American Economic Journal: Applied Economics 2009, 1:2, 1–34
http://www.aeaweb.org/articles.php?doi=10.1257/app.1.2.1

Are There Missing Girls in the United States? Evidence from Birth Data†

By Jason Abrevaya*

We offer evidence of gender selection within the United States. Analysis of comprehensive birth data shows unusually high boy-birth percentages after 1980 among later children (most notably third and fourth children) born to Chinese and Asian Indian mothers. Based upon linked data from California, Asian Indian mothers are found to be significantly more likely to have a terminated pregnancy and to give birth to a boy when they have previously only given birth to girls. The observed boy-birth percentages are consistent with over 2,000 "missing" Chinese and Indian girls in the United States between 1991 and 2004. (*JEL* J11, J16)

> *It is not possible to assess how popular sex-determination tests and gender-selection techniques might be among Indian-Americans or any other group. There are no official statistics, and people who wish to choose the sex of their child do not wish to discuss it publicly…*
>
> —*New York Times* article, August 15, 2001

Amartya Sen (1990, 1992) coined the term "missing women" to illustrate differential mortality rates experienced by women in several Asian countries. Sen (1990, 1992) has estimated 80–100 million "missing women" in Asia[1] and has pointed to gender selection as one contributing factor:

> *Given a preference for boys over girls that many male-dominated societies have, gender inequality can manifest itself in the form of the parents' wanting the new born to be a boy rather than a*

* Department of Economics, University of Texas at Austin, Austin, TX 78712 (e-mail: abrevaya@eco.utexas.edu). The author is grateful to Jan Christensen, Karl Halfman, and Roxana Killian of the California Department of Health Services (CDHS) for their assistance during the data acquisition process. The California birth data used in this paper cannot be released due to a confidentiality agreement with the CDHS. The federal birth data and census data used in this paper were obtained from the Inter-University Consortium for Political and Social Research (ICPSR) and the National Bureau of Economic Research (NBER). Ken Chay, David Hummels, Jeffrey Kubik, Steve Levitt, and seminar participants at the University of Georgia, Purdue University, University of Oklahoma, and Syracuse University also provided helpful comments. Jack Barron provided invaluable computer assistance. Dudley Poston, Jr. kindly provided data on Chinese and South Korean sex ratios at birth. An earlier version of this paper (Abrevaya 2005) was based on fewer years of birth data and inadvertently included multiple births (e.g., twins) in some of the empirical analysis.

† To comment on this article in the online discussion forum visit the articles page at: http://www.aeaweb.org/articles.php?doi=10.1257/app.1.2.1.

[1] Several studies (e.g., Stephan Klasen and Claudia Wink 2002, 2003) re-examined the level and trend of "missing women" in Asia.

> *girl. There was a time when this could be no more than a wish (a daydream or a nightmare, depending on one's perspective), but with the availability of modern techniques to determine the gender of the fetus, sex-selective abortion has become common in many countries. It is particularly prevalent in East Asia, in China and South Korea in particular, but also in Singapore and Taiwan, and it is beginning to emerge as a statistically significant phenomenon in India and South Asia as well. (Sen 2001)*

The existing evidence on gender-selective abortion in Asia is primarily indirect based upon unusually high percentages of boys being born.[2] In particular, several Asian countries, including China, India, South Korea, and Taiwan, have seen significant increases in the percentage of boys at birth since the 1970s and 1980s, when ultrasound technology (and to a lesser extent amniocentesis technology) became available and affordable to women (see, for example, Dudley L. Poston, Jr., Julie Luan Wu, and Han Gon Kim 2003; Robert D. Retherford and T. K. Roy 2003; and Poston and Karen S. Glover 2006). To illustrate these trends, Figure 1 provides a plot of boy-percentages-at-birth for China, India, South Korea, and the United States.[3] Whereas the likelihood of a male birth has remained at just above 51 percent in the United States, the percentage of male births rose above 53 percent in China, India, and South Korea in the late 1980s and early 1990s.

Recent research has pointed to more subtle forms of gender bias (specifically, bias favoring sons) in the United States. Shelly Lundberg and Elaina Rose (2003) find that single mothers are more likely to marry a child's biological father if the child is a boy. Gordon B. Dahl and Enrico Moretti (2008) find that parents with sons are less likely to be divorced and that divorced fathers are more likely to have custody of their sons. One previous study that considers the effect of gender bias on prenatal (rather than postnatal) outcomes is Aparna Lhila and Kosali Simon (2008), who find no evidence from federal birth data that gender is related to quality of prenatal care (e.g., prenatal visits, smoking, etc.).

While the boy-percentage trend for the United States in Figure 1 certainly doesn't provide evidence of gender-selective practices in the aggregate, evidence for gender selection may exist at a more disaggregated level.[4] One might suspect, for instance, that those races associated with the Asian countries in Figure 1 (Chinese, Indian, Korean) would be more likely to practice gender selection due to cultural biases.[5] This idea has been suggested by others, including John A.

[2] An exception is Baochang Gu and Yongping Li (1994), who examine the sex ratio of aborted fetuses in southern Zhejiang province. They find a significantly larger proportion of female fetuses aborted after daughters are born.

[3] Three-year moving averages are plotted at each year. Sources: China and South Korea, Poston and Glover (2006); India, Office of the Registrar General of India (2001); United States federal birth data (see Section I).

[4] The slow decline in boy-birth percentages within the United States over the last three decades has been noted in previous research (Devra Lee Davis, Michelle B. Gottlieb, and Julie R. Stampnitzky 1998; Michele Marcus et al. 1998). This decrease has also been observed in other countries, including Canada (Bruce B. Allan et al. 1997), Denmark (Henrik Møller 1996), and the Netherlands (K. M. van der Pal-de Bruin, S. P. Verloove-Vanhorick, and N. Roeleveld 1997).

[5] The term "Indian" will be used to mean "Asian Indian" (rather than "American Indian") throughout this paper.

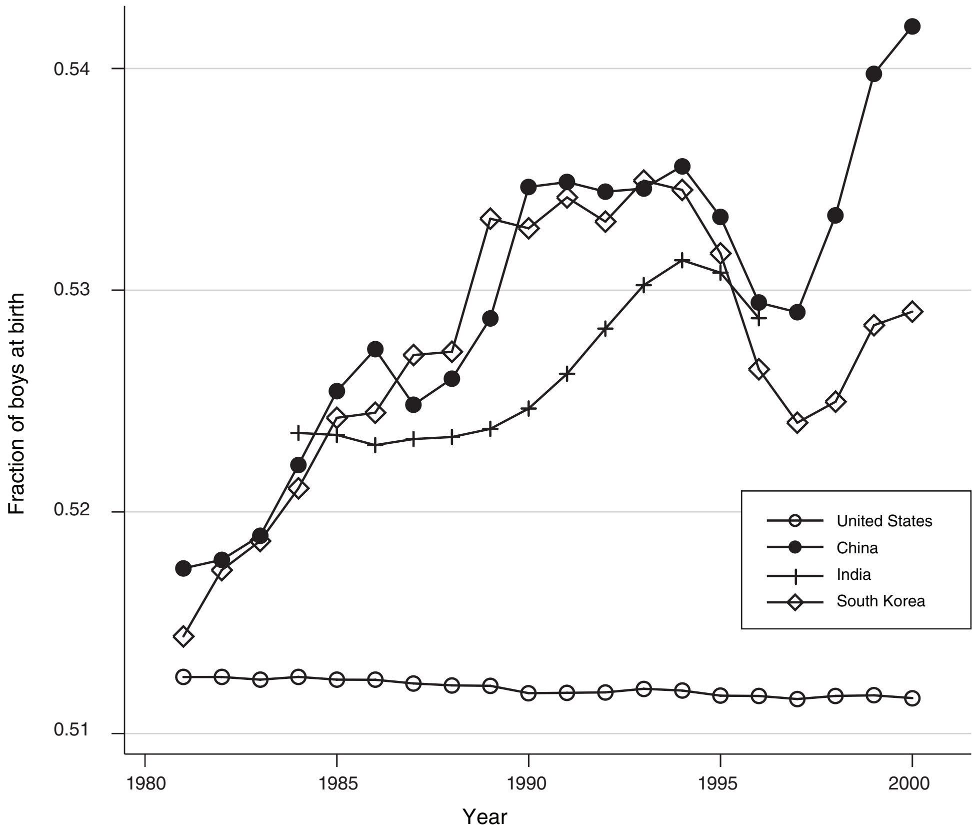

Figure 1. Likelihood of a Male Birth, by Country

Robertson (2001). "Until they are more fully assimilated, immigrant groups in Western countries may retain the same gender preferences that they would have held in their homelands." As anecdotal evidence to this point, a recent *New York Times* article (Susan Sachs 2001) described efforts by several companies to directly market gender identification and pre-conceptive selection products to Indian expatriates in North America:

> *"Desire a Son?" asked an advertisement in recent editions of India Abroad, a weekly newspaper for Indian expatriates in the United States and Canada. "Choosing the sex of your baby: new scientific reality," declared another in the same publication. A third ad ran in both India Abroad and the North American edition of The Indian Express. "Pregnant?" it said. "Wanna know the gender of your baby right now?"*

The incentives for gender selection depend not only on gender preferences but also family size (i.e., number of children already born). Even in the absence of exogenous family-size limits (such as the Chinese "One Child Policy"), gender-selection incentives (in the presence of gender bias) become stronger as a family approaches its own size limit. For instance, consider a family that has a strong preference for having at least one son and is willing to have at most two children. If the first child is a boy, this family might stop having children. If the first child is a girl, the family would have another child and a greater incentive (than in the first pregnancy) to determine gender and, perhaps, undertake a gender-selective procedure. If there were many such families, the data in the aggregate would indicate a higher percentage of boys among second births (as compared to first births) due to the combination of fertility stopping (by families with first-born sons) and gender determination/selection (by families with first-born daughters).[6] More generally, as a family has more children, the incentives for gender selection increase as the opportunity cost of having a child of the less-preferred gender increases.[7]

The foregoing argument suggests that son-biased gender selection is most likely to manifest itself through unusually high boy-birth rates at later births and unusually high boy-birth rates following daughters.

As such, this paper will investigate whether these two irregularities in boy births are present among specific races within the United States. Moreover, if gender selection is arising from cultural biases, we would expect the timing of these irregularities to mirror those in the parents' home countries. The use of higher parity and conditional-upon-previous-gender boy-birth percentages has been considered in several previous studies of Asian countries (see, for example, Chai Bin Park and Nam-Hoon Cho 1995, Retherford and Roy 2003, Prabhat Jha et al. 2006, and Avraham Y. Ebenstein 2007). The recent study by Sylvie Dubuc and David Coleman (2007) found that the likelihood of male births to India-born mothers in the United Kingdom had an overall upward trend since the 1980s and is significantly higher at third and later births after 1990.[8]

If parents wish to select their baby's gender, there are currently three options in the United States: gender-selective abortion, gender-selective in vitro fertilization (IVF), or sperm sorting. The latter two options are performed prior to pregnancy. Gender-selective IVF is a modified version of the traditional IVF procedure, in which fertilized embryos are transferred into the mother's uterus. For gender-selective IVF, however, embryos are genetically tested ("preimplantation genetic diagnosis") to determine gender and chosen accordingly. Such testing is nearly 100 percent accurate for gender determination and, when done for gender reasons only (rather than avoiding a genetic disease), has been banned in many countries. Although a very effective means of gender selection, the IVF procedure is expensive (\$10,000–\$20,000 per implantation cycle).[9] Sperm sorting, on the other hand, is less expensive (costing a few thousand dollars) but also less effective. The procedure

[6] Fertility stopping by itself has no impact on boy-birth percentages, but the sample of families having second children are over represented by those families with first-born daughters. As such, the second-born boy-birth percentage would be even higher than it would have been if all families with first-born sons had also had a second child.

[7] Jinyojng Kim (2005), Jason Abrevaya (2005), and Ebenstein (2007) develop dynamic models of gender selection.

[8] Dubuc and Coleman (2007) did not control for potential confounders (maternal characteristics, prenatal variables) that might affect the likelihood of a male birth.

[9] Insurance coverage for IVF is currently mandated in only a handful of states (Connecticut, Illinois, Massachusetts, New Jersey, and Rhode Island). See http://www.asrm.org/Patients/insur.html.

involves selecting sperm from a given sperm sample in order to increase the probability of the desired gender when the egg is fertilized.[10] Although both gender-selective IVF and sperm sorting may be options for gender selection, these two procedures would likely only account for a very small proportion of the gender-selective procedures that might have occurred in the United States in the past few decades. The reasons for this include their recent introduction, their high expense, and the limited number of doctors willing to perform such procedures. As such, this study will focus primarily on abortion as the means for gender selection. On the other hand, when thinking about the future of gender selection, these more advanced technologies will play an increasingly important role.

Turning to gender-selective abortion, the introduction of ultrasound and amniocentesis in the 1970s made such a procedure a possibility. Although neither technology was introduced for the explicit purpose of determining the gender of a fetus, both technologies are capable of this determination during the first half of pregnancy. Amniocentesis, generally performed between the fourteenth and eighteenth weeks of pregnancy, is nearly 100 percent accurate in determining gender but has a small risk (0.5–1.0 percent) of miscarriage associated with it. Ultrasound, which can usually be used to detect gender between the sixteenth and twentieth weeks of pregnancy, is safer than amniocentesis but is somewhat less accurate in gender determination.[11] If either ultrasound or amniocentesis is used as a precursor to gender-selective abortion, the abortion would most likely occur during the second trimester of pregnancy. Although most abortions in the United States occur prior to the second trimester, there are a large number of abortions that do occur during the second trimester and later. Table 1 provides some summary statistics on abortions in the United States in 1980, 1990, and 2000, as reported by the Centers for Disease Control (CDC 2003). Since 1980, roughly 5 percent of abortions have occurred at 16 weeks or later. These numbers, of course, do not imply gender selection. They merely indicate that a nonnegligible fraction of abortions occur after the point that gender determination is possible. Another interesting fact from Table 1 is that a large percentage of abortions are associated with women who have previously had live births (41.6 percent in 1980, 54.8 percent in 1990, and 60.0 percent in 2000). According to Stanley K. Henshaw and Lawrence B. Finer (2003), the average cost of an abortion at 20 weeks of gestation was just over $1,000 in 2001. For abortions that are not "medically necessary," this cost is most likely paid out of pocket.

The outline of the paper is as follows. Section I describes the different data sources (federal birth data, California birth data, and census data) used in the empirical analysis. Section II reports the empirical results. Wherever possible,

[10] One company that offers sperm sorting in the United States (Microsort) claims a success rate of 92 percent (668 out of 726) for couples who desired a girl and 81 percent (172 out of 211) for couples who desired a boy. These success rates were reported for pregnancies through January 1, 2007 and are on the company's Web site (http://www.microsort.com). Scientific evidence of the technology's effectiveness has existed for more than a decade (e.g., Lawrence A. Johnson et al. 1993).

[11] Chorionic villus sampling (CVS) can also be used for gender determination. CVS is performed at 10–13 weeks and is nearly 100 percent accurate. However, CVS carries a greater risk of fetal loss than amniocentesis and is rarely performed. For example, use of CVS during pregnancy was reported for only 0.1 percent of births in California between 2000 and 2003.

Table 1—Summary Statistics on Abortion in the United States

	1980	1990	2000
Reported # of legal abortions	1,297,606	1,429,247	857,475
Panel A: Weeks of gestation			
8 weeks or less	51.7%	51.6%	58.1%
9–10	26.2%	25.3%	19.8%
11–12	12.2%	11.7%	10.2%
13–15	5.1%	6.4%	6.2%
16–20	3.9%	4.0%	4.3%
21 weeks or more	0.9%	1.0%	1.4%
Panel B: Previous live births			
Zero	58.4%	46.2%	40.0%
One	19.4%	25.9%	27.7%
Two or more	22.2%	27.9%	32.3%

Source: Centers for Disease Control and Prevention (2003).

results are reported separately for the following racial groups: whites (specifically, non-Hispanic whites), Chinese, Indian, Japanese, and Korean.[12] The sample of white births is extremely large and therefore allows very precise estimates of boy-birth percentages and their determinants. We view the white sample as a "control" for comparison with the Chinese, Indian, and Korean samples since there is likely to be minimal gender selection among whites. The Japanese sample serves as another "control" for comparison since Japan has not exhibited the gender-selective trend observed in other Asian countries. First, we analyze the federal and California birth data to determine the factors associated with a baby's gender. The statistical analysis on births after 1980, both with and without controls, indicates that Chinese and Indian mothers are significantly more likely to have sons at higher birth parities (third and fourth children) than for their first child. Second, we analyze a maternally linked version of the California data. This version allows us to condition upon gender of previous children and to determine whether the current baby's gender and terminated pregnancies are systematically related to previous children's gender. We find that Indian mothers are significantly more likely to have a terminated pregnancy and to give birth to a boy when they have previously only given birth to girls. Third, we use a simple framework to infer the degree of gender selection that would explain the unusual boy-birth percentages observed in the data. Finally, we briefly consider evidence from census data on race-specific gender preferences (specifically, the decision to have a second or third child based upon the gender of previous children) and also the likelihood of sons conditional upon previous children's gender. Section III concludes.

[12] Results for other racial groups (the largest being black, Hispanic, Vietnamese, and Filipino) are available from the author. We find no convincing evidence consistent with gender selection among other racial groups. Although Vietnam has seen a recent increase in its boy-birth percentage (Institute for Social Development Studies 2007), this increase has been far less dramatic (and occurred later) than the increase in China, India, and South Korea. The black and Hispanic samples, like the white sample, are extremely large and offer precise estimates. Qualitatively, these estimates are extremely similar to those for white mothers, with the overall percentage of male births slightly lower than within the white sample. For instance, see the results for black mothers in Figure 2.

I. Data Sources

Unfortunately, existing abortion data in the United States are inadequate for analyzing evidence of gender-selective practices. First, gender is not recorded in the two primary abortion surveys in the United States, conducted by the CDC and the Alan Guttmacher Institute. Second, although information on the number of previous live births is available in these surveys, there is no information on the gender of a mother's existing children. Third, not all states have abortion data available. Ted Joyce et al. (2005), who have compiled the most comprehensive data on abortions to date, indicate that 19 states (including populous states such as California, Florida, and Illinois) had data unavailable "due to statutory restrictions or inadequate data collection and/or storage." Fourth, when women are asked about the reason(s) for having an abortion, gender preference is rarely mentioned (see, for example, Aida Torres and Jacqueline Darroch Forrest 1998).

Therefore, rather than using abortion data, we consider three different comprehensive data sources that allow us to analyze trends and potential irregularities in male births: federal birth data (annual files from 1971 to 2004) from the National Center for Health Statistics (NCHS); California birth data (annual files from 1970 to 2005) from the California Department of Health Services (CDHS); and, the 5-percent public-use microdata samples (PUMS) of the United States census (1980, 1990, and 2000). The federal birth data and census data are publicly available, whereas the California birth data contain personal identifiers and are subject to confidentiality restrictions. As discussed in more detail below, the personal identifiers were used to maternally link births and identify siblings.

Table 2 provides a summary of the three data sources to clarify the advantages and disadvantages of each. Further details for each of the three data sources are given below.

Federal Birth Data.—These annual data files contain information on births occurring within the United States, obtained from birth certificates filed in individual states. Since 1985, a 100 percent sample of birth certificates has been used to compile these data. In 1971, a 50 percent sample of birth certificates was used. From 1972 to 1984, a 100 percent sample was used for states participating in the Vital Statistics Cooperative Program (with the number of such states increasing from 6 to 46 during the period), and a 50 percent sample was used for other states. Each record in the federal birth data contains detailed information about the birth (including gender and parity), maternal characteristics (including age, education, and race), and prenatal care (including month of first prenatal visit). Each birth record also indicates the number of previous terminated pregnancies a mother has experienced and, from 1989 on, whether ultrasound and/or amniocentesis were used during pregnancy. The number of terminated pregnancies includes both voluntary and involuntary terminations but does not specify the type(s) of termination(s). The federal data has two important limitations. First, detailed Asian races (including Indian and Korean) were only recorded in the data starting in 1992. Prior to 1992, the only specific Asian races recorded were Chinese and Japanese.[13] Second, due to lack

[13] Depending on the year, other Asian races are categorized as "Other" or "Other Asian or Pacific Islander."

Table 2—Summary of Data Sources

Dataset	Federal birth data	California birth data	Census (5% PUMS) data
Years	1971–2004	1970–2005	1980, 1990, 2000
Sample	1971–1984: 50–100% of births 1985–2004: 100% of births	All California births	5% of US population
Asian race information	Chinese and Japanese in all years; detailed races available from 1992 on	Chinese and Japanese in all years; detailed races available from 1982 on	Detailed races available
Able to link siblings	No	Yes, using maternal identifiers	Yes, using household identifiers
Prenatal care data	Yes, with ultrasound and amniocentesis usage available from 1989 on	Yes, with ultrasound and amniocentesis usage available from 1989 on	No
Data on previous terminated pregnancies	Yes, all years	Yes, all years	No

of personal identifiers, there is no way to reliably link siblings together. Therefore, although the birth-order and gender of a given child are observed, one cannot relate birth outcomes to the gender of a mother's previous children.

California Birth Data.—The California birth data contain information on all births in California between 1970 and 2005 (a total of over 16.9 million births, accounting for roughly 10 percent of all births in the United States). The California data overcome the two limitations of the federal data mentioned above. First, detailed Asian races are available starting in 1982. Second, personal identifiers (specifically, mother's maiden name and mother's birthdate) enable accurate matching of a given mother's births. This paper will use both an unlinked version and a linked version of the data. The unlinked version, which makes no use of the personal identifiers, serves as a useful complement to the federal data since the Indian and Korean identifiers are absent from the federal data between 1982 and 1991. The linked version is used in order to analyze birth and pregnancy outcomes for a mother's second and/or third child, conditioning on gender of previous children. Additional details on the linking algorithm are provided in Appendix A.

Census Data.—While birth data are ideal for examining prenatal gender selection, the census data will also be considered for complementary evidence. Despite allowing sibling linkages, the census data has several drawbacks relative to the linked California birth data:

- smaller sample sizes for the Asian races,
- lack of prenatal data, and
- observation of household's gender composition after births (which may be affected by childhood deaths, household dissolution, etc.). Since all family

members (with age and gender) in a household are observed, the census data are suitable for examining how fertility decisions depend on the gender mix of previous children. This idea has been pursued by others (e.g., Dahl and Moretti forthcoming) but not at the detailed racial level considered here. Details on construction of the census samples are provided in Appendix B.

For an overview of the birth data and a comparison of their sample sizes, Table 3 reports sample averages of the variables that will be used in the analysis in the next section. Results are broken down by mother's race and reported for 1992–2004, the years for which information is available in both datasets for the five races considered. The last row reports the percentage of US births that occurred in California for each of the racial categories. For the purposes of this study, an appealing feature of the California data is the disproportionately large number of Asian births. The percentages of births occurring in California for the four Asian racial categories range from 27.9 percent for Indian mothers to 44.7 percent for Korean mothers, which are far greater than the 10 percent overall percentage of the country's births.

The percentage of foreign-born mothers among Chinese, Indian, and Korean births is extremely high, about 90 percent for Chinese mothers and 95 percent for both Indian and Korean mothers. The percentage of births to fathers of the same race is also very high for these races, between 70 and 80 percent for Chinese and Korean births and nearly 90 percent for Indian births. As a comparison, for the Japanese sample, the percentage of foreign-born Japanese mothers (57.5 percent) and same-race fathers (39.8 percent) are significantly lower. The high percentage of foreign-born Asian mothers and same-race fathers suggests that cultural influences could play a role in fertility decisions. Table 3 also indicates several differences between the Asian mothers and non-Asian mothers. Compared to white mothers, Asian mothers are, on average, older when they give birth, more educated, more likely to have first-trimester prenatal care, less likely to have had a previous termination, and more likely to have a boy.

II. Empirical Results

A. *Boy-birth Percentages at Later Births*

In this section, we analyze the trends in boy-birth percentages at later births. First, time-series plots from the California birth data are provided. Then, we document boy-birth percentages by birth parity in both the California and federal data. Finally, to establish a more convincing link between birth parity and male births, we use regression analysis to control for observable maternal characteristics and prenatal care variables that might influence the likelihood of a male birth. Most of the results will be described in terms of boy-birth percentages. At times, however, we will also mention the associated sex ratio at birth (SRB) (defined as the number of boys born per 100 girls) since this measure is commonly used in the demographic literature.

Figure 2, based upon the California birth data, plots the time series of boy-birth percentages within California broken down by birth parity (two categories:

Table 3—Descriptive Statistics for Federal Birth Data, 1992–2004

	Chinese	Indian	Japanese	Korean	White
Boy birth	0.519	0.516	0.514	0.520	0.513
Birth parity	1.92	2.01	1.96	2.12	2.27
Mother's age	30.21	27.35	30.60	29.16	26.78
Mother's education	14.20	14.62	14.67	14.72	13.45
Foreign-born mother	0.904	0.950	0.575	0.945	0.055
Same-race father	0.772	0.886	0.398	0.719	0.875
Father's race missing	0.037	0.042	0.035	0.031	0.095
No prenatal care	0.003	0.009	0.005	0.005	0.007
1st-trimester initial visit	0.845	0.809	0.873	0.818	0.858
2nd-trimester initial visit	0.106	0.121	0.080	0.116	0.102
3rd-trimester initial visit	0.021	0.030	0.018	0.032	0.018
Prior terminated pregnancy	0.228	0.202	0.254	0.229	0.258
Ultrasound during pregnancy	0.636	0.660	0.664	0.576	0.687
Amniocentesis during pregnancy	0.057	0.029	0.079	0.035	0.032
Number of US births	381,034	258,871	112,448	122,169	30,706,760
Number of California births	152,435	72,088	39,239	54,601	2,308,550
Percentage of US births occurring in California	40.0%	27.8%	34.9%	44.7%	7.5%

Notes: Sample averages are reported. "Birth parity" is the number of previous births plus one. "Same-race father" is 1 if mother and father have the same reported race. The "initial visit" indicator variables are based upon the timing of first prenatal care visit (1–3 months, 4–6 months, and 7+ months, respectively).

first/second births and third/fourth births) and mother's race.[14] Results for black mothers are also included for comparison. For black and white mothers, the third/fourth boy-birth percentages track the first/second boy-birth percentages very closely. For white mothers, the third/fourth boy-birth percentages are slightly lower (0.1–0.2 percentage points) than the first/second birth percentages. For the four Asian categories, the first/second boy-birth percentages remain stable over the time period shown on the graphs. Looking at third/fourth boy-birth percentages, however, the pattern observed for Indian mothers is striking. During the 1980s, the boy percentage among third/fourth births was 1–2 percentage points higher than among first/second births. Starting around 1990, this difference rose dramatically, with the level of boy-birth percentages among third/fourth births reaching around 58 percent (SRB of 138) in the mid-1990s. During this same time period, the level of boy-birth percentages among first/second births to Indian mothers remained stable and consistently under 52 percent. The time-series pattern at third/fourth births roughly coincides with the overall boy-birth pattern seen in India at the same time (see Figure 1).

For Chinese and Korean mothers, there is some evidence of a difference in boy-birth percentages at later births. Among Korean mothers, the third/fourth boy-birth percentage was mostly 1–2 percentage points higher since the mid-1980s, although the difference disappears in the last several years of the time series plot. Among Chinese mothers, the difference appears primarily between the late 1980s and the mid 1990s, with essentially no difference before or afterward.[15] The

[14] Each point represents a seven-year moving average.

[15] Strangely, there is a very low likelihood of boy births among third/fourth Chinese births in the early and mid 1970s. Although we do not have a good explanation for this occurrence, this period predates the availability of ultrasound technology and, therefore, is likely not related to gender selection.

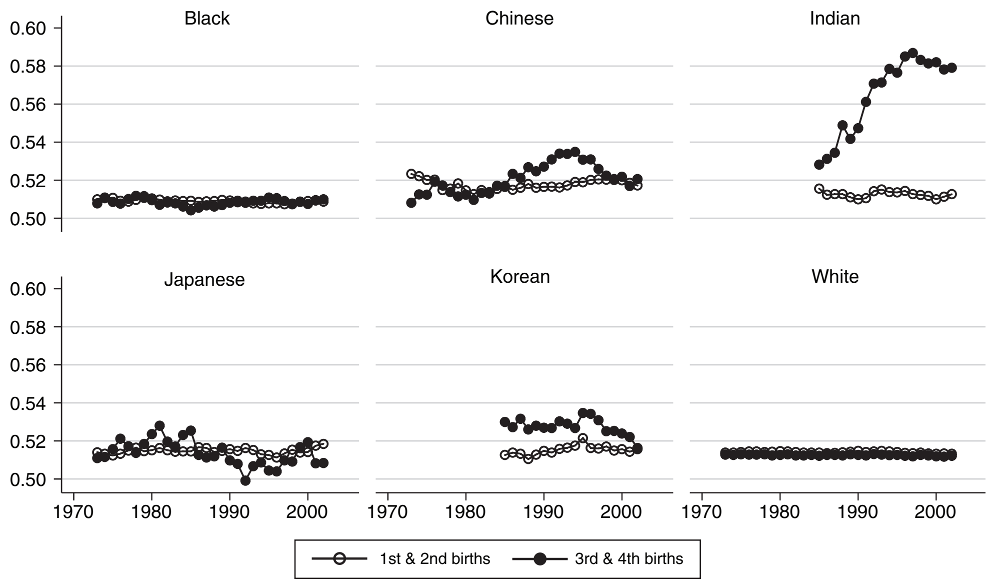

FIGURE 2. BOY-BIRTH LIKELIHOODS BY BIRTH PARITY AND RACE, CALIFORNIA

pattern for births to Japanese mothers is a bit more erratic, with the boy-birth percentages at later births moving both above and below the boy-birth percentages at earlier births.

To provide a more detailed look at the trends among Chinese, Indian, and Korean mothers, Figure 3 plots the time series at first, second, third, and fourth births (rather than combining first/second and third/fourth as in Figure 2). White mothers are included as a comparison. As before, the boy-birth percentages for white mothers have little relation to birth parity and exhibit no visible trends (with "normal" SRB levels of 105–106). For the Asian races, however, the fourth-birth percentages are higher than the lower-parity percentages. The fourth-birth percentage for Chinese mothers peaked around 56 percent (SRB of 127) in 1996, whereas the fourth-birth percentage for Indian mothers was near 60 percent (SRB of 150) in the early 1990s and continued near this level after 2000. For Indian mothers, there is also a noticeable difference between first-birth and second-birth percentages during the 1990s. We note, however, that the Indian second-birth percentages are quite similar to those of Chinese and Korean mothers, while the Indian first-birth percentages dipped below the usual level during this period.

Table 4 provides a breakdown of the likelihood of boy births by time period (1970–1980, 1981–1990, 1991–2005), by race, and by parity of the child. Results for the federal birth data (panel A) and California birth data (panel B) are reported. Each cell has a boy-birth percentage with its associated standard error. The total number of births (parity one through four) is reported in the last column. The pattern among white births is that higher-parity births are slightly less likely to be boys, which (as discussed below) would be expected if later births are more common among women with lower socioeconomic status and lower-quality prenatal care.

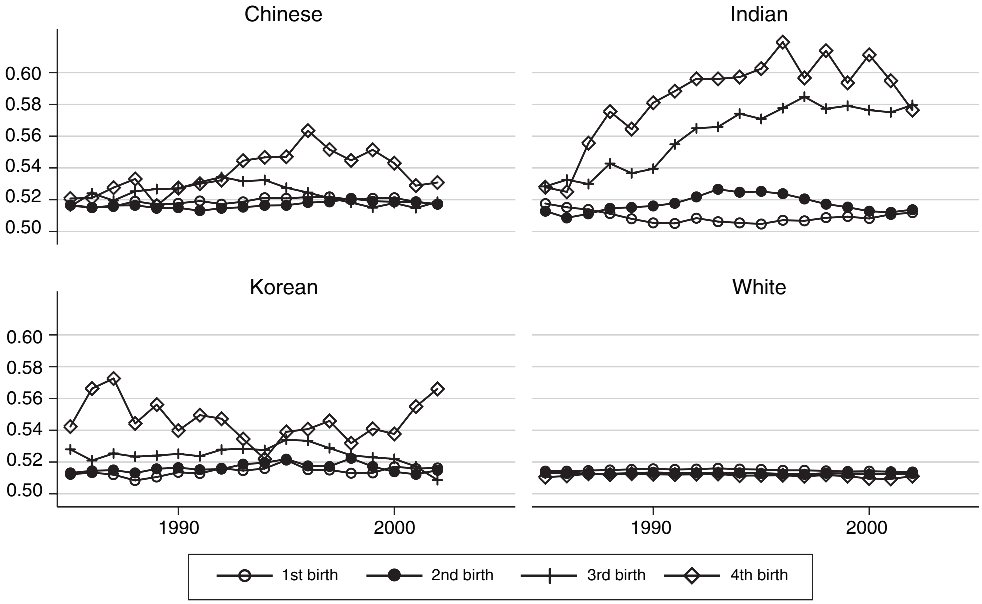

Figure 3. Boy-birth Likelihoods by Birth Parity and Race, California

The first-child boy percentages among the Asian races are roughly the same as those for whites. For later children in later time periods (since 1980), higher boy percentages among Chinese, Indian, and Korean births begin to emerge. In the latest time period for the US data, the boy-birth percentages for Chinese mothers are 53 percent (SRB of 113) and 54 percent (SRB of 118) at third and fourth births, respectively; both of these percentages are significantly different from the first-boy-birth percentage (at a 5 percent level). For Indian mothers in the US between 1992 and 2004, we observe similar boy-birth percentages at third and fourth births (54.4 and 53.5 percent, respectively) that are also significantly different from the first-boy-birth percentage. The higher percentages for Korean mothers at later births are not statistically different within the US sample, although there is evidence of a statistically significant difference at fourth births within the California sample. The higher boy-birth percentages among Indian mothers at higher parity are more dramatic in the California sample than in the US sample. For the 1991–2005 period, where there are significant differences at higher parity, the boy-birth likelihood is 57.5 percent (SRB of 135) and 59.0 percent (SRB of 144) for third and fourth births, respectively.

To control for other factors (such as mother's age, prenatal care, etc.) that might affect the likelihood of having a boy, Table 5 reports regression results using these same data sources and time periods. The reported results are from linear probability models, where the dependent variable is an indicator variable equal to one for boy births. Heteroskedasticity-robust standard errors are reported. The linear probability model is particularly appropriate for this application since the fitted probabilities are very close to 50 percent; probit estimation yields nearly identical results in all cases. To make this table (and Tables 6–8) easier to read, all

Table 4—Boy-birth Likelihoods, US and California Birth Data

		1st birth	2nd birth	3rd birth	4th birth	Sample size
Panel A: Federal natality data						
Chinese	1971–1980	0.519	0.513	0.513	**0.478**	53,879
		(0.003)	(0.004)	(0.006)	(0.011)	
	1981–1990	0.517	0.517	**0.526**	0.525	151,925
		(0.002)	(0.002)	(0.004)	(0.008)	
	1991–2004	0.518	0.518	**0.530**	**0.540**	399,820
		(0.001)	(0.001)	(0.003)	(0.006)	
Indian	1992–2004	0.510	**0.516**	**0.544**	**0.535**	255,610
		(0.001)	(0.002)	(0.003)	(0.007)	
Japanese	1971–1980	0.507	**0.520**	0.519	0.517	42,997
		(0.004)	(0.004)	(0.007)	(0.014)	
	1981–1990	0.513	0.514	0.512	0.517	72,201
		(0.003)	(0.003)	(0.005)	(0.011)	
	1991–2004	0.513	0.514	0.519	0.521	119,267
		(0.002)	(0.002)	(0.004)	(0.009)	
Korean	1992–2004	0.519	0.519	0.527	0.529	121,021
		(0.002)	(0.002)	(0.004)	(0.011)	
White	1971–1980	0.515	**0.514**	**0.513**	**0.512**	17,519,422
		(0.000)	(0.000)	(0.000)	(0.001)	
	1981–1990	0.514	**0.514**	**0.513**	**0.513**	24,497,438
		(0.000)	(0.000)	(0.000)	(0.000)	
	1991–2004	0.514	**0.513**	**0.512**	**0.511**	32,250,458
		(0.000)	(0.000)	(0.000)	(0.000)	
Panel B: California natality data						
Chinese	1970–1980	0.520	0.516	0.516	0.497	33,416
		(0.004)	(0.005)	(0.008)	(0.013)	
	1981–1990	0.516	0.516	0.517	0.524	78,792
		(0.003)	(0.003)	(0.005)	(0.010)	
	1991–2005	0.519	0.517	0.525	**0.539**	174,772
		(0.002)	(0.002)	(0.004)	(0.009)	
Indian	1982–1990	0.515	0.512	0.530	0.543	17,026
		(0.006)	(0.006)	(0.011)	(0.023)	
	1991–2005	0.509	**0.518**	**0.575**	**0.590**	82,999
		(0.002)	(0.003)	(0.006)	(0.013)	
Japanese	1970–1980	0.518	0.512	0.512	0.517	25,478
		(0.005)	(0.005)	(0.009)	(0.017)	
	1981–1990	0.512	0.520	0.513	0.531	30,834
		(0.004)	(0.005)	(0.009)	(0.018)	
	1991–2005	0.516	0.512	0.511	0.496	45,017
		(0.003)	(0.004)	(0.007)	(0.015)	
Korean	1982–1990	0.511	0.514	0.530	**0.559**	33,670
		(0.004)	(0.004)	(0.009)	(0.022)	
	1991–2005	0.517	0.517	0.520	**0.550**	63,726
		(0.003)	(0.003)	(0.006)	(0.016)	
White	1970–1980	0.515	0.513	**0.512**	0.513	2,426,607
		(0.000)	(0.001)	(0.001)	(0.001)	
	1981–1990	0.514	0.513	0.513	0.512	2,260,572
		(0.000)	(0.001)	(0.001)	(0.002)	
	1991–2005	0.515	**0.513**	**0.513**	**0.511**	2,613,136
		(0.000)	(0.001)	(0.001)	(0.001)	

Notes: Each race/time period cell reports the fraction of male births, with standard error in parentheses. Bold indicates a significant difference (at the 5 percent level) from the first-birth boy likelihood.

Table 5—Boy-birth Regressions

		Federal data			California data		
Race	Parity	1971–1980	1981–1990	1991–2004	1970–1980	1981–1990	1991–2005
Chinese	2nd child	−0.395	−0.068	0.002	−0.533	−0.092	−0.139
		(0.532)	(0.295)	(0.180)	(0.642)	(0.410)	(0.268)
	3rd child	−0.758	0.923**	1.176**	−0.592	0.150	0.750*
		(0.779)	(0.452)	(0.304)	(0.925)	(0.614)	(0.445)
	4th child	−3.422**	0.742	2.250**	−2.966**	0.593	2.058**
		(1.358)	(0.830)	(0.644)	(1.479)	(1.072)	(0.923)
Indian	2nd child			0.791**		0.025	0.990**
				(0.233)		(0.898)	(0.403)
	3rd child			3.575**		2.076	6.658**
				(0.380)		(1.309)	(0.693)
	4th child			2.481**		3.227	7.942**
				(0.722)		(2.502)	(1.390)
Japanese	2nd child	1.432**	0.020	0.143	−0.849	0.648	−0.525
		(0.591)	(0.423)	(0.332)	(0.724)	(0.647)	(0.531)
	3rd child	1.629*	−0.371	0.581	−1.081	0.082	−0.712
		(0.852)	(0.616)	(0.504)	(1.060)	(0.979)	(0.824)
	4th child	1.378	0.106	0.527	−0.971	1.465	−2.627
		(1.607)	(1.180)	(0.950)	(1.890)	(1.937)	(1.622)
Korean	2nd child			0.326		0.277	−0.174
				(0.333)		(0.628)	(0.451)
	3rd child			1.254**		1.712	−0.027
				(0.526)		(1.063)	(0.717)
	4th child			1.154		4.360*	3.011*
				(1.168)		(2.369)	(1.632)
White	2nd child	−0.114**	−0.111**	−0.067**	−0.141*	−0.138*	−0.185**
		(0.031)	(0.024)	(0.021)	(0.077)	(0.078)	(0.073)
	3rd child	−0.232**	−0.183**	−0.144**	−0.269**	−0.161	−0.194**
		(0.043)	(0.032)	(0.028)	(0.106)	(0.105)	(0.097)
	4th child	−0.407**	−0.215**	−0.204**	−0.285*	−0.259	−0.320**
		(0.064)	(0.051)	(0.043)	(0.156)	(0.166)	(0.147)

Notes: Each estimate is from a linear regression with boy birth as the dependent variable where the sample consists of singleton births (first through fourth children) to mothers of a given race. Heteroskedasticity-robust standard errors are reported in parentheses. Estimates and standard errors have been multiplied by 100 and should be interpreted as differences in boy-birth percentage from first-child births. The specification includes birth year, a full set of mother's age dummies, and indicator variables for foreign-born mother, same-race father, father's race missing, no prenatal care, second trimester initial visit, third trimester initial visit, and previous terminated pregnancy. The 1991–2004/5 regressions also include mother's education and indicators for ultrasound and amniocentesis use during pregnancy.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

estimates have been scaled up by a factor of 100 so that they can be interpreted as percentage-point effects; for instance, an estimate of 1 would correspond to an increase of one percentage point in the boy birth probability.

For each time period and race considered in Table 5, results are reported for a regression specification that includes parent-related and pregnancy-related control variables. The covariates are birth year, a full set of mother's age dummies, and indicators for foreign-born mothers, same-race father, father's race missing, no prenatal care, initial prenatal visit in second trimester, initial prenatal visit in third trimester, and previous terminated pregnancy. For the most recent time period (1991–2004, US; 1991–2005, California), the covariates also include mother's education and indicators for ultrasound and amniocentesis usage during pregnancy. The indicator variable for first-child births is the "omitted category," so that the estimates for the three birth-parity indicators ("second child," "third child," and "fourth child") should be interpreted as a difference in boy-birth likelihood from first-child births. For instance, in the US sample of Chinese births for 1991–2004, the regression indicates that, holding parental and prenatal characteristics fixed, the fourth child is 2.250 percentage points more likely to be male than the first child.

For white births, the likelihood of a boy becomes slightly lower at higher parity, even when other variables are included as controls. This finding holds during the 1970s, the period in which gender determination would have been either impossible or very unlikely, and then continues in the later periods. Note that the magnitudes of the birth-parity effects are quite low for white births (for example, between 0.067 and 0.407 percentage points in the federal data regressions with control variables), but the huge sample sizes allow these effects to be precisely estimated.

For Chinese births, statistical evidence of higher boy percentages for third and fourth children is seen in the 1991–2004 federal sample and the 1991–2005 California sample (just over 2 percentage points more likely to have a fourth-child boy than a first-child boy). The evidence of higher boy percentages at later births is even stronger among Indian parents, with larger effects seen for the third child (3.6 percentage points in the federal sample and 6.7 percentage points in the California sample) and the fourth child (roughly 8 percentage points in the California sample). For Indian parents, even the second-child boy percentage is significantly higher than the first-child boy percentage (at a 5 percent level), with a magnitude of around one percentage point. For Korean parents, nearly all of the estimates on the third- and fourth-order births are positive but only a few are statistically significant at the 5 percent level. For Japanese births, there is no significant effect from birth parity found in any of the results after 1980.

We stress the importance of controlling for observable differences in parents' characteristics and prenatal behavior as these factors can affect the likelihood of a male birth. Male fetuses have a more difficult time surviving pregnancy (e.g., Thomas T. Perls and Ruth C. Fretts 1998, and Reiko Mizuno 2000).[16] Since the birth data contain only live births, a "survival bias" could contaminate the estimated birth-parity effects if other biological factors are not considered. All things being equal, one would expect that male births are more likely in the presence of favorable demographics (younger mothers, higher education) and quality prenatal care (earlier

[16] Although the data on fetal deaths in the United States are limited, the existing information indicates that the percentage of male fetal deaths is significantly higher than the percentage of male live births. For instance, according to data from the NCHS, 53.3 percent of the 214,043 fetal deaths that occurred after 20 weeks of gestation between 1995 and 2002 were male. Gender is usually not recorded for fetal deaths prior to 20 weeks of gestation. Of the 21,399 fetal deaths where gender was recorded between 1995 and 2002, 66.9 percent were identified as male. *Source*: National Center for Health Statistics, Perinatal Mortality Data, 1995–2002. Data was obtained from the National Bureau of Economic Research.

prenatal visits, ultrasound usage). Although it is tempting to infer gender-selective practices from the estimates on the ultrasound, amniocentesis, and previous-termination indicator variables, we strongly caution against doing so. Ultrasound is primarily a proxy for quality prenatal care, whereas both amniocentesis and previous termination are both primarily proxies for pregnancy complications. As such, even in the absence of gender selection, ultrasound usage would be expected to be positively associated with male births and amniocentesis and previous termination to be negatively associated with male births. In the interest of space, we do not report the full set of estimates for the control variables. Appendix C reports results for the sample of white mothers, where significant effects are found for several control variables. However, as seen above, these effects have no important impact on the conclusions regarding the birth-parity estimates (and their magnitudes are very small in comparison to the size of the birth-parity differences found among Chinese and Indian mothers).

B. *California Birth Outcomes Conditional upon Previous Gender*

This section utilizes the maternally linked California birth data to analyze the relationship between previous gender(s) of a mother's child(ren) and subsequent birth outcomes. The analysis focuses on second- and third-birth outcomes, as the number of births for the Asian races becomes too small at higher birth parities. In addition, we focus on the time period 1982–2005 for which information is available on all races.

Table 6 reports the boy-birth regression results for the samples of second- and third-child births. The second-child regressions include an indicator variable for a first-born girl child and the set of control variables considered in the regressions of Section IIA. Only the coefficient estimate of the first-born girl indicator variable is reported. For the third-child regressions, the same control variables are included and coefficient estimates are reported for a no-sons indicator variable and a one-son indicator variable (relative to the omitted category of two sons).

For both Chinese and Indian mothers, there is a significantly positive effect of first-child gender on second-child gender. Chinese and Indian mothers were, respectively, 0.9 and 2.8 percentage points more likely to have a boy as their second child if their first child was a girl. The associated z-statistics are 2.20 and 3.63, respectively. The positive effects of previous female births among Chinese and Indian mothers are also seen in the third-child regression results. While the estimates for Chinese mothers are not statistically significant, the estimate of the effect of the no-sons indicator variable (relative to two sons) for Indian mothers is extremely large and statistically significant (11.3 percentage points, with a z-statistic of 3.68). This estimate implies that Indian mothers with no previous sons are roughly 20 percent more likely to have a third-child son than Indian mothers with two previous sons.

No significant effects of previous gender are found for Japanese or Korean mothers. White mothers, however, are slightly more likely to give birth to sons after daughters are born. The magnitudes of these effects are of an order of magnitude different from that found for Indian mothers. While it is conceivable that these effects are the result of gender selection, the lack of other systematic evidence for white

Table 6—Boy-birth Regressions, Conditional on Previous Gender(s), California Birth Data

	2nd-child regression	3rd-child regression	
	Coefficient on first-born girl indicator variable	Coefficient on no-sons indicator variable	Coefficient on one-son indicator variable
Chinese	0.932**	2.165	0.807
	(0.424)	(1.354)	(1.287)
Indian	2.766**	11.256**	5.486*
	(0.761)	(3.059)	(3.002)
Japanese	0.215	−0.272	2.981
	(0.758)	(2.221)	(1.945)
Korean	0.887	0.115	−0.439
	(0.736)	(2.537)	(2.360)
White	0.187**	0.394*	0.532**
	(0.091)	(0.218)	(0.191)

Notes: Estimates are from linear regression models with boy birth as the dependent variable. The 2nd-child (3rd-child) regressions are for the sample of second (third) births to mothers of a given race between 1982 and 2005. Heteroskedasticity-robust standard errors are reported in parentheses. Estimates and standard errors have been multiplied by 100. The other regression covariates are mother's age, mother's age squared, birth year, and indicators for foreign-born mother, same-race father, and father's race missing.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

mothers would make such a conclusion unwarranted. Moreover, the largest effect on white male-birth likelihood is found after a gender mix (one son, one daughter), which is unlikely to be driven by gender selection. The most plausible explanation for this association is biological in nature; for instance, recent research (Henriette Savarre Nielsen et al. 2008) suggests that male births have small (but significant) negative effects upon future birth outcomes.

To focus on the previous-gender effects found for Indian mothers, Figures 4 and 5 provide time-series plots of the boy-birth likelihoods for second and third births, respectively, conditional on gender(s) of previous child(ren).[17] Figure 4 shows a consistently higher likelihood (since the late 1980s) of a second-birth boy when there is a first-born girl, although the time series shows evidence that this difference has narrowed since the mid-1990s. Throughout the time period shown, the percentage of sons born after first-born daughters averaged around 54 percent (SRB of 117). In Figure 5, the association of third-child boy births with previous gender mix is quite dramatic. The boy-birth percentage among Indian mothers with no boys (first two children were female) began to increase sharply after 1990, reaching a peak of around 68 percent (SRB of 213) in 1995 and 1996, and thereafter decreasing to a level of about 58 percent (SRB of 138). Also, note that the boy-birth percentages for one previous boy and two previous boys track each other fairly closely until 2000, where an increase in third-child boy-birth percentages is observed among Indian mothers with a boy-girl mix.

Using the linked nature of the California data, we are also able to construct a "termination-since-last-birth" indicator variable by comparing the number of previously

[17] Each point represents a seven-year moving average.

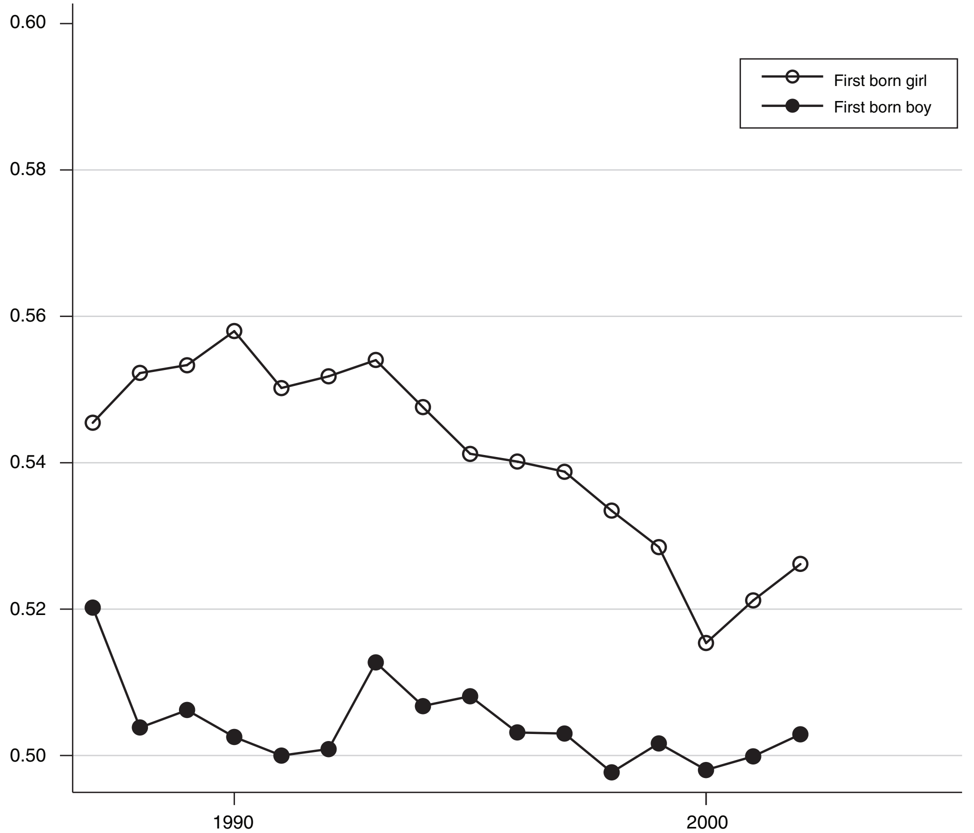

Figure 4. Boy-birth Likelihoods for Second Children of Indian Mothers in California

terminated pregnancies reported in two successive pregnancies. Specifically, the variable was defined to be equal to one if the number reported at the later pregnancy was larger than the number reported at the previous pregnancy, and zero otherwise. Although clearly still imperfect as an indicator of gender-selective practices (due to termination proxying for a difficult pregnancy), this variable is a better proxy for gender-selective practices since it focuses on the time period just before the birth in question.

With the constructed termination-since-last-pregnancy indicator as the dependent variable, Table 7 reports the conditional-upon-previous-gender regression results. We caution that these estimates are not direct evidence of gender-selective behavior since we have no measure of intent. For white births, there is a small, positive association (0.14 percentage points) between a first-born girl and a terminated pregnancy between the first and second birth. The overall percentage of white mothers that have a termination between their first and second pregnancy is just over 14 percent, so the 0.14 percentage-point difference represents only about a 1 percent difference relative to the baseline. For Indian second births, the estimated positive association is larger (0.97 percentage points (p-value = 0.04)). Based on the overall percentage (11.3 percent) of Indian mothers having a termination between their

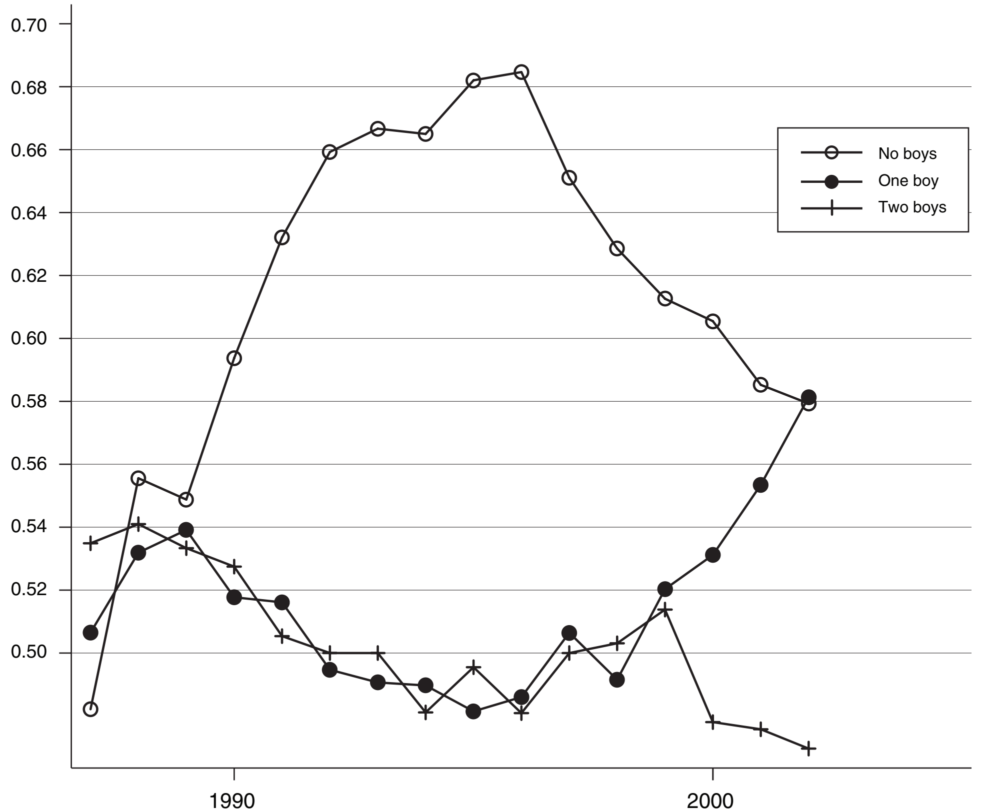

Figure 5. Boy-birth Likelihoods for Third Children of Indian Mothers in California

first and second pregnancies, this effect means that Indian mothers with a first-born daughter are nearly 10 percent more likely to have a terminated pregnancy prior to their second birth than Indian mothers with a first-born son.

For Indian third births, the estimated difference in the likelihood of a termination between the second and third births is 5.56 percentage points (s.e. 1.86) higher for mothers with no sons as compared to mothers with two sons. The magnitude of this difference is extremely large, relative to the overall likelihood (11.6 percent) of a terminated pregnancy between second and third births among Indian mothers. The unconditional (without control variables) percentages for Indian mothers with two previous daughters and two previous sons are 14.2 percent and 8.3 percent, respectively. This difference implies that Indian mothers with two previous daughters are 71 percent more likely to have a termination prior to their third birth than Indian mothers with two previous sons.

As a reality check on the results for Indian mothers, an anonymous referee suggested the following falsification exercise. For the third-birth results, if the dependent variable (an indicator of termination between second and third births) is replaced by an indicator of termination between first and second births, we should see no effect of gender. Indeed, this is what happens; a regression with the alternative termination

Table 7—Termination-Since-Last-Pregnancy Regressions, Conditional on Previous Gender(s), California Birth Data

	2nd-child regression	3rd-child regression	
	Coefficient on first born girl indicator variable	Coefficient on no-sons indicator variable	Coefficient on one-son indicator variable
Chinese	−0.121	0.923	0.374
	(0.262)	(0.836)	(0.784)
Indian	0.972**	5.559**	2.672
	(0.479)	(1.864)	(1.731)
Japanese	0.112	−2.464	−2.441*
	(0.516)	(1.564)	(1.392)
Korean	0.027	−0.858	−0.197
	(0.477)	(1.686)	(1.591)
White	0.140**	−0.136	0.196
	(0.064)	(0.155)	(0.136)

Notes: Estimates are from linear regression models with termination-since-last-pregnancy as the dependent variable. The 2nd-child (3rd-child) regressions are for the sample of second (third) births to mothers of a given race between 1982 and 2005. Heteroskedasticity-robust standard errors are reported in parentheses. Estimates and standard errors have been multiplied by 100. The other regression covariates are mother's age, mother's age squared, birth year, and indicators for foreign-born mother, same-race father, and father's race missing.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

indicator yields an estimate on the no-sons indicator of 2.94 (s.e. 1.86). Similarly, for the second-birth results, we replaced the dependent variable (an indicator of termination between first and second births) by an indicator of termination prior to the first birth. In this case, the regression coefficient on the first-born-daughter indicator variable was negative (−0.27) and insignificant (s.e. 0.47).

C. *A More Detailed Look at Chinese and Indian Subsamples*

The results of the previous sections are based on samples that pool together all mothers of a given race. In this section, we consider finer subsamples of the data for Chinese and Indian mothers. Specifically, we are interested in examining whether the birth-parity and conditional-upon-previous-gender effects are more prevalent among those with stronger cultural ties (specifically, births to parents of the same race) or depend on age or education.

Table 8 summarizes the results. The regression specifications are identical to those used in Tables 5 and 6. The first column of the table reports the original estimate (for the full sample). The remaining columns consider five different subsamples of births: same-race fathers, mothers younger than 30, mothers age 30 or older, mothers with high school education or below, and mothers with more than high school education.

For the same-race father subsamples, every estimate indicates a stronger effect of birth parity upon boy-birth likelihood. In the federal data, the fourth-child effect becomes around 3 percentage points for both Chinese and Indian same-race parents. The previously large birth-parity effects for Indian mothers in California

Table 8—Breakdown of Samples for Chinese and Indian Mothers

			Subsample with:				
		Original sample	Same-race father	Mother's age < 30	Mother's age ≥ 30	Mother HS education or less	Mother beyond HS education
Panel A: Federal data (1991–2004 results, Table 5)							
Chinese	2nd-child indicator	0.002	0.103	0.095	−0.146	0.090	−0.074
	3rd-child indicator	1.176**	1.476**	2.596**	0.596*	1.591**	0.805**
	4th-child indicator	2.250**	3.065**	1.526	2.338**	3.205**	1.148
Indian	2nd-child indicator	0.791**	0.927**	0.808**	0.723*	1.322**	0.497*
	3rd-child indicator	3.575**	4.068**	2.894**	3.896**	3.406**	3.468**
	4th-child indicator	2.481**	3.094**	2.236*	2.459**	3.503**	0.967
Panel B: California data (1991–2005 results, Table 5)							
Chinese	2nd-child indicator	−0.139	0.038	0.170	−0.316	−0.199	−0.157
	3rd-child indicator	0.750*	1.035**	1.225	0.573	0.479	0.875
	4th-child indicator	2.058**	2.753**	1.502	2.187**	3.887**	0.647
Indian	2nd-child indicator	0.990**	1.103**	1.217**	0.862	2.520**	0.665
	3rd-child indicator	6.658**	7.279**	4.742**	7.465**	5.459**	7.505**
	4th-child indicator	7.942**	9.387**	9.001**	7.596**	8.713**	7.013**
Panel C: California data (2nd-child results, Table 6)							
Chinese	First-born-girl indicator	0.943**	1.162**	1.462*	0.772	1.610*	0.751
Indian	First-born-girl indicator	2.766**	3.113**	2.417**	2.848**	2.836	2.714**
Panel D: California data (3rd-child results, Table 6)							
Chinese	No-sons indicator	2.165	2.638*	−4.735	3.195**	−1.195	3.142**
	One-son indicator	0.807	0.056	−4.763	1.661	−6.488**	2.980**
Indian	No-sons indicator	11.256**	13.659**	1.214	14.925**	15.927**	10.230**
	One-son indicator	5.486*	7.584**	−6.168	9.749**	4.579	6.187*

Note: Aside from choice of subsamples, the regression specifications are identical to those used in Tables 6 and 7.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

become even larger when looking only at births to Indian fathers (from 6.7 to 7.3 at third births and from 7.9 to 9.4 at fourth births). The effect of previous daughters upon boy-birth likelihood is also stronger for same-race Chinese and Indian parents. Among Chinese third births, the coefficient estimate on the no-sons indicator becomes significant at the 10 percent level (with a magnitude of 2.6 percentage points). For Indian third births, the magnitudes on both the no-sons and one-son indicator estimates increase, and the one-son estimate is now significant at a 5 percent level (relative to the two-sons category).

With respect to age, the birth-parity effects in the US and California data are evident for both younger (age less than 30) and older (age 30 or greater) mothers. The biggest distinction between younger and older mothers appears in the third-child conditional-upon-previous-gender California results (bottom panel of Table 8). Specifically, the no-sons indicator estimates are statistically significant for older Chinese and Indian mothers but insignificant for younger mothers. While this difference is consistent with higher opportunity costs later in a mother's fertility years, the difference in statistical significance may be driven by the relatively smaller sample size of young mothers (1,306 younger Chinese mothers and 582 younger Indian mothers, as compared with 7,971 and 1,383 older mothers, respectively).

Finally, the breakdown by education level yields mixed conclusions. The birth-parity effects for Chinese mothers are larger and more significant among mothers with 12 years or less of education. For Indian mothers, however, both education categories exhibit similar birth-parity effects, although the second-child effect seems to be more pronounced (and significant) among less-educated Indian mothers (2.5 percentage points in the California sample, relative to the overall estimate of 1 percentage point). The conditional-upon-previous-gender estimates for Indian mothers are quite similar in magnitude for less-educated and more-educated mothers. These results strongly suggest that the unusual boy-birth pattern for Indian mothers is not a phenomenon isolated among women with less education.

D. *Inferring the Prevalence of Gender Selection from Boy-birth percentages*

In this section, the following question is considered: If unusual boy-birth percentages are the result of gender-selective abortions, what does the observed boy-birth percentage imply about the prevalence of both gender determination and gender selection? We consider the case where the gender bias favors sons and gender-selective abortion is only chosen when the female gender is revealed.[18] Let p denote the "natural" probability of a boy birth. Let g denote the probability that a woman has a gender-determinative procedure (meaning that gender-selective abortion would be chosen if female gender is revealed).[19] Finally, let $\tilde{p}$ denote the boy-birth probability in the presence of gender selection. The probability $\tilde{p}$ is the quantity corresponding to the boy-birth percentage observed in the data. Note that $\tilde{p}$ is related to p and g as follows:

$$\tilde{p} = \frac{\Pr(\text{boy birth})}{\Pr(\text{live birth})} = \frac{p}{1 - g(1-p)}. \tag{1}$$

Equivalently, g can be written in terms of the probabilities p and $\tilde{p}$ as follows:

$$g = \frac{\tilde{p} - p}{\tilde{p}(1-p)}. \tag{2}$$

To infer anything about the prevalence of gender determination/selection, a value for the "natural" boy birth probability (p) is needed. A very conservative choice of p, based upon the first-boy-birth percentages reported in Table 4, is $p = 0.52$. For this value of p and realized boy-birth probabilities $(\tilde{p})$ ranging from 0.52 to 0.65, Figure 6 shows the implied probabilities of gender determination and gender-selective

[18] To the extent that the reverse is true for a subgroup of the population (daughter bias and gender-selective abortion only for males), the prevalence of gender determination/selection discussed below would be a lower bound on the actual prevalence.

[19] If all pregnant women had a gender-revealing ultrasound performed, g would represent the fraction of women who would have a gender-selective abortion if a female is revealed.

abortion. As an illustration, consider the boy-birth percentages for Indian births reported in Table 4. In the federal birth data, the fraction of boy births among third and fourth children is approximately 0.54. If this higher percentage is the result of gender selection, Figure 6 indicates that the probability of gender determination is approximately 8 percent.[20] For the 1991–2005 California estimates (57.5 percent boy-birth percentage for third children and 59.0 percent for fourth children), the implied gender-determination probabilities are much higher—about 18 percent for third births and 24 percent for fourth births.

How do these implied gender-selection probabilities relate to the number of implied abortions? As an example, again consider the 1991–2005 sample of California births to Indian women, for which there were a total of 7,102 third births and 1,428 fourth births. The implied gender-selection probabilities (18 percent and 24 percent) from above would correspond to roughly 850 abortions during this time period. If the unusual boy-birth percentages among Indian births are truly the result of gender-selective abortion, this represents a crude estimate of the number of "missing girls" within California between 1991 and 2005. Table 9 provides similar estimates of nationwide abortion numbers for third and fourth births to Chinese (1991–2004) and Indian (1992–2004) mothers. The table reports results for natural boy-birth probabilities (p) of 0.52 and 0.515. For the conservative $p = 0.52$ choice, the number of implied "missing girls" among 1991–2004 Chinese third and fourth births is just over 900. The estimate for 1992–2004 Indian third and fourth births is nearly 1,300. Overall, then, the boy-birth percentages at higher parity are consistent with more than 2,000 "missing" Chinese and Indian girls in the United States between 1991 and 2004.

E. *Census Data: Gender Preferences and Boy Births*

In this section, we briefly consider an analysis of the Census PUMS data as a complement to our birth data analysis. First, we consider the decision of families to have either a second or third child based on the gender(s) of their previous child(ren) and how this decision has changed over time. This fertility-stopping analysis is similar to that undertaken by Dahl and Moretti (2008), although they pool Asian races together in their results. Second, analogous to the analysis of the linked California birth data, we consider the likelihood of having a son conditional on the gender(s) of previous child(ren).[21]

Table 10 summarizes fertility-stopping behavior by race. Among families with at least one child, the table reports the percentage of families that had a second child within five years of the birth of the first child. Similarly, for every family with at least two children, the table reports the percentage of families that had a third child within five years of the birth of their second child. Results are provided for two time periods (1966–1979 and 1980–1994), with observations categorized by first- (or

[20] If p is taken to be 0.51, which is closer to the observed percentage of first-birth boys for Indian parents in the federal and California samples, the implied probability of gender determination would be higher.

[21] This analysis is a revised version of my 2005 working paper (Abrevaya 2005). Douglas Almond and Lena Edlund (2008) also show male-biased sex ratios following girls among Chinese, Koreans, and Asian Indians in the 2000 census data.

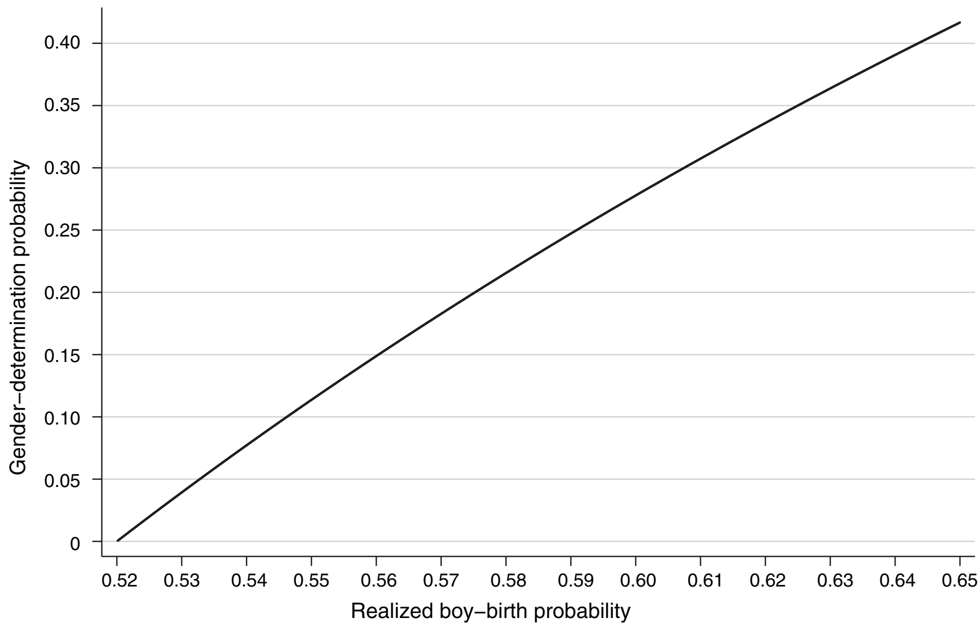

Figure 6. Implied Prevalence of Gender Determination (for p = 0.52)

second-) child birth year and previous gender mix ("girl"/"boy" for second births and "0 boys"/"1 boy"/"2 boys" for third births). In addition, the table reports the change in these percentages over the two time periods.

Overall, gender of the first child does not appear to play an important role in determining whether a family has a second child. In 1980–1994, there is slight evidence of son preference among Indian families (more likely to have a second child if the first was a girl) and daughter preference among Japanese families. A significant son preference in Chinese families is observed in the earlier 1966–1979 period. Across all races, there is a decrease in the likelihood of having a second child from the earlier to the later time period. The largest decreases are observed among Chinese and Korean families.

The third-child results highlight much larger gender-preference differences between races. For white and Japanese families, the overall preference is for a gender mix: families are most likely to have a third child if the gender of the previous two was the same (either two sons or two daughters) and least likely to have a third child if they have had a son and a daughter. For Chinese, Indian, and Korean families, there is a definite bias toward having a son:

- families with two daughters are far more likely (about 10–16 percentage points) to have a third child than families with one or two sons; and,

- families with two sons are about equally likely to have a third child as families with a son and a daughter, although the overall likelihood of having a third child drops in the later time period across all races.

Table 9—Number of Abortions Consistent with High Boy-birth Percentages

			If natural boy-birth probability = 52.0 percent:		If natural boy-birth probability = 51.5 percent:	
Sample	Fraction of boy births	Number of births	Implied probability of female fetus abortion	Implied number of abortions	Implied probability of female fetus abortion	Implied number of abortions
Chinese 3rd children (1991–2004, US)	0.530	36,018	3.8%	680	5.6%	1,019
Chinese 4th children (1991–2004, US)	0.540	6,802	7.4%	252	9.2%	315
Indian 3rd children (1992–2004, US)	0.544	25,172	8.8%	1,111	10.6%	1,342
Indian 4th children (1992–2004, US)	0.535	5,815	5.6%	163	7.4%	217
Indian 3rd children (1991–2005, California)	0.575	7,102	19.2%	679	20.8%	741
Indian 4th children (1991–2005, California)	0.590	1,428	23.8%	169	25.4%	182

The pattern of gender-mix preferences remains fairly similar across the two time periods for each race. The drop in fertility (with respect to third children) is far more pronounced among the Asian races than whites. As discussed earlier in this paper, smaller family sizes lead to increased opportunity costs of having a child of the less-preferred gender and, thus, to greater incentives for gender determination.

Table 11 provides a breakdown of boy-birth percentages from the census PUMS data based upon race, time period, and previous gender. The only statistically significant difference (at the 5 percent level) among second births is for Chinese families in 1980–1994 (53.4 percent chance of a son following a daughter versus 49.1 percent chance of a son following a son). Similar differences are found for Indian and Korean families in this later time period, but neither is significant at the 5 percent level due to the relatively small sample sizes. For third births in 1980–1994, the boy-birth percentages for Chinese, Indian, and Korean families are highest after two previous daughters (57.1 percent, 57.4 percent, and 57.1 percent, respectively). Statistically speaking, however, there is no compelling evidence that these percentages are significantly larger than the one-previous-son and two-previous-son percentages for any of these individual races. Pooling the three races together, as in Almond and Edlund (2008), would yield statistical significance. Given the small sample size issue, future research might focus on the 100 percent census sample in order to investigate child gender sequences within families. While these data would still be subject to the drawbacks (relative to birth data) discussed in Section I, such research would complement the analysis of the California linked data since the census data covers all states.

III. Conclusion and Discussion

This study has offered evidence consistent with gender selection at later births within the United States. For Chinese and Indian parents, the likelihood of having

Table 10—Fertility-Stopping for Second and Third Children, Census Data

Race	Time period	Fraction of families having a 2nd child if first child is a:		Fraction of families having a 3rd child given number of previous boys:			Number of families with at least one child	Number of families with at least two children
		Girl	Boy	0 boys	1 boy	2 boys		
Chinese	1966–1979	**0.664**	**0.622**	**0.383**	**0.296**	**0.261**	3,791	1,345
		(0.011)	(0.011)	(0.021)	(0.014)	(0.018)		
	1980–1994	0.544	0.537	**0.315**	0.173	0.201	8,265	3,698
		(0.008)	(0.008)	(0.018)	(0.010)	(0.014)		
Indian	1966–1979	0.649	0.627	0.382	0.219	0.272	2,548	816
		(0.014)	(0.013)	(0.026)	(0.015)	(0.025)		
	1980–1994	**0.626**	**0.587**	**0.333**	0.200	0.209	5,543	2,750
		(0.009)	(0.009)	(0.020)	(0.012)	(0.017)		
Japanese	1966–1979	0.630	0.625	0.312	**0.221**	0.332	2,869	1,056
		(0.013)	(0.013)	(0.024)	(0.015)	(0.024)		
	1980–1994	**0.604**	**0.639**	0.264	0.226	0.275	2,965	1,538
		(0.013)	(0.012)	(0.028)	(0.017)	(0.026)		
Korean	1966–1979	0.682	0.666	**0.346**	0.220	0.208	2,685	1,041
		(0.013)	(0.013)	(0.024)	(0.015)	(0.020)		
	1980–1994	0.624	0.598	**0.276**	**0.122**	**0.178**	4,585	2,438
		(0.010)	(0.010)	(0.021)	(0.011)	(0.017)		
White	1966–1979	**0.631**	**0.634**	**0.407**	**0.327**	**0.389**	659,456	254,305
		(0.001)	(0.001)	(0.002)	(0.001)	(0.002)		
	1980–1994	**0.616**	**0.621**	**0.371**	**0.294**	**0.363**	752,237	399,923
		(0.001)	(0.001)	(0.002)	(0.001)	(0.002)		

Notes: "Having a 2nd (3rd) child" means that the second (third) child is born within five years of the first (second) child. Standard errors are reported in parentheses. Bold indicates an estimate is significantly different (at the 5 percent level) from the other category in the second-child results or both of the other two categories in the third-child results.

a son is significantly higher for third-born and fourth-born children as compared to first-born children.[22] Controlling for maternal characteristics, prenatal-care variables, and time trends, the increase in boy birth likelihood explained by birth parity is extremely significant and of an order of magnitude larger than other determinants. On the other hand, slight evidence of birth-parity effects is found among Korean births (specifically, fourth births within California) and no evidence is found among Japanese births.

The evidence from the California birth data is particularly striking for Indian births between 1991 and 2005: third and fourth children are 6.7 and 7.9 percentage points more likely to be sons, respectively. Moreover, Indian mothers were significantly more likely to have a son and a terminated pregnancy since last birth if they had only daughters previously. For third births, Indian mothers with two daughters

[22] Although it is also possible that gender selection occurs among first-born children, the existing data do not support this conclusion. For Chinese births (see Table 7), there has been almost no change since 1971 in the boy-birth percentage among first-born and second-born children. Unfortunately, such a time-series comparison is infeasible for Indian and Korean births since data is not available prior to 1992 at the federal level and 1982 at the California level.

TABLE 11—BOY-BIRTH LIKELIHOODS CONDITIONAL ON PREVIOUS GENDER, CENSUS DATA

Race	Time period	Fraction of families having a second-born son if first child is a:		Fraction of families having a third-born son given number of previous boys:		
		Girl	Boy	0 boys	1 boy	2 boys
Chinese	1966–1979	0.526	0.521	0.522	0.511	0.510
		(0.014)	(0.014)	(0.043)	(0.033)	(0.050)
	1980–1994	**0.534**	**0.491**	0.571	**0.483**	0.560
		(0.011)	(0.010)	(0.030)	(0.027)	(0.035)
Indian	1966–1979	0.499	0.472	0.506	0.607	0.517
		(0.018)	(0.018)	(0.055)	(0.047)	(0.066)
	1980–1994	0.532	0.502	0.574	0.558	**0.457**
		(0.012)	(0.012)	(0.032)	(0.030)	(0.042)
Japanese	1966–1979	0.502	0.492	0.478	0.468	0.479
		(0.017)	(0.017)	(0.052)	(0.045)	(0.052)
	1980–1994	0.527	0.502	0.528	0.533	0.473
		(0.017)	(0.016)	(0.053)	(0.039)	(0.048)
Korean	1966–1979	0.526	0.516	0.460	0.441	0.521
		(0.017)	(0.017)	(0.050)	(0.041)	(0.060)
	1980–1994	0.526	0.500	0.571	**0.462**	0.563
		(0.013)	(0.013)	(0.040)	(0.042)	(0.049)
White	1966–1979	0.513	0.512	0.514	0.518	0.512
		(0.001)	(0.001)	(0.003)	(0.002)	(0.003)
	1980–1994	0.511	0.513	**0.504**	0.515	0.518
		(0.001)	(0.001)	(0.003)	(0.002)	(0.003)

Notes: Standard errors are reported in parentheses. Bold indicates an estimate is significantly different (at a 5 percent level) from the other category in the 2nd-child results or both of the other two categories in the 3rd-child results. Sample sizes are reported in Table 10.

were roughly 20 percent more likely to have a son than Indian mothers with two sons and 70 percent more likely to have a terminated pregnancy (in between the second and third birth).

The use of an extensive set of control variables in the boy-birth regression analyses rules out any simple biological explanations for the observed irregularities in boy-birth percentages. As such, gender selection stands out as the most logical explanation of the observed irregularities. This conclusion is further supported by the observed timing of the irregularities, concurrent with the increased availability of ultrasound and amniocentesis technologies. The third-birth and fourth-birth trends among Chinese and Indian mothers (Figure 3) match closely with the corresponding trends seen in China and India (Figure 1). Moreover, the trend among Indian mothers is extremely similar to that found in the United Kingdom by Dubuc and Coleman (2007).

The simple framework of Section IID suggests that the unusually high boy percentages among third- and fourth-born Indian children in California would be consistent with gender-determination rates of around 20 percent (i.e., 20 percent of female fetuses being aborted at these higher parities). Combined, the estimates for Chinese and Indian births (Table 9) are consistent with over 2,000 "missing girls" in the United States between 1991 and 2004.

Future research might focus on the underlying motives for gender selection within the United States. Common explanations for the trends in Asian countries, such as exogenously imposed child limits or extensive dowry systems, should not be relevant.[23] For Indian mothers, we found no evidence that the observed boy birth irregularities were isolated among less-educated (or more-educated) mothers. Gender-selection motives may simply stem from overriding cultural son biases that remain with immigrants who come to the United States. Since such son bias has been previously documented to vary over different regions in China and India, it would be interesting to relate the likelihood of male births within the United States to the specific regions from which Chinese and Indian mothers immigrated.

Overall, the empirical findings are in line with the gender preferences seen in the census data and the stronger incentives for gender selection that arise at later births. For Chinese, Indian, and Korean families, the census data indicate a strong son bias in the decision to have a third child, with a much higher likelihood of having a third child among families with two daughters. In contrast, the third-child outcomes from the census data indicate a preference for a gender mix among white and Japanese families. Despite the gender-mix preference that appears in the fertility decisions for these races, the empirical results do not suggest that gender selection is being used to achieve a gender mix. For example, the aggregate birth-parity effects for white parents (estimated in Section IIA) do not change much from the 1971–1980 period to later time periods.

Several factors could lead to an increase in the prevalence of gender selection within the United States. First, if the declining trend in family size continues, there would be increased incentives (holding gender preferences fixed) for gender selection. Second, introduction of technologies that can reliably and safely detect gender at an earlier stage in pregnancy (than amniocentesis or ultrasound) would reduce the "cost" of abortion by allowing women to have early-term rather than late-term abortions. Third, the availability of improved preconceptive gender-selective technologies at lower costs will tend to increase the prevalence of gender selection.[24] Most importantly, a preconceptive gender selection method would entirely eliminate the need for a gender-based abortion, which involves prohibitive costs (including moral and ethical costs) for most parents.

Although the predominant gender-mix preference in the United States is not likely to change much in the near future, it is possible that the son bias observed among some of the Asian races (Chinese, Indian, and Korean) could diminish. Such a change could occur for a variety of reasons, including reduced cultural bias toward sons and increasing proportions of second- and third-generation Asian mothers in the United States.

Given that the predominant preference within the United States is for a gender mix, an increase in gender selection would not lead to a gender-imbalance problem in

[23] It is unclear how prevalent dowries are within the United States, as we could not find any evidence on this point.

[24] The CDC (2004) documented the increased use of "assisted reproductive technology" (defined as fertility treatments involving both sperm and eggs, predominantly IVF). The number of live-birth deliveries using this technology increased steadily from 14,507 in 1996, to 33,141 in 2002 (roughly 1 percent of live births in the United States).

the aggregate. Such a gender imbalance could, however, arise among subpopulations with a bias toward sons or daughters. The effect on family size would be ambiguous: although families could achieve gender mix with fewer children, some families would be willing to have additional children if they could choose gender. Given that gender-selective procedures are not currently banned in the United States, the most predictable effect of increased gender selection would be the ensuing debate on the surrounding moral and ethical issues and potentially the fight over regulation.[25]

Appendix A: Construction of the Maternally Linked California Birth Data

The CDHS provided data for every birth that occurred in California between 1970 and 2005. The total number of birth records during the 36-year period was 16,932,031. In addition to the publicly available data, the author was provided with data on mother's first name, mother's maiden name (surname), and mother's date of birth. The first name and birth date items were available for all births after 1981 and 1988, respectively. A full name for each mother was created by concatenating the first name and maiden name together (with a space in between). Any records that had missing values for mother's name, mother's age, mother's birth date (for births after 1988), or total number of previous live births were dropped, leaving 16,799,227 observations.

For any two births in the sample, the pair of births is considered a potential match if all of the following conditions are met:

- An exact match on mother's full name (or mother's maiden name if one of the births occurred before 1982).
- An exact match between the month and year of the earlier birth and the month-of-last-birth and year-of-last-birth reported at the later birth.
- Consistency of the total-previous-live-births variable (meaning an increase of one from the earlier birth to the later birth).
- Consistency of mother's age information, meaning:
    (a) if both births occurred after 1988, an exact match on mother's birthdate.
    (b) if at least one birth occurred between 1970 and 1988, the reported difference between the mother's age at the earlier birth and her age/birthdate at the later birth was possible given the number of months between the two births.

After all potential matches are recorded, a pair of births is then considered an actual match if

- the earlier birth is not a potential match with any other later births, and

[25] The President's Council on Bioethics considered some of these issues at its October 2002 meeting. Full transcripts are available at http://www.bioethics.gov/transcripts/oct02/index.html.

- the later birth is not a potential match with any other earlier births.

To link more than two births for a given mother together, additional linkages are made based on the actual matches of the birth pairs. For instance, suppose that three births are denoted A, B, and C, in chronological order. If both pairs A-B and B-C represent actual matches, then the birth sequence A-B-C would be linked together. Additional births could be added to this sequence if A is an actual match with an earlier birth or if C is an actual match with a later birth. This process is continued until all matched birth sequences are constructed.

The matching algorithm resulted in 9,821,455 births (58.0 percent of the total) being part of a matched birth sequence. The remainder of the births consisted of

- only children,
- births that could not be uniquely matched together,
- births that could not be matched due to the mother's other births not being in the sample (e.g., because they occurred before 1970 or outside of California), or
- births that could not be matched due to coding errors (e.g., misspelled name or incorrect age).

Table A1 provides a racial breakdown of the birth sequences used in the analysis, reporting the number of mothers for whom the first two (three) births are observed and the second (third) birth occurs between 1982 and 2005. The first column corresponds to the sample sizes for analysis that conditions on the gender of the first child, whereas the second column corresponds to the sample sizes for analysis that conditions on the gender mix of the first two children. Since race itself is not used to maternally link the data, we were able to link post-1982 births of Indian and Korean mothers to pre-1982 births of these same mothers. The linked data regressions (Tables 6 and 7) use only observations from 1982–2005 to avoid underrepresentation of those mothers who stopped having children before 1982 (when their race would have been identified).

Appendix B: Details on 5 Percent PUMS Census Data Analysis

The 1980, 1990, and 2000 editions of the 5 percent PUMS census data were used. The racial category was determined by the reported race of the mother. In 2000, the census questionnaire allowed respondents to also indicate "secondary" racial categories. For the 2000 sample, the categorization was based upon the primary racial category reported for the mother.

In order to condition upon gender of first child or first two children, it is necessary to identify mothers for whom first child information is available. Although the 1980 and 1990 data contain an item related to a mother's fertility (specifically, the "number of children ever born"), we decided to use the same method of family construction for each of the three samples. Specifically, a family was only retained

Table A1—Sample Sizes by Race, California Data

Race	Number of mothers with first two births observed	Number of mothers with first three births observed
Chinese	60,391	10,888
Indian	20,886	3,107
Japanese	19,690	4,585
Korean	20,174	3,257
White	1,277,114	421,561

in the sample if the oldest child in the household was 13 years of age or younger. This choice would misclassify birth order for families with older children that have left the household, but the cutoff of 13 was chosen to minimize this possibility. Other cutoff choices yielded extremely similar results, although choosing a lower cutoff reduces the sample size available for analysis. We dropped any families for which the sex or age of any child was "allocated" in the data.

Each child's age (in years) is reported in the census data, taking on values between 0 and 17. The birth year of a child was calculated by subtracting the reported age (plus one) from the census year. This birth year is used to categorize families into the time periods in Tables 10 and 11 (based on first child's birth year and second child's birth year, respectively). Table 10 reports the likelihood of having an additional (second or third) child within five years of the previous child. For the second-child outcomes, the families considered are those whose oldest child is at least five years of age. Similarly, for the third-child outcomes, the families considered are those whose second oldest child is at least five years of age. A family is recorded as "having an additional child" if the difference in ages between the previous child and the "additional child" is less than or equal to five years. Finally, the earliest birth year considered is 1966, which corresponds to 13-year-old children from the 1980 sample, and the latest birth year considered is 1994, which corresponds to 5-year-old children from the 2000 sample.

Appendix C: Detailed Boy-birth Regression Results for White Mothers

In the interest of space, coefficient estimates for the boy birth regressions in Table 5 were reported only for the birth-parity indicator variables. To show the association of male births with other observable maternal and pregnancy related variables, we provide the complete set of estimates (Table A2) for white mothers in the federal data between 1991 and 2004. The sample size is huge (over 30 million births), which allows for precise estimation of the effects. Mother's education has a positive association with male births (0.03 percentage points per year of education). Mothers with first-trimester initial prenatal visits are least likely to have sons, holding all else fixed, which indicates that the other prenatal visit categories largely proxy for problem-free pregnancies. Finally, the previous termination, ultrasound, and amniocentesis indicator variables all have the expected signs. Mothers with a previous terminated pregnancy or an amniocentesis are less likely to have boys (0.16 percentage points for

Table A2—Boy-birth Regression for White Mothers, US Data, 1991–2004

	Coefficient estimate (s.e.)
2nd child	−0.0666**
	(0.0212)
3rd child	−0.1444**
	(0.0278)
4th child	−0.2037**
	(0.0432)
Birth year	−0.0034
	(0.0024)
Foreign-born mother	0.0809**
	(0.0406)
Same-race father	0.4128**
	(0.0538)
Father's race missing	0.0363
	(0.0610)
Mother's education	0.0303**
	(0.0047)
No prenatal care	0.2381**
	(0.1209)
2nd-trimester initial visit	0.5172**
	(0.0310)
3rd-trimester initial visit	0.3574**
	(0.0706)
Previous terminated pregnancy	−0.1642**
	(0.0211)
Ultrasound during pregnancy	0.0704**
	(0.0196)
Amniocentesis during pregnancy	−0.5046**
	(0.0529)
Age dummies?	Yes
Observations	30,723,930

** Significant at the 5 percent level.
* Significant at the 10 percent level.

termination, 0.50 percentage points for amniocentesis), as these indicators proxy for pregnancy problems. In contrast, mothers who have an ultrasound are 0.07 percentage points more likely to have a boy.

REFERENCES

Abrevaya, Jason. 2005. "Are There Missing Girls in the United States? Evidence on Gender Preference and Gender Selection." Unpublished.

Allan, Bruce B., Rollin Brant, Judy E. Seidel, and John F. Jarrell. 1997. "Declining Sex Ratios in Canada." *Canadian Medication Association Journal*, 156(1): 37–41.

Almond, Douglas, and Lena Edlund. 2008. "Son-biased Sex Ratios in the 2000 United States Census." *Proceedings of the National Academy of Sciences*, 105(15): 5681–82.

Bankole, Akinrinola, Susheela Singh, and Taylor Haas. 1998. "Reasons Why Women Have Induced Abortions: Evidence from 27 Countries." *International Family Planning Perspectives*, 24(3): 117–52.

Centers for Disease Control and Prevention. 2001. *Vital Statistics of the United States, 1999 Volume I, Natality*. Hyattsville, MD.

Centers for Disease Control and Prevention. 2003. "Abortion Surveillance—United States, 2000." *Morbidity and Mortality Weekly Report*, 52(SS–12).

Centers for Disease Control and Prevention. 2004. *2002 Assisted Reproductive Technology Success Rates: National Summary and Fertility Clinic Reports*. Atlanta, GA.

Dahl, Gordon B., and Enrico Moretti. 2008. "The Demand for Sons." *Review of Economic Studies*, 75(4): 1085–1120

Davis, Devra Lee, Michelle B. Gottlieb, and Julie R. Stampnitzky. 1998. "Reduced Ratio of Male to Female Births in Several Industrial Countries: A Sentinel Health Indicator?" *Journal of the American Medical Association*, 279(13): 1018–23.

Dubuc, Sylvie, and David Coleman. 2007. "An Increase in the Sex Ratio of Births to India-born Mothers in England and Wales: Evidence for Sex-Selective Abortion." *Population and Development Review*, 33(2): 383–400.

Ebenstein, Avraham Y. 2007. "Fertility Choices and Sex Selection in Asia: Analysis and Policy." Unpublished.

Gu, Baochang, and Yongping Li. 1994. "Sex Ratio at Birth and Son Preference in China." Paper presented at the UNFPA/KIHASA International Symposium on Issues Related to Sex Preference for Children in the Rapidly Changing Demographic Dynamics in Asia, Seoul, Republic of Korea.

Henshaw, Stanley K., and Lawrence B. Finer. 2003. "The Accessibility of Abortion Services in the United States, 2001." *Perspectives on Sexual and Reproductive Health*, 35(1): 16–24.

Institute for Social Development Studies. 2007. "New 'Common Sense': Family-Planning Policy and Sex Ratio in Vietnam." 4th Asia Pacific Conference on Reproductive and Sexual Health and Rights, Hyderabad, India.

Jha, Prabhat, Rajesh Kumar, Priya Vasa, Neeraj Dhingra, Deva Thiruchelvam, and Rahim Moineddin. 2006. "Low Male-to-Female Sex Ratio of Children Born in India: National Survey of 1.1 Million Households." *Lancet*, 367: 211–18.

Johnson, Lawrence A., Glenn R. Welch, Keyvan Keyvanfar, Andrew Dorfmann, Edward F. Fugger, and Joseph D. Schulman. 1993. "Preimplantation Diagnosis: Gender Preselection in Humans? Flow Cytometric Separation of X and Y Spermatozoa for the Prevention of X-linked Diseases." *Human Reproduction*, 8(10): 1733–39.

Joyce, Ted, Robert Kaestner, Sanders Korenman, and Stanley Henshaw. 2005. "Family Cap Provisions and Changes in Births and Abortions." *Population Research and Policy Review*, 23(5): 475–511.

Kim, Jinyoung. 2005. "Sex Selection and Fertility in a Dynamic Model of Conception and Abortion." *Journal of Population Economics*, 18(1): 41–67.

Klasen, Stephan, and Claudia Wink. 2002. "A Turning Point in Gender Bias in Mortality? An Update on the Number of Missing Women." *Population and Development Review*, 28(2): 285–312.

Klasen, Stephan, and Claudia Wink. 2003. "'Missing Women': Revisiting the Debate." *Feminist Economics*, 9(2–3): 263–99.

Lhila, Aparna, and Kosali Simon. 2008. "Prenatal Health Investment Decisions: Does the Child's Sex Matter?" *Demography*, 45(4): 885–905

Lundberg, Shelly, and Elaina Rose. 2003. "Child Gender and the Transition to Marriage." *Demography*, 40(2): 333–49.

Marcus, Michele, John Kiely, Fujie Xu, Michael McGeehin, Richard Jackson, and Tom Sinks. 1998. "Changing Sex Ratio in the United States, 1969–1995." *Fertility and Sterility*, 70(2): 270–73.

Mizuno, Reiko. 2000. "The Male/Female Ratio of Fetal Deaths and Births in Japan." *Lancet*, 356(9231): 738–39.

Møller, Henrik. 1996. "Change in Male:Female Ratio among Newborn Infants in Denmark." *Lancet*, 348(9030): 828–29.

Nielsen, Henriette Svarre, Laust Mortensen, Ulrikka Nygaard, Ole Schnor, Ole Bjarne Christiansen, and Anne-Marie Nybo Andersen. 2008. "Brothers and Reduction of the Birth Weight of Later-born Siblings." *American Journal of Epidemiology*, 167(4): 480–84.

Office of the Registrar General of India. 2001. *Census of India: 2001*. http://www.censusindia.gov.in/ (accessed December 14, 2004).

Park, Chai Bin, and Nam-Hoon Cho. 1995. "Consequences of Son Preference in a Low-Fertility Society: Imbalance of the Sex Ratio at Birth in Korea." *Population and Development Review*, 21(1): 59–84.

Perls, Thomas T., and Ruth C. Fretts. 1998. "Why Women Live Longer than Men." *Scientific American: Women's Health*, 100–103.

Poston, Dudley L., Jr., and Karen S. Glover. 2006. "China's Demographic Destiny: Marriage Market Implications for the Twenty-first Century." In *Fertility, Family Planning, and Population Policy in China*, ed. Dudley L. Poston, Jr., Che-Fu Lee, Chiung-Fang Chang, Sherry L. McKibben, and Carol S. Walther, 172–86. Abingdon, UK: Routledge Publishers.

Poston, Dudley L., Jr., Baochang Gu, Peihang Liu, and Terra McDaniel. 1997. "Son Preference and the Sex Ratio at Birth in China: A Provincial Level Analysis." *Social Biology*, 44: 55–76.

Poston, Dudley L., Jr., Julie Luan Wu, and Han Gon Kim. 2003. "Patterns and Variation in the Sex Ratio at Birth in the Republic of Korea." *Development and Society*, 32: 47–60.

Retherford, Robert D., and T. K. Roy. 2003. "Factors Affecting Sex-selective Abortion in India and 17 Major States." National Family Health Survey Subject Reports, no. 21.

Robertson, John A. 2001. "Preconception Gender Selection." *American Journal of Bioethics*, 1(1): 2–9.

Sachs, Susan. 2001. "Clinics' Pitch to Indian Émigrés: It's a Boy." *New York Times*, August 15.

Sen, Amartya. 1990. "More than 100 Million Women are Missing." *New York Review of Books*, 37(20): 61–66.

Sen, Amartya. 1992. "Missing Women." *British Medical Journal*, 304: 587–88.

Sen, Amartya. 2001. "The Many Faces of Gender Inequality." The New Republic, 9: 35–40.

Torres, Aida, and Jacqueline Darroch Forrest. 1988. "Why Do Women Have Abortions?" *Family Planning Perspectives*, 20(4): 169–76.

Trivers, Robert L., and Dan E. Willard. 1973. "Natural Selection of Parental Ability to Vary the Sex Ratio of Offspring." *Science*, 179(4068): 90–92.

van der Pal-de Bruin, K. M., S. P. Verloove-Vanhorick, and N. Roeleveld. 1997. "Change in Male:Female Ratio among Newborn Babies in Netherlands." *Lancet*, 349(9044): 62.