations between fertility and labor supply arises in part from the fact that there are strong theoretical reasons to believe that fertility and labor supply are jointly determined (see, e.g., Schultz, 1981, or Goldin, 1990). In fact, this endogeneity is reflected in the academic research agenda. On one hand, papers on labor supply often treat child-status variables as regressors in hours of work equations, while on the other hand, economic demographers and others discuss regressions and models that are meant to characterize the impact of wages or measures of labor-force attachment on fertility. Since fertility variables cannot be both dependent and exogenous at the same time, it seems unlikely that either sort of regression has a causal interpretation.[3]

This paper focuses on the causal link running from fertility to the work effort of both men and women. Our main contribution is the use of a new instrumental variables (IV) strategy based on the sibling sex mix in families with two or more children. This instrument exploits the widely observed phenomenon of parental preferences for a mixed sibling-sex composition. In particular, parents of same-sex siblings are significantly and substantially more likely to go on to have an additional child.[4] Because sex mix is virtually randomly assigned, a dummy for whether the sex of the second child matches the sex of the first child provides a plausible instrument for further childbearing among women with at least two children. Moreover, in spite of the fact that the sibling sex mix is obviously a function of the sex of both children, we present a range of evidence which suggests that there is little possibility that any secular impact of the sex of offspring contaminates the IV estimates.

We also compare results generated using sex mix as an instrument to results generated using twins to construct instruments. Twinning at first birth has been used in a number of previous studies to estimate the relationship between

ply is discussed by, among others, Orley Ashenfelter and James J. Heckman (1974), Heckman and Thomas E. MaCurdy (1980), and Jonathan Gruber and Julie Berry Cullen (1996).

[2] The survey by Alice Nakamura and Masao Nakamura (1992) argues that a search for exogenous variation is so difficult it is not even fruitful (pp. 60–61).

[3] In the chapter on models of marital status and childbearing in the *Handbook of Labor Economics*, Mark R. Montgomery and James Trussel (1986 p. 205) note:

> "One of the rites of passage for a labor economist involves the estimation of a Probit model for female labor force participation. It is standard practice for the Probit equation to include some indicators for a woman's marital status and the number and age distribution of her children. In estimating such a model, the labor economist veers dangerously close to a theory of householdforma tion, childbearing, and labor supply; namely, that household formation and fertility can be safely taken as exogenous with respect to a woman's supply of hours."

Many of the papers cited in the *Handbook* chapter on female labor supply (Mark R. Killingsworth and Heckman, 1986) fit this description. A widely cited paper that discusses regressions of fertility measures on measures of earnings is Jacob Mincer (1963). A more recent example is Heckman and James R. Walker (1990), who use nonlinear techniques. See also Schultz (1981 p. 171) who asks: "What is cause and what is effect?" in a discussion of William P. Butz and Michael P. Ward (1979), and other demographic studies of the relationship between fertility and labor-force participation.

[4] Charles F. Westoff et al. (1963) were among the first to report preferences for a mix. In a survey of desired fertility and a follow-up study of actual fertility among couples with two children, they found that parents of two boys or two girls both desired and ultimately had more children than parents of mixed pairs. See Nancy E. Williamson (1976) for an international review. After completing the first draft of this paper in July 1996, we learned of concurrent work using sex-preference instruments to estimate the effect of fertility on female labor supply in the United Kingdom (Maria Iacovou, 1996).

childbearing and labor-market outcomes. Examples include Stephen G. Bronars and Jeff Grogger (1994) and Jaisri Gangadharan and Joshua L. Rosenbloom (1996), both of which use large Census samples, and Rozenzweig and Wolpin (1980a, b). Here, our interest is primarily in the comparison between twins estimates and results using the same-sex instruments. We therefore focus on multiple second births, so that both twinning and the sex-mix instrument identify the impact of moving from the second to the third child. Because third-born children who are twins are older than other third-born children, the juxtaposition of estimates based on twinning and the sex mix allows us to compare the effects of fertility on labor supply when the children are different ages. By combining the twins and same-sex instruments, we can estimate the time it takes for the labor-supply consequences of childbearing to disappear.

Section I discusses the data and the sex-mix instruments' first stage, and briefly describes how sex mix can be incorporated into standard economic models of fertility. Section II presents the main set of empirical results on fertility and labor supply, including an analysis of effects in subgroups defined by husbands' earnings and mothers' schooling. Section III compares the estimates using sex-mix instruments to estimates based on twins. Section IV discusses the empirical findings in light of recent trends in female labor-force participation, and Section V concludes.

I. Data, Descriptive Statistics, and First-Stage Relationships

A. *Data and Descriptive Statistics*

The sex-mix estimation strategy is implemented using information on labor supply, the sex of mothers' first two children, and an indicator of multiple births in the 1980 and 1990 Census Public Use Micro Samples (PUMS). To motivate the empirical work, Table 1 reports labor-force participation rates and the probability of additional childbearing among women aged 21–35 and aged 36–50 in the 1970, 1980, and 1990 PUMS. Data for 1970 are from the 1/100 state file (U.S. Department of Commerce, Bureau of the Census, 1972); data for 1980 and 1990 are from the 5-percent samples (U.S. Department of Commerce, Bureau of the Census, 1983, 1995). The table shows substantial declines in fertility and increases in labor supply in both age-groups. Statistics for women aged 21–35 with at least two children show a similar pattern, as do the statistics for women aged 36–50.

There is no retrospective fertility information in the PUMS data sets other than the number of children ever born. We therefore matched children to mothers within households in a manner similar to that described in the appendix to Angrist and Evans (1996a). Briefly, we attached people in a household labeled as "child" in the primary relationship code to a female householder or the spouse of a male householder. In households with multiple families, detailed relationship codes as well as subfamily identifiers were used to pair children with mothers. We deleted any mother for whom the number of children in the household did not match the reported number of children ever born.[5] Using the sex of the oldest two children, we defined same-sex sibling pairs in both Censuses.

Because the Census does not track children across households, the sample is limited to mothers aged 21–35 whose oldest child was less than 18 years of age at the time of the Census. Few women younger than age 21 have two children, while a child over age 17 is increasingly likely to have moved to a different household. Restricting the women's age-group to less than or equal to 35 means the age 18 cutoff for firstborn children does not generate a highly selected sample. Data from the Fertility, Birth Expectations, and Marital History Supplement to the June 1990 Current Population Survey (CPS) show that among women aged 35 with two or more children, at least 93 percent have an oldest child younger than age 18. This fraction falls to 85 percent at age 36 but is equal to 100 percent for women aged 32 or younger. Although women aged 21–35 with at least two children may appear to constitute an unusually young high-fertility group,

[5] Note also that the sample is restricted to women for whom the reported values of age and sex of their two oldest children were not allocated by the U.S. Bureau of the Census.

TABLE 1—FERTILITY AND LABOR-SUPPLY MEASURES

Sample	1970 PUMS	1980 PUMS	1990 PUMS
Women aged 21–35			
Mean children ever born	1.78	1.27	1.18
Percent with 2 or more children	52.10	40.40	37.60
Percent worked last year	60.00	73.40	79.30
Observations	203,918	1,326,631	1,478,546
Women aged 36–50			
Mean children ever born	2.85	2.86	2.15
Percent with 2 or more children	73.40	78.50	68.90
Percent worked last year	57.30	66.70	78.50
Observations	181,502	852,204	1,253,095
Women aged 21–35 with 2 or more children			
Mean children ever born	3.06	2.61	2.57
Percent with more than 2 children	55.60	39.90	39.10
Percent worked last year	44.80	58.00	66.60
Observations	106,239	535,587	577,397
Married women aged 21–35 with 2 or more children			
Mean children ever born	3.02	2.58	2.53
Percent with more than 2 children	54.90	39.00	37.50
Percent worked last year	41.80	55.80	67.50
Observations	91,286	436,483	439,408

Notes: The 1970 PUMS data are from the 1/100 state file. The 1980 and 1990 data are from the 5-percent PUMS. Calculations from the 1990 PUMS use sample weights. The married samples include women married at the time of the Census.

our tabulations of the June 1990 CPS show that over half of all women aged 28–35 fall into this group. The proportion is lower for women aged 21–27 but still includes at least one-quarter of the entire age cohort.[6]

The empirical analysis is conducted on two subsamples from each Census data set. The first includes all women with two or

[6] It is also worth noting that a substantial fraction of the change in completed family size between 1970 and 1990 occurred at parities greater than 2. Our tabulations of 1970 and 1990 Census data show that about 71 percent of women aged 40–55 in both years had two or more children. The proportion having three or more children, however, fell from about 0.47 in 1970 to 0.39 in 1990.

more children. The second includes only married women because this is the sample that many economic theories of household production (e.g., Gronau, 1973) are meant to describe. The married sample is also used to explore the impact of children on fathers' labor supply.[7] The 1980 married sample is restricted to couples who were married at the time of the Census, married only once, and married at the time of their first birth. There are 394,835 observations in the full 1980 sample and 254,654 observations in the 1980 married sample (64 percent of the total). Information on the timing of first marriage and the number of marriages is not available in the 1990 PUMS, so that the 1990 married sample includes all women who were married at the time of the Census. The full 1990 sample includes 380,007 women and the married 1990 sample includes 301,588 women (79 percent of the total—higher than for 1980 because the 1990 sample-selection rule is less restrictive).

Descriptive statistics and variable definitions for covariates, instruments, and dependent variables are given in Table 2. The covariate of primary interest in our labor-supply models is the indicator *More than 2 children.* The first instrumental variable for *More than 2 children* is the indicator *Same sex.* The table also shows averages for the two components of *Same sex,* the indicators *Two boys* and *Two girls.* Among all women with two children in 1980, 40.2 percent had a third child. The corresponding figure for 1990 is 37.5. In both samples, just over 50 percent of all two-child families had children of the same sex and just over 51 percent of first births were boys.

Labor-supply estimates are also computed using multiple second births to generate instruments. In the 1980 PUMS, multiple births are defined as siblings having the same age and quarter of birth. The mean for this indicator of twin births, which we call *Twins-2,* is 0.0085 in the 1980 full sample and 0.0083 in the 1980 married sample.[8] For purposes of comparison, we drew a sample of all second births born to women aged 21–35 from the 1976 Vital Statistics Natality Data tapes (National Center for Health Statistics). This data set contains a 50-percent sample of all births in the country and should provide an accurate estimate of twinning probabilities for the women in our sample. Data from 1976 offer a useful comparison since roughly 40 percent of second children in our Census sample were born 1976–1979. The vital statistics data imply a second-birth twinning probability of 0.0079, just slightly lower than the probability we estimate using 1980 Census data.

Quarter of birth is not reported in the 1990 PUMS, so multiple births in 1990 were defined as children reported to be of the same age. Using this procedure, we calculated that 1.2 percent of all second births in 1990 were multiple births. This naturally produces a much higher estimate of the number of twins since two children born in the same 12-month period are classified as twins. We used data from the 1980 PUMS to estimate the error in twin rates calculated using age in years only. Using age in years to define twins in the 1980 data generates an estimated twin rate of 0.01185, which is 35 percent larger than the value we calculate using age in years and quarter of birth. We therefore restricted the analysis using twins to data from the 1980 PUMS.

Demographic and labor-supply variables, described in the lower half of Table 2, include measures of mother's age, age at first birth, years of education, and indicators for race and ethnic background. We also report values for the husbands of women in the married sample. The labor-supply variables are based on Census questions concerning work in 1979 or 1989. These variables measure whether respondents *Worked for pay,* their *Weeks worked,* usual *Hours/week,* and annual *Labor*

[7] One reason estimates are presented for the full sample as well as for the married sample is that conditioning on marital status raises the possibility that selection bias affects estimates in the selected sample.

[8] Of the 3,356 multiple second births in the all-women sample, only 23 were triplets or higher-plurality births. Therefore, we use the terms multiple births and twins interchangeably.

TABLE 2—DESCRIPTIVE STATISTICS, WOMEN AGED 21–35 WITH 2 OR MORE CHILDREN

	Means and (standard deviations)					
	1980 PUMS			1990 PUMS		
		Married couples			Married couples	
Variable	All women	Wives	Husbands	All women	Wives	Husbands
Children ever born	2.55 (0.81)	2.51 (0.77)	—	2.50 (0.76)	2.48 (0.74)	—
More than 2 children (=1 if mother had more than 2 children, =0 otherwise)	0.402 (0.490)	0.381 (0.486)	—	0.375 (0.484)	0.367 (0.482)	—
Boy 1st (s_1) (=1 if first child was a boy)	0.511 (0.500)	0.514 (0.500)	—	0.512 (0.500)	0.514 (0.500)	—
Boy 2nd (s_2) (=1 if second child was a boy)	0.511 (0.500)	0.513 (0.500)	—	0.511 (0.500)	0.512 (0.500)	—
Two boys (=1 if first two children were boys)	0.264 (0.441)	0.266 (0.442)	—	0.264 (0.441)	0.265 (0.441)	—
Two girls (=1 if first two children were girls)	0.242 (0.428)	0.239 (0.427)	—	0.241 (0.428)	0.239 (0.426)	—
Same sex (=1 if first two children were the same sex)	0.506 (0.500)	0.506 (0.500)	—	0.505 (0.500)	0.503 (0.500)	—
Twins-2 (=1 if second birth was a twin)	0.0085 (0.0920)	0.0083 (0.0908)	—	0.012 (0.108)	0.011 (0.105)	—
Age	30.1 (3.5)	30.4 (3.4)	33.0 (4.6)	30.4 (3.5)	30.7 (3.3)	33.4 (4.8)
Age at first birth (parent's age in years when first child was born)	20.1 (2.9)	20.8 (2.9)	24.0 (4.0)	21.8 (3.5)	22.4 (3.5)	25.1 (4.7)
Worked for pay (=1 if worked for pay in year prior to census)	0.565 (0.496)	0.528 (0.499)	0.977 (0.150)	0.662 (0.473)	0.667 (0.471)	0.968 (0.175)
Weeks worked (weeks worked in year prior to census)	20.8 (22.3)	19.0 (21.9)	48.0 (10.5)	26.2 (22.9)	26.4 (22.9)	47.1 (12.0)
Hours/week (average hours worked per week)	18.8 (18.9)	16.7 (18.3)	43.5 (12.3)	22.5 (19.1)	22.2 (18.9)	44.0 (13.3)
Labor income (labor earnings in year prior to census, in 1995 dollars)	7,160 (10,804)	6,250 (10,211)	38,919 (25,014)	9,550 (13,071)	9,616 (13,238)	36,623 (30,283)
Family income (family income in year prior to census, in 1995 dollars)	42,342 (26,563)	47,646 (25,821)	—	42,558 (34,692)	49,196 (34,740)	—
Non-wife income (family income minus wife's labor income, in 1995 dollars)	—	41,635 (24,734)	—	—	39,580 (31,892)	—
Number of observations	394,835	254,654	254,654	380,007	301,588	301,588

Notes: The samples include women aged 21–35 with two or more children except for women whose second child is less than a year old. In the 1980 PUMS, the married women sample refers to women who were married at the time of their first birth, married at the time of the survey, and married once. In the 1990 PUMS, the married women are those married at the time of the Census.

income. The latter three variables are set to zero for those who did not work for pay during the year. The final two variables in the table are measures of *Family income* and, for the married sample, a variable called *Non-wife income* computed as family income minus the wife's labor income.[9] The descriptive statistics show that women's labor-force participation rates, weeks and hours worked, and age at first birth increased between 1980 and 1990. Women's real (1995 dollar) earnings increased substantially as well, especially for married women, while real *Non-wife income* declined. It should also be noted that the variances of earnings, husband's earnings, and family income increased substantially over this period.

Finally, note that husband's age at first birth was calculated assuming that the husband is the father of the children in the household. All of the husbands' variables are computed based in this assumption, which seems plausible for the 1980 data since women in the 1980 married women sample were married only once and they were married at the time of their first birth. Only 4.7 percent of the husbands in this sample were married before, and very few children live with their fathers after divorce. On the other hand, the 1990 match is probably not as good as the 1980 match. We therefore confirmed the basic first-stage relationships used in this paper with June CPS data, which includes true retrospective fertility information. See the appendix to our earlier paper (Angrist and Evans, 1996a) using a Census household match for more on data problems and issues related to the match.

B. *Sex Mix and Fertility*

The phenomenon of parental preferences for a mixed sibling-sex composition has been documented in a number of studies. For example, Yoram Ben-Porath and Finis Welch (1976) found that in the 1970 Census, 56 percent of families with either two boys or two girls had a third birth, whereas only 51 percent of families with one boy and one girl had a third child.

The theoretical impact of sex mix on fertility can be captured in the standard quantity/quality model of fertility, originally outlined by Becker and Gregg H. Lewis (1973) and Becker and Nigel Tomes (1976), and extended in detail by Rosenzweig and Wolpin (1980a). In these models, parents derive utility from the number of children and a complementary good, "child quality," which enters the utility function and budget constraint in proportion to the number of children. Child quality is generated through the purchase of inputs and the expenditure of parents' time in home production. Ben-Porath and Welch (1980) describe the sex mix as something that determines child quality in quantity/quality of models. Alternately, the impact of sex preferences can be modeled using state-dependent utility. Suppose a mother already has $n_x \geq 1$ children and she is trying to decide how many additional children to have (n_c). If parents prefer a mixed sibling-sex composition, then a same-sex sibling composition reduces the utility from n_x. This in turn raises the marginal utility of n_c, increasing the chances that parents will try to have additional children. Twinning can similarly be incorporated into this model as a shock to n_x that cannot be fully offset by future fertility choices. For a more detailed theoretical discussion, see our working paper (Angrist and Evans, 1996b).

Table 3 reports estimates of the impact of child sex and the sex mix on fertility similar to those in Ben-Porath and Welch (1976). The first panel looks at sex preferences in families with one or more children by showing the fraction of women with at least one child who had a second child, conditional on the sex of the first child. The third row of this panel shows the difference by sex. In spite of the fact that attitudinal surveys suggest many couples would prefer more boys than girls, or prefer their firstborn child to be male (see, e.g., Williamson, 1976), there is only one subsample (all women in the 1990 PUMS) where subsequent fertility is a function of the sex of the first child. Even in this case, the impact of the sex of the firstborn on fertility is very small.

The second panel of Table 3 documents the relationship between the fraction of women who have a third child and the sex of the first two children. The first three rows from this

[9] In the few cases where there were negative or zero family-income values, we set the variables equal to one when computing logs. Family income and person wage and salary income are top coded at $75,000 in the 1980 Census. In the 1990 Census, family income is top coded at $999,999 and individual wage and salary income is top coded at $140,000, with state medians of top-coded values substituted for the top code.

TABLE 3—FRACTION OF FAMILIES THAT HAD ANOTHER CHILD BY PARITY AND SEX OF CHILDREN

	All women				Married women			
	1980 PUMS (649,887 observations)		1990 PUMS (627,362 observations)		1980 PUMS (410,333 observations)		1990 PUMS (477,798 observations)	
Sex of first child in families with one or more children	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child
(1) one girl	0.488	0.694 (0.001)	0.489	0.665 (0.001)	0.485	0.720 (0.001)	0.487	0.698 (0.001)
(2) one boy	0.512	0.694 (0.001)	0.511	0.667 (0.001)	0.515	0.720 (0.001)	0.513	0.699 (0.001)
difference (2) − (1)	—	0.000 (0.001)	—	0.002 (0.001)	—	0.000 (0.001)	—	0.001 (0.001)

	All women				Married women			
	1980 PUMS (394,835 observations)		1990 PUMS (380,007 observations)		1980 PUMS (254,654 observations)		1990 PUMS (301,588 observations)	
Sex of first two children in families with two or more children	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child
one boy, one girl	0.494	0.372 (0.001)	0.495	0.344 (0.001)	0.494	0.346 (0.001)	0.497	0.331 (0.001)
two girls	0.242	0.441 (0.002)	0.241	0.412 (0.002)	0.239	0.425 (0.002)	0.239	0.408 (0.002)
two boys	0.264	0.423 (0.002)	0.264	0.401 (0.002)	0.266	0.404 (0.002)	0.264	0.396 (0.002)
(1) one boy, one girl	0.494	0.372 (0.001)	0.495	0.344 (0.001)	0.494	0.346 (0.001)	0.497	0.331 (0.001)
(2) both same sex	0.506	0.432 (0.001)	0.505	0.407 (0.001)	0.506	0.414 (0.001)	0.503	0.401 (0.001)
difference (2) − (1)	—	0.060 (0.002)	—	0.063 (0.002)	—	0.068 (0.002)	—	0.070 (0.002)

Notes: The samples are the same as in Table 2. Standard errors are reported in parentheses.

section show the sample characteristics of women in the following groups: those with one boy and one girl, those with two girls, and those with two boys. The next two rows report estimates for women with two children of the same sex and for women with one boy and one girl. The final row reports the differences between the same-sex and mixed-sex group averages.

Both data sets and samples suggest that women with two children of the same sex are much more likely to have a third child than the mothers of one boy and one girl. For example, in the 1980 all-women sample, only 37.2 percent of women with one boy and one girl have a third child, compared to 43.2 for women with two girls or two boys. The relationship between sex mix and the probability of additional childbearing is even larger for married women, reaching a precisely estimated 7-percentage-point difference in the 1990 Census. This is approximately 21 percent of the rate of additional childbearing among women with one boy and one girl. Finally, we note that the relationship between sex mix and childbearing is confirmed in data from the fertility supplements to the June 1980, 1985, and 1990 CPS. This is important because, unlike the Census where information about children is partly based on our household match, the

June CPS contains detailed fertility histories for each woman, including information on the dates of birth and sex of each child.

The virtual random assignment of *Same sex* makes it very likely that reduced-form regressions of fertility and labor-supply outcomes on the instruments have a causal interpretation. One simple check on this claim is to compare the demographic characteristics of people who have same-sex and mixed-sex sibling compositions. Table 4 reports *Same sex* contrasts for mother's age, age at first birth, race, ethnicity, and years of education in the 1980 and 1990 all-women samples. Even in these large samples, none of the contrasts is significantly different from zero at the 5-percent level. The magnitude of the differences by *Same sex* is also very small. For example, the difference of −0.0028 for years of education in 1980 represents 0.02 percent of the sample mean years of schooling, which is about 13.

In contrast with the small and insignificant differences in demographic characteristics by *Same sex,* there are some large and precisely estimated differences in mean demographic variables by twin status. The estimates in the final column of Table 4 replicate the well-known result that twins are more likely for older women (John A. H. Waterhouse, 1950) and for blacks (Ntinos Myrianthopoulos, 1970). Women with more years of schooling are also more likely to have twins, although this probably reflects more childbearing at older ages among more educated women.

II. Fertility and Labor Supply

A. *Wald Estimates*

Because sibling-sex composition is virtually randomly assigned, simple statistical techniques can be used to illustrate how the sex-mix IV strategy identifies the effect of fertility on parents' labor supply. Consider the bivariate regression model,

$$y_i = \alpha + \beta x_i + \varepsilon_i \tag{1}$$

where y_i is a measure of labor supply and x_i is the endogenous fertility measure of interest. Let z_i denote the binary instrument, *Same sex.* The IV estimate of β in this equation is

$$\beta_{IV} = (\bar{y}_1 - \bar{y}_0)/(\bar{x}_1 - \bar{x}_0), \tag{2}$$

where $\bar{y}_1$ is the mean of y_i for those observations with $z_i = 1$ and other terms are similarly defined. The numerator and denominator capture the reduced-form relationships between y_i and z_i and between x_i and z_i. The IV method attributes any effect of z_i on y_i to the effect of z_i on x_i.

Although equation (1) is written as a bivariate regression with constant coefficients, Guido W. Imbens and Angrist (1994) have shown that β_{IV} can be interpreted as a local average treatment effect specific to the instrument, z_i. In this case, β_{IV} estimates the average effect of x_i on y_i for individuals whose fertility has been affected by their children's sex mix. Similarly, when z_i is an indicator of multiple births at the second pregnancy, *Twins-2,* the IV estimates reflect the effect of children on labor supply for those who have had more children than they otherwise would have because of twinning. For this reason, the *Twins-2* and *Same sex* instruments do not necessarily identify the same average effect.

The first six columns of Table 5 report the components of β_{IV} when *Same sex* is used as the instrument in the all-women samples from the 1980 and 1990 PUMS. The last three columns report corresponding results from the 1980 PUMS using the *Twins-2* instrument. The first two rows of the table show the denominator of the Wald estimate, $\bar{x}_1 - \bar{x}_0$, for two possible choices of x_i. One is an indicator for having had a third child, *More than 2 children.* The other is total *Number of children.* The effect of the *Same sex* instrument on *More than 2 children,* equal to the difference in means reported at the bottom of Table 3, is 0.06 in 1980 and 0.063 in 1990. The effect of *Same sex* on *Number of children* is 0.077 in 1980 and 0.084 in 1990. The effect of *Twins-2* on the probability of having a third birth in 1980 is 0.60, and the effect of *Twins-2* on *Number of children* is 0.81.

Below the estimates of $\bar{x}_1 - \bar{x}_0$, columns (1) and (4) of Table 5 report $\bar{y}_1 - \bar{y}_0$ for alternative outcomes using the *Same sex* instrument. These results show that in addition

TABLE 4—DIFFERENCES IN MEANS FOR DEMOGRAPHIC VARIABLES BY *SAME SEX* AND *TWINS*-2

	Difference in means (standard error)		
	By *Same sex*		By *Twins-2*
Variable	1980 PUMS	1990 PUMS	1980 PUMS
Age	−0.0147	0.0174	0.2505
	(0.0112)	(0.0112)	(0.0607)
Age at first birth	0.0162	−0.0074	0.2233
	(0.0094)	(0.0114)	(0.0510)
Black	0.0003	0.0021	0.0300
	(0.0010)	(0.0011)	(0.0056)
White	0.0003	−0.0006	−0.0210
	(0.0012)	(0.0013)	(0.0066)
Other race	−0.0006	−0.0014	−0.0090
	(0.0005)	(0.0009)	(0.0041)
Hispanic	−0.0014	−0.0007	−0.0069
	(0.0009)	(0.0010)	(0.0047)
Years of education	−0.0028	0.0100	0.0940
	(0.0076)	(0.0074)	(0.0415)

Notes: The samples are the same as in Table 2. Standard errors are reported in parentheses.

to having more children than women with one boy and one girl, women with two children of the same sex have a lower probability of working, work fewer weeks per year and fewer hours per week, and have lower annual earnings and lower family income. All but the final result is statistically significant in both Census years.

The Wald estimates for 1980 calculated by dividing $\bar{y}_1 - \bar{y}_0$ by $\bar{x}_1 - \bar{x}_0$ when x_i is *More than 2 children* imply that having more than two children reduced labor supply by 13.3 (−0.008/0.06) percentage points, weeks worked by about 6.4 weeks, hours of work per week by 5.2, and labor income by just over \$2,200 per year. The results for 1990 are also negative, though (with the exception of family income) somewhat smaller. The Wald estimates calculated using the effect of *Same sex* on total *Number of children* put these effects in per-child terms. In per-child terms, the estimates are about 0.78 as large in 1980 and 0.75 as large in 1990 as the estimates produced with *More than 2 children* in the denominator.

The last three columns in the table show that women whose second pregnancy resulted in twins are also less likely to work. With the exception of the estimate for family income, which is not very precise, the Wald estimates generated by *Twins-2,* reported in column (6), are lower than the Wald estimates based on *Same sex.* In Section III, we explore the comparison between *Same sex* and *Twins-2* estimates further and show how they can be reconciled.

As with the *Same sex* estimates, *Twins-2* estimates in per-child terms are necessarily smaller than estimates treating the indicator *More than 2 children* as the endogenous regressor. But the factor of proportionality connecting the per-child and *More than 2 children* estimates using *Twins-2* is also 0.75. It therefore makes little difference which denominator is used because estimates based on *More than 2 children* can always be converted into per-child estimates by multiplying by 0.75. We chose to discuss estimates treating *More than 2 children* as the endogenous regressor in the re-

TABLE 5—WALD ESTIMATES OF LABOR-SUPPLY MODELS

	1980 PUMS			1990 PUMS			1980 PUMS		
		Wald estimate using as covariate:			Wald estimate using as covariate:			Wald estimate using as covariate:	
Variable	Mean difference by *Same sex*	*More than 2 children*	*Number of children*	Mean difference by *Same sex*	*More than 2 children*	*Number of children*	Mean difference by *Twins-2*	*More than 2 children*	*Number of children*
More than 2 children	0.0600 (0.0016)	—	—	0.0628 (0.0016)	—	—	0.6031 (0.0084)	—	—
Number of children	0.0765 (0.0026)	—	—	0.0836 (0.0025)	—	—	0.8094 (0.0139)	—	—
Worked for pay	−0.0080 (0.0016)	−0.133 (0.026)	−0.104 (0.021)	−0.0053 (0.0015)	−0.084 (0.024)	−0.063 (0.018)	−0.0459 (0.0086)	−0.076 (0.014)	−0.057 (0.011)
Weeks worked	−0.3826 (0.0709)	−6.38 (1.17)	−5.00 (0.92)	−0.3233 (0.0743)	−5.15 (1.17)	−3.87 (0.88)	−1.982 (0.386)	−3.28 (0.63)	−2.45 (0.47)
Hours/week	−0.3110 (0.0602)	−5.18 (1.00)	−4.07 (0.78)	−0.2363 (0.0620)	−3.76 (0.98)	−2.83 (0.73)	−1.979 (0.327)	−3.28 (0.54)	−2.44 (0.40)
Labor income	−132.5 (34.4)	−2208.8 (569.2)	−1732.4 (446.3)	−119.4 (42.4)	−1901.4 (670.3)	−1428.0 (502.6)	−570.8 (186.9)	−946.4 (308.6)	−705.2 (229.8)
ln(*Family income*)	−0.0018 (0.0041)	−0.029 (0.068)	−0.023 (0.054)	−0.0085 (0.0047)	−0.136 (0.074)	−0.102 (0.056)	−0.0341 (0.0223)	−0.057 (0.037)	−0.042 (0.027)

Notes: The samples are the same as in Table 2. Standard errors are reported in parentheses.

mainder of the paper because this emphasizes the fact that the fertility increment induced by either instrument is a move from two to more than two children.

B. *Two-Stage Least-Squares Estimation*

While the Wald estimates provide a simple illustration of how the instruments identify the effect of children on labor supply, the rest of the paper discusses two-stage least-squares (2SLS) and ordinary least-squares (OLS) estimates of regression models relating labor-market outcomes to fertility and a variety of exogenous covariates. 2SLS estimation allows us to accomplish three things. First, even if there is no association between the instrument and exogenous covariates, as suggested by Table 4, controlling for exogenous covariates can lead to more precise estimates if the treatment effects are roughly constant across groups.

Second, we can use 2SLS to control for any secular additive effects of child sex when using *Same sex* as an instrument. This is desirable because *Same sex* is an interaction term involving the sex of the first two children, and therefore potentially correlated with the sex of either child. To see this, let s_1 and s_2 be indicators for male firstborn and second-born children. The instrument can be written as

$$\textit{Same sex} = s_1 s_2 + (1 - s_1)(1 - s_2). \tag{3}$$

Assuming that child sex is independent and identically distributed (i.i.d.) over children, the population regression of *Same sex* on either s_j produces a slope coefficient equal to $2E[s_j] - 1$, which is zero only if $E[s_j] = 1/2$.[10] Since the probability of giving birth to a male child is 0.51, there is a slight positive association between *Same sex* and the sex of each child. This correlation is a concern only if s_j

[10] Proof: Assuming child sex is i.i.d., we have $E[s_1] = E[s_2]$ and $E[s_1 s_2] = E[s_j]^2$. Therefore, $\mathrm{Cov}(\textit{Same sex}, s_j) = E[s_j](E[s_j] - E[\textit{Same sex}])$. Some manipulation gives $E[s_j] - E[\textit{Same sex}] = (1 - E[s_j])(2E[s_j] - 1)$. Since the variance of s_j is $E[s_j](1 - E[s_j])$, the regression coefficient is $(2E[s_j] - 1)$.

affects labor supply for reasons other than family size. Such effects could arise if the sex of offspring affects the father's commitment to the family (see, e.g., Philip S. Morgan et al., 1988) or changes the way parents treat their children (Kristin F. Butcher and Anne Case, 1994; Duncan Thomas, 1994). Secular effects of sex mix on labor supply could also be generated by the fact that boys are more likely than girls to have disabilities (see, e.g., Angrist and Victor Lavy, 1996) since having a disabled child might change parents' behavior. Adding s_1 and s_2 as regressors to the estimating equations reduces the likelihood of omitted-variables bias from these sources.

Of course, controls for additive effects can only eliminate bias from omitted variables with effects that are additive in the number of children. However, a third advantage of the 2SLS framework is that it allows us to exploit the fact that the *Same sex* instrument can be decomposed into two instruments, leading to an overidentified model. In particular, the separate indicators, *Two boys* $[s_1 s_2]$ and *Two girls* $[(1 - s_1)(1 - s_2)]$, are both available as potential instruments. This is useful because bias from any secular effects of child sex on labor supply should be different for these two instruments, while the labor-supply consequences of childbearing seem likely to be independent of whether *Same sex* equals *Two boys* or *Two girls*. A natural specification test is therefore the conventional instrument-error overidentification test statistic for 2SLS estimation using both instruments, since this is the same as a test for whether the *Two boys* and *Two girls* instruments give the same estimate when used separately.[11]

The following regression models are used to link labor-supply variables for husbands and wives to the endogenous *More than 2* variable, x_i, and the list of exogenous covariates, including additive effects for the sex of each child:

$$(4) \quad y_i = \alpha_0' \mathbf{w}_i + \alpha_1 s_{1i} + \alpha_2 s_{2i} + \beta x_i + \varepsilon_i,$$

where $\mathbf{w}_i$ is a vector of demographic variables, and s_{1i} and s_{2i} are indicators for the sex of the first two children of mother i. Initially, $\mathbf{w}_i$ is limited to variables that are clearly exogenous to fertility: mother's age and age at first birth, plus race and Hispanic indicators. In the just-identified model where *Same sex* is the only instrument, the first-stage equation relating *More than 2 children* to sex mix is

$$(5) \quad x_i = \pi_0' \mathbf{w}_i + \pi_1 s_{1i} + \pi_2 s_{2i} + \gamma(\textit{Same sex}_i) + \eta_i,$$

where γ is the first-stage effect of the instrument.

The alternative identification strategy uses the two components of *Same sex*—*Two boys* and *Two girls*—as instruments for *More than 2 children.* In this case, however, either s_{1i} or s_{2i} must be dropped from the list of covariates because s_{1i}, s_{2i}, $s_{1i}s_{2i}$, and $(1 - s_{1i})(1 - s_{2i})$ are linearly dependent. We chose to drop s_{2i} (the results are not sensitive to this choice, or to the elimination of both s_{1i} and s_{2i}, as we show below). In this case, the equation of interest becomes

$$(6) \quad y_i = \alpha_0' \mathbf{w}_i + \alpha_1 s_{1i} + \beta_i x_i + \varepsilon_i.$$

The first-stage relationship between x_i and sex mix is

$$(7) \quad x_i = \pi_0' \mathbf{w}_i + \pi_1 s_{1i} + \gamma_0(\textit{Two boys}_i) + \gamma_1(\textit{Two girls}_i) + \eta_i,$$

where $\textit{Two boys}_i = s_{1i}s_{2i}$ and $\textit{Two girls}_i = (1 - s_{1i})(1 - s_{2i})$.

C. 2SLS Results

The first-stage results linking sex mix and fertility are reported in Table 6. In the top half of the table, we report results from the 1980 PUMS. These estimates show that women in 1980 with same-sex children are estimated to be 6.2 percentage points more likely to have a third child in a model with covariates. The corresponding estimate for married women is 6.9 percent. The

[11] See Whitney K. Newey and Kenneth D. West (1987) for this interpretation of overidentification tests. Angrist (1991) discusses the dummy instrument case.

TABLE 6—OLS ESTIMATES OF *MORE THAN 2 CHILDREN* EQUATIONS

	All women			Married women		
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)
1980 PUMS						
Boy 1st	—	−0.0080 (0.0015)	0.0001 (0.0021)	—	−0.0111 (0.0018)	−0.0016 (0.0026)
Boy 2nd	—	−0.0081 (0.0015)	—	—	−0.0095 (0.0018)	—
Same sex	0.0600 (0.0016)	0.0617 (0.0015)	—	0.0675 (0.0019)	0.0694 (0.0018)	—
Two boys	—	—	0.0536 (0.0021)	—	—	0.0598 (0.0026)
Two girls	—	—	0.0698 (0.0021)	—	—	0.0789 (0.0026)
With other covariates	no	yes	yes	no	yes	yes
R^2	0.004	0.084	0.084	0.005	0.078	0.078
1990 PUMS						
Boy 1st	—	−0.0081 (0.0015)	−0.0083 (0.0022)	—	−0.0097 (0.0017)	−0.0086 (0.0024)
Boy 2nd	—	0.0002 (0.0015)	—	—	−0.0011 (0.0017)	—
Same sex	0.0628 (0.0016)	(0.0623) (0.0015)	—	0.0702 (0.0018)	0.0703 (0.0017)	—
Two boys	—	—	0.0624 (0.0021)	—	—	0.0692 (0.0023)
Two girls	—	—	0.0621 (0.0022)	—	—	0.0714 (0.0024)
With other covariates	no	yes	yes	no	yes	yes
R^2	0.004	0.082	0.082	0.005	0.082	0.082

Notes: Other covariates in the models are indicators for *Age, Age at first birth, Black, Hispanic,* and *Other race.* The variable *Boy 2nd* is excluded from columns (3) and (6). Standard errors are reported in parentheses.

t-statistics for these first-stage effects are well over 30. As in Table 3, the estimates for the 1990 PUMS (reported in the lower half of Table 6) are somewhat larger in both the full and married women samples.

Table 6 also provides some evidence of an association between having a male child and reduced childbearing at higher parities. Note, however, that the effect of *Boy 1st* in the 1980 data is explained entirely by the difference in the effect of *Two boys* and *Two girls* when these regressors are entered separately. In other words, when the effects of sex mix are allowed to differ by sex, there is no relationship between *Boy 1st* and fertility, although the effect of *Same sex* on fertility in 1980 is larger for boys than for girls. The *Boy 1st* effects for 1990 remain significant in all specifications, but they are very small, especially in comparison to the effects of the sex mix.

Next, we use the sex mix to estimate the effect of *More than 2 children* on measures of employment and earnings in 1980. Table 7 reports a set of OLS estimates and two sets of 2SLS estimates using *Same sex* and the pair of dummies *Two boys* and *Two girls* as instruments. The exogenous regressors are the same as in Table 6 (coefficients not reported). In models that use *Two boys* and *Two girls* as instruments, we dropped the *Boy 2nd* variable from the list of covariates. The first three columns show results for the full sample, the next three columns show results for married women, and the last three columns show results for the husbands of married women.

OLS estimates in both the full and married women sample suggest that the presence of a third child reduces the probability of working by about 17 percentage points, and causes weeks worked to fall by about 8–9 per year, hours per week to fall by 6–7, and family income to fall by about 13 percent. OLS estimates of earnings effects are \$3,166 in the married sample and \$3,768 in the full sample. Not surprisingly, all of these OLS estimates are very precisely estimated.

In contrast with the results for women, OLS estimates of the effect of *More than 2 children* on husbands' labor supply are small. Having a third child is estimated to reduce the probability a husband worked for pay by less than one percentage point. The impact of a third child on other measures of husbands' labor supply is also small, though precise enough to be significantly different from zero. The estimated effect on annual weeks worked is −0.90 and the estimate for hours per week is 0.25. The effect on husbands' earnings appears substantial (−\$1,506), but this amount is still only about 3.9 percent of the average earnings of men in the sample.

The first set of 2SLS estimates uses *Same sex* alone as an instrument. In the full sample, the estimates (standard errors) for the dependent variables *Worked for pay, Weeks worked, Hours/week,* and *Labor income* models are −0.12 (0.025), −5.7 (1.1), −4.6 (0.95), and −1,961 (542). These results suggest that having a third child causes a 20–30-percent reduction in women's labor supply and earnings. One important finding is that estimates using *Same sex* as an instrument are smaller than the corresponding OLS estimates. This is true for the labor-supply estimates in the married women's sample as well, although here the gap between 2SLS and OLS estimates is not as large. Overall, however, the OLS estimates appear to exaggerate the causal effect of fertility on female labor supply.

It is also worth noting that the relationship between the OLS and 2SLS estimates is similar when the estimates are converted into per-child units. Above, we noted that 2SLS estimates treating *More than 2 children* as the endogenous regressor should be multiplied by about 0.75 to obtain estimates in per-child terms (i.e., with *Number of children* as the endogenous regressor). It turns out that the OLS estimates can be converted into per-child units by multiplying by about 0.7 using either 1980 or 1990 data. Thus, regardless of whether the estimates are cast in terms of the effect of having more than two children or in per-child units, the OLS estimates are considerably smaller than the 2SLS estimates.

In addition to differing from the OLS estimates, the estimates using *Same sex* as an instrument also differ from most of the 2SLS estimates previously reported in the literature on children and labor supply. In his review article, Browning (1992 p. 1469) notes that, "There is one salient difference between studies that take fertility as exogenous and those that take it as endogenous. In many of the latter it is found that fertility either has no effect on labor supply ... or it has a positive effect." Browning also points out that it is not clear from these estimates whether children really have no effect on female labor supply, or whether the instruments are too weak or simply poorly chosen. While the 2SLS estimates generated by *Same sex* are smaller than the corresponding OLS estimates, they are still negative, precise, and of a plausible magnitude.

In contrast to the female labor-supply estimates, there is little evidence of a relationship between having a third child and family income. Given the strong labor-supply effects, the weak impact on family income may seem surprising. There are a few potential explanations for this result. First, the lost income due to a reduction in mothers' labor supply could be made up by other family members. The small and statistically insignificant 2SLS estimates in the ln(*Non-wife income*) equations suggest that this is not the case. The most likely explanation is that the instrument is not powerful enough to detect the family-income consequences of childbearing. For example, in the married women sample, the third child reduces female earnings by about 21 percent and female labor income is (on average) 13 percent of total family income. If the third child does not alter husbands' labor supply, we would expect an estimated effect of *More than 2 children* in the ln(*Family income*) equations of roughly $-0.21 \times 0.13 = -0.027$, which is close to the reported estimate of -0.05. But the standard error for this estimate is slightly higher than 0.05, so that impacts this small cannot be precisely measured.

Gronau (1977 p. 1102) reports results suggesting that husbands increase their work effort in response to an increase in family size. Table 7 also reports estimates of the impact of the third child on husbands' labor supply in the 1980 married sample. While the OLS estimates show a small but significant (negative) relationship between husbands' labor supply and additional childbearing, estimates constructed using *Same sex* as an instrument generate no evidence of any effects on the labor supply of men. The standard errors on the 2SLS estimates for husbands' variables *Worked for pay, Weeks worked*, and *Hours/week* are actually smaller than the corresponding standard errors for women, and they are small enough so that modest positive or negative effects could be detected if they existed.

The labor-supply effects estimated using 1990 data are remarkably similar to those estimated for 1980. This can be seen in Table 8, which reports OLS estimates and 2SLS estimates for 1990 using *Same sex* and *Two boys* and *Two girls* as instruments. Some of the estimated effects are slightly smaller in 1990 than in 1980, but these differences are not statistically meaningful. One difference between the 1980 and 1990 results that does seem noteworthy is the larger negative impact of childbearing on married women's earnings in 1990, perhaps because of an increase in women's wages. This result may also be attributable to the fact that married women are delaying childbearing (average age at first birth increased from 20.8 in 1980 to 22.4 in 1990 for this group), and therefore they have more years of experience and higher wages when they exit the workforce due to childbirth.

Table 6 shows that mothers of two girls are more likely than mothers of two boys to have a third child. So the first-stage relationship differs for these two instruments. However, the 2SLS estimates in Tables 7 and 8 show that the additional predictive power provided by separating the two components of *Same sex* does not change the coefficient estimates very much or lead to an appreciable increase in precision using either the 1980 or 1990 data.

We noted above that the overidentification test statistic associated with the use of *Two boys* and *Two girls* as instruments jointly tests for a difference between 2SLS estimates computed using only *Two boys* and 2SLS estimates computed using only *Two girls*. The p-values for this test are reported in square brackets in both Tables 7 and Table 8. The p-values for the 1990 estimates suggest that it does not matter which instrument is used. In fact, the 2SLS estimates using *Two boys* and *Two girls* in 1990 are remarkably close. On the other hand, the 1980 2SLS estimates are consistently smaller when *Two girls* alone is used as the instrument. Moreover, some of the p-values for estimates computed using 1980 data indicate a significant contrast between the *Two girls* and *Two boys* instruments, although no marginal significance level is below 1 percent. As in the 1990 data, however, in the 1980 data both instruments are always associated with more children and reduced labor supply.

D. *Other Specification Issues*

Other specification issues considered here include the robustness of the results, the generality of the results, and the validity of the instruments. Because sex mix is essentially randomly assigned, the results reported in Tables 7 and 8 are unchanged by altering the basic set of covariates. For example, using data for married women from the 1980 PUMS, we estimated models adding the following covariates to the vector $\mathbf{w}_i$: linear and

TABLE 7—OLS AND 2SLS ESTIMATES OF LABOR-SUPPLY MODELS USING 1980 CENSUS DATA

	All women			Married women			Husbands of married women		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimation method	OLS	2SLS	2SLS	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Instrument for *More than 2 children*	—	*Same sex*	*Two boys, Two girls*	—	*Same sex*	*Two boys, Two girls*	—	*Same sex*	*Two boys, Two girls*
Dependent variable:									
Worked for pay	−0.176	−0.120	−0.113	−0.167	−0.120	−0.113	−0.008	0.004	0.001
	(0.002)	(0.025)	(0.025)	(0.002)	(0.028)	(0.028)	(0.001)	(0.009)	(0.008)
			[0.013]			[0.013]			[0.013]
Weeks worked	−8.97	−5.66	−5.37	−8.05	−5.40	−5.16	−0.82	0.59	0.45
	(0.07)	(1.11)	(1.10)	(0.09)	(1.20)	(1.20)	(0.04)	(0.60)	(0.59)
			[0.017]			[0.071]			[0.030]
Hours/week	−6.66	−4.59	−4.37	−6.02	−4.83	−4.61	0.25	0.56	0.50
	(0.06)	(0.95)	(0.94)	(0.08)	(1.02)	(1.01)	(0.05)	(0.70)	(0.69)
			[0.030]			[0.049]			[0.71]
Labor income	−3768.2	−1960.5	−1870.4	−3165.7	−1344.8	−1321.2	−1505.5	−1248.1	−1382.3
	(35.4)	(541.5)	(538.5)	(42.0)	(569.2)	(565.9)	(103.5)	(1397.8)	(1388.9)
			[0.126]			[0.703]			(0.549)
ln(*Family income*)	−0.126	−0.038	−0.045	−0.132	−0.051	−0.053	—	—	—
	(0.004)	(0.064)	(0.064)	(0.004)	(0.056)	(0.056)			
			[0.319]			[0.743]			
ln(*Non-wife income*)	—	—	—	−0.053	0.023	0.016	—	—	—
				(0.005)	(0.066)	(0.066)			
						[0.297]			

Notes: The table reports estimates of the coefficient on the *More than 2 children* variable in equations (4) and (6) in the text. Other covariates in the models are *Age, Age at first birth,* plus indicators for *Boy 1st, Boy 2nd, Black, Hispanic,* and *Other race.* The variable *Boy 2nd* is excluded from equation (6). The *p*-value for the test of overidentifying restrictions associated with equation (6) is shown in brackets. Standard errors are reported in parentheses.

quadratic terms in the wife's education, quadratic terms in wife's age, age at first birth, linear and quadratic terms in husband's age, husband's age at first birth and education, linear and quadratic terms in husband's labor income, and a full set of state dummy variables.[12] In these models, the 2SLS estimates (standard errors) of the *More than 2 children* coefficient have the following values: *Worked for pay,* −0.122 (0.027); *Weeks worked,* −5.45 (1.18); *Hours/week,* −5.04 (0.99); *Labor income,* −1,390 (555). All of these values are within 5 percent of the corresponding estimates from Table 7.

A referee and others who read earlier versions of this paper expressed concern about whether the results are likely to be representative of the impact of childbearing in general since the sample is restricted to women with two or more children and to women in a relatively young age-group. Estimates of the effect of going from two to more than two children do not necessarily generalize. On the other hand, we believe these results are likely to be of general interest because a significant fraction of the change in fertility between 1970 and 1990 was due to reductions in the number of women having more than two children. As noted in Section I, this fact is apparent in Census data on completed family size.

[12] Two of these covariates, years of education and husband's earnings, are potentially endogenous because they may be partly determined by fertility. For this reason, they were excluded from the main set of estimates.

TABLE 8—OLS AND 2SLS ESTIMATES OF LABOR-SUPPLY MODELS USING 1990 CENSUS DATA

	All women			Married women			Husbands of married women		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimation method	OLS	2SLS	2SLS	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Instrument for *More than 2 children*	—	*Same sex*	*Two boys, Two girls*	—	*Same sex*	*Two boys, Two girls*	—	*Same sex*	*Two boys, Two girls*
Dependent variable:									
Worked for pay	−0.155	−0.092	−0.092	−0.147	−0.104	−0.104	−0.102	0.017	0.017
	(0.002)	(0.024)	(0.024)	(0.002)	(0.024)	(0.024)	(0.001)	(0.009)	(0.009)
			[0.743]			[0.576]			[0.989]
Weeks worked	−8.71	−5.66	−5.64	−8.25	−5.76	−5.76	−1.03	1.01	1.01
	(0.08)	(1.16)	(1.16)	(0.09)	(1.15)	(1.15)	(0.05)	(0.63)	(0.63)
			[0.391]			[0.670]			[0.708]
Hours/week	−6.80	−4.08	−4.10	−6.39	−3.94	−3.95	−0.06	0.85	0.83
	(0.07)	(0.98)	(0.98)	(0.07)	(0.96)	(0.96)	(0.05)	(0.69)	(0.69)
			[0.489]			[0.665]			[0.180]
Labor income	−3984.4	−2099.6	−2096.2	−3753.9	−2457.5	−2456.3	929.7	1348.7	1354.8
	(44.2)	(664.0)	(663.8)	(50.7)	(669.7)	(669.7)	(114.9)	(1536.0)	(1535.9)
			[0.830]			[0.893]			[0.711]
ln(*Family income*)	−0.119	−0.124	−0.122	−0.103	−0.054	−0.054	—	—	—
	(0.005)	(0.071)	(0.071)	(0.004)	(0.051)	(0.051)			
			[0.270]			[0.878]			
ln(*Non-wife income*)	—	—	—	−0.004	0.020	0.020	—	—	—
				(0.005)	(0.068)	(0.068)			
						[0.452]			

Notes: The table reports the coefficient on the *More than 2 children* variable in equations (4) and (6) in the text estimated with 1990 Census data. Other covariates in the models are *Age, Age at first birth,* plus indicators for *Boy 1st, Boy 2nd, Black, Hispanic,* and *Other race.* The variable *Boy 2nd* is excluded from equation (6). The *p*-value for the test of overidentifying restrictions associated with equation (6) is shown in brackets. Standard errors are reported in parentheses.

The sample was restricted to women under the age of 35 because nearly all children born to women in this age-group are still at home. Relaxing this restriction, a greater fraction of women with two or more children are lost because the oldest child is increasingly unlikely to be at home. We note, however, that the results are not very sensitive to this sample-selection rule. For example, expanding the age-group to include women up to age 45 in the 1980 data, the sample size increases to 552,606 observations. The resulting 2SLS estimate (standard error) of the effect of the *More than 2 children* variable in the *Worked for pay* model is −0.096 (0.021) in the all-women sample.

A final point is that because *Same sex* is an interaction term, the 2SLS estimates were computed using a model that controls for additive effects of s_1 and s_2. This specification was motivated by a concern with the validity of the instruments and possible omitted variables bias. It is useful to know whether the control for these additive effects is important because if it is, identification turns on our ability to distinguish additive effects from the interaction term. Moreover, when using *Two boys* and *Two girls* as separate instruments, we must drop one of the additive effects. As it turns out, the 2SLS estimates are virtually invariant to the inclusion of regressors that control for the sex of each child.[13] The coefficients (standard errors) on the *More than 2 children* variable in the *Worked for pay, Weeks worked,*

[13] The Wald estimates in Table 5 also constitute 2SLS estimates with no controls for main effects.

and *Hours/week* equations change by no more than 2 percent. There is also little evidence of an association between having male children and labor supply. The only significant sex effects are for s_2 in the 1980 sample, but these are small.[14]

E. *Heterogeneity in the Impact of Children on Labor Supply*

A number of theoretical models describe how the impact of children on labor supply might vary with the wages or schooling of husbands or wives. For example, Angrist and Evans (1996b) outline a version of Gronau's (1977) model of market work and home production that incorporates child quality effects of the sort discussed by Becker and Lewis (1973). This model predicts that higher own wages of either partner magnify the labor-supply consequences of childbearing, although there are no cross-wage effects.[15] Gronau's survey paper (1986 p. 287) refers to a number of empirical studies consistent with this prediction, showing that the labor supply of more educated women is more sensitive to the presence of children than the labor supply of less educated women. Earlier, Gronau (1973 p. S170) reported finding that the effect of a child on his mother's value of time increases with the mother's education. On the other hand, an assumption implicit in most empirical labor-supply models, where the focus is on wage effects and not the consequences of childbearing, is that there are no interactions in the effect of wages and the number of children (see, e.g., Thomas A. Mroz, 1987).

We use the *Same sex* instrument to explore the question of how the labor-market consequences of childbearing varies with the earnings or earnings potential of husbands and wives. Panel A of Table 9 reports OLS and 2SLS estimates of the effect of *More than 2 children* on married women, conditional on the position of their husbands in the husbands' earnings distribution. The first column shows the first-stage relationship between *More than 2 children* and *Same sex,* interacted with dummies that indicate whether husbands' earnings are in the upper third, middle third, or lower third of the earnings distribution. These estimates show that the effect of *Same sex* on fertility is increasing in husbands' earnings. For women with high-wage husbands, however, 2SLS estimates of labor-supply effects are smaller and they are not significantly different from zero. Note that average participation rates do not decline enough with husbands' earnings to account for the decline in the magnitude of the coefficients among women with high-wage husbands. It is also worth noting that the OLS estimates do not decline nearly as much with husbands' earnings as do the 2SLS estimates.

It is not possible to analyze labor-supply effects conditional on women's wages because wages are unobserved for women who do not work. But we can condition on schooling, which is an important predictor of individual earnings potential. This is done in Panel B of Table 9 for married women in the 1980 sample with less than a high-school education (18 percent of the sample), high-school graduates (49 percent of the sample), and more than a high-school education (33 percent of the sample). The reduced forms show a strong association between *Same sex* and fertility in each schooling group, although the effect is about 1 percentage point smaller for mothers in the highest education category. The 2SLS estimates suggest that women with relatively low levels of schooling experience the largest effects of children on labor supply. In contrast, there is no statistically significant association between additional childbearing and labor supply for women with more than a high-school education. As with the estimates that condition on husbands' earnings, the variation in 2SLS estimates by schooling group differs from the variation in OLS estimates,

[14] For example, the coefficient (standard error) on *Boy 2nd* in the *Worked for pay* model is −0.0038 (0.0015). Estimates for this variable in the *Weeks worked* and *Hours/week* equations are −0.164 (0.069) and −0.127 (0.059), respectively.

[15] The explanation for this is that in equilibrium, the marginal returns to an hour spent at home are higher for high-wage people than for low-wage people. Although the effects of children are generally ambiguous, in the Angrist and Evans (1996b) model, increasing the number of children increases time spent at home because of returns to scale in parental time spent on child-rearing. Returns to scale are larger when the marginal return to hours (and hence wages) are higher.

TABLE 9—2SLS ESTIMATES OF LABOR-SUPPLY MODELS WITH INTERACTION TERMS USING 1980 CENSUS DATA

Sample/variables	*More than 2 children* First stage	*Worked for pay* Mean of dependent variable	OLS	2SLS	*Weeks/year* Mean of dependent variable	OLS	2SLS
A. Results for wives by *husband's earnings:*							
Bottom third of husband's earnings distribution	0.057 (0.003)	0.570	−0.186 (0.003)	−0.122 (0.060)	21.1	−9.23 (0.15)	−7.55 (2.60)
Middle third of husband's earnings distribution	0.072 (0.003)	0.569	−0.165 (0.004)	−0.185 (0.047)	20.8	−8.31 (0.15)	−7.11 (2.04)
Top third of husband's earnings distribution	0.079 (0.003)	0.448	−0.152 (0.003)	−0.078 (0.042)	15.2	−6.76 (0.15)	−3.17 (1.82)
B. Results for wives by *wife's education:*							
Wife < high-school graduate	0.071 (0.004)	0.468	−0.150 (0.005)	−0.121 (0.064)	16.1	−7.30 (0.20)	−7.12 (2.80)
Wife high-school graduate	0.073 (0.003)	0.524	−0.156 (0.003)	−0.147 (0.038)	19.2	−7.74 (0.13)	−6.42 (1.65)
Wife > high-school graduate	0.063 (0.003)	0.567	−0.179 (0.004)	−0.082 (0.054)	20.4	−8.33 (0.15)	−2.93 (2.33)
C. Results for wives by *wife's education* for women whose *husband's earnings* are in middle third:							
Wife < high-school graduate	0.079 (0.008)	0.481	−0.138 (0.009)	−0.275 (0.109)	16.7	−7.10 (0.38)	−10.2 (4.83)
Wife high-school graduate	0.076 (0.004)	0.551	−0.157 (0.003)	−0.189 (0.060)	20.3	−8.33 (0.21)	−7.78 (2.64)
Wife > high-school graduate	0.062 (0.006)	0.640	−0.184 (0.006)	−0.125 (0.098)	23.7	−9.07 (0.28)	−3.98 (4.30)
D. Results for husbands by *wife's education:*							
Wife < high-school graduate	0.071 (0.004)	0.945	−0.014 (0.001)	−0.013 (0.020)	44.5	−1.36 (0.10)	−0.21 (1.37)
Wife high-school graduate	0.074 (0.003)	0.981	−0.005 (0.001)	0.005 (0.012)	48.4	−0.53 (0.06)	0.92 (0.81)
Wife > high-school graduate	0.063 (0.003)	0.987	−0.002 (0.001)	0.009 (0.016)	49.2	−0.23 (0.08)	0.25 (1.14)

Notes: The table reports estimates of the coefficient on *More than 2 children* in equation (4) in the text, modified to allow interactions with wives' schooling and husbands' education as indicated. Main effects for each interaction variable (*husband's earnings distribution* and *wife's education*) are included in the equation. Other covariates in the models are those listed in the notes to Table 7. Data are from the 1980 married women and husband samples. Standard errors are reported in parentheses.

which show effects that increase in magnitude as schooling increases.

Because mothers' education and husbands' wages are correlated, it is not clear whether a set of estimates that condition on husbands' earnings and a set of estimates that condition on mothers' education are capturing distinct phenomena. We therefore present estimates by mothers' education group in a sample restricted to women whose husbands have earnings in the middle third of the earnings distribution. Again, the 2SLS estimates suggest that the impact of childbearing on labor supply declines as education rises, contradicting a theoretical prediction and the OLS estimates, both of which suggest that the labor supply of more educated women responds more to the presence of children. This finding is even more remarkable when viewed in light of the fact that participation rates increase with mothers' schooling. It should be noted, however, that while the results by education group differ substantially, and the pattern of differences is consistent across outcomes, the estimates for subgroups are not very precise.

The last panel in Table 9 (Panel D) reports estimates for husbands, conditional on wives' education. OLS estimates of labor-supply effects for husbands are small and negative, and they decrease in magnitude as wives' schooling increases. As with the overall estimates for husbands in Table 7, the 2SLS estimates for husbands in Table 9 are small and not significantly different from zero.

Table 10 reports estimates conditional on husbands' earnings and wives' schooling using 1990 data. These results are largely similar to those for 1980, showing 2SLS estimates that decline in magnitude with husbands's earnings and wives' schooling, while the OLS estimates for wives by schooling group are stable or increasing. One interesting difference, however, is that the 1990 results show some small, but statistically significant, *positive* effects of childbearing on the labor supply of the husbands of less educated women. The estimated effect (standard error) on participation rates is 0.031 (0.013) for husbands of women who are high-school graduates and 0.053 (0.023) for husbands of women who did not graduate high school. The effect on weeks worked is also significantly different from zero for the husbands of high-school graduates. These estimates suggest a possible change in husbands' labor-supply response to childbearing between 1980 and 1990, at least for some groups. On the other hand, the 1990 husband effects are still less than half the size of most of the corresponding estimates for women, and they appear even smaller when viewed in light of the greater degree of labor-force attachment among husbands.

III. Comparison with Estimates Using Multiple Births

The most important source of exogenous variation in fertility used in fertility research to date is twinning. Rosenzweig and Wolpin (1980b) used 87 U.S. twin pairs to estimate labor-supply effects, and Rosenzweig and Wolpin (1980a) used 25 twin pairs from India to estimate the effect of family size on school progress. Bronars and Grogger (1994) were the first to study the consequences of multiple births with Census data. They used twins in the 1970 and 1980 PUMS to estimate the effect of additional childbearing on mothers' labor-market status, though most of their estimates are for unwed mothers. Gangadharan and Rosenbloom (1996) also used Census twins to estimate the reduced-form effect of twinning on labor-supply variables, but they fail to scale the reduced-form effects of twinning into effects of childbearing. These studies focused almost exclusively on twinning at first birth. An exception is the Bronars and Grogger study, which also briefly discusses (p. 1149) some estimates using twins at second birth.

A twin second birth is similar to the *Same sex* instrument in that it can be used to measure the consequences of moving from two to three children. We noted in the discussion of Table 4, however, that the use of twins as an instrument may be problematic since twinning probabilities appear to be correlated with some observed characteristics of the mother. On the other hand, if the demographic characteristics associated with twinning are all observed, then these factors can be controlled in 2SLS estimation.

2SLS estimates using *Same sex* and *Twins-2* are compared in Table 11. As in Table 7, the models used to produce these estimates

TABLE 10—2SLS ESTIMATES OF LABOR-SUPPLY MODELS WITH INTERACTION TERMS USING 1990 CENSUS DATA

Sample/variables	*More than 2 children* First stage	*Worked for pay* Mean of dependent variable	OLS	2SLS	*Weeks/year* Mean of dependent variable	OLS	2SLS
A. Results for wives by *husband's earnings:*							
Bottom third of husband's earnings distribution	0.064 (0.003)	0.668	−0.160 (0.003)	−0.129 (0.045)	26.3	−8.8 (0.15)	−5.99 (2.18)
Middle third of husband's earnings distribution	0.076 (0.003)	0.728	−0.133 (0.003)	−0.151 (0.039)	29.8	−8.09 (0.15)	−8.37 (1.88)
Top third of husband's earnings distribution	0.071 (0.003)	0.61	−0.137 (0.003)	−0.029 (0.040)	23.6	−7.27 (0.14)	−2.74 (1.93)
B. Results for wives by *wife's education:*							
Wife < high-school graduate	0.069 (0.004)	0.531	−0.145 (0.004)	−0.257 (0.061)	19.2	−7.34 (0.20)	−12.9 (2.91)
Wife high-school graduate	0.078 (0.003)	0.661	−0.140 (0.003)	−0.100 (0.035)	26.3	−8.07 (0.14)	−5.57 (1.67)
Wife > high-school graduate	0.064 (0.002)	0.718	−0.147 (0.003)	−0.058 (0.038)	29.1	−8.43 (0.13)	−3.60 (1.84)
C. Results for wives by *wife's education* for women whose *husband's earnings* are in middle third:							
Wife < high-school graduate	0.073 (0.008)	0.579 —	−0.128 (0.008)	−0.279 (0.097)	21.7	−6.92 (0.37)	−15.4 (4.85)
Wife high-school graduate	0.082 (0.004)	0.707 —	−0.122 (0.005)	−0.204 (0.052)	28.8	−7.62 (0.23)	−9.20 (2.58)
Wife > high-school graduate	0.071 (0.004)	0.795 —	−0.130 (0.005)	−0.071 (0.060)	33.3	−8.40 (0.28)	−6.05 (2.98)
D. Results for husbands by *wife's education:*							
Wife < high-school graduate	0.069 (0.004)	0.919 —	−0.033 (0.002)	0.053 (0.023)	42.3	−2.36 (0.11)	1.68 (1.57)
Wife high-school graduate	0.076 (0.003)	0.971 —	−0.007 (0.001)	0.031 (0.013)	47.3	−0.70 (0.07)	3.05 (0.91)
Wife > high school graduate	0.064 (0.002)	0.982 —	−0.004 (0.001)	−0.014 (0.014)	48.7	−0.41 (0.07)	−1.53 (0.99)

Notes: The table reports estimates of the coefficient on *More than 2 children* in equation (4) in the text, modified to allow interactions with wives' schooling and husbands' education as indicated. Main effects for each interaction variable (*husband's earnings distribution* and *wife's education*) are included in the equation. Other covariates in the models are those listed in the notes to Table 8. Data are from the 1990 married women and husband samples. Standard errors are reported in parentheses.

TABLE 11—COMPARISION OF 2SLS ESTIMATES USING *SAME SEX* AND *TWINS-2* INSTRUMENTS IN 1980 CENSUS DATA

	All women		Married women		Husbands	
Model	(1)	(2)	(1)	(2)	(1)	(2)
Instrument for *More than 2 children*	*Same sex*	*Twins-2*	*Same sex*	*Twins-2*	*Same sex*	*Twins-2*
Dependent variable:						
Worked for pay	−0.125	−0.079	−0.123	−0.087	0.004	−0.001
	(0.026)	(0.013)	(0.028)	(0.017)	(0.009)	(0.005)
Weeks worked	−5.82	−3.64	−5.47	−4.21	0.65	−0.35
	(1.15)	(0.60)	(1.23)	(0.72)	(0.61)	(0.36)
Hours/week	−4.76	−3.33	−4.91	−3.49	0.57	−0.49
	(0.98)	(0.51)	(1.03)	(0.61)	(0.71)	(0.42)
Labor income	−1961.7	−1262.2	−1329.8	−1453.1	−1194.8	616.8
	(560.5)	(292.8)	(579.1)	(339.8)	(1421.4)	(836.9)
ln(*Family income*)	−0.021	−0.071	−0.049	−0.025	—	—
	(0.067)	(0.035)	(0.057)	(0.033)		
ln(*Non-wife income*)	—	—	0.026	0.051	—	—
			(0.068)	(0.040)		

Notes: The table reports 2SLS estimates of the coefficient on *More than 2 children* in equation (4) in the text using *Same sex* and *Twins-2* as instruments. Other covariates in the models are *Age, Age at first birth,* ages of the first two children, plus indicators for *Boy 1st, Boy 2nd, Black, Hispanic,* and *Other race.* Data are from the 1980 Census. Standard errors are reported in parentheses.

include exogenous covariates to control for mothers' age, race, age at first birth, and the sex of the first two children. Additional covariates included in these models are the ages of the first and second child in quarters. The estimates of female labor-supply effects produced using *Twins-2* are consistently smaller than the corresponding estimates using *Same sex.* Although the contrast between *Same sex* and *Twins-2* coefficient estimates is not large enough to be statistically significant for many of the individual coefficients, the comparison of estimates strongly suggests these two shocks have different effects.

A likely explanation for the smaller *Twins-2* effects is that, conditional on the age of the second child, a third child who is born as a consequence of twinning is necessarily older than a third child who is born for other reasons. This is because third children who are born as twins are exactly the same age as second children, while at least a year and usually longer must go by between the second child's birth and the birth of a non-twin third child. In the 1980 sample, for example, the average age of third children who are twins is 6.4 years while the average age of other third children is five years. Regression-adjusting for the covariates used to construct the estimates in Table 10, the age gap between twins and other third children grows to about 2.5 years. This difference in ages has implications for labor-supply estimates constructed using *Same sex* and twins instruments if the effect of children on labor supply is larger when the children are younger.

We use the following model to check whether differences in the *Same sex* and *Twins-2* 2SLS estimates can be explained by differences in the ages of third children. The equation of interest is

$$(8) \quad y_i = \alpha_0' \mathbf{w}_i + \alpha_1 s_{1i} + \alpha_2 s_{2i} + \alpha_3 a_{1i} + \alpha_4 a_{2i} + \beta_i x_i + \varepsilon_i ,$$

TABLE 12—2SLS AND 3SLS ESTIMATES OF TWO-PARAMETER LABOR-SUPPLY MODELS USING 1980 CENSUS DATA

	Worked for pay		*Weeks/year*		*Hours/week*		*Labor income*	
Variable	Not restricted	Restricted	Not restricted	Restricted	Not restricted	Restricted	Not restricted	Restricted
A. Instruments: *Same sex* and *Twins-2*								
β_0	−0.191	−0.178	−8.94	−8.24	−6.79	−7.22	−2959	−2827
	(0.066)	(0.059)	(2.91)	(2.72)	(2.48)	(2.38)	(1423)	(1002)
β_1	0.015	0.013	0.724	0.616	0.473	0.540	232	211
	(0.096)	(0.009)	(0.429)	(0.398)	(0.366)	(0.348)	(210)	(139)
a^*	12.4	13.4	12.3	13.4	14.4	13.4	12.8	13.4
	(3.69)	(4.38)	(3.42)	(4.38)	(6.03)	(4.38)	(5.62)	(4.38)
B. Instrument: *Same sex* (restricted $a^* = 13.4$)								
β_0	—	−0.184	—	−8.58	—	−7.01	—	−2891
		(0.038)		(1.69)		(1.44)		(827)
C. Instrument: *Twins-2* (restricted $a^* = 13.4$)								
β_0	—	−0.174	—	−8.02	—	−7.35	—	−2787
		(0.030)		(1.32)		(1.13)		(646)

Notes: Panel A of the table reports 2SLS and 3SLS parameter estimates for equation (10) in the text. *Same sex* and *Twins-2* are both used as instruments. The restricted models in Panel A force the parameter a^* (the age at which labor-supply effects decay to zero) to be the same for all four dependent variables in joint estimation using nonlinear 3SLS. Panels B and C report 2SLS estimates of equation (11) using the *Same sex* and *Twins-2* instruments separately. Other covariates in the models are listed in the notes to Table 11. The data are from the 1980 Census all-women sample. Standard errors are reported in parentheses.

where a_{1i} and a_{2i} are the ages of the first two children. The coefficient β_i is now an individually varying causal effect that depends on the age of the third child. In particular, we assume

$$(9) \qquad \beta_i = \beta_0 + \beta_1 a_{3i},$$

where a_{3i} is equal to the age of the mothers' third child for women who have a third child and is equal to zero otherwise. Combining (8) and (9) generates the estimating equation,

$$(10) \quad y_i = \alpha_0' \mathbf{w}_i + \alpha_1 s_{1i} + \alpha_2 s_{2i} + \alpha_3 a_{1i} + \alpha_4 a_{2i} + \beta_0 x_i + \beta_1 (a_{3i} x_i) + \varepsilon_i .$$

Assuming that differences in a_{3i} are the only reason why the *Same sex* and *Twins-2* instruments generate different estimates, we can use both instruments to estimate the coefficients on the two endogenous regressors in (10), x_i and $a_{3i} x_i$.

2SLS estimates of β_0 and β_1 are reported in Table 12 for the full 1980 sample, where a_{3i} was measured to the nearest quarter for the purposes of estimation. All of the estimates of β_0 are negative and all of the estimates of β_1 are positive, suggesting that the negative impact of childbearing declines as the third child ages. The table also reports estimates of the value of a_{3i} at which $\beta_i = 0$; this is $a^* \equiv -\beta_0/\beta_1$. Estimates of a^* are 12.4 years for effects on *Worked for pay,* 12.3 years for effects on *Weeks worked,* 14.4 years for effects on *Hours/week,* and 12.8 years for effects on *Labor income.*[16] We also estimated a^* under the

[16] The linear model for β_i is obviously an approximation since it implies that the effects of childbearing on

restriction that $\beta_1 a^* \equiv -\beta_0$ with the same value of a^* across all four dependent variables. The estimation method for this model is three-stage least squares (3SLS). The restricted estimate is 13.4 with standard error of 4.4. Imposing this restriction leads to slightly smaller standard errors for the estimates of β_0 and β_1. The test statistic for this restriction, distributed as $\chi^2(3)$ under the null hypothesis that the restrictions are satisfied, takes on the value 0.79, which has a p-value of about 0.86.

To further illustrate how this model reconciles the *Same sex* and *Twins-2* estimates, note that if $\beta_1 = -\beta_0/a^*$, we have

$$(11)\quad y_i = \alpha_0'\mathbf{w}_i + \alpha_1 s_{1i} + \alpha_2 s_{2i} + \alpha_3 a_{1i} + \alpha_4 a_{2i} + \beta_0(1 - a_{3i}/a^*)x_i + \varepsilon_i.$$

Substituting the pooled estimate of a^* into (11), we can use the *Same sex* and *Twins-2* instruments to construct separate estimates of β_0 in (11) by treating $(1 - a_{3i}/a^*)x_i$ as the single endogenous regressor. The results when $a^* = 13.4$, reported in Panels B and C of Table 12, show that the *Same sex* and *Twins-2* instruments generate very similar estimates of β_0 for all dependent variables. This suggests that the model of the effect of childbearing embodied in (9), combined with the restriction that β_i decays to zero at age 13.4, does a good job of reconciling the *Same sex* and *Twins-2* estimates.[17]

Differences in the age of the third child constitute one of many possible explanations for the contrast between the *Same sex* and *Twins-2* estimates. For example, there may be economies of scale in parenting two children of the same age. On the other hand, closely spaced young children may require more attention than an older child and a younger child. It is worth noting, however, that the model outlined in this section also serves to explain why the twins estimates reported by Bronars and Grogger (1994) for married mothers are smaller than the *Same sex* estimates reported here. When Bronars and Grogger use twins to estimate the effects of childbearing conditional on the age of the first child (and hence on the age of the twin), they find effects on the labor-force participation rates of mothers of children aged 0–3 remarkably similar to the *Same sex* estimates, with no effects for the mothers of children aged 10–13 (see Table 4 in Bronars and Grogger, 1994). Because second-born twins are younger, on average, than firstborn twins, age differences could also explain why the effects of twins at second birth briefly mentioned by Bronars and Grogger in the text of their article (p. 1149) are larger than their estimates of the effects of twins at first birth.[18]

IV. Implications for the Increase in Female Labor Supply

At the turn of the century, less than 20 percent of all workers were women. Today, women make up almost half the workforce (Goldin, 1990). A number of researchers have attempted to decompose the rise in the female labor-force participation rates into components attributable to demand and supply shifts. For example, Mincer (1962) concluded that 90 percent of the rise in postwar labor-force participation of married women can be attributed to an increase in demand. James P. Smith and Ward (1984) also found that demand characteristics can explain a majority of the increase in total hours worked by all women in between 1850 and 1980. In contrast, Goldin (1990) argues that shifts in supply explain about half of the change in female

labor supply become positive once the third child is older than a^*. This approximation seems harmless since only about 2 percent of third children are older than 13 in our data (because the oldest mother is aged 35).

[17] If we set $a^* = -\beta_0/\beta_1$ using the coefficient estimates from each equation, then the *Same sex* and *Twins-2* estimates of (11) for any equation will necessarily be identical. The point of estimation with a^* fixed at 13.4 is to show how one extra free parameter reconciles *all four* of the *Same sex* and *Twins-2* estimates.

[18] The fact that the Bronars and Grogger estimates are less precise than the twins estimates reported here is likely due to their having drawn a 1-percent sample of singleton births for the comparison sample. Note also that Bronars and Grogger's estimates are for the reduced-form impact of twinning; for purposes of comparison, these estimates should be scaled up by dividing by the twins first-stage effect (about 0.68 in their data).

labor-force participation between 1960 and 1980.

Declining fertility represents a potentially important supply shifter that might account for some of the increase in female labor-force attachment. How much of the trend in labor-force attachment in the population we have studied can be accounted for by reduced childbearing beyond the second child? Table 1 showed that the probability of having more than two children for women aged 21–35 with at least two children fell by 16.5 percentage points between 1970 and 1990, a drop of about 30 percent. At the same time, labor-force participation rates rose by 21.8 percentage points, a 49-percent increase. Similar statistics for other groups reported in Table 1 show that our sample is not unusual in experiencing these trends.

Using the *Same sex* 2SLS estimate of the impact of *More than 2 children* on *Worked for pay* from Table 7 (−0.119), declining fertility can account for an employment increase equal to 0.165×0.12, which is about 2 percentage points. This calculation suggests that even though childbearing clearly affects labor supply, the increase in female labor-force participation has been so large that declining fertility can explain only a small fraction of the overall change.

V. Conclusions

Economic models of household behavior generate a rich variety of predictions and theoretical relationships, few of which have been confronted with credible empirical evidence. The evidence reported here is unique in that it derives from plausibly exogenous sources of variation in family size. However, the empirical results probably raise as many questions as they answer.

2SLS and IV estimates that exploit the fertility consequences of sibling sex composition and twinning both confirm the OLS estimates showing that children lead to a reduction in female labor supply, although the OLS estimates appear to exaggerate the causal effect of children. This is probably not too surprising, at least not to the mothers of small children. What is surprising is that the effects of children on labor supply appear to be much smaller and possibly even absent among college-educated women and women whose husbands have high wages. This result contradicts the predictions of some theories of household time allocation as well as the OLS results, which suggest that the labor-supply consequences of childbearing are larger for more educated women. Our estimates consistently show that the labor-market consequences of childbearing are more likely to be severe for poor and less educated women.

Equally important is the finding that husbands change their labor-market behavior very little in response to a change in family size. Even the husbands of well-educated and relatively well-paid women do not change their work habits in response to the birth of a child. Thus, families absorb the cost of caring for a third child either through a reduction in wife's earnings or by purchasing child-care services from nonfamily providers. If additional childbearing does lead husbands to put additional time into home production of child care, this is done at the expense of the husband's leisure time and not through a reduction in his work effort. We also find little evidence of an increase in husbands' earnings that would offset the decline in wives' earnings. In spite of the increase in women's wages and labor-force participation rates during the period studied here, the labor-market behavior of most married men appears to have remained largely insensitive to the number of children.

REFERENCES

Angrist, Joshua D. "Grouped-Data Estimation and Testing in Simple Labor Supply Models." *Journal of Econometrics*, February–March 1991, *47*(2–3), pp. 243–66.

Angrist, Joshua D. and Evans, William N. "Schooling and Labor-Market Consequences of the 1970 State Abortion Reforms." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5406, January 1996a.

______. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." National Bureau of Economic Research (Cambridge, MA)

Working Paper No. 5778, September 1996b.

Angrist, Joshua D. and Lavy, Victor. "The Effect of Teen Childbearing and Single Parenthood on Childhood Disabilities and Progress in School." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5807, October 1996.

Ashenfelter, Orley and Heckman, James J. "The Estimation of Income and Substitution Effects in a Model of Family Labor Supply." *Econometrica*, January 1974, *42*(1), pp. 73–85.

Becker, Gary S. "Human Capital, Effort, and the Sexual Division of Labor." *Journal of Labor Economics*, January 1985, Pt. II, *3*(1), pp. S33–58.

Becker, Gary S.; Landes, Elisabeth M. and Michael, Robert T. "An Economic Analysis of Marital Instability." *Journal of Political Economy*, December 1977, *85*(6), pp. 1141–87.

Becker, Gary S. and Lewis, Gregg H. "On the Interaction Between the Quantity and Quality of Children." *Journal of Political Economy*, March–April 1973, Pt. II, *81*(2), pp. S279–88.

Becker, Gary S. and Tomes, Nigel. "Child Endowments and the Quantity and Quality of Children." *Journal of Political Economy*, August 1976, Pt. 2, *84*(4), pp. S143–62.

Ben-Porath, Yoram and Welch, Finis. "Do Sex Preferences Really Matter?" *Quarterly Journal of Economics*, May 1976, *90*(2), pp. 285–307.

______. "On Sex Preferences and Family Size," in Julian L. Simon and Julie DeVanzo, eds., *Research in population economics*. Greenwich, CT: JAI Press, 1980, pp. 387–99.

Blau, Francine and Grossberg, Adam J. "Maternal Labor Supply and Children's Cognitive Development." *Review of Economics and Statistics*, August 1992, *74*(3), pp. 474–81.

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

Browning, Martin. "Children and Household Economic Behavior." *Journal of Economic Literature*, September 1992, *30*(3), pp. 1434–75.

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

Butz, William P. and Ward, Michael. P. "The Emergence of Countercyclical U.S. Fertility." *American Economic Review*, June 1979, *69*(3), pp. 318–28.

Coleman, Mary T. and Pencavel, John. "Trends in Market Work Behavior of Women Since 1940." *Industrial and Labor Relations Review*, July 1993, *46*(4), pp. 653–76.

Gangadharan, Jaisri and Rosenbloom, Joshua L. "The Effects of Child-bearing on Married Women's Labor Supply and Earnings: Using Twin Births as a Natural Experiment." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5647, July 1996.

Goldin, Claudia. *Understanding the gender gap*. New York: Oxford University Press, 1990.

______. "Career and Family: College Women Look to the Past." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5188, July 1995.

Gronau, Reuben. "The Effect of Children on the Housewife's Value of Time." *Journal of Political Economy*, March–April 1973, Pt. II, *81*(2), pp. S168–99.

______. "Leisure, Home Production, and Work—The Theory of the Allocation of Time Revisited." *Journal of Political Economy*, December 1977, *85*(6), pp. 1099–124.

______. "Home Production—A Survey," in Orley Ashenfelter and Richard Layard, eds., *Handbook of labor economics*, Vol. 1. Amsterdam: North-Holland, 1986, pp. 273–304.

______. "Sex-Related Wage Differentials and Women's Interrupted Careers—The Chicken or the Egg." *Journal of Labor Economics*, July 1988, *6*(3), pp. 277–301.

Gruber, Jonathan and Cullen, Julie Berry. "Spousal Labor Supply as Insurance: Does Unemployment Insurance Crowd Out the Added Worker Effect?" National Bureau of

Economic Research (Cambridge, MA) Working Paper No. 5608, June 1996.

Heckman, James J. and MaCurdy, Thomas E. "A Life-Cycle Model of Female Labor Supply." *Review of Economic Studies*, January 1980, *47*(1), pp. 47–74.

Heckman, James J. and Walker, James R. "The Relationship Between Wages and Income and the Timing and Spacing of Births: Evidence from Swedish Longitudinal Data." *Econometrica*, November 1990, *58*(6), pp. 1411–441.

Iacovou, Maria. "Fertility and Female Labor Force Participation." Unpublished manuscript, University College, London, July 1996.

Imbens, Guido W. and Angrist, Joshua D. "Identification and Estimation of Local Average Treatment Effects." *Econometrica*, March 1994, *62*(2), pp. 467–76.

Killingsworth, Mark R. and Heckman, James J. "Female Labor Supply: A Survey," in Orley Ashenfelter and Richard Layard, eds., *Handbook of labor economics*, Vol. 1. Amsterdam: North-Holland, 1986, pp. 103–204.

Korenman, Sanders and Neumark, David. "Marriage, Motherhood, and Wages." *Journal of Human Resources*, Spring 1992, *27*(2), pp. 233–55.

Mincer, Jacob. "Labor Force Participation of Married Women: A Study of Labor Supply," in H. Gregg Lewis, ed., *Aspects of labor economics*. Princeton, NJ: Princeton University Press, 1962, pp. 63–105.

______. "Market Prices, Opportunity Costs, and Income Effects," in Carl F. Christ, ed., *Measurement in economics*. Stanford, CA: Stanford University Press, 1963, pp. 67–82.

Montgomery, Mark R. and Trussell, James. "Models of Marital Status and Childbearing," in Orley Ashenfelter and Richard Layard, eds., *Handbook of labor economics*, Vol. 1. Amsterdam: North-Holland, 1986, pp. 205–71.

Morgan, Philip S.; Lye, Diane N. and Condran, Gretchen A. "Sons, Daughters, and the Risk of Marital Disruption." *American Journal of Sociology*, July 1988, *94*(1), pp. 110–29.

Mroz, Thomas A. "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions." *Econometrica*, July 1987, *55*(4), pp. 765–99.

Myrianthopoulos, Ntinos C. "An Epidemiologic Survey of Twins in a Large Prospectively Studied Population." *American Journal of Human Genetics*, November 1970, *22*(6), pp. 611–29.

Nakamura, Alice and Nakamura, Masao. "The Econometrics of Female Labor Supply and Children." *Econometric Reviews*, 1992, *11*(1), pp. 1–71.

National Center for Health Statistics. *Vital statistics of the United States: Natality*. Hyattsville, MD: National Center for Health Statistics, various years.

Newey, Whitney K. and West, Kenneth D. "Hypothesis Testing with Efficient Method of Moments Estimation." *International Economic Review*, October 1987, *28*(3), pp. 777–87.

Pencavel, John. "Labor Supply of Men: A Survey," in Orley Ashenfelter and Richard Layard, eds., *Handbook of labor economics*, Vol. 1. Amsterdam: North-Holland, 1986, pp. 3–101.

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

______. "Life-Cycle Labor Supply and Fertility: Causal Inferences From Household Models." *Journal of Political Economy*, April 1980b, *88*(2), pp. 328–48.

Schultz, Paul T. *Economics of population*. Reading, MA: Addison-Wesley, 1981.

______. "Testing the Neoclassical Model of Family Labor Supply and Fertility." *Journal of Human Resources*, Fall 1990, *25*(4), pp. 599–634.

Smith, James P. and Ward, Michael P. *Women's wages and work in the twentieth century*. Santa Monica, CA: Rand Corporation, 1984.

Stafford, Frank P. "Women's Work, Sibling Competition, and Children's School Performance." *American Economic Review*, December 1987, *77*(5), pp. 972–80.

Thomas, Duncan. "Like Father, Like Son; Like Mother, Like Daughter: Parental Resources and Child Height." *Journal of*

Human Resources, Fall 1994, *29*(4), pp. 950–89.

U.S. Department of Commerce, Bureau of the Census. *Census of population and housing, 1970 [United States]: Public use microdata samples 1/100, 5-percent file* [computer file]. Washington, DC: U.S. Department of Commerce, Bureau of the Census [producer], 1972.

______. *Census of population and housing, 1980 [United States]: Public use microdata sample, 5-percent sample* [computer file]. Washington, DC: U.S. Department of Commerce, Bureau of the Census [producer], 1983.

______. *Census of population and housing, 1990 [United States]: Public use microdata sample, 5-percent sample* [computer file]. Washington, DC: U.S. Department of Commerce, Bureau of the Census [producer], 1995.

Waterhouse, John A. H. "Twinning in Twin Pedigrees." *British Journal of Social Medicine*, October 1950, *5*, p. 197.

Westoff, Charles F.; Potter, Robert G. and Sagi, Phillip C. *The third child: A study in the prediction of fertility*. Princeton, NJ: Princeton University Press, 1963.

Williamson, Nancy E. *Sons or daughters: A cross-cultural survey of parental preferences*. Beverly Hills, CA: Sage, 1976.

Willis, Robert J. "What Have We Learned from the Economics of the Family?" *American Economic Review*, May 1987 (*Papers and Proceedings*), *77*(2), pp. 68–81.