THE REVIEW OF ECONOMIC STUDIES

Review of Economic Studies (2008) **75**, 1085–1120 0034-6527/08/00451085$02.00

The Demand for Sons

GORDON B. DAHL
University of California, San Diego, and NBER

and

ENRICO MORETTI
University of California, Berkeley, and NBER

First version received September 2005*; final version accepted December* 2007 (*Eds.*)

Do parents have preferences over the gender of their children, and if so, does this have negative consequences for daughters versus sons? In this paper, we show that child gender affects the marital status, family structure, and fertility of a significant number of American families. Overall, a first-born daughter is significantly less likely to be living with her father compared to a first-born son. Three factors are important in explaining this gap. First, women with first-born daughters are less likely to marry. Strikingly, we also find evidence that the gender of a child *in utero* affects shotgun marriages. Among women who have taken an ultrasound test during pregnancy, mothers who have a girl are less likely to be married at delivery than those who have a boy. Second, parents who have first-born girls are significantly more likely to be divorced. Third, after a divorce, fathers are much more likely to obtain custody of sons compared to daughters. These three factors have serious negative income and educational consequences for affected children. What explains these findings? In the last part of the paper, we turn to the relationship between child gender and fertility to help sort out parental gender bias from competing explanations for our findings. We show that the number of children is significantly higher in families with a first-born girl. Our estimates indicate that first-born daughters caused approximately 5500 more births per year, for a total of 220,000 more births over the past 40 years. Taken individually, each piece of empirical evidence is not sufficient to establish the existence of parental gender bias. But taken together, the weight of the evidence supports the notion that parents in the U.S. favour boys over girls.

1. INTRODUCTION

Do parents have preferences over the gender of their children? In this paper, we study the effect of having a girl on family structure and fertility behaviour in the U.S. We find that having girls has significant effects on marriage, shotgun marriage when the sex of the child is known before birth, divorce, child custody, and fertility. Taken individually, each piece of evidence is not sufficient to establish the existence of parental gender bias. But taken together, the weight of the evidence supports the notion that parents in the U.S. favour boys over girls.

We begin by documenting the effect of having a first-born girl on the probability that a child grows up without a father in the household.[1] We find that fathers are significantly less likely to be living with their children if they have daughters versus sons. The effect is quantitatively substantial, accounting for a 3·1% lower probability of a resident father for families with a first-born girl. The gender differential in living without a father is substantial in every year since 1960 and remains sizable today. We estimate that in any given year, roughly 52,000 first-born daughters younger than 12 years (and all their siblings) would have had a resident father if they had been boys.

1. Throughout the paper, we focus on the effect of having a first-born daughter versus a first-born son. Other comparisons, such as having two girls versus two boys, are also possible and show even larger effects on family composition. However, since we show that fertility is endogenous to the sex of the first-born child, these estimates are harder to interpret causally.

Three factors are important in explaining why fathers are more likely to live with their sons than their daughters. First, women with first-born daughters are more likely to have never been married than those with first-born sons, particularly in more recent censuses. Even more striking evidence comes from the analysis of shotgun marriages using birth certificate data. We test whether a first-born child's gender affects marital status at delivery when gender is known in advance because the mother has taken an ultrasound test during pregnancy. Among women who have had an ultrasound test, we find that those who have a girl are less likely to be married at delivery than those who have a boy. This evidence suggests that couples who conceive a child out of wedlock and find out that it will be a boy are more likely to marry before the birth of their baby. Second, parents who have first-born girls are significantly more likely to be divorced. The probability that a first marriage ends in divorce for a family with a first-born daughter is 2·2% higher compared to a family with a first-born son. Third, divorced fathers are much more likely to obtain custody of sons compared to daughters.

Growing up without a father has important negative consequences for children. We find that first-born girl families have lower income and higher poverty rates. These effects are large: for children in families with an absentee father due to the first-born daughter effect, family income is reduced by 50% and the chances of poverty are increased by 34%. Notably, children whose first-born sibling is a girl have lower educational achievement.

How should these findings be interpreted? One possible explanation is that parents prefer sons over daughters. According to this interpretation, parents—most likely fathers—have a preference for living with their sons over their daughters. Because of this preference, fathers are more likely to marry their partner, less likely to divorce, and more likely to fight for child custody in case of divorce if they have sons versus daughters. But gender bias is not the only possible explanation. It is possible that parents have unbiased gender preferences, but they recognize that the lack of a male role model is more harmful for sons versus daughters or that fathers have other forms of comparative advantage in raising sons. Another alternative explanation is that the monetary or time cost of raising girls is higher, and this discourages fathers of girls from marrying their partners, remaining in marginal marriages, or taking custody of a child after a divorce. While our results on family living arrangements are consistent with the gender bias hypothesis, in isolation they do not rule out the possibility of these alternative hypotheses.

We examine how child gender affects fertility decisions to help sort out these alternative explanations. Based on a simple model, if parents are biased towards boys, the probability of having additional children should be *higher* for all-girl families than for all-boy families. On the other hand, if parents are unbiased but one of the alternative hypotheses mentioned above is true, we should see that the probability of having additional children is *equal or lower* for all-girl families compared to all-boy families. We find that families with a first-born girl are significantly more likely to have additional children. In families with a first-born daughter, the total number of children rises by 0·3%. This effect is sizable. It implies that first-born daughters caused approximately 5500 more births per year, or 220,000 additional births during our sample period 1960–2000. Together with the results on marriage, shotgun marriage, divorce, and custody, this finding on fertility is most consistent with a preference for boys.

Our findings are important for several reasons. First, regardless of how one interprets our findings on family structure and fertility, we show that child gender matters. The results on the educational and economic outcomes indicate that the negative effects on children living in families where the first-born child is a girl are substantial. While our findings indicate that some of the negative consequences of a first-born daughter affect younger siblings of both genders, girls are overall more likely to be exposed to these negative effects. Moreover, if there is evidence of parental sex bias in family living arrangements and fertility decisions, it may be indicative of other ways in which parents treat boys and girls unequally. For example, even in families where the parents are married, parents who prefer boys may give less attention and nurturing to their daughters. They may also devote fewer financial resources to their education and health. In this sense, our results are related to the existing literature that documents an unequal intra-household allocation of resources.[2]

Understanding the magnitude of parental sex bias and the way it changes over time may also have important policy implications. Rapid progress in sex selection technologies promises to make it increasingly possible for couples to choose the sex of their children. Although these techniques are used by a negligible number of couples due to their high cost, they are expected to become substantially cheaper and more reliable in the near future.[3] High-tech sex selection poses a range of difficult policy dilemmas and is currently extremely controversial. While it is currently legal in the U.S., it is controversial, and other countries have outlawed it.[4] Obviously, the use of gender selection technologies becomes a more pressing issue if parents have strong preferences for one gender. As these technologies become more widely used, strong preferences for boys could, in the long run, lead to imbalances in the population gender ratio.

The remainder of the paper is organized as follows. In Section 2, we present evidence on the differences between girls and boys in the probability of living without a father. In Sections 3 and 4, we discuss competing interpretations of our findings and present additional empirical evidence. Conclusions are given in Section 5.

2. CHILD GENDER AND FAMILY STRUCTURE

Child gender may affect family structure through three channels. First, child gender may affect the probability that a mother never marries but rather remains single. Second, conditional on marriage, child gender may affect the probability of divorce. Third, conditional on divorce, child gender may affect custody arrangements, that is which parent obtains custody of the child.

At any moment in time, parents can be married, divorced, or never married. (For brevity, when we refer to divorce we also include separations.) The overall gender differential in the probability that a first-born child lives without his or her father can be written as:

$$\Pr(\text{No Dad}|G) - \Pr(\text{No Dad}|B) = [\Pr(\text{NM}|G) - \Pr(\text{NM}|B)] + [\Pr(\text{MC}|G) - \Pr(\text{MC}|B)], \quad (1)$$

where G and B denote whether the first child is a girl or a boy, NM stands for never married, and MC stands for mother custody. Equation (1) is simply an accounting identity and reflects the fact that children can live without their father either because their mother never married or because a divorced mother has custody of her children.[5]

2. Many studies document a persistent gender gap in outcomes related to the labour market, most notably education and wages (Blau, 1998; Goldin, 2002). Because parental sex preferences are not easy to control for in wage equations, the finding of lower wages for women may in part reflect parental bias for boys that results in unequal intra-household allocation of psychological and material resources (Butcher and Case, 1994; Thomas, 1990, 1996; Case and Deaton, 2002).

3. Some analysts estimate that the market for sperm sorting in the U.S. will be "between $200 million and $400 million" in the near future. The current price for this procedure in the U.S. is $15,000–$20,000.

4. Canada, the U.K., and some Australian states have outlawed sex selection technologies for non-medical reasons. Sex selection is also forbidden by the European Convention on Bioethics. The American Society of Reproductive Medicine supports sperm sorting but is opposed to the use of pre-implantation genetic diagnosis, in which embryos are produced outside the body and then screened and selected for certain genetic characteristics. It argues that these technologies have "the potential to reinforce gender bias in a society".

5. For simplicity, throughout the paper, we ignore single, never-married fathers with children. Their number is negligible because in virtually all cases of births out of wedlock, children stay with the mother. Due to data limitations in the U.S. Census, we also ignore non-marital unions. The number of children in households with partners of the opposite sex sharing living quarters is small; according to Current Population Survey (CPS) data, this number is less than 4% of all children younger than 18 years in couple households in the year 2000 and much smaller in previous years.

Because custody arises only in the case of divorce, and likewise divorce arises only if a mother has ever been married, the second difference on the R.H.S. of equation (1) reflects a combination of three separate components: the ever-married rate (*i.e.* one minus the never-married rate), the divorce rate conditional on ever having been married, and the custody rate conditional on a divorce having occurred. Hence, an alternative way of writing equation (1) is:

$$\Pr(\text{No Dad}|G)-\Pr(\text{No Dad}|B)=[\Pr(\text{NM}|G)-\Pr(\text{NM}|B)]+[\Pr(\text{EM}|G)$$
$$\times\Pr(D|\text{EM},G)\times\Pr(\text{MC}|D,\text{EM},G)-\Pr(\text{EM}|B)\times\Pr(D|\text{EM},B)\times\Pr(\text{MC}|D,\text{EM},B)], \quad (2)$$

where EM stands for ever married and D stands for divorce. Equation (2) makes clear that marriage, divorce, and custody rates all combine to influence the probability that a child lives without his or her father. A simple decomposition of equation (2) identifies how these three channels impact the chances a daughter versus a son will live without their father:

$$\Pr(\text{No Dad}|G)-\Pr(\text{No Dad}|B)=w_1[\Pr(\text{NM}|G)-\Pr(\text{NM}|B)]+w_2[\Pr(D|\text{EM},G)$$
$$-\Pr(D|\text{EM},B)]+w_3[\Pr(\text{MC}|D,\text{EM},G)-\Pr(\text{MC}|D,\text{EM},B)], \quad (3)$$

where w_1, w_2, and w_3 are weights equaling $1-\Pr(D|\text{EM},G)\times\Pr(\text{MC}|D,\text{EM},B)$, $\Pr(\text{EM}|B)\times\Pr(\text{MC}|D,\text{EM},B)$, and $\Pr(\text{EM}|G)\times\Pr(D|\text{EM},G)$, respectively. As before, equation (3) is simply an accounting identity that makes explicit that the overall difference between daughters and sons living without a father is the sum of three components. The first component is the difference in the probability of having a single mother who has never been married. The second component is the difference in the probability of divorce, conditional on ever having been married. The third component is the difference in the probability that, after divorce, custody of the children is assigned to the mother.

2.1. *Overall effect of child gender on the probability of an absentee father*

We begin our empirical analysis by documenting how child gender affects the probability that a child lives in a household without a father, that is $\Pr(\text{No Dad}|G)-\Pr(\text{No Dad}|B)$ in equation (3). We then assess the empirical relevance of three main channels that may generate gender differences in the probability of an absentee father.

We use data from the 1960 to 2000 U.S. Censuses and include the set of all parents who have children living at home with them. To minimize the probability that some of the children might have left the household, our Census sample includes all families with a mother and/or a father between the ages 18 and 40, with children younger than 12 years living with them.[6] The unit of observation is the household. The models we estimate are linear probability models, where the

6. The 12-year cutoff is probably conservative but minimizes the probability that some of the children have left home. If we examine the spacing between the first and the second child for mothers whose first child is younger than 8 years (and therefore more certain to be at home), we find that for 96% of mothers, the spacing between the first and the second child is 5 years or less. If most children do not leave home until they are 17 or 18 (or even later), using our 12-year cutoff, we are including virtually all children ever born to a mother. We have experimented with alternative cutoff ages for children (8 and 10 years old) and found similar results. An alternative would be to use children of any age and restrict the sample to mothers whose "number of children ever born" is equal to the number of children observed in the household. However, this cannot be done consistently because the variable "number of children ever born" is not available in the 2000 Census. In addition, the "number of children ever born" variable is not available for fathers, so any analysis of divorce or custody that includes fathers could not be based on this variable. Using this alternative definition for the 1960–1990 Census years for mothers only yields similar results to the definition we use (when compared to results based on mothers only).

TABLE 1

First-child gender and the probability of living without a father; U.S. Census data and CPS Fertility Supplements

		Channels for living without a father				
				First marriage ended in divorce		
	Living without a father (1)	Mother never married (2)	Current divorce or separation (3)	CPS Fertility Supplements (4)	Census 1960–1980 (5)	Maternal custody after divorce (6)
First child a girl	0·0050**	0·0009**	0·0016**	0·0103**	0·0045**	0·0250**
	(0·0006)	(0·0004)	(0·0006)	(0·0030)	(0·0007)	(0·0017)
Controls?	Yes	Yes	Yes	Yes	Yes	Yes
First boy baseline	0·162	0·062	0·125	0·32	0·206	0·853
Percent effect	3·1	1·4	1·3	3·2	2·2	2·9
Observations	4,681,967	4,681,967	4,329,737	96,859	1,932,964	588,164

Notes: S.E. are given in parentheses. Data are from the 1960 to 2000 U.S. Censuses for columns (1), (2), (3), and (6), and from the 1960, 1970 form 1, and 1980 Censuses for column (5). Column (4) uses data from the 1980, 1985, 1990, and 1995 CPS Fertility Supplements. The basic sample includes all households with parents who are between the ages 18 and 40, and who have children living at home between 0 and 12 years old (excluding widows). In column (1), the dependent variable is a dummy equal to 1 if the children live without the father—because the parents are divorced and the children live with the mother, the mother has never been married or the father is dead. In column (2), the dependent variable is a dummy equal to 1 if the mother has never married. In column (3), the basic sample is further restricted to households with ever-married parents and the dependent variable is a dummy equal to 1 if the parent is divorced or separated at the time of the survey. In column (4), the CPS sample includes all ever-married mothers between the ages 20 and 70 who report having ever had at least one child and has the advantage that it is based on a full report of fertility and marital histories. Column (5) restricts the basic sample to all ever-married mothers. The sample in column (5) is restricted to 1960–1980 data because information on whether an individual has been married more than once is not available in the 1990 and 2000 Censuses. The dependent variable for both columns (4) and (5) is a dummy equal to 1 if the respondent's first marriage ended in divorce. Column (6) further restricts the basic sample to all households with divorced or separated fathers and mothers and the dependent variable is a dummy equal to 1 if the parent living with the children is the mother. Control variables include a cubic in age, and dummies for race (White people, Black people, Asian, other), education (less than high school, high school, college), region of residence (nine regions), and cohort of birth (10-year birth cohorts). The first boy baseline is calculated as the average predicted probability of the outcome variable of interest for first-born boy families using the estimated coefficients on the control variables. The percent effect is the increase in the probability of living without a father for a first-born girl family to a first-born boy family. **Statistically significant from 0 at the 5% level; *statistically significant from 0 at the 10% level.

relevant outcome is regressed on a dummy, indicating whether the first-born child is a girl.[7] All the models in the paper are weighted and control for a vector of parents' characteristics, including a cubic in age, and dummies for race, education, region of residence, and decade of birth.

Column (1) in Table 1 reports the effect of child gender on the overall probability of having a father in the household. The outcome variable is a dummy equal to 1 if the children are living without a father, defined by whether the children do not live with their biological or adopted father at the time of the Census. The independent variable of interest is a dummy equal to 1 if the first child is a girl. The coefficient indicates that families in which the first child is a girl are 0·50 percentage points more likely to have an absentee father than families in which the first child is a boy. This estimate provides the total effect on the probability of living without a father when the first child is a girl versus a boy, including any indirect effect that operates through subsequent fertility choices.

To help assess the magnitude of the estimated marginal effect, throughout the paper, we also report the "first-born boy baseline", which is a measure of the fraction of first-born boy

7. Probit models yield virtually identical marginal effects compared to linear probability models, and almost all the predicted values for the various outcomes lie between 0 and 1. The one exception is the never-married outcome, where the marginal effect from a probit model is slightly smaller, but still statistically significant, and where 24% of the predicted values lie below 0. Since we take advantage of linearity in our decomposition, we present the OLS estimates.

families without a resident father. In the absence of covariates, this would just be the intercept term, and adding the estimated marginal effect to this baseline would give the fraction of first-born girl families without a father. Since we are including covariates, the first-born boy baseline is calculated as the average predicted probability of having a non-resident father for first-born boy families using the estimated coefficients on the control variables. This average predicted probability is very close to the raw fraction of first-born boy families without a father. We also report the "percent effect", which is the increase in the probability of the outcome of interest for a first-born girl family compared to a first-born boy family; that is, it is the ratio of the coefficient for a first-born girl family to the first-born boy baseline. This percent effect is simply the odds ratio minus 1. In column (1), the percent effect indicates that the probability of living without a father increases by 3·1% when comparing a family whose first child was a boy to a family whose first child was a girl. We view this as a surprisingly large effect. It implies that in any given year, approximately 52,000 first-born daughters age 12 and below (and all their siblings) would have had a father present in the household had they been first-born sons instead.[8]

Although not reported in the tables, models can also be estimated based on the sex mix of the first two and three children as well.[9] In the sample of families with at least two children, the probability of living without a father is 0·38 percentage points higher (S.E. = 0·09 percentage points) if the first two children are boys versus girls. In the sample of families with at least three children, the probability of living without a father is 0·47 percentage points higher (S.E. = 0·19 percentage points) when the first three children are daughters versus sons. While interesting in their own right, these estimates are harder to interpret. The number of boys and girls after the first-born child is no longer random because fertility decisions and marriage, divorce, and custody decisions are likely to be endogenous.

For completeness, in Table A1, we report the effects of a first-born daughter broken down by Census year, decade of birth, race, and education. These estimates by subgroup are necessarily less precise because they are based on smaller samples and therefore are not the main focus of our analysis.

2.2. *Channel 1: Marriage and shotgun marriage*

Having documented sizable differences between boys and girls in the probability of living without a father, we turn to the three possible channels that may be driving them identified in equation (3). Gender differences in the rate at which mothers do not marry represents a first channel through which gender bias may manifest itself. In terms of equation (3), this gender difference is $\Pr(\mathrm{NM}|G) - \Pr(\mathrm{NM}|B)$. Here, we empirically investigate the effect of the sex composition of children on two marriage outcomes. First, we test whether mothers who have boys are more likely to ever have been married than those who have girls. Second, we investigate the effect of the sex of the first child on shotgun marriage for first-time mothers.

2.2.1. Marriage. In column (2) of Table 1, we document the relationship between children's gender and the probability that a mother has never married.[10] We use all mothers between

8. We calculate this as the number of first-born girls aged 0–12 (approximately 10·4 million in an average year during our sample period) multiplied by the coefficient estimate from Table 1, column (1) (0·005).

9. Interestingly, there is a small and significant positive correlation in sibling gender. For example, the second child in a first-born girl family is 1·25 percentage points more likely to be a girl. Ben-Porath and Welch (1976) also notice this correlation in sibling gender but argue that it is small enough to be negligible for parents' fertility decisions.

10. The only existing evidence that we are aware of on this topic is a paper by Lundberg and Rose (2003) based on the Panel Study of Income Dynamics. Using a competing risks analysis, they find that the birth of a son speeds the transition into marriage when the child is born before a mother's first marriage.

age 18 and 40 with children younger than 12 years in the 1960–2000 Censuses. The dependent variable is a dummy equal to 1 if the mother has never been married at the time of the Census. Among women who have one child or more, those whose first baby is a girl are 0·09 percentage points more likely to never have been married than women whose first baby is a boy. The percent effect indicates that having a first-born girl increases the probability of never marrying by 1·4%.

2.2.2. Shotgun marriage. We turn to the effect of child sex on marital status at the time of birth of the baby. For this analysis, we use data from all birth certificates of first-time mothers from the California Birth Statistical Master File, for 1989–1994.[11] We begin by asking whether marital status *at the time of birth* is correlated with the sex of the child. The first column in Table 2 shows that at delivery, gender of the first child is not correlated with marital status. This is reassuring because for a majority of parents in our sample, gender of the first child is unknown until birth. When we control for mother characteristics in column (2)—including race, education, age, Hispanic status, immigrant status, and year—the coefficient flips sign, and remains virtually 0.

The most interesting results of the table are in columns (3) through (6). Here, we test whether gender of the child matters when the mother has taken an ultrasound test during pregnancy and therefore knows with high probability the gender of the baby in advance. Although the main medical reason for taking the test is not disclosure of the child's sex, but is diagnostic, ultrasound tests are typically able to reveal the sex of the baby very accurately by the 12th–14th week of pregnancy.[12]

We regress marital status on a dummy equal to 1 if the child is a girl, a dummy for ultrasound, and the interaction of the girl dummy and the ultrasound dummy. The interaction of the girl dummy and the ultrasound dummy is negative and statistically significant. The coefficient in column (3) suggests that women who take the test and have a girl are 0·37 percentage points less likely to be married at delivery than those who take the test and have a boy. Because we control for the ultrasound main effect, these estimates are not driven by differences in the probability of ultrasound across mothers.[13] When we also condition on mothers' characteristics in column (4), the marginal effect is 0·30 percentage points, slightly smaller than the unconditional one, but still significant. As one might expect, the S.E. in the conditional model is slightly lower than that in the unconditional model.

We interpret these findings as evidence that a child's gender matters for shotgun marriages. Fathers who find out during pregnancy that their child will be a boy are more likely to marry their partner before delivery than those who find out that their child will be a girl. This finding is particularly relevant for the interpretation of our results, as it indicates that the effect of gender on marriage takes place even *before* parents have a chance to interact with their baby.

11. Vital Statistics imputes marital status for this data set based on whether there is a father listed on the birth certificate with the same last name. This imputation is likely to introduce some measurement error in the dependent variable and could lead to inconsistent estimates if the likelihood that a woman adopts her partner's surname is correlated with whether she has an ultrasound and parental gender preferences, as captured in the interaction term.

12. Whitlow, Lazanakis and Economides (1999) find that the test is 90% accurate in determining the sex of the baby at 14 weeks, 79% at 13 weeks, 75% at 12 weeks, and 46% at 11 weeks. Efrat (1999) finds an accuracy of more than 98% by 12 weeks. See also Sicherman and Divon (1999). Approximately 38% of mothers in our sample have had an ultrasound test during pregnancy, while the fraction of mothers taking the test in more recent years is higher.

13. The coefficient on the ultrasound main effect is large and positive, indicating that married women are more likely to take the test. In general, women with higher socioeconomic status are more likely to have an ultrasound. When we control for mother characteristics in column (4), the coefficient on ultrasound drops by more than half.

TABLE 2

First-child gender and the probability of a shotgun marriage; California Birth Certificate data

			Ultrasound during pregnancy				Ultrasound during labour	
			Unweighted		Weighted			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First child a girl	−0·0003	0·0007	0·0010	0·0019**	0·0014	0·0021**	−0·0004	0·0007
	(0·0008)	(0·0007)	(0·0010)	(0·0009)	(0·0010)	(0·0009)	(0·0008)	(0·0007)
First girl × ultrasound		−0·0037**	−0·0030**	−0·0046**	−0·0039**	−0·0024	−0·0007	
			(0·0016)	(0·0014)	(0·0017)	(0·0015)	(0·0055)	(0·0049)
Ultrasound			0·0657**	0·0303**	0·0521**	0·0300**	−0·0003	−0·0057
			(0·0011)	(0·0010)	(0·0012)	(0·0011)	(0·0038)	(0·003)
Controls?	No	Yes	No	Yes	No	Yes	No	Yes
First boy baseline			0·084	0·084	0·084	0·084		
Percent effect			−4·4	−3·6	−5·5	−4·6		
R^2	0·00	0·22	0·01	0·23	0·01	0·19	0·00	0·22
Observations	1,403,601	1,403,601	1,403,601	1,403,601	1,403,601	1,403,601	1,403,601	1,403,601

Notes: S.E. are given in parentheses. Data are from the 1989–1994 California Birth Statistical Master File. The sample includes all first-time mothers. The dependent variable is a dummy equal to 1 if the mother is married when the baby is born. Models in columns (2), (4), (6), and (8) control for mother's race (three groups), mother's education, mother's age, mother's immigrant status, mother's Hispanic status, and year. The weighted regressions use as weights the predicted probability that a woman is unmarried at the time of birth, using mother's race, mother's age, mother's immigrant status, mother's Hispanic status, year, and all their interactions. The first boy baseline is the probability that an unmarried woman gets pregnant and marries before the birth of her first child, calculated using the 1980, 1985, 1990, and 1995 CPS Fertility Supplements. Percent effect is the decrease in the probability of being married at the time of birth for the mother of a first-born girl compared to the mother of a first-born boy. **Statistically significant from 0 at the 5% level; *statistically significant from 0 at the 10% level.

How large is this effect? For the regressions in columns (3) and (4), the percent effect—that is the ratio of the coefficient and the baseline—is around 4%.[14] We view this as a large effect. Moreover, this is likely to be a lower bound because the coefficients in columns (3) and (4) are likely to be downward biased estimates of the effect of child gender for mothers who are not married at conception. The reason is that in our data, we do not know whether a mother was married at the time of conception. Ideally, we exclude women who were already married when they got pregnant from the sample since information from an ultrasound is unlikely to affect their marriage probability at delivery.[15] *A priori*, we expect women who are already married to automatically have a zero coefficient on the interaction term "first girl × ultrasound". In contrast, unmarried women who get pregnant have the potential for this interaction term to influence whether they are married at delivery. By mixing women who are married and unmarried at conception in the sample, we are biasing the coefficient estimate of interest (*i.e.* the coefficient for mothers who are not married at conception) towards 0. Around 66% of women are married at the birth of their baby, indicating that this concern has the potential to severely bias the coefficient on "first girl × ultrasound" towards 0.

In columns (5) and (6), we attempt to correct for some of this downward bias by weighting observations by the predicted probability that the mother is unmarried at conception. The idea is that although we cannot get rid of always-married mothers in the sample, we can probabilistically identify which mothers are most likely to be unmarried at conception. The timing of pregnancy and marriage is recorded in the 1980, 1985, 1990, and 1995 CPS Marriage and Fertility Supplements. We use as weights the predicted probability that a woman is unmarried at the time of birth using mother's race, age, immigrant status, Hispanic status, and year, as well as the interactions of these variables, obtained from a regression based on CPS data.[16] Consistent with our interpretation, the weighted estimates are larger than the unweighted estimates. In the regression without controls, the effect increases 24% to −0·46 percentage points (column (5)), and with controls, the effect increases 30% to −0·39 percentage points (column (6)). The corresponding percent effects are −5·5% and −4·6%.

The estimates in Table 2 can be used to compute how much of the increase in out-of-wedlock births of girls can be accounted for by the increased fraction of mothers who have an ultrasound test. Between 1989 and 2000, the fraction of mothers who had an ultrasound test increased from 27% to 59%, and the fraction of out-of-wedlock births increased from 28% to 34%. In that period, there were on average about 2 million births of girls per year in the U.S. If the effect in Table 2 is time invariant and the effect for first-born daughters also applies to later-born daughters, then the number of out-of-wedlock births of girls due to an ultrasound test was 1998 (2,000,000 × 0·27 × 0·0037) in the year 1989 and 4366 (2,000,000 × 0·59 × 0·0037) in the year 2000. The total number of out-of-wedlock births increased from 560,000 to 680,000 over this same time period.

14. Ideally, the first-born boy baseline in Table 2 would be the fraction of first-time mothers of boys who get pregnant when they are unmarried, get an ultrasound sometime during their pregnancy, and then get married before the baby is born. Unfortunately, our birth certificate data do not have information on whether the mother was married at the time of conception. Fortunately, timings of pregnancy and marriage are recorded in the 1980, 1985, 1990, and 1995 CPS Marriage and Fertility Supplements. While not ideal (since we do not know whether a woman had an ultrasound), we use this data to calculate an imputed baseline for the probability that an unmarried woman gets pregnant and marries before the birth of her first child.

15. Using the 1980, 1985, 1990, and 1995 CPS Marriage and Fertility Supplements, we find that the probability that a couple is married at conception and decides to get divorced before the birth of the baby is extremely rare: it occurs in only 0·1% of cases.

16. The weights exhibit the expected patterns. For example, younger women and Black women have a higher predicted probability of being unmarried. Using the predicted probability that a woman is unmarried at the time of birth (instead of conception) as weights is not ideal. However, we suspect that the difference is trivial, as the predicted probability that a woman is unmarried at the time of birth is likely to be highly correlated with the predicted probability that a woman is unmarried at the time of conception.

Therefore, 1·9% of the increase in out-of-wedlock births of girls can be accounted for by the increase in the fraction of mothers who take the ultrasound test.

The confidential California Birth Statistical Master File also reports whether the mother has had an ultrasound *during labour*.[17] We use this information to perform a specification check. Unlike for ultrasound during pregnancy, we expect the interaction of ultrasound during labour and child gender not to matter for marital status at birth. The coefficient on the interaction in column (7) is 0·0024, but it is very imprecisely estimated and is not statistically significant from zero. When we control for mother's characteristics in column (6), the coefficient on the interaction drops to virtually 0.

Finally, we note that selective abortion based on information generated by the ultrasound is in theory possible. While we cannot completely rule out the possibility of selective abortion based on child sex, we find little evidence of it. In our sample, the gender ratio is the same among babies whose parents took the ultrasound test and babies whose parents did not. Notably, even if it existed, selective abortion driven by a strong preference for boys would imply that the parents with the strongest preference for boys—that is those who are willing to abort a child based on her gender—are not in our sample, so that our measured effect may be smaller than the true effect.

2.3. *Channel 2: Divorce*

In this section, we investigate the relationship between the sex composition of children and the probability of divorce. Divorce represents a second channel through which the gender of children affects family structure. In terms of equation (3), we estimate $[\Pr(D|\mathrm{EM},G) - \Pr(D|\mathrm{EM},B)]$. We find that parents who have girls first have historically been more likely to be divorced than those who have boys first.

In contrast to the marriage, shotgun marriage, and custody outcomes, the existing literature on the effect of child gender on divorce is sizable. Pioneering work in the 1980's sociology literature by Spanier and Glick (1981) and Morgan, Lye and Condran (1988) documented a correlation between child gender and divorce. Later research by Katzev, Warner and Acock (1994); Mott (1994); and Morgan and Pollard (2002) also finds an effect, although Morgan and Pollard find no effect for more recent years. In contrast, other researchers find no statistically significant correlation between child gender and divorce (Mauldon, 1990; Bracher *et al.*, 1993; Devine and Forehand, 1996; Wu and Penning, 1997; Diekmann and Schmidheiny, 2004).

One limitation of many previous studies is that they pool together families of different size, thus confounding the effect of gender on divorce with the effect of gender on family size. As shown in Section 4, the gender of existing children affects future fertility decisions. Because of this endogeneity of family size, many of the existing estimates are difficult to interpret. Another advantage of our analysis is the large sample size of our Census data. One possible explanation for why several studies fail to find a statistically significant effect is the use of small samples. With the exception of Morgan and Pollard (2002), sample sizes in recent sociological studies range between 140 and 10,000 observations.

In column (3) of Table 1, the dependent variable is a dummy equal to 1 if the children currently reside with a divorced or a separated mother or father, and 0 if the children live with parents who are currently married.[18] The third column reports the estimated coefficient when this divorce dummy is regressed on the gender of the first child, for parents with one child or more.

17. Ultrasound scans during labour and birth are used to diagnose major complications, such as uterine bleeding (to visualize the placenta). Only 2·2% of mothers in our sample have an ultrasound during labour.

18. Our analysis implicitly assumes that all the children from a divorced family live with either their father or their mother. Using administrative data on divorce from 1989 Vital Statistics records, we find that the children are divided in only 2% of cases.

The estimate is not affected by the endogeneity of family size. It provides the total effect on divorce of having a girl versus a boy for the first child, including any indirect effects that operate through differences in subsequent fertility, gender birth order, or gender mix. The coefficient indicates that parents whose first child is a girl are 0·16 percentage points more likely to be divorced than those whose first child is a boy. The percent effect indicates that the probability of divorce when moving from a family whose first child was a boy compared to a family whose first child was a girl increases by around 1·3%.

As the size of the family increases, the divorce effect becomes even larger. While not reported in the table, we find a 0·19 percentage point effect for having two daughters first, or 1·9% of the baseline all-boy divorce rate. When we turn to families with three or more children, we find even larger effects, with a 0·31 percentage point effect for having three daughters versus three sons first, or 3·2% effect. The endogeneity of family size makes these estimates more difficult to interpret, however. When examining separate estimates by education category, the divorce effect declines monotonically with education level, so that mothers with a college education have almost no difference in divorce rates based on the gender of their first-born child (Table A1). When broken down by race, the effects for minorities are too imprecisely estimated to be informative.[19]

A limitation of the divorce results presented in column (3) is that due to the nature of the Census data, the dependent variable is whether a parent is divorced or separated at the time of the Census survey. However, many divorcees remarry. If having daughters versus sons affects the probability of remarriage, this could confound the effect of having girls on divorce. To gauge the potential impact the divorce definition has on our estimates, in columns (4) and (5), we examine the probability that a respondent's first marriage ended in divorce using two alternative data sets. We first use the 1980, 1985, 1990, and 1995 CPS Marriage and Fertility Supplements, which report the complete fertility and marital history of respondents in detail. Column (4) reports larger percentage point effects for first-born girl families than those in column (3); these effects are around 1 percentage point. The implication is that parents with first-born girls are 3·2% more likely to have their first marriage end in divorce. However, column (4) also makes clear the limitations of CPS data. Even with 4 years of pooled CPS data, the S.E. on the estimates are large. We have also used the fact that in the 1960–1980 Censuses, one can back out whether a respondent's first marriage ended in divorce. Using this as the dependent variable, the effect in column (5) is 0·45 percentage points, which is almost three times larger than the result reported in column (3). As we see in Section 2.6, part of the reason for the increase is that the divorce effect (using the definition of column (3)) is stronger from 1960 to 1980 compared to 1990–2000. However, a majority of the increase appears to be due to a larger effect on first marriages, consistent with column (4).

2.4. *Channel 3: Custody*

We next turn to the effect of child gender on the custody arrangements of children with divorced mothers and fathers. Custody represents a third channel for child gender to affect family structure. In terms of equation (3), we estimate $\Pr(\mathrm{MC}|D,\mathrm{EM},G) - \Pr(\mathrm{MC}|D,\mathrm{EM},B)$.

In column (6) of Table 1, we present estimates of the probability that a divorced mother has custody based on the gender of the children. The sample includes only divorced mothers and

19. The divorce estimates for Asians are negative, although insignificant. This counterintuitive result has a simple explanation. For all racial groups, the effect for divorce is large in the earlier decades and declines in later decades. As it turns out, the number of Asians in the Census is very small in earlier years (only 1·5% of the sample from 1960 to 1980), and much larger in more recent years (3·8% of the sample from 1990 to 2000). As a consequence, the estimate for Asians is picking up mostly the effect in later years, which are smaller for all groups. When we re-estimate our models by race and period, we find that Asians have the largest percent effect from 1960 to 1980.

fathers in the Census, where, as before, the sample is restricted to parents between ages 18 and 40 who have children younger than 12 years living with them. We define maternal custody as whether the children are living with a divorced mother at the time of the Census. The variable we call "custody" is therefore a broad measure of the mother's versus a father's access to and time spent with her children and is likely to reflect joint custody and visitation rights in addition to sole maternal or paternal custody. This is exactly the type of variable we want if we are thinking about day-to-day access to children.

We find that children of divorced parents are more often assigned to the mother, but when the father does obtain custody, he is more likely to have custody of his sons. Specifically, column (6) indicates that divorced mothers with a first-born daughter are 2·5 percentage points more likely to have their children living with them. This amounts to a 2·9% effect. Put somewhat differently, column (6) implies that divorced fathers obtain custody 14·7% of the time when they have a first-born son, but only 12·2% of the time when they have a first-born daughter.

2.5. *Direct and indirect effects of child gender and timing*

We focus attention on the sex of the first child because this has the cleanest causal interpretation since whether the first child is a boy or a girl can arguably be viewed as random.[20] Estimates based only on the sex of the first child provide the total effect on the relevant outcome of having a first-born girl versus a first-born boy. As such, our casual estimates capture several effects—the *direct* effect of a first-born girl on family structure, an *indirect* effect due to subsequent differential fertility, and additional *indirect* effects due to gender birth order and gender mix.[21]

While it would be interesting to isolate the direct effect of the gender of the first-born child, this is not fully possible with our Census data. To aid in the interpretation of our findings, we present additional estimates in Table 3.[22] In the first panel, we restrict the sample to families with only one child and re-estimate the models in Table 1. This holds constant family size, although we emphasize that this restricted sample consists of parents who endogenously chose to have only one child. For this restricted sample of one-child families, the probability of an absentee father for a first-born girl versus boy rises to 0·74 percentage points (versus 0·50 percentage points in column (1) of Table 1). The estimates for most of the other outcomes (with the exception of "current divorce") are similar to, or larger, than the analogous estimates in Table 1.

One way to benchmark our estimates is to compare the first-born girl effect versus the effect of family size on family structure. In the second panel of Table 3, we include the number of children as an additional explanatory variable and instrument for it with the occurrence of twins at first birth. Several points are worth mentioning. First, other researchers have instrumented for family size with twin births and found similar estimates for the divorce outcome (Jacobsen, Pearce and Rosenbloom, 2001; Cáceres-Delpiano, 2006; Jena, 2006). Second, the estimated coefficients for "first child a girl" do not appreciably change with the addition of the family size

20. Recent medical literature documents that natural methods based on timing of intercourse have no significant effect on offspring sex (Wilcox, Weinberg and Baird, 1995). Furthermore, although the technology is evolving rapidly, clinical methods to influence offspring gender are currently used by an insignificant fraction of the population and are still not very accurate. We have calculated the fraction of first-born children who are girls in each Census and found that it is roughly constant over time (0·491, 0·488, 0·487, 0·486, and 0·489, for 1960–2000, respectively). The fraction of second-born children who are girls, conditional on the sex of the first child, is also relatively stable over time (0·498, 0·500, 0·495, 0·498, 0·492 conditional on a first-born girl, and 0·490, 0·482, 0·484, 0·486, 0·486 conditional on a first-born boy for 1960–2000, respectively). These trends are consistent with the notion that improvements in technology in gender selection have not yet received widespread adoption.

21. We note, however, that the effect of a first-born girl does not include the demand for variety in its simplest form, where parents have a preference for at least one child of each gender. For example, parents deciding to have a second child have a roughly 50% probability of achieving gender variety, regardless of their first-born's sex.

22. We thank an anonymous referee for suggesting these additional sets of estimates.

TABLE 3

Additional estimates of first-child gender and the probability of living without a father; U.S. Census data and CPS Fertility Supplements

	Living without a father (1)	Channels for living without a father: Mother never married (2)	Current divorce or separation (3)	First marriage ended in divorce: CPS Fertility Supplements (4)	First marriage ended in divorce: Census 1960–1980 (5)	Maternal custody after divorce (6)
	Effects for families with only one child					
First child a girl	0·0074**	0·0021**	−0·0007	0·0150**	0·0039**	0·0401**
	(0·0011)	(0·0008)	(0·0010)	(0·0031)	(0·0011)	(0·0027)
Observations	1,822,578	1,822,578	1,599,206	20,166	711,889	287,912
	Effects controlling for number of children (number of children instrumented with twins at first birth)					
First child a girl	0·0050**	0·0009**	0·0016**	—	0·0045**	0·0249**
	(0·0006)	(0·0004)	(0·0006)		(0·0007)	(0·0017)
No. children	0·0044	0·0017	0·0018	—	0·0116**	0·0086
	(0·0039)	(0·0022)	(0·0038)		(0·0056)	(0·0097)
Observations	4,681,967	4,681,967	4,329,737	—	1,932,964	588,164
	Effects by age of first child					
First child a girl × first child age	0·00039**	−0·00034**	0·00028**	0·00024	0·00038**	0·00217**
	(0·00015)	(0·00011)	(0·0014)	(0·00023)	(0·00017)	(0·00051)
First child a girl	0·0025**	0·0029**	−0·0002	0·0054	0·0020*	0·0092**
	(0·0010)	(0·0008)	(0·0009)	(0·0057)	(0·0011)	(0·0039)
First child age	0·0105**	0·0014**	0·0092**	0·0174**	0·0116**	0·0154**
	(0·0001)	(0·0001)	(0·0001)	(0·0003)	(0·0002)	(0·0004)
Observations	4,681,967	4,681,967	4,329,737	96,859	1,932,964	588,164

Notes: S.E. are given in parentheses. See notes to Table 2 for data samples, variable definitions, and included control variables. The first panel is restricted to families with one child. The second panel uses twins at first birth as an instrument for the number of children. Twins at first birth is imputed based on the first and second child being the same age in years and born in the same quarter of the year. The first-stage estimate for column (1) indicates that first-born twins raise total fertility by 0·697 children (S.E. = 0·009). First-stage estimates for the other columns are similar. Results are not reported for the CPS sample in the second panel due to the small number of twins in that data set. **Statistically significant from 0 at the 5% level; *statistically significant from 0 at the 10% level.

variable (although not shown, this is also true without using the twins instrument). Third, the estimates suggest that additional children destabilize families. The estimates are imprecisely estimated and not statistically significant except for the "first marriage ended in divorce" outcome. For this outcome, having an additional child is associated with a 1·2 percentage point increase in the chance of a first marriage ending in divorce, or a 5·5% increase. Fourth, the estimates suggest that the first-born girl effect and the additional child effect are roughly of the same order of magnitude for the various outcomes. Finally, using the additional child estimates, one can calculate the indirect effect of increased fertility due to a first-born daughter. These indirect fertility effects are negligible, as the number of additional births due to a first-born daughter is small (Section 4.1). For any of the outcomes, we calculate that the indirect fertility effect due to a first-born daughter accounts for less than 2% of the total effect of a first-born daughter.

In the final panel of Table 3, we explore the timing of the gender effects. From the never-married results (Table 1) and the shotgun marriage results (Table 2), it appears that some couples *anticipate* the net benefits of marriage to be lower with a daughter versus a son. The estimates for the divorce channel in Table 1 are purged of these couples who choose not to marry in the first place. Therefore, one interpretation of the positive divorce estimates in Table 1 is that they reflect parents' learning about the true costs and benefits of children within marriage. To examine this more closely, we added the age of the first-born child into all of our regressions, and more importantly, an interaction term for the age and gender of the first-born child. The results indicate that there is a quarter-point percentage gap between living with a father for boys versus girls younger than 1 year. This gap appears to be driven in large part from the gap in the never-married rates for newborns in column (2). The absentee father gap between girls and boys increases with a child's age. Assuming a linear specification, each additional year increases the gap by 0·04 percentage points, so that by the time children are 12 years old, the gap rises to 0·72 percentage points. This result is statistically significant. The same pattern also exists for the divorce and custody channels. Fathers are more likely to stay with a son, and the effect magnifies as a child ages. These results provide suggestive evidence that parents learn over time about their preferences for sons versus daughters. Of course, alternative mechanisms could also describe the rising gap over time, including age-specific gender effects, family size effects, and subsequent child gender order and mix. Regardless of interpretation, however, we note that the estimates are still casual.

2.6. *Relative importance of the marriage, divorce, and custody channels*

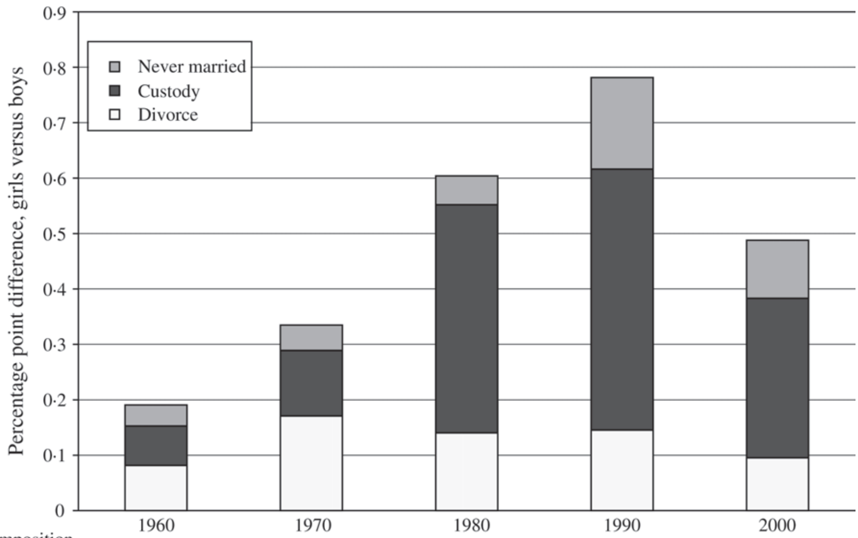
We have documented that child gender affects the probability of growing up without a father. We have also shown that gender differences in marriage, divorce, and custody all appear to be significant channels, although we have not said anything yet about their relative importance. In this section, we use equation (3) to decompose the overall difference in the fraction of girls versus boys who live without their father into its three component parts. The goal of this exercise is purely descriptive and the evidence we present should be considered simply as an *accounting* decomposition.

Figure 1 graphs the difference in the probability that a father is present in the household based on the sex of the first-born child, for each Census year. The height of the bars indicates that in each Census year, there are sizable gender differences in $\Pr(\text{No Dad}|G) - \Pr(\text{No Dad}|B)$. An interesting pattern emerges: the gap in the fraction of daughters versus sons not living with their fathers increases monotonically from 1960 to 1990. Although it declines in the 2000 Census, it remains substantial. In 1960, a father is 0·2 percentage points more likely to live with his children if his first child is a boy. By 1990, the gap widens to almost 0·8 percentage points before falling to 0·5 percentage points in 2000.

The three divisions within each bar represent the contribution of the divorce, custody, and never-married channels to the overall effect. The decomposition is based on equation (3), which puts the effects on the same scale. Specifically, the never-married portion of each bar is the first difference on the R.H.S. of equation (3), $w_1[\Pr(\text{NM}|G) - \Pr(\text{NM}|B)]$. The divorce and custody portions of each bar are calculated as $w_2[\Pr(D|\text{EM},G) - \Pr(D|\text{EM},B)]$ and $w_3[\Pr(\text{MC}|D,\text{EM},G) - \Pr(\text{MC}|D,\text{EM},B)]$, respectively.[23]

Figure 1 reveals that in the early years, the gender difference in divorce rates is the most important factor in explaining the gender difference in the probability of living without a father. It accounts for 43% and 51% of the total difference in the probability of living without a father in 1960 and 1970. By 1980, the divorce effect only accounts for 23% of the difference. In 1990

23. The decomposition in equation (3) is not unique. Specifically, equation (3) is derived by adding and subtracting the term $[\Pr(\text{EM}|G)+\Pr(\text{EM}|B)]\times\Pr(D|\text{EM},G)\times\Pr(\text{MC}|D,\text{EM},B)$ to equation (2). One could also add and subtract other terms and get slightly different expressions for the weights. This is analogous to the "index number problem" encountered in a standard Oaxaca decomposition. Empirically, this is not an issue because the alternative weights all yield virtually identical results for the decomposition.



Decomposition

	1960	1970	1980	1990	2000
Never married (%)	20	14	9	21	22
Custody (%)	37	35	68	60	59
Divorce (%)	43	51	23	19	20

FIGURE 1

Child gender and the probability of living without a father by Census year, with a decomposition into the component channels

and 2000, the divorce effect explains around 20% of the difference, although the underlying divorce coefficients are no longer significant in these later years. As the importance of differential divorce rates declines, the relative contribution of differential custody rates in the decomposition increases. Not only are fathers increasingly likely to obtain custody of their children in general (from 7% in 1960 to 22% in 2000), but they are also increasingly more likely to obtain custody of their sons relative to their daughters. According to the decomposition, the custody channel accounts for slightly over one-third of the total difference in 1960 and 1970 but increases substantially to more than 60% in 1980, 1990, and 2000. One possible explanation for this trend is that fathers may now be using the custody margin as a substitute for the divorce margin as a way to retain access to their sons. The third part of the decomposition is the difference in the probability that a mother has never been married. As mothers with children have become increasingly likely to be married (from 1% in 1960 to 14% in 2000), the percentage point gap between mothers with first-born girls and those with first-born boys has also been increasing. This effect explains a larger proportion of the gender difference in 1990 and 2000 compared to 1970 and 1980.

We draw three conclusions from Figure 1. First, unequal access of fathers to their sons versus daughters has not disappeared but has actually increased from 1960 to 1990 until dropping somewhat in 2000. Second, the method by which fathers gain differential access to sons versus daughters has altered in the past half century. In the early years of our data, couples were substantially more likely to remain married (*i.e.* not divorce) if their first child was a boy. However, as paternal custody rates have risen three-fold over time, custody has replaced divorce as the primary avenue by which fathers have differential access to their sons. This change may reflect no change in the preferences of parents, but rather in the custodial laws in place or the behaviour of judges (*e.g.* Brinig, 2005). In addition, as single motherhood has become more common, an increasing number of mothers and fathers have chosen not to marry if they have a girl. Third, we conclude that examining each of the outcomes in isolation would yield an incomplete and misleading picture. For example, by examining differential divorce rates alone, one might be tempted to say that gender differences in family structure are getting smaller over time, while in fact these rates are only part of the story.

Do the trends documented in Figure 1 mean that parents have become more gender biased over time? Not necessarily. Secular changes in family structure make it difficult to compare the amount of gender bias over time. In particular, the never-married, divorce, and paternal custody rates have all risen over this same time period, irrespective of child gender. This suggests that the opportunities and relative costs for parents to express their gender preferences through various living arrangements may have changed over time. If gender preferences remained constant over time but the "costs" to fathers of gaining access to sons decreased—for example because paternal custody became more widespread, irrespective of child gender—we would still expect to see a change in the number of sons versus daughters living with their fathers. In the end, we cannot say whether gender bias has increased or decreased over time. What we can say is that, if what drives the gender differences in Figure 1 is indeed gender bias, it has not disappeared by the year 2000.

2.7. *The effect of child gender on economic and educational outcomes*

We have shown that girls are more likely to live without their father because a first-born daughter decreases the probability of marriage, increases the probability of divorce, and decreases the probability of paternal custody. In this section, we are interested in assessing whether these changes result in negative consequences for children. It is natural to think that the absence of a father could have important economic and emotional consequences. Indeed, a large body of academic research points to a negative association between absentee fathers and children's outcomes.[24] However, these associations are often difficult to interpret because they could in theory be explained by unobserved heterogeneity. Families with an absentee father are likely to differ from intact families in many unobservable dimensions. To the extent that these unobservables are correlated with children outcomes, naïve comparisons will suffer from spurious correlation.

Here, we directly test whether children in families where the first-born child is a girl experience worse economic and educational outcomes than children in families where the first-born child is a boy.[25] For the economic outcomes, we use several variables measured at the family level in our U.S. Census data spanning 1960–2000. We find that families with first-born girls appear to be significantly worse off financially than families with first-born boys. For example, after conditioning on age, race, age, education, region, and decade of birth, family income is 0·44 (S.E. = 0·1) percentage points lower for families where the first-born child is a girl relative to families where the first-born is a boy. Notably, the drop in income associated with child gender appears to be particularly important at the bottom of the income distribution. Children in families

24. For example, growing up in a single-mother household is associated with higher probability of living in poverty, dropping out of high school, becoming a teenage parent, and experiencing unemployment. See McLanahan and Sandefur (1994), Haveman and Wolfe (1994), and Case and Paxson (2001).

25. In related research, Lundberg and Rose (2002*a*) find that men's labour supply and wage rates increase more with the birth of sons versus daughters, Lundberg and Rose (2002*b*) examine the differential investment in sons versus daughters, and Mammen (2003) finds little evidence that boys and girls fare differently in child support receipt or amounts if their parents are divorced.

with first-born girls are 0·12 (S.E. = 0·06) percentage points more likely to be below the poverty line and 0·10 (S.E. = 0·04) percentage points more likely to be on welfare. These first-born girl families are also 0·17 (S.E. = 0·07) percentage points less likely to own a home. These effects are all statistically significant.

While small on average, these effects are substantial for those families where the presence of a first-born girl results in the absence of the father. For example, the effect for family income implies that family income is on average $128 lower for families with first-born girls. This is the average effect for all first-born girl families, regardless of whether the family is intact or not. If having a first-born girl affects family income only through the probability of having an absentee father, then the income drop is concentrated on just a few families: those where the father left the house because he had a first-born daughter. The magnitude of this effect for affected families is sizeable and significant. When translated into dollar terms, it implies that family income drops by $18,000, or approximately 50%. Similar estimates for poverty imply that children with absentee fathers due to the first-born daughter effect are 24% more likely to be poor. (These estimates are instrumental variable estimates where having an absentee father is instrumented for by child gender, and are all statistically significant.)[26]

Does the fact that girls are more likely to grow up without a father, and to experience worse economic outcomes, translate in lower schooling achievement for girls? It is difficult to compare educational outcomes of first-born boys with first-born girls since gender differences in achievement are likely to dominate any psychological or economic effect that may arise from the absence of the father. Instead, we compare the outcomes for second-born children whose first-born sibling is female with the outcomes for second-born children whose first-born sibling is male. We use two alternative educational outcomes that work for young children. First, we use an indicator for whether the child is not on grade, that is whether the child lags behind relative to his or her peers.[27] Being on grade is defined as a dummy equal to 1 if the child is at or above the median grade achieved by other children with the same state, age (using age in years and quarter of birth), gender, and Census year. We find that second-born children whose older sibling is a girl are 0·36 (S.E. = 0·10) percentage points less likely to be on grade compared to second-born child whose older sibling is a boy, after controlling for whether the first-born sibling is on grade. Because we are comparing a child's grade with the median grade achieved in the same state, age, quarter of birth, and gender cell, gender differences in grade progression are not polluting our estimates. Second, we use an indicator for whether a 3- or 4-year-old child is enrolled in pre-school. We find that second-born children whose older sibling is a girl are 0·42 (S.E. = 0·19) percentage points less likely to be enrolled in pre-school compared to second-born children whose older sibling is a boy.

Overall, these estimates suggest that the presence of a first-born girl in the family is associated with generally worse outcomes. First-born girls are more likely to grow up economically disadvantaged, presumably because they are less likely to live in intact families. For families with more than one child, these negative consequences extend to younger siblings.

3. INTERPRETATION

While the evidence presented in Section 2 reveals a clear pattern of gender differences in family structure, there are several alternative interpretations for why these differences may arise. First,

26. This result is consistent with contemporaneous papers by Bedard and Deschenes (2005) and Ananat and Michaels (2007), who use child gender as an instrumental variable to study women's labour supply and family income.

27. Failure to be on grade has been used by others as an indicator of grade repetition and generally lagging educational performance (Currie and Yellowitz, 2000; Oreopoulos, Page, and Stevens, 2003; Cascio, 2004).

there is the possibility that parents are "gender biased", that is they prefer living with sons over daughters. This explanation is purely based on tastes: for whatever reason, one of the two parents (or both) derives more utility from living with boys than girls. In the Appendix, we model this possibility by assuming that the utility function of one of the two parents (or both) depends on the number of children and on their gender.

Alternatively, it is possible that parents are unbiased but that each parent has a comparative advantage in raising children of their own gender. For example, child psychologists argue that the presence of the father is relatively more important for boys than for girls.[28] We call this the "role model" hypothesis. According to this hypothesis, parents care about the well-being of their children, and when deciding whether to marry, divorce, bargain over custody arrangements, or have more children, take into account the asymmetric impact of a father's presence on boys and girls. In the Appendix, we model this possibility by assuming that one of the arguments in the parent's utility function is their children's utility, and that boys suffer a larger utility loss if they are separated from their father than girls.

A closely related explanation is that, for *technological* reasons, fathers have a comparative advantage at raising boys. This explanation posits that men are relatively more efficient at taking care of and raising sons than daughters.[29] It is similar in spirit to the role model hypothesis, although it enters through the budget constraint instead of the utility function. It implies that the "price" of a boy is lower if the father is present in the household. We call this the "technology" hypothesis. This hypothesis and the role model hypothesis are similar to the "child production" effects discussed in Lundberg (2005).

A fourth possible explanation is that the monetary or time cost of raising girls is, for exogenous reasons, different than the cost of raising boys. For example, the cost for girls could be higher than that for boys.[30] We call this the "differential cost" hypothesis. In the Appendix, we model this possibility by letting the budget constraint depend on child gender. In particular, we assume that a parent's budget constraint includes, among other items, the number of girls times their price plus the number of boys times their price, and we allow the price of girls to differ from that of boys.

Another possible explanation is that boys are harder to look after than girls—possibly because they have more health and behavioural problems—and fathers are more likely to stay with sons because they care about their children's outcomes. In this story, altruistic fathers realize that growing up in an intact family is relatively more important for children who have more problems, and boys have more problems. We call this the "compensatory behaviour" hypothesis.

Each of these hypotheses could potentially explain the evidence in Section 2. In the next section, we provide the intuition on why this is so. We also explain the motivation for three additional pieces of evidence that may help us test for gender bias. In the Appendix, we develop a more formal framework for the interested reader. Of course, we do not think of the five explanations presented above as mutually exclusive. More than one explanation could very well be at play. While our goal is to test for the existence of gender bias, we will not be able to rule out

28. See Lamb (1997, 1987) for a survey of the literature, which concludes that fathers play a larger role in the development of sons than of daughters. Fathers spend more time with their sons (Morgan *et al.*, 1988; Lamb, 1997), and longitudinal data on child development show that the absence of a father has more severe and enduring impacts on boys than on girls (Hetherington, Cox and Cox, 1978). Several papers in economics examine the effect of absentee fathers on children of both genders (*e.g.* Gruber, 2000; Page and Stevens, 2004).

29. A 2000 Gallup Poll survey of 1026 adults indicates that 63% of men and 42% of women think that boys are easier to raise than girls (with no difference also being a valid response). Interestingly, women who have never been married are slightly more likely to think that girls are easier to raise, but women who have been married overwhelmingly believe that boys are easier to raise.

30. Olsen (1983) estimates that for one-child families, a girl costs around $900 more each year to raise up to the age of 18 compared to a boy.

the alternative stories as contributing factors. Moreover, the explanations that we have chosen are intended to be fairly general. Several variants of these stories could readily fit within our framework, as long as one can model child gender as affecting parents' utility, child utility functions, or the budget constraint.[31]

3.1. *What each hypothesis implies for marriage, divorce, and custody*

Consider first the results for marriage and shotgun marriage. One natural interpretation is that parents, and perhaps fathers in particular, have a preference for living with and spending time with sons over daughters (*i.e.* the gender bias hypothesis). This could persuade some couples near the margin to marry if they have boys since married fathers are likely to have significantly more interaction with their children. If parents have a net bias for boys, marriages with a son versus a daughter generate a larger combined parental utility gain. Therefore, couples who give birth to sons are more likely to marry.

Alternative explanations are also consistent with the evidence on marriage and shotgun marriage. It is possible that unbiased fathers decide to marry their partner if she has boys, not because they prefer boys over girls but because they believe that an absentee father hurts sons more than daughters (role model hypothesis) or that fathers have other comparative advantages in raising sons versus daughters (technology hypothesis) so that the surplus associated with a potential marriage is larger with sons. It is also possible that the time or monetary cost of raising girls is higher than the cost of raising boys (differential cost hypothesis). If parents are unbiased, parents do not care about the role model effect, and girls are more expensive, some fathers may be more willing to marry their partner if she has a boy because the cost is lower. Notably, the shotgun marriage results suggest that all these effects would need to be anticipated before either parent interacts with the child outside the womb. Finally, it is possible that unbiased parents realize an intact family is relatively more important for children who have more health or behavioural problems (compensatory behaviour hypothesis). It is therefore possible that unbiased, altruistic fathers are more likely to marry their partner if she has boys because boys happen to have more health or behavioural problems and fathers are forward-looking.

The possible interpretations of the divorce finding are similar. If fathers like living with sons more than with daughters, for example, they might be more likely to stay in a marginal marriage (rather than divorce) if they have boys than girls. Historically, in most divorce cases, fathers lost day-to-day access to children. If a couple has a net bias for sons (regardless of whether it is primarily due to a father's preferences), the combined utility loss that occurs with a divorce is larger for boy families compared to girl families. On the other hand, it is possible that unbiased parents in marginal marriages decide to stay together if they have boys because they realize that the presence of the father at home is more important for boys than for girls, that fathers are more

31. While other hypotheses are in theory possible, we focus on these throughout the rest of the paper. For example, daughters may be more vulnerable to physical abuse from their fathers than sons. Although we recognize this as a possible contributing factor, we do not believe that it is the main force at play. Moreover, while this hypothesis may explain the documented gender differences in marriage, divorce, and custody, it is unlikely to explain our evidence on fertility in the next section. Alternatively, some evolutionary biology theories may explain part of the documented gender differences in marriage and divorce. In particular, certain animal species appear to have adapted to regulate the sex of their offspring according to environmental conditions (Charnov *et al.*, 1982). Norberg (2004) argues that natural selection adjusts the sex ratio at birth based on whether partners are married or not prior to the child's conception. While this hypothesis can explain some of our findings, we do not believe that it is the most natural explanation for all of the evidence taken as a whole. For example, it is unlikely to explain the results on shotgun marriage and sonograms or the evidence on fertility and stated preferences in the next section.

efficient at taking care of boys, or that girls cost more to raise in time and money.[32] The custody outcome follows a similar logic.

3.2. *What each hypothesis implies for fertility and health shocks*

To make further progress in differentiating between the alternative hypotheses, in the following section, we estimate how gender composition affects fertility and whether fathers are more likely to stay with those children who are born with health problems. Consider first what each hypothesis implies for fertility. Couples with a gender bias for boys will be more likely to have another child if they do not yet have any boys because they keenly want a son. For example, if parents prefer boys, a couple with two girls should be more likely to have a third child than an otherwise identical couple with two boys. If couples prefer sons, one way to think about it is that the *effective* number of children is larger in an all-boy family. Since the marginal utility of an additional child is decreasing in the number of effective children, all-boy families have a lower marginal utility for an additional child.

On the other hand, in the case where parents are only concerned about the role model effect and they value additional children equally regardless of sex, there is no reason for fertility decisions to depend on the gender of their existing children. If parents take into account the effect of their children's gender on future divorce, the prediction of the role model hypothesis is even stronger: having girls should *decrease* the probability of having another child (see the Appendix for details). Alternatively, suppose parents are unbiased, parents do not care about the role model effect, and girls are more expensive to raise than boys. In this case, controlling for family size, couples who have only girls will be less likely to have another child compared to couples who have only boys. The reason is that, as long as children are normal goods, *ceteris paribus*, all-girl families are poorer than all-boy families. Due to a pure income effect, couples with girls will have fewer children. For the technology hypothesis, if the presence of a father in a married household makes it cheaper to raise boys compared to girls, couples with girls should also have fewer children due to an income effect. This prediction is strengthened if parents are forward-looking and take into account future divorce decisions.

To sum up: if parents are biased towards boys, we should see that the probability of having another child is higher for families that have all girls than for families that have all boys. On the other hand, if parental preferences are unbiased and the role model hypothesis is true, the technology hypothesis is true, or girls are more expensive than boys, we should see that the probability of having another child is lower or equal for families that have all girls compared to families that have all boys.

In order to separate gender bias from compensatory behaviour, the fertility results are not enough because the fertility implications of the compensatory behaviour hypothesis are similar to the implications of the gender bias hypothesis. We therefore also examine an additional test based on health shocks to differentiate the gender bias hypothesis from the compensatory behaviour hypothesis. If the compensatory behaviour hypothesis is true, we should see that fathers are more likely to stay with those children who are more difficult to raise because they are born with health problems. We should also see that the effect of health problems on the presence of the father at home is stronger for boys than for girls.

32. For the differential cost hypothesis to explain the findings, one needs to assume that in the married state, girls cost more, but in the divorced state, *for the father*, child support and time costs do not vary based on the sex of the child. Even if husbands and wives can make transfers to each other, some mothers may not be able to provide a big enough transfer to get a marginal father to remain in the marriage (or enter into marriage) if she has a girl. We stress that this is a *corner solution*, where husbands with daughters are more likely to be non-resident since marriage versus divorce (or, similarly, versus never marrying in the first place) is a relatively more expensive option.

So far, we have provided the intuition for each of the outcomes *in isolation*. However, it is unlikely that couples make choices on marital status, living arrangements, and fertility in isolation. It is more likely that they *simultaneously* determine marital status, living arrangements, and fertility. In the Appendix, we present a simple two-period model, where parents have transferable utility functions and are forward-looking. To keep things relatively simple, the Appendix focuses on divorce and fertility and the primary hypotheses, although other components could easily be added to the model. We show that even when marital status and fertility are simultaneously determined, the predictions outlined in this section continue to hold under fairly general conditions.

4. DIFFERENTIATING BETWEEN ALTERNATIVE HYPOTHESES

4.1. *Evidence from fertility*

Here, we present estimates of the effect of child gender on fertility using the sample of married women from the 1960–2000 Censuses used previously. Early work on fertility in the U.S. was done by Ben-Porath and Welch (1976) using 1970 data. Subsequent work has focused on intergenerational transmission of fertility preferences (Anderton *et al.*, 1987) and the demand for variety in offspring sex (Angrist and Evans, 1998). In this paper, we do not focus on the effect of gender mix. We simply note that a demand for variety in the sex mix of children can coexist with a demand for one gender over another. For example, parents can prefer a mix of one boy and one girl over two boys and at the same time also prefer two boys over two girls.

In dealing with the question of fertility, we face two problems. First, most couples plan on having more than one child, regardless of the sex of their first child. For this reason, the *measured* fertility effect of having a second child should be small based on the sex of the first child. In the extreme case where all couples want at least two children regardless of the sex of the first child, the sex of the first child should have no effect on the probability of having a second child. In the more realistic case where most couples' *ex ante* desired number of children is two or more, one should expect to find a small effect for the probability of having a second child. Using data from the 1985–1995 CPS Fertility Supplements, we calculate that 85% of mothers more than 40 have two or more children. In earlier years, we expect this fraction might be even higher.

It is important to recognize, however, that the sex of the *first* child can still affect the likelihood of having three or more children, even if most families have at least two children. To see this, compare two families that plan on having at least two children, where one family has a first-born girl, while the other has a first-born boy. After the birth of the second child, assuming gender-birth order does not matter, the only difference is that one family has a 50% chance of having two girls, while the other has a 50% chance of having two boys.[33] Both families have a 50% chance of having one child of each gender. Therefore, if the family with two girls is more likely to have an additional child compared to the family with two boys, this effect will load onto the sex of the first child, which is arguably random.

A second problem is caused by data limitations. In the Census, only *current* marital status is reported consistently in all the years. Although our sample for the fertility results includes only currently married women, a fraction of these women have been previously divorced. This is a problem because women whose first child is a girl are more likely to divorce (as documented in Table 1), and divorced women have lower fertility. For this reason, our estimates of the

33. Empirically, it is appears to be true that gender-birth order does not affect fertility decisions. Families with a first-born boy and a second-born girl have the same probability of having a third child as families with a first-born girl and a second-born boy.

TABLE 4

First-child gender and fertility; U.S. Census data

	Total number of children (1)	Breakdown by number of children: Two or more children (2)	Three or more children (3)	Four or more children (4)	Five or more children (5)
		All couples			
First child a girl	0·0024	−0·0012	0·0014**	0·0013**	0·0006**
	(0·0015)	(0·0008)	(0·0007)	(0·0004)	(0·0002)
Controls?	Yes	Yes	Yes	Yes	Yes
First boy baseline	2·01	0·653	0·246	0·075	0·022
Percent effect	0·1	−0·2	0·6	1·8	2·8
Observations	3,741,573	3,741,573	3,741,573	3,741,573	3,741,573
		Couples in first marriage			
First child a girl	0·0070**	0·0008	0·0031**	0·0019**	0·0008**
	(0·0018)	(0·0008)	(0·0008)	(0·0005)	(0·0003)
Controls?	Yes	Yes	Yes	Yes	Yes
First boy baseline	2·03	0·657	0·254	0·082	0·025
Percent effect	0·3	0·1	1·2	2·4	3·2
Observations	1,507,291	1,507,291	1,507,291	1,507,291	1,507,291

Notes: S.E. are given in parentheses. The sample in the top panel includes all currently married mothers between the ages of 18 and 40 with children living at home between 0 and 12 years old in the 1960–2000 U.S. Censuses. The sample in the bottom panel is restricted to 1960–1980 data because information on whether an individual has been married more than once is not available in the 1990 and 2000 Censuses. The sample in the bottom panel is restricted to all currently married mothers in their first marriage. The dependent variable in column (2) is a dummy equal to 1 if the family has two or more children; the dependent variables in columns (3)–(5) are similarly defined. Control variables include a cubic in age, and dummies for race (White people, Black people, Asian, other), education (less than high school, high school, college), region of residence (nine regions), and cohort of birth (10-year birth cohorts). The first boy baseline is calculated as the average predicted probability of the outcome variable of interest for first-born boy families using the estimated coefficients on the control variables. The percent effect is the increase in the probability of living without a father for a first-born girl family to a first-born boy family. **Statistically significant from 0 at the 5% level; *statistically significant from 0 at the 10% level.

relationship between child gender and fertility are biased towards finding a *negative* relationship between the first-girl indicator and fertility.

With these two caveats in mind, we turn to Table 4. The top panel is based on data for all couples. The first column suggests that families where the first child is a girl end up having more children than families where the first child is a boy, although the difference is not significant. Column (2) indicates that the probability of having a second child is actually negative, although not statistically significant, if the first child is a girl. The estimates in column (3) reveal the probability of having three or more children is 0·14 percentage points higher when the first child is a girl. In other words, first-born girl families are 0·6% more likely to have three or more children compared to first-born boy families. Significant positive effects are also found for the probability of four or more and five or more children when the first-born child is a girl.

We suspect that the statistically insignificant effect in column (1) may be the result of the negative bias resulting from divorce described above. If we could observe the entire marital history of respondents, we could account for this bias. Although this is not possible for the entire sample, the Censuses for 1980 and earlier years do report if a woman has been married more than once. Our preferred estimates are therefore based on the 1960–1980 Censuses and on the sample of women in their first marriage. Using this sample, the first-born girl coefficient for total number of children increases almost three-fold to 0·007 and is now statistically significant. Interestingly, in column (2), the negative coefficient for having two or more children flips sign once we restrict the sample to couples in their first marriage. Similarly, the coefficient estimates for having three or more, four or more, and five or more children are also larger when we focus on couples in their first marriage. The differences between the top and the bottom panels are consistent with the notion that mothers with first-born girls are more likely to divorce and that mothers who experience a divorce spell have fewer children.[34]

To help in interpreting the magnitude of the fertility effect, we calculate how many additional children were born because of first-born girls. Given that approximately 800,000 first-born daughters were born each year on average from 1960 to 2000, we estimate that first-born daughters caused approximately 5500 more births each year, or 220,000 additional births over the past 40 years. One interpretation of the fertility response to a first-born girl documented in Table 4 is that parents have preferences over the sex mix of their children. This interpretation suggests that a large number of girls would likely not have been born if parents could perfectly choose the gender of their children with zero cost.

Notice that we focus on the gender of the first child to preserve the causal interpretation of our estimates. Whether the first child is a boy or a girl can arguably be viewed as random. However, whether the first two or three children are boys versus girls is no longer random since fertility decisions and divorce decisions are both endogenous. While not shown, the effects of the gender mix of the first two and three children also show some interesting patterns. For example, in families with two girls born first, fertility is 0·017 (S.E. = 0·003) children higher compared to families with two boys born first. By way of comparison, in families where the first two children are a boy and then a girl or a girl and then a boy, fertility is −0·060 (S.E. = 0·003) and −0·057 (S.E. = 0·003) children higher, respectively. This effect for mixed gender is often referred to as the demand for variety (Angrist and Evans, 1998). The estimates imply that the demand for variety is about three to four times as large as the demand for sons. Examining the gender mix of the first three children, families with three girls born first are the most likely type of family to have another child, and statistically so.

Table A1 also reveals some interesting patterns when the data are broken down into specific subgroups. Since our preferred estimates are based on the 1960–1980 data set where we can condition on whether mothers are in their first marriage, we cannot say much about trends by Census year or decade of birth. However, the results by race and education are interesting. The effect on fertility of having a first-born daughter is almost seven times larger for Asians compared to White people, with the effect for Black people in the middle. Table A1 also shows that the fertility effect of a first-born daughter declines monotonically with education, with college graduates showing essentially no fertility effect.

Overall, it appears clear that families with a first-born girl have higher fertility than families with a first-born boy. In light of the discussion in Section 3, we can start to draw some conclusions. It would be difficult to explain our findings on fertility if there was no parental gender preference and only the role model or technological hypothesis was at work. If parents are unbiased, and the role model or technological hypothesis is true, we should see that the probability of having additional children is *lower or equal* for families that have first-born girls compared

34. One limitation of the Census data used in Table 4 is that we observe family size at the time of the Census and cannot distinguish between completed and uncompleted fertility. Because some mothers have not finished bearing children, the estimates in Table 4 are a weighted average of the effect for mothers with completed fertility and the effect for mothers with uncompleted fertility. They are therefore likely to *underestimate* the effect of child gender composition on lifetime fertility. To address this problem, we have repeated the analysis using a unique panel of California mothers obtained by longitudinally linking birth certificates for the years 1989–2001. Our estimates for the effect of gender composition on the probability of having another child, based on the longitudinal California sample, are qualitatively similar, although larger than the estimates in Table 4 (see Dahl and Moretti, 2004).

to families that have first-born boys. It would still be possible to explain the fertility findings *in isolation* under a differential cost hypothesis. If parents are unbiased, but raising boys is *more* expensive than raising girls, families with boys are "poorer" than families with girls, and therefore have fewer children. However, a scenario where parents are unbiased and raising boys is relatively more expensive is at odds with the marriage, divorce, and custody results, which require the cost of raising girls to be higher than the cost of raising boys if differential costs are the only force at play.

4.2. *Evidence from health shocks*

In this section, we test two predictions of the compensatory behaviour hypothesis. In particular, the hypothesis suggests that (i) fathers are more likely to stay with those children who are more difficult to raise because they are born with health problems and (ii) the probability that a father stays with a child born with a health problem is higher if the child is a boy. We use data from the National Health Interview Survey for the years 2003–2005, which have very detailed information on many congenital health conditions that affect children.

Table 5 shows the results. The first row in Table 5 indicates that boys who have any health condition are significantly more likely to live without a father (column (1)) and that the effect for girls is not statistically different from the effect for boys (column (3)). Both these findings appear inconsistent with the compensatory behaviour hypothesis. Column (2) confirms that girls are more likely to live without a father. This is consistent with our earlier findings in Table 1, although we note that in this data set, we do not know the birth order of the children.

The remainder of the table repeats the same test using specific health conditions, such as mental retardation, chronic heart condition, partial blindness, or walking impairment, that require a wheelchair or a brace. In general, these estimates confirm that if a child is born with a serious health condition, the father is less likely (in some cases, equally likely) to be present. Notably, this effect is the same irrespective of whether the child is a boy or a girl. Of course, the interpretation of these findings is complicated by the possible presence of unobserved heterogeneity. However, at least some of the conditions in Table 5 are congenital and therefore unlikely to be very sensitive to parental behaviour (although we recognize that there could be unobservables that drive both the child's health status at birth and family structure). Furthermore, all the models control for economic and demographic characteristics, such as parental income, race, region, and age. More importantly, because child gender is arguably random, the coefficients in column (3) are unlikely to suffer from spurious correlation. The findings in Table 5 appear inconsistent with the predictions of the compensatory behaviour hypothesis.

In conclusion, we emphasize that it is difficult to interpret our results in Section 2 in isolation. Taken separately, each piece of evidence has multiple interpretations. But together with our fertility and health shock results, the findings in Section 2 document a pattern that is consistent with the presence of a gender bias for sons among U.S. parents.

4.3. *Evidence from stated preferences*

Finally, we measure gender bias using data from a survey that directly asks parents whether they would prefer a son or a daughter. If parents, when directly asked, say that they do not care about the gender of their children, this would cast doubt on the interpretation of our findings as evidence of gender bias. Moreover, stated preferences are useful in identifying whether the preference for sons is due to the father, the mother, or both parents. In a transferable utility framework, revealed preferences cannot distinguish whether parental sex bias for boys comes from fathers or mothers.

TABLE 5

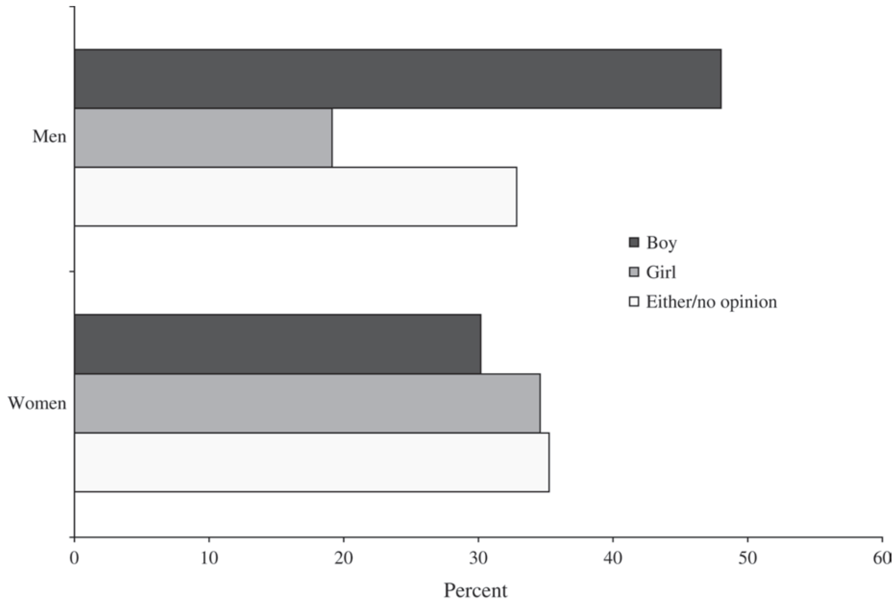
Child health, child gender, and the probability of living without a father; National Health Interview Survey

	Coefficient on condition (1)	Coefficient on girl (2)	Coefficient on condition × girl (3)
Any condition	0·053**	0·014**	−0·009
	(0·006)	(0·006)	(0·009)
Selected conditions			
Developmental retardation	0·046**	0·009**	0·0001
	(0·017)	(0·004)	(0·0297)
Attention deficit disorder	0·100**	0·014**	−0·004
	(0·014)	(0·005)	(0·025)
Down's syndrome	−0·051	0·008**	−0·056
	(0·081)	(0·004)	(0·124)
Chronic heart condition	0·082**	0·008**	−0·079
	(0·035)	(0·004)	(0·048)
Trouble seeing	0·045*	0·007*	0·023
	(0·023)	(0·004)	(0·034)
Impairment that requires	0·062**	0·008**	−0·024
wheelchair or brace	(0·025)	(0·004)	(0·037)
Seizure	0·045	0·008**	−0·095
	(0·041)	(0·004)	(0·061)
Diabetes	0·007	0·007*	0·216
	(0·085)	(0·004)	(0·133)
Asthma	0·081**	0·013**	−0·011
	(0·009)	(0·005)	(0·014)
Respiratory allergy	0·057**	0·011**	−0·020
	(0·010)	(0·005)	(0·015)

Notes: S.E. are given in parentheses. Data are from the 2003–2005 National Health Interview Survey. Each row is a separate model. The dependent variable is a dummy equal to 1 if the child lives without the father, either because the parents are divorced and the children live with the mother or because the mother has never been married. In each household, data were collected on one random child. Column (1) reports the coefficient on a dummy for whether the relevant child has the specified condition. Column (2) reports the coefficient on a dummy for whether the relevant child is a girl. Column (3) reports the coefficient on the interaction between the condition dummy and the girl dummy. The entry in the first row is a dummy equal to 1 if the child has any of the following conditions: mental retardation, any other developmental delay, attention deficit hyperactivity disorder (only defined for children 2 and older), asthma, Down's syndrome, cerebral palsy, muscular dystrophy, cystic fibrosis, sickle cell anaemia, autism, diabetes, chronic arthritis, chronic congenital heart disease, other chronic heart conditions, hay fever, any kind of respiratory allergy, any kind of food or digestive allergy, eczema or any kind of skin allergy, frequent or repeated diarrhea or colitis, anaemia, three or more ear infections, seizures, trouble seeing, or impairment that requires wheelchair or braces. Entries in the remaining rows show coefficients for selected conditions. For space considerations, we do not present results for all available conditions. Controls include mother's race, income, region of residence, and Hispanic origin. Sample size in row (1) is 22,047. Sample sizes in other rows are similar. **Statistically significant from 0 at the 5% level; *statistically significant from 0 at the 10% level.

Take, for example the finding that having girls increases the probability of having an extra child relative to having boys. If this result is indeed explained by gender bias on the part of one of the parents, it could be coming from either the father or the mother. In the case where an extra child generates enough surplus for the father, he can transfer utility to the mother to make her at least as well off with the extra child. Alternatively, the mother could make transfers to the father.

Starting in 1941 and continuing to the present, Americans have stated in Gallup Polls that they would prefer to have sons over daughters, with very little change in these preferences over time (Dahl and Moretti, 2004). More recent Gallup Poll surveys help determine whether this gender bias is due to a father's versus a mother's tastes. The surveys asked 2129 adults in 2000 and 2003: "Suppose you could only have one child. Would you prefer that it be a boy or a girl?" Figure 2 shows that women seem to have only a slight preference for a daughter: 35% say "girl"



Question: *Suppose you could only have one child. Would you prefer that it be a boy or a girl?*

Note: Surveys were conducted on nationally representative random samples of adults (18 years or older). There were 1067 female respondents and 962 male respondents across the two survey years

FIGURE 2

Preference for boys versus girls, for men and women; 2000 and 2003 Gallup Poll data

and 30% say "boy". In sharp contrast, men express an overwhelming preference for a son: 19% say "girl", and 48% say "boy". The fact that men prefer sons by a 2·5 to one margin suggests that the marriage, shotgun marriage, divorce, and fertility findings may largely be due to fathers' gender biases.

Finally, Table 6 investigates how the preference for a son versus a daughter depends on a respondent's characteristics. For individuals who stated a preference, we ran a probit regression where the dependent variable was equal to 1 if the respondent stated that they preferred a boy. Men are approximately 23 percentage points more likely to indicate a preference for a son, controlling for covariates. Black people are about 12 percentage points more likely to prefer a boy compared to White people. Liberal respondents and those who attend church are also more likely to report a preference for a daughter. Education and income have no statistically significant impact. As individuals get older, they prefer boys less, perhaps indicating that life's experiences (including raising girls) help temper any bias. While these results are interesting, in isolation they are not definitive, especially since we do not have any information on the characteristics of the respondent's partner.

5. CONCLUSIONS

This paper shows that child gender affects the marital status and fertility decisions of American parents and that there are serious negative effects for children in households with first-born girls. Overall, we find that fathers of boys are substantially more likely to be living with their children

TABLE 6

Stated preferences for a boy or a girl; Gallup Poll data

Variable	Probit (1 = Prefer Boy, 0 = Prefer Girl) Coefficient (1)	S.E. (2)	Marginal effect (3)
Male	0·6128	0·0752	0·2335
Age	−0·0335	0·0129	−0·0130
Age squared	0·0003	0·0001	0·0001
Race			
White	—	—	—
Black	0·3200	0·1315	0·1188
Other	0·0029	0·1583	0·0011
Race unknown	−0·0811	0·3836	−0·0317
Hispanic	0·0854	0·1394	0·0331
Education			
High school dropout	0·0142	0·1108	0·0055
High school	—	—	—
Some college or more	0·0329	0·1180	0·0127
Region			
East	—	—	—
Midwest	−0·0570	0·1052	−0·0222
South	−0·1111	0·0989	−0·0432
West	0·0656	0·1094	0·0253
Urban/rural status			
Urban	—	—	—
Suburban	0·0449	0·0848	0·0174
Rural	0·1150	0·1050	0·0442
Income			
Income <30K	—	—	—
$30\text{K} \leq$ income $< 50\text{K}$	−0·0973	0·1850	−0·0381
$50\text{K} \leq$ income $< 75\text{K}$	0·1193	0·1011	0·0459
Income ≥ 75K	0·1822	0·1165	0·0695
Refused	0·0550	0·1141	0·0212
Marital status			
Divorced or separated	0·1899	0·1306	0·0721
Married	0·1025	0·1022	0·0397
Widowed	0·0111	0·1785	0·0043
Never married	—	—	—
Attend church often	−0·2098	0·0767	−0·0810
Ideology			
Conservative	—	—	—
Moderate	−0·0464	0·0829	−0·0180
Liberal	−0·3693	0·1010	−0·1452
Refused	−0·0095	0·2859	−0·0037
Constant	0·5490	0·3944	
Log likelihood	−837·12		
Observations	1325		

Notes: Data are from the 2–4 December, 2000, and 18–20 July, 2003, Gallup Poll surveys. The question asked of respondents was "Suppose you could only have one child. Would you prefer that it be a boy or a girl?" The survey data are based on telephone interviews of a national random sample of adults aged 18 years and older. This table includes only individuals with a stated preference for a boy or a girl; multinomial logit results including "no preference" as a third option are similar.

compared to fathers of girls. A first-born daughter is 3·1% less likely to live with her father compared to a first-born son. This gender differential impacts a significant number of children. We estimate that in any given year, roughly 52,000 first-born daughters younger than 12 years and their siblings would have had a resident father if they had been boys. Three factors are important in explaining why fathers are more likely to live with their sons than with their daughters. First, mothers of girls are substantially more likely to have never been married than those of boys. Surprisingly, we also find evidence that when the gender of a child is known *in utero*, it affects shotgun marriages. Among women who have taken an ultrasound test during pregnancy, and therefore probabilistically know the gender of their child in advance, we find that mothers who have a girl are less likely to be married *at delivery* than those who have a boy. Second, parents of girls are significantly more likely to be divorced than those of boys. Third, after a divorce, fathers are much more likely to have custody or visitation rights for their sons than for their daughters.

Taken individually, each piece of evidence on family living arrangements is consistent with, but does not necessarily imply, parental gender bias. We turn to evidence on fertility to help us differentiate between possible explanations. We find that, consistent with parental preference for sons, fertility is significantly higher for families where the first-born child is a girl. Our estimate indicates that first-born daughters caused approximately 5500 more births each year, or a total of 220,000 more births to occur during our sample period 1960–2000. Survey data confirm parental preference for boys and indicate that such preference is much stronger among men. Of course, the existence of gender bias does not rule out the possibility that other concerns also play a part in parents' decisions, as many explanations may be simultaneously responsible. But the weight of the evidence does not allow us to reject the hypothesis that parents, most likely fathers, have a stronger demand for sons than for daughters. Taken together, our findings suggest that the age-old favouring of boys is not confined to the past. It is subtle and manifests itself in different ways over time in the U.S. and remains significant today.

Regardless of how one interprets our findings on family structure and fertility, however, the serious negative economic and educational effects for children whose first-born sibling is a girl are interesting in their own right. Because of the effects of gender on family structure, first-born girls and their siblings live in families where income is lower, the poverty rate is higher, welfare participation is higher, home ownership is lower, and child support payments following a divorce are lower. In addition, second-born children in first-born girl families have lower educational attainment compared to second-born children in first-born boy families. While our results indicate that some of the negative consequences of a first-born daughter affect younger siblings of both genders, girls are overall more likely to be exposed to these negative effects.

APPENDIX

This appendix develops a simple two-period, forward-looking model for divorce and fertility decisions. While for simplicity, we do not develop predictions for marriage and custody outcomes, we note that those predictions are similar to the one for divorce. In the model, parents have transferable utility functions and experience stochastic shocks that influence the quality of a marriage. This stylized model aids in understanding the implications that (i) gender bias, (ii) role model, and (iii) differential costs have for divorce and fertility stopping rules. While in Section 4, we provide the intuition for a static context where parents choose marital status and fertility separately, here, we consider a two-period context where parents choose marital status and fertility simultaneously, and are forward-looking.

Fortunately, the intuition developed in Section 4 carries through in this context. In particular, the model illustrates that a gender bias for sons, a role model effect for sons, and a higher cost of raising girls all have the same prediction for divorce: parents are more likely to divorce if they have a daughter versus a son. However, the models have different testable implications for fertility under fairly general conditions. With only gender bias, parents will be more likely to have an additional child if their first child was a girl. In contrast, with a pure role model story or if the cost raising girls is higher, the opposite is true.

A.1. *Utility as a function of the sex of children*

Our model for divorce and fertility decisions is grounded in the work of Becker (1973, 1974), which assumes that marriage markets are cleared by transfers between spouses. Specifically, we assume that both the husband and the wife have transferable utility functions (*i.e.* quasi-linear) of the general form $h(B_t, G_t, C_t) + X_t$ where the subscript t denotes time, B_t and G_t are the number of boys and girls in the family, C_t is non-transferable consumption, and X_t is consumption that is transferable. Transferable utility functions have been widely used in bargaining models in addition to the marriage context. The advantage of transferable utility is that when considering divorce, marriage, and fertility decisions, one only needs to compare the sum of the husband's and wife's utility. There is no need to consider the allocation of consumption goods in the marriage or determine which spouse has more power in the marriage. Because of this assumption, we do not refer separately to the husband's and wife's utility functions in what follows.

A marriage occurs when there is utility created from the union, although the future value of a marriage is not known with certainty. We model uncertainty in a marriage as follows. In each period, there is a shock to the marriage which is mean zero and independently and identically distributed. We assume that the shock is normally distributed: $\varepsilon_t \sim N(0,\sigma^2)$. The shock is marriage specific, so if the couple separates, the shock is not present.[35] We assume that the shock is independent of the gender composition and size of the family. To facilitate discussion, we separate out the value of this shock when writing the utility function.

The couple's aggregated utility in period t can therefore be written as:

$$U(B_t, G_t, C_t, i) + X_t + I[i = M] \times \varepsilon_t \qquad i = M, D, \tag{A.1}$$

where I is an indicator equal to 1 if the couple chooses to stay married after the realization of the shock and 0 otherwise. Throughout, we use the letters M and D (usually as superscripts) to refer to the married and divorced states, respectively. The combined period budget constraint is:

$$pB_t + qG_t + rC_t + X_t = Y_t, \tag{A.2}$$

where p, q, and r are the prices of boys, girls, and non-transferable consumption, and Y_t is combined income for the married couple. We do not allow prices or income to differ in the married and divorced states. While not very realistic, this simplification does not change the main insights of the model. The numeraire good is transferable utility. For simplicity, we assume that the budget constraint holds with equality in each period with no borrowing or saving.

To make definitions of the gender bias, role model, and differential costs hypotheses more concrete, let children enter as a single, additive argument in the subutility function, and let children be the only arguments whose value interacts with whether the couple is married or divorced, so that:

$$U(B_t, G_t, C_t, i) = U(\tilde{K}_t^i, C_t) \qquad i = M, D, \tag{A.3}$$

where $\tilde{K}_t^i$ represents the effective number of children. If parents value children equally, the effective number of children is an equally weighted sum of the number of boys and girls. However, parents may value boys or girls more, implying unequal weights. They may also have a different value for the effective number of children when married versus divorced. To allow for these possibilities, we write the effective number of children in the married and divorced states as $\tilde{K}_t^i = \alpha^i B_t + \beta^i G_t, i = M, D$, where α^i and β^i are positive scalars that weight how much parents value girls compared to boys in the married and divorced states. We assume throughout that $\alpha^M > \alpha^D$ and $\beta^M > \beta^D$, so that children provide more utility in the married state, *ceteris paribus*. The rationale for this assumption is that one of the parents (most likely the husband) has limited access to the children—and therefore "consumes" less of them—in the divorced state.

In this setup, boys and girls are perfect substitutes. It follows that there is a quality-quantity tradeoff between the gender and the number of children. We assume that utility increases at a decreasing rate as the number of effective children increases. That is:

$$\partial U/\partial \tilde{K}_t^i > 0, \quad \partial^2 U/\partial (\tilde{K}_t^i)^2 < 0 \qquad i = M, D. \tag{A.4}$$

It is now a simple matter to describe what we mean by a gender bias, a role model, and a cost differential. A *gender bias* for boys occurs when $\alpha^M > \beta^M$, meaning that a girl is valued at some fraction of a boy in the married state. This is because either the husband or the wife derives more utility from living with boys than with girls. Holding the number of children fixed, the effective number of children in the married state increases with the number of boys. For simplicity, we also assume that in the divorced state, boys and girls are valued equally: $\alpha^D = \beta^D$. The assumption that $\alpha^D = \beta^D$ simplifies the model but can easily be relaxed. (One set of assumptions that would imply $\alpha^D = \beta^D$ is that the husband is biased, the wife is unbiased, after a divorce the husband loses access to the children, and that children provide utility only to the parent they live with. In this setup, in the divorced state, the father gets zero utility from his non-resident children and the wife gets equal utility from her children.)

35. It is easiest to think of the shock as normalized so that it is measured on the same scale as the linear component of the utility functions. In what follows the normal assumption is not crucial. Other distributions such as a uniform distribution or a logistic distribution yield the same general implications but do not have as nice an interpretation. A shock that occurs in the divorced state for each spouse could also be added to the model, but this complication does not change any of the key predictions.

If parents are unbiased, a *role model* for boys exists if $\alpha^D < \beta^D$ and $\alpha^M = \beta^M$, which implies that boys provide less utility compared to girls in the divorced state. To better understand why α^D might be less than β^D, consider the simple case where children are assigned to mothers irrespective of their gender after a divorce. The reason boys provide less utility to parents in the divorced state is that altruistic parents take into account the happiness of their children. Under the role model hypothesis, the absence of a father has a larger negative impact on sons, so that boys suffer more after a divorce. The utility functions of altruistic parents take into account this differential utility loss for boys versus girls.[36]

Finally, *differential costs* favouring boys occurs if $p<q$ so that boys are cheaper than girls. If a father divorces and leaves the family, we assume that boys and girls cost the same for the father since courts are unlikely to order very different child support and alimony payments for fathers with boys and fathers with girls. Any asymmetry in costs in the divorced state is borne by the mother, who retains custody of the children.

A.2. *Two-period model*

Using the utility functions described above, we develop a simple two-period model for divorce and fertility decisions. The model is forward-looking, with couples making decisions in the first period before knowing the value of the marriage-specific shock in the second period.

To make things simple, in period 1, we start with couples who are married and already have one child. The couple makes two choices in the first period: whether to divorce and whether to have an additional child. The decisions are made sequentially, with the couple first deciding whether to divorce after realizing the value of the first-period marriage-specific shock. If the couple chooses to stay together, they then decide whether to have an additional child, which will be born in the second period. In the second period, the couple observes two new pieces of information: the sex of the additional child (if they chose to have an additional child in the first period) and the value of the second-period marriage-specific shock. Given this information, the couple then decides whether to remain married or divorce as before. The second period is the terminal period in the model. The divorce decision in period 1 takes into account the number and sex composition of children, as well as the expected future benefit of remaining married. Likewise, the additional child decision in period 1 takes into account the probability of divorce in the future. The expected benefit of remaining married today for a given family type is a function of the optimal choice of whether to have a child in the next period, which in turn is based on the expected benefits of another child (which might be a boy or a girl), given that you have a probability of divorce in the next period (as a function of the sex composition and the realization of the shock). In other words, the expected benefit of staying married today for tomorrow's utility incorporates the possibility of additional children and future divorce.

To figure out divorce and fertility decisions, we need to know the expected utility of various family compositions in the future. To economize on notation, it is helpful to use a shorthand label for various family types. Let B stand for boy and G stand for girl. In our two-period model, there are six possible family compositions: *B, G,* BB, BG, GB, and GG. We refer to utility and other relevant variables that are a function of family type using these abbreviations as subscripts. We also abbreviate the utility function with a superscript to indicate whether the couple is currently married or divorced. For example, utility in the married state for a family with one boy and one girl is written as $U^M_{\text{BG}} = U(\alpha^M + \beta^M, C_t) + X_t + \varepsilon_t$. For clarity, we continue to separately write out the shock, ε, and do not include it in the shorthand definition. From this point on, we omit the time subscript on variables when it is clear which time period is the relevant one.

We now define some useful expressions that are a function of utility in the married and divorced states. The probability that the marriage will survive in period 2 is denoted by:

$$\pi_c = 1 - F(U_c^D - U_c^M), \tag{A.5}$$

where the subscript c denotes family composition (*i.e. B, G,* BB, BG, GB, or GG) and $F(\cdot)$ is the cumulative distribution function of the random variable ε, which by our previous assumption is normally distributed. The expected value of the shock conditional on staying in the marriage (*i.e.* the truncated mean of the shock) is given by:

$$\lambda_c^M = E(\varepsilon_{t+1} | \varepsilon_{t+1} > U_c^D - U_c^M). \tag{A.6}$$

Similarly, define λ_c^D as the expected value of the marriage-specific shock conditional on divorce.

For those who remain married in the first period, we can now write the expected utility (evaluated at period t) for a family with composition c in period t+1, as:

$$\theta_c = \pi_c (U_c^M + \lambda_c^M) + (1 - \pi_c) U_c^D = \theta(\tilde{K}^M, \tilde{K}^D). \tag{A.7}$$

36. We could explicitly model the utility of parents as a function of the utility of children and allow the utility loss experienced by boys in the divorced state to be larger than the utility loss experienced by girls. This would generate the same definition of a role model for boys but would complicate the notation. Of course a role model for girls could exist as well; for example, if fathers were assigned custody of children after a divorce, it is easy to see why α^D might be greater than β^D.

Equation (A.7) is a function of U_c^M and U_c^D, which are in turn functions of the effective number of children in the married and divorced states, $\tilde{K}^M$ and $\tilde{K}^D$. Equation (A.7) plays a pivotal role in determining forward-looking divorce and fertility decisions.

A.3. *The effect of sex composition on divorce*

We first examine divorce predictions based on child gender in a forward-looking context and then turn to the fertility predictions. It is easiest to discuss predictions by comparing a family whose first child is a girl to a family whose first child is a boy. In the discussion, we assume without loss of generality a gender bias for boys, a role model effect for boys, and that boys cost less than girls. (One could easily assume the opposite effects, in which case the predictions would be reversed.) To focus on the effects of each hypothesis, we examine each case separately. For example, when considering a gender bias for boys, we assume no role model effect and equal costs. While more than one effect may be present, the point is to show which effect dominates.

For a family with one boy in the first period, the probability of divorce in period 1 is:

$$\Pr\left(\varepsilon_t < (U_B^D - U_B^M) + \gamma\,(U_B^D - \max\left\{\tfrac{1}{2}\theta_{\mathrm{BB}} + \tfrac{1}{2}\theta_{\mathrm{BG}}, \theta_B\right\})\right), \tag{A.8}$$

where γ is a discount factor. A similar expression for a family with one girl in the first period is:

$$\Pr\left(\varepsilon_t < (U_G^D - U_G^M) + \gamma\,(U_G^D - \max\left\{\tfrac{1}{2}\theta_{\mathrm{GB}} + \tfrac{1}{2}\theta_{\mathrm{GG}}, \theta_G\right\})\right). \tag{A.9}$$

These equations capture the fact that individuals make divorce decisions based on current comparisons of utility as well as the forward-looking option value of remaining married. Comparisons of equations (A.8) and (A.9) reveal whether families with boys or families with girls are more likely to divorce. All three hypotheses (gender bias, role model, and differential costs) have the same prediction: in the first period, families with a girl are more likely to divorce compared to families with a boy. The intuition is given in Section 2.

The proofs for the gender bias and role model results rely on utility being an increasing, concave function in the effective number of children, properties of the normal distribution, and revealed preference. Detailed proofs are available on request. The differential cost result is proven by first remembering that we assume if a father divorces and leaves the family, boys and girls cost the same for the father since child support payments and other monetary costs are likely to be the same. Any asymmetry in costs in the divorced state is borne by the mother, who retains custody of the children. In a transferable utility setting, when there is an interior solution, there is no effect of differential cost on divorce. However, if the couple reaches a situation where the mother cannot transfer enough utility to her husband, we end up with a corner solution where the father opts out of the marriage.

A.4. *The effect of sex composition on fertility*

We now consider the fertility predictions of gender bias, role model, and differential costs. As before, we focus the discussion by comparing a family which has one girl to a family which has one boy and assume without loss of generality a gender bias for boys, a role model effect for boys, and that boys cost less than girls. The fertility decision depends on the probability that the couple will have a girl versus a boy, the probability that the couple will remain married as a function of sex composition, and the expected value of the shock if they remain married.

After choosing whether to divorce in period 1, couples who choose to stay together make their fertility decision. For a family with one boy in the first period, the couple will choose to have another child if:

$$\tfrac{1}{2}\theta_{\mathrm{BB}} + \tfrac{1}{2}\theta_{\mathrm{BG}} - \theta_B > 0, \tag{A.10}$$

and similarly, a family with one girl will have another child if:

$$\tfrac{1}{2}\theta_{\mathrm{GB}} + \tfrac{1}{2}\theta_{\mathrm{GG}} - \theta_G > 0. \tag{A.11}$$

Comparing the L.H.S. of these inequalities reveals whether boy or girl families have higher fertility under the three different hypotheses.

Fertility predictions are slightly more involved than the divorce predictions. Without saying something about the curvature of the utility functions, it is hard to compare the expected value of an additional child for boy versus girl families. It follows immediately that if θ, as described in equation (A.7), is an increasing, concave function of both $\tilde{K}^M$ and $\tilde{K}^D$, the following is true: (i) gender bias predicts higher fertility for girl families and (ii) role model predicts higher fertility for boy families. If θ is an increasing, convex function of $\tilde{K}^M$ and $\tilde{K}^D$, the opposite is true.

It is easy to show that θ is an increasing, concave function of both $\tilde{K}^M$ and $\tilde{K}^D$ if:

$$\lambda_c^i (U_c^i)' \leq (U_c^i)''/(U_c^i)' \qquad i = M, D, \tag{A.12}$$

TABLE A1

Estimates based on the sex of the first child for specific subgroups; U.S. Census data

	Non-resident father			Mother never married			Divorce			Mother custody			Fertility (No. children)		
	Coefficient (1)	Baseline (2)	% (3)	Coefficient (4)	Baseline (5)	% (6)	Coefficient (7)	Baseline (8)	% (9)	Coefficient (10)	Baseline (11)	% (12)	Coefficient (13)	Baseline (14)	% (15)
Model 1: By Census year															
1960	0·0019			0·0004			0·0009			0·0136**			0·0066		
	(0·0011)	0·054	3·5	(0·0004)	0·007	5·2	(0·0012)	0·052	1·8	(0·0063)	0·910	1·5	(0·0067)	2·3	0·3
1970	0·0034**			0·0005			0·0019**			0·0142**			0·0048		
	(0·0006)	0·102	3·3	(0·0003)	0·026	2·1	(0·0006)	0·084	2·3	(0·0019)	0·923	1·5	(0·0035)	2·2	0·2
1980	0·0060**			0·0005			0·0017**			0·0304**			0·0083**		
	(0·0007)	0·174	3·5	(0·0004)	0·057	0·9	(0·0007)	0·142	1·2	(0·0017)	0·873	3·5	(0·0020)	1·9	0·4
1990	0·0077**			0·0019**			0·0021			0·0323**					
	(0·0013)	0·187	4·1	(0·0006)	0·069	2·8	(0·014)	0·155	1·3	(0·0035)	0·821	3·9	—	—	—
2000	0·0047**			0·0012			0·0012			0·0210**					
	(0·0024)	0·241	1·9	(0·0019)	0·134	0·9	(0·0023)	0·157	0·5	(0·0062)	0·786	2·7	—	—	—
Model 2: By decade of birth															
1930's	0·0031*			0·0005			0·0019			0·0164**			0·0110		
	(0·0018)	0·057	5·5	(0·0007)	0·008	5·7	(0·0017)	0·053	3·7	(0·0080)	0·930	1·8	(0·0089)	2·2	0·5
1940's	0·0014			0·0004			0·0008			0·0038*			0·0046		
	(0·0009)	0·107	1·3	(0·0005)	0·027	1·5	(0·0008)	0·087	0·9	(0·0022)	0·949	0·4	(0·0042)	2·0	0·2
1950's	0·0051**			0·0011*			0·0023**			0·0153**			0·0056**		
	(0·0010)	0·193	2·6	(0·0007)	0·072	1·5	(0·0010)	0·142	1·6	(0·0020)	0·913	1·7	(0·0026)	1·7	0·3
1960's	0·0092**			0·0037**			0·0035*			0·0264**					
	(0·0020)	0·236	3·9	(0·0012)	0·111	3·3	(0·0021)	0·160	2·2	(0·0044)	0·878	3·0	—	—	—
1970's	0·0036			−0·0000			0·0039			0·0104					
	(0·0042)	0·323	1·1	(0·0036)	0·222	−0·0	(0·0040)	0·153	2·6	(0·0099)	0·849	1·2	—	—	—

TABLE A1—*Continued*

Model 3: By race														
White	0·0046** (0·0006)	0·121	3·8	0·0004 (0·0004)	0·031	1·3	0·0016** (0·0006)	0·110	1·5	0·0266** (0·0020)	0·845	3·1	0·0058** (0·0018)	2·02	0·3
Black	0·0073** (0·0023)	0·455	1·6	0·0039* (0·0021)	0·277	1·4	0·0013 (0·0024)	0·274	0·5	0·0174** (0·0034)	0·895	1·9	0·0109 (0·0076)	2·18	0·5
Asian	0·0031 (0·0031)	0·065	4·8	0·0028 (0·0019)	0·022	13·1	−0·0028 (0·0029)	0·055	−5·1	0·0530** (0·0200)	0·806	6·6	0·0402** (0·0116)	1·92	2·1
Other	0·0101* (0·0059)	0·252	4·0	0·0007 (0·0048)	0·143	0·5	0·0071 (0·0054)	0·162	4·4	0·0306** (0·0144)	0·788	3·9	0·0103 (0·0220)	2·14	0·5
Model 4: By education														
Less than high school	0·0061** (0·0013)	0·216	2·8	0·0004 (0·0010)	0·093	0·5	0·0029** (0·0012)	0·157	1·9	0·0243** (0·0031)	0·864	2·8	0·0185** (0·0045)	2·27	0·8
High school	0·0054** (0·0009)	0·159	3·4	0·0010* (0·0006)	0·061	1·7	0·0022** (0·0008)	0·122	1·8	0·0248** (0·0027)	0·853	2·9	0·0060** (0·0025)	2·00	0·3
College	0·0039** (0·0010)	0·136	2·9	0·0010 (0·0006)	0·046	2·3	0·0003 (0·0009)	0·111	0·3	0·0260** (0·0032)	0·845	3·1	−0·0005 (0·0028)	1·91	−0·0

Notes: S.E. are given in parentheses. Each entry is a separate regression. Models in columns (1), (4), (7), and (10) are analogous to models in columns (1), (2), (3), and (6) in Table 2, respectively. The model in column (13) is analogous to column (1) in the bottom panel of Table 5 and uses only data from the 1960–1980 Censuses. For the models by decade of birth, the samples are limited to mothers between the ages 20 and 30. See notes to Tables 2 and 5 for further details. **Statistically significant from 0 at the 5% level; *statistically significant from 0 at the 10% level.

where the derivatives are taken with respect to $\tilde{K}^i$ for $i = M, D$, holding the number of children fixed. The proof for these concavity results can be seen by taking first and second derivatives and using properties of normal density functions, normal distribution functions, and truncated normal distributions.

Since the predictions hinge on the convexity or concavity of θ, it is important to understand the two conditions described in equation (A.12). Both conditions indicate that the expected future benefit of a change in family composition, c, depends on the amount of curvature in the utility function. In the current setting, the terms λ_c^M and λ_c^D are simply two different inverse Mill's ratios or hazard functions (*i.e.* ratios of density functions to survivor functions). This quantity multiplied by the marginal utility of an effective child in the married and divorced states must be less than what is often referred to as the coefficient of absolute risk aversion, an expression that describes the curvature of the utility function.

To make things more concrete, consider a family with one boy versus one girl and let $i = M$. Suppose parents have a gender bias for boys. The L.H.S. of equation (A.12) captures the idea that boy families are more likely to remain married in period 2 and therefore more likely to enjoy the benefits of an additional child in the married state. The R.H.S. of equation (A.12) captures the idea that additional children have potentially rapidly decreasing marginal utility in the married state. Since couples prefer sons, and because sons and daughters are perfect substitutes, the effective number of children is larger in a boy family that remains married. A boy family has more effective children in the married state, so the value of an additional child is lower compared to a girl family by an amount that depends on the curvature of the utility function. For concavity of θ to hold, there must be enough curvature in the utility function so that the benefit due to a lower divorce probability for boy families is smaller than the increase in marginal utility from an additional child for girl versus boy families in the married state. The intuition behind concavity when considering the role model hypothesis and $i = D$ follows similar logic. To summarize, if there is sufficient curvature in the utility function so that equation (A.12) holds, gender bias predicts girl families are relatively more likely to have an additional child. In contrast, the role model hypothesis predicts that boy families are relatively more likely to have an additional child.

Regardless of the concavity of θ, the differential cost hypothesis (when girls are more expensive) predicts higher fertility for boy families. The intuition is that having a girl versus a boy can be thought of as a pure income effect. If children are normal goods and girls have a higher price than boys, then couples whose first child is a girl are poorer. The income effect reduces the demand for additional children as well as other consumption goods that are normal goods.[37] Of course, if boys are more expensive, the opposite is true, and girl families have higher fertility.

Acknowledgements. We are grateful to Michael Ash, Mark Bils, David Card, Hal Cole, Janet Currie, Nora Gordon, Jeff Grogger, Steven Landsburg, David Levine, Matt Neidell, Duncan Thomas, Miguel Urquiola, and seminar participants at the NBER Fall Labor Studies Meetings, Syracuse University, Texas A&M, University of California, Los Angeles, University of Rochester, and University of Texas at Austin, for providing useful suggestions. Ben Bolitzer, Fernando Vendramel Ferreira, Fernando Levia, and Marina Renzo provided excellent research assistance. Moretti thanks the California Center for Population Research for generous financial support.

REFERENCES

ANANAT, E. and MICHAELS, G. (2007), "The Effect of Marital Breakup on the Income Distribution of Women with Children" (CEPR Discussion Paper No. DP6228).

ANDERTON, L. D., NORIKO, O. T., LEE, L. B. and GERALDINE, P. M. (1987), "Intergenerational Transmission of Relative Fertility and Life Course Patterns", *Demography*, **24** (4), 467–480.

ANGRIST, J. and EVANS, W. (1998), "Children and Their Parents Labor Supply: Evidence from Exogenous Variation in Family Size", *American Economic Review*, **88** (3), 450–477.

BECKER, G. (1973 and 1974), "A Theory of Marriage: Part I", *Journal of Political Economy*.

BEDARD, K. and DESCHENES, O. (2005), "Marital Instability and Female Labor Supply: Evidence from Exogenous Variation in Firstborn Sex", *Journal of Human Resources*, **40**, 411–434.

BEN-PORATH, Y. and WELCH, F. (1976), "Do Sex Preferences Really Matter?", *Quarterly Journal of Economics*, **90**, 285–307.

BLAU, F. (1998), "Trends in the Well-Being of American Women, 1970-1995", *Journal of Economic Literature*, **36**, 112–165.

BRACHER, M., SANTOW, G., MORGAN, S. P. and Tressell, D. (1997), "Marriage Dissolution in Australia—Models and Explanations", *Population Studies*, **47**, 403–425.

BRINIG, M. F. (2005), "Does Parental Autonomy Require Equal Custody at Divorce?", *Louisiana Law Review*, **65**, 126–139.

37. Note that no price effect arises because gender is revealed *after* parents have made the decision to have another child. Leung (1991) shows that in a setting where fertility is stochastic and influenced by the parent's precautionary or proactive measures to have a child, and where consumption and children are perfect substitutes, the opposite result can be true. In Leung's setting, the gender bias and differential cost hypotheses are not separately identifiable.

BUTCHER, K. F. and CASE, A. (1994), "The Effect of Sibling Composition on Women's Education and Earnings", *Quarterly Journal of Economics*, **109** (3), 242–259.

CÁCERES-DELPIANO, J. (2006), "The Impacts of Family Size on Investment in Child Quality", *Journal of Human Resources*, **41** (4), 934–754.

CASCIO, E. U. (2004), "Schooling Attainment and the Introduction of Kindergartens into Public Schools" (manuscript, University of California at Davis).

CASE, A. and DEATON, A. (2002), *Consumption, Health, Gender, and Poverty* (Policy Research Working Paper Series 3020, The World Bank).

CASE, A. and PAXSON, C. (2001), "Mothers and Others: Who Invests in Children's Health?", *Journal of Health Economics*, **20**, 301–328.

CHARNOV, E. L., LOS-DEN HARTOGH, R. L., JONES, W. T. and VAN DEN ASSE, J. (1981), "Sex Ratio Evolution in a Variable Environment", *Nature*, **289**, 27–33.

CURRIE, J. and YELLOWITZ, A. (2000), "Are Public Housing Projects Good for Kids?", *Journal of Public Economics*, **75** (1), 99–124.

DAHL, G. and MORETTI, E. (2004), "The Demand for Sons" (NBER Working Paper No. 10281).

DEVINE, D. and FOREHAND, R. (1996), "Cascading Toward Divorce: The Roles of Marital and Child Factors", *Journal of Consulting and Clinical Psychology*, **64**, 424–427.

DIEKMANN, A. and SCHMIDHEINY, K. (2004), "Do Parents of Girls Really Have a Higher Risk of Divorce?", *Journal of Marriage and Family*, **66** (3), 651–660.

EFRAT, Z., AKINFENWA, O. and NICOLAIDES, K. (1999), "First-Trimester Determination of Fetal Gender by Ultrasound", *Ultrasound in Obstetrics and Gynecology*, **13**, 305–307.

GOLDIN, C. (2002), "The Rising (and Then Declining) Significance of Gender" (NBER Working Paper No. 8915).

GRUBER, J. (2000), "Is Making Divorce Easier Bad for Children?" (NBER Working Paper No. 7968).

HAVEMAN, R. and WOLFE, B. (1994), *Succeeding Generations: On the Effects of Investments in Children* (New York: Russell Sage Foundation).

HETHERINGTON, E., COX, M. and COX, R. (1978), "The Aftermath of Divorce in Mother-Child, Father-Child Relationships", in J. Stevens and M. Matthews (eds.) (Washington, DC: National Association for the Education of Young Children) 121–132.

JACOBSEN, J. P., PEARCE, J. W. and ROSENBLOOM, J. L. (2001), "The Effects of Child-Bearing on Women's Marital Status: Using Twin Births as a Natural Experiment", *Economics Letters*, **70** (1), 133–138.

JENA, A. (2006), "Do Children Cause Divorce? Evidence from a Natural Experiment Using Twins" (Working Paper, University of Chicago).

KATZEV, A. R., WARNER, R. I. and ACOCK, A. C. (1994), "Girls or Boys? Relationship of Child Gender to Marital Instability", *Journal of Marriage and the Family*, **56**, 89–100.

KEVANE, M. and LEVINE, D. (2001), "The Changing Status of Daughters in Indonesia" (Working Paper No. 74, Institute of Industrial Relations, University of California, Berkeley).

LAMB, M. (1987), *The Father's Role: Cross Cultural Perspectives* (New York: Lawrence Erlbaum Associates).

LAMB, M. (1997), "The Role of the Father: An Overview", in M. Lamb (ed.) *The Role of the Father in Child Development* (New York: Wiley), 1–36.

LEUNG, S. F. (1991), "A Stochastic Dynamic Analysis of Parental Sex Preferences and Fertility", *Quarterly Journal of Economics*, **121**, 1063–1088.

LUNDBERG, S. (2005), "Sons, Daughters, and Parental Behavior", *Oxford Review of Economic Policy*, **21** (3), 340–356.

LUNDBERG, S. and ROSE, E. (2002*a*), "The Effect of Sons and Daughters on Men's Labor Supply and Wages", *Review of Economics and Statistics*, **84** (2), 251–268.

LUNDBERG, S. and ROSE, E. (2002*b*), "Investment is Sons and Daughters: Evidence from the Consumer Expenditure Survey", Paper presented at the Joint Center for Poverty Research, Chicago, IL.

LUNDBERG, S. and ROSE, E. (2003), "Child Gender and the Transition to Marriage", *Demography*, **40** (2), 333–349.

MAMMEN, K. (2003), "The Effect of Children's Gender on Divorce and Child Support" (Unpublished Paper, Barnard College).

MAULDON, J. (1990), "The Effect of Marital Disruption on Children's Health", *Demography*, **27** (3), 135–157.

McLANAHAN, S. and SANDEFUR, G. (1994), *Growing Up with a Single Parent: What Hurts, What Helps* (Cambridge, MA: Harvard University).

MORGAN, S. P., LYE, D. N. and CONDRAN, G. A. (1988), "Sons, Daughters and the Risk of Marital Disruption", *American Journal of Sociology*, **94** (1), 110–129.

MORGAN, S. P. and POLLARD, M. (2002), "Do Parents of Girls Really Have a Higher Risk of Divorce?" (mimeo, Center for Demographic Studies).

MICHAEL, S. and MORGAN, S. P. (2003), "Emerging Parental Gender Indifference? Sex Composition of Children and the Third Birth", *American Sociological Review*, **67** (4), 23–45.

MOTT, F. L. (1994), "Sons, Daughters and Father's Absence—Differentials in Father-Leaving Probabilities and In-Home Environments", *Journal of Family Issues*, **15**, 97–128.

NORBERG, K. (2004), "Partnership Status and the Human Sex Ratio at Birth" (NBER Working Paper No. 10920).

OLSEN, L. (1983), *Costs of Children* (New York: Heath and Company).

OREOPOULOS, P., PAGE, M. E. and STEVENS, A. H. (2003), "Does Human Capital Transfer from Father to Son?" (NBER Working Paper No. 10164).

PAGE, M. E. and HUFF STEVENS, A. (2004), "The Economic Consequences of Absent Parents", *Journal of Human Resources*, **39** (1), 80–107.
SICHERMAN N. and DIVON, K. (1999), "Measurement Errors in Ultrasound and the Assessment of Fetal Growth" (mimeo, Columbia University).
SPANIER, G. B. and GLICK, P. C. (1981), "Marital Instability in the United States: Some Correlates and Recent Changes", *Family Relations*, **31**, 329–338.
TEACHMAN, J. D. and SCHOLLAERT, P. T. (1989), "Gender of Children and Birth Timing", *Demography*, **26** (3), 189–199.
THOMAS, D. (1990), "Intra-Household Resource Allocation: An Inferential Approach", *Journal of Human Resources*, **25** (4), 635–664.
THOMAS, D. (1996), "Like Father, Like Son; Like Mother, Like Daughter", *Journal of Human Resources*, **29** (4), 950–986.
WHITLOW, B., LAZANAKIS, M. and ECONOMIDES, D. (1999), "The Sonographic Identification of Fetal Gender from 11 to 14 Weeks of Gestation", *Ultrasound in Obstetrics and Gynecology*, **13**, 301–304.
WILCOX, A. J., WEINBERG, C. R. and BAIRD, D. D. (1995), "Timing of Sexual Intercourse in Relation to Ovulation Effects on the Probability of Conception, Survival of the Pregnancy, and Sex of the Baby", *New England Journal of Medicine*, **333**, 1517–1521.
WU, Z. and PENNING, M. J. (1997), "Marital Instability After Midlife", *Journal of Family Issues*, **18**, 459–478.