Spring 2019

Missing girls in the United States: Literature review and some analysis

Bryan Jackson

Missing Girls in the United States: Literature Review and Some Analysis

Bryan Jackson

Table of Contents

Introduction

The topic of the missing girls phenomenon in many countries such as India, China and South Korea is well known, and was noted first and most famously by Amartya Sen (Sen, 1990). Following the introduction of ultrasound and amniocentesis technology in the 1970s, boy-girl birth ratios in these countries rose dramatically, and the consensus of interested researchers is that this increase was due to gender-selective abortion (Das Gupta, 2005). While research on this topic is mostly confined to nations other than the United States, there exists a small literature on the topic of missing girls within the US. The goal of this paper is to provide a brief summary of the literature on the missing girls phenomenon as it pertains to nations other than the US, to provide a comprehensive review of the literature on the subject as it pertains to the US, and finally to extend an empirical study by Jason Abrevaya on missing girls in the US by utilizing recently available data.

Section 1: Global Perspective

There is a wealth of literature on the missing girls phenomenon in a variety of countries, primarily in east and southeast Asia (Das Gupta, 2005). Because the analysis in this paper will focus primarily on the Chinese-, Korean-, and Indian-American communities, we will focus our attention mostly on China, Korea and India for the purpose of discussing the missing girls phenomenon outside the United States.

1.1: China

In China, a noticeable increase in the boy-girl birth ratio became apparent starting in the early 1980s according to Chinese Census data, with the number of boys born per one hundred girls born increasing from 108.5 in 1982 to 116.9 in 2000, the highest recorded level in the world at the time (Scharping, 2003). This was following the legalization of abortion in 1956 and the introduction of the One Child Policy in 1979, following which abortion became an essential component of family planning policy. While the One Child Policy is often mentioned in relation to the missing women phenomenon in China, literature on the subject pays special heed to the implementation of ultrasound technology from approximately the period 1979-1995 as a causal factor (Ebenstein, 2010). The missing women of China are expected to comprise 51% of the world's total by the year 2050 (Guilmoto & Bongaarts, 2015).

Because the introduction of ultrasound technology in China happened at different times in different areas, it is possible to conduct a panel data analysis of the various counties of China to determine the effect of the introduction of ultrasound technology on the boy-girl birth ratio, and therefore to infer more carefully the effect of fetal gender-determining technology on the woman's choice of whether to abort a female fetus. Cheng, Li, and Meng (2013) study this phenomenon using a nationally representative dataset with microdata describing the birth history of about 560,000 women in China; they also use a dataset which describes the year in which individual counties in China introduced ultrasound technology. They provide the following model:

$$\mathrm{Boy}_{\mathrm{ict}} = \beta_1(1st + ultrasound_{ct}) + \beta_2(2nd * ultrasound_{ct}) + \beta_3(3rd^{+} * ultrasound_{ct}) + \beta_4 * 2nd + \beta_5 * 3rd^{+} + X_{ict} * \gamma + \mu_c + \nu_t + \mu_c * t + \epsilon_{ict}$$

Where *i* indexes individuals, *c* indexes counties, and *t* indexes years. Boy_{ict} is a binary variable equal to 1 if the child is male and 0 if the child is female; 1st, 2nd, and 3rd+ are binary variables equal to 1 if the birth parity of the child is one, two, or higher, respectively, and 0 otherwise; $ultrasound_{ct}$ is a binary variable equal to 1 if ultrasound technology has already been introduced in county *c* in year *t* and 0 otherwise; X_{ict} is a vector of individual-specific control variables, such as maternal age, ethnicity, and pre-natal care; μ_c is a vector of variables equal to 1 if the birth was in county *c* and 0 otherwise; ν_t is a vector of dummy variables equal to 1 if the birth was in year *t* and 0 otherwise; and $\mu_c * t$ is a vector of county-specific linear time trends (Cheng, Li, & Meng, 2013).

Using a variety of robustness checks, Cheng, Li, and Meng find that access to ultrasound technology in a county is strongly related to an increased boy-birth ratio in a county. Further, they find that this effect is especially strong in counties where the One Child Policy is enforced most stringently. They also find that the effect of ultrasound technology on the sex ratio becomes markedly stronger at later birth parities, especially when at earlier birth parities the mother gave birth to girls.

1.2: South Korea

In South Korea, the boy-girl birth ratio began to increase dramatically in the year 1985, increasing from about 108 boys born per 100 girls in 1980 to 112 boys born per 100 girls in

1990 (Park & Cho, 1995), indicating a son preference in South Korea. This son preference has a long history in South Korea, originating from the Choson Dynasty (reigned from 1392 to 1910) which instituted a patriarchal system based on Confucian values. The Confucian system in South Korea instituted a strictly patriarchal system, whereby property ownership and influential social contacts were exclusively limited to men. This strongly male-oriented societal structure created a powerful social pressure for families to have a boy to continue the family lineage and provide for the parents in their old age (Chung & Das Gupta, 2007).

The strong desire to have a male child became particularly pertinent in South Korea in the 1980s when the birthrate declined precipitously. Since parents had the ability to determine the sex of their baby before its birth, the ingrained desire for a male child meant that parents had an incentive to choose to abort a fetus if they had not had a male child yet and they did not plan on having any additional children (Chung & Das Gupta, 2007).

The extent of son preference in South Korea, and the effect of son preference on actual fertility rates, is easier to determine than in the case of China since the South Korean government has conducted surveys in which women were asked about both their fertility preferences and the actual sex of their offspring. The Korea Institute for Health and Social Affairs published the 1991 National Fertility and Family Health Survey, which surveyed 7,000 currently or previously married women aged 15-49 about the sex of each of their children and whether they felt that they “must have a son,” among many other questions, contains some very revealing information about both the effect and temporal trajectory of son preference in South Korea.

The survey found that women born in the time period 1955-1964 had a 35.0 percent chance of stating they “must have a son,” versus only 27.1 percent of women born in the time period 1965-1974, indicating a negative son preference time trajectory. Among those women who stated they “must have a son,” the male-female birth ratio of children born at the second birth parity or later was 194.1 males per 100 females, while for those who replied negatively to the question of strong son preference the same ratio was 105 males per 100 females. These numbers indicate a very powerful effect of son preference on actual birth ratios. The apparent decrease over time in stated son preference, along with the strong effect of son preference on birth ratios, have been used to argue that the decrease in the boy birth ratio in South Korea which occurred in the late 1990s and early 2000s was the consequence of changing attitudes in South Korean society toward son preference (Chung & Das Gupta, 2007).

1.3: India

In India, the boy-girl birth ratio increased from 1039 boys born per girl in 1981 to 1079 boys born per girl in 2001 (Office of the Registrar General & Census Commissioner, India, 2002); this increase occurred in spite of a 1994 law which outlawed sex-selective abortion, a law which has often been ignored (Jha P, 2006). This increase in the boy-girl birth ratio was also more pronounced in urban India (boy-girl birth ratio increased from 1042 in 1981 to 1104 in 2001) than in rural India (birth ratio increased from 1038 in 1981 to 1071 in 2001). By contrast, among third-parity births where the mothers’ previous two births were girls, the birth ratio was 1391 boys per 1000 girls in 2001 (Office of the Registrar General & Census Commissioner, India, 2002).

According to one study, sex selection in India has caused at least half a million abortions in the period from 1981-2001 (Jha P, 2006). The missing women of India are estimated to comprise about 5% of the actual Indian population by the year 2050. The missing women of India are projected to comprise 28% of the total missing women in the world by the year 2050 (Guilmoto & Bongaarts, 2015).

Section 2: United States

Overall, the boy-girl birth ratio in the United States has remained near the biological norm (Almond & Edlund, 2008) of about 1050 boys born per 1000 girls for at least the past seventy years, since before abortion was universally legal and since before sexual determination techniques became sophisticated enough to be reliable and widely available. On the surface, this fact would seem to indicate that the missing girls phenomenon is not prevalent within the United States. However, some research has been done to attempt to disaggregate birth data in order to isolate racial groups which correspond to diaspora communities from countries with high boy-girl birth ratios.

Apparently, the first article written about this issue was by Jason Abrevaya, who looked specifically at Chinese-, Indian- and Korean-American births in the United States to determine whether there was evidence for the existence of the missing girls phenomenon within these communities. Abrevaya produced an early paper on this issue in 2005 which remains unpublished; Almond and Edlund (2008) followed up on Abrevaya's analysis with a brief analysis which focused primarily on issues originally raised by Abrevaya. In 2009, Abrevaya

published a pruned and updated version of his original paper in the American Economic Journal. This section of this paper will analyze Abrevaya's methods and arguments from his 2009 paper, along with those of some of the (few) other academics who have written about the missing girls issue as it pertains to the United States.

Prior to Abrevaya's analysis, some work was done on the issue of son preference in the United States, although these earlier studies did not focus on sex-selective abortion. Lundberg and Rose (2003), using a longitudinal study of families (the Panel Study of Income Dynamics [PSID]), found that unmarried couples are more likely to marry following the birth of a child if that child is a boy. Furthermore, the PSID has been used to show that American fathers spend more time with their sons than with their daughters (Yeung, Sandberg, Davis-Kean, & Hofferth, 2001).

Dahl and Moretti (2004) present an econometric study of US births using US census data and the Current Population Survey (CPS) Fertility Supplement. They find that while some circumstantial evidence exists to suggest son preference, such as that families with sons are more likely to remain married than families without sons and that families which have a first-born son have significantly lower fertility than families with a first-born daughter, no evidence exists to suggest the existence of the missing girls phenomenon in the United States.

One problem with Dahl and Moretti's analysis, however, is that Census data is limited to a 5% sample, which given the already small number of births among certain Asian communities in the US makes the much-reduced sample size difficult to use when trying to differentiate genuinely anomalous boy-girl birth proportions from random statistical noise. In the 2000

PUMS census data, for example, there are fewer than two thousand observations from Indian-American mothers (Almond & Edlund, 2008). The issue of small sample sizes becomes particularly pertinent when considering subsets of women; for example, the numbers of women in these three groups (the Chinese-American, Korean-American, and Indian-American) from the census data who birthed at least two girls prior to having a third child is vanishingly small.

Abrevaya (2009) attempts to discern the effect of a mother's race on the probability that the mother's child will be a boy. To do this, he uses three sources of data:

1. Federal natality data (years 1971-2004) from the Centers for Disease Control (CDC), which is acquired from parents during the birth certification process after a child is delivered in a US hospital. This is population level data and includes information about, inter alia: mother's race, father's race, child's gender, child's birth parity, state of birth, and prenatal care. According to Lhila and Simon (2008), prenatal care factors such as smoking and number of prenatal care visits, which are measured in federal birth data, have no measurable impact on the gender of a baby. Siblings cannot be linked using this data. All three racial groups under consideration are included as a racial classification in this data.
2. PUMS 5% census data (years 1980, 1990, 2000), which is used as supporting, rather than primary, evidence due to its relatively small number of observations from Asian mothers. Data regarding race can be disaggregated based on whether the mother belongs to one of the three racial categories under consideration.

3. California birth data (years 1970-2005) from the California Department of Health Services (CDHS). Abrevaya obtained permission from the state of California to use this data, which is population level birth data from the state of California. This data contains a greater number of variables than federal natality data. Siblings may be linked using this data. Data can be disaggregated according to race and each of the three racial categories under consideration is included as a classification.

Abrevaya notes that about 90% of Chinese-American mothers and about 95% of Korean-American and Indian-American mothers in the federal natality data were born outside of the United States; and approximately 70%, 80%, and 90%, respectively, of Chinese-American-, Korean-American-, and Indian-American-mothered children were fathered by members of their mother's race. These figures make it probable that there is some cultural continuity among these groups from their countries of origin, making fetal sexual selection a reasonable possibility for mothers from these groups.

When presenting statistics from the data, Abrevaya alternatively uses babies born to white American, black American, or Japanese-American mothers as a point of reference (Japanese-Americans were used largely because Japan is an example of an Asian society where the missing girls phenomenon does not appear to exist). Focusing on federal natality data for the moment: For black, white, and Japanese mothers, Abrevaya notes that the boy-birth ratio remains virtually constant for each birth parity. For Indian, Korean, and Chinese mothers, however, there is a much different pattern. Indian mothers have a consistent boy-birth ratio among all birth parities up until the mid-1980s, after which time first and second parity boy birth ratios remained constant

at about 1040 boys born per 1000 girls but third and fourth boy birth ratios increased to about 1160 boys born per 1000 girls on average in the years from 1985 to 2007.

The situation with Korean-American mothers is very similar to that of Indian-American mothers up until the mid-1980s: first and second parity boy birth ratios are essentially the same as third and fourth parity boy birth ratios. However, following 1985, the third and fourth parity boy birth ratio for Korean-American mothers increased to where it was approximately three percent higher than the first and second parity boy birth ratio until about 1995, at which point the difference in boy birth ratios between first and second birth parities, and third and fourth birth parities declined significantly. This situation mirrors findings from a study by Chung and Das Gupta (2007) which studies births to families in Korea and finds that the boy-birth ratio in Korea declined significantly during the 1990s due to changing attitudes in Korean society toward the necessity of giving birth to a boy.

The situation with Chinese-American mothers is very similar to Korean-American mothers. Until the mid-1980s, there was very little difference between first and second birth parity boy birth ratios and third and fourth parity boy birth ratios among Chinese-American mothers. Following 1985, the third and fourth birth parity boy birth ratios were 2-3% higher than the first and second parity boy birth ratios and remained that way for about ten years, at which point the difference declined until it was negligible by the mid-2000s.

Another interesting finding from the federal natality data is the effect of gender in early birth parities on whether a mother has an additional child. If mothers are more likely to have an additional child when their earlier children were all girls, one possible interpretation is that families are "trying again" in the hopes of giving birth to a boy rather than a girl.

Among white women in the United States, there was no significant and consistent change in the boy-birth ratio in the third and fourth birth parities compared to the boy-birth ratio in the first and second birth parities during the period within Abrevaya's analysis. For example, among white mothers the percentage of boys born for the fourth birth parity or higher was 51.1%; the percentage of boys born for the first birth parity was 51.4%, a very small effect giving no evidence of a preference for boys over girls. The same result broadly holds true for black American mothers.

There are some fluctuations in the percentage of boys born in the third and fourth birth parities to Japanese-American mothers, but Abrevaya notes that these fluctuations are inconsistent and follow no discernible pattern (in some years, the percentage of boys born at third and fourth birth parities is higher than the percentage of boys born at first and second birth parities; in some years it is lower). Abrevaya suggests that this fluctuation is probably related to the relatively small number of babies born to Japanese-American mothers compared to the number born to white American and black American mothers; this difference may lead to statistically interesting but ultimately meaningless fluctuations.

Among Indian American women, 53.5% of children born in the fourth birth parity or higher were boys compared to just 51.0% in the first birth parity. Among Chinese women, 54.0% of children born in the fourth birth parity were boys compared to only 51.8% in the first birth parity. Among Korean women, 52.9% of children in the fourth birth parity were boys compared to 51.9% of children in the first birth parity.

Abrevaya notes some interesting differences between findings in the federal natality data and the findings in the California birth data. For example, third parity births to Indian-American women in the federal data are boys 54.5 percent of the time, compared to 57.5 percent of the time in the California data. These differences exist also among children born to Korean-American and Chinese-American women, although those differences are less striking. The differences do not appear to exist among black American, white American, or Japanese-American children. One possible interpretation of Abrevaya's findings regarding apparent differences in son preference between mothers in California and mothers in the US more broadly is that in California, where there is a higher concentration of members of the Asian community, there is a greater continuity of cultural values from the native country than elsewhere in the United States.

Abrevaya goes on to conduct a simple regression analysis to determine the effect of birth parity on boy-girl birth ratio. He uses the federal natality data and conducts separate regressions for each racial group. The paradigm used is a linear probability model, with boy birth used as the dependent variable for each regression and the principle independent variables of interest as second, third, and fourth parity birth or

higher. Control variables include birth year, dummies for same-race father, foreign-born mother, prenatal care, second trimester initial visit, previous terminated pregnancy, mother's age, mother's education, and indicator variables for ultrasound and amniocentesis use.

The proper interpretation of the results is to view the coefficient on the birth parity variables as the increase in probability of boy birth that the observed child was in the respective birth parity. For white American mothers, we find that the coefficient on second, third and fourth birth parity or higher are all positive and significant. However, each of the coefficients is smaller than .2, indicating that there is a very small effect of birth parity on boy birth percentage.

For Chinese-American mothers, the effect of second birth parity is not statistically significant. Effects of third and fourth birth parity are statistically significant, with third birth parity having a positive effect of 1.18 and fourth birth parity having a positive effect of 2.25. This effect is consistent with the findings of Scharping (2003) about mothers in China.

For Indian American mothers, the effect of each birth parity is significant; the second birth parity has an effect of 0.791. The third birth parity has an effect of 3.575. The fourth birth parity or higher has an effect of 2.481.

For Korean American mothers, only the second birth parity has a significant effect. The second birth parity has an effect of 1.254.

Section 3: Brief Analysis

Our analysis will follow the same basic format of Abrevaya's simple regression analysis from the last section. We will run four regressions: One for children born to white mothers, another for children born to Indian mothers, another for children born to Chinese mothers, and another for children born to Korean mothers. We will use for our data the federal natality birth data for the years 2005-2017. Our dependent variable will be a variable equal to 1 if the child is a boy and 0 if the child is a girl; therefore, this is a linear probability model. Our independent variables of interest are dummy variables for birth parity, equal to one if the child was born in the respective parity and zero otherwise. Our control variables are the same ones from Abrevaya's analysis: birth year, dummies for same-race father, foreign-born mother, prenatal care, second trimester initial visit, previous terminated pregnancy, mother's age, mother's education, and indicator variables for ultrasound and amniocentesis use.

The results of each of the four regression are printed below. The coefficients on the birth parity variables are given with heteroskedasticity robust standard errors in parentheses next to the coefficients. A single star (*) is printed next to the coefficient if the result is significant at the 10% level; two stars (**) are printed next to the coefficient if the result is significant at the 5% level. Results are rounded to the nearest thousandth. Results can be interpreted as the percentage change in probability of boy birth given that a child was born in the respective parity.

Race	Parity	Coefficient	Heteroskedasticity-Robust Standard Error
White	Second birth	-0.041**	0.019
	Third birth	-0.127**	0.023
	Fourth birth	-0.186**	0.044
Indian	Second birth	0.724**	0.215
	Third birth	2.879**	0.331
	Fourth birth	2.858**	0.862
Chinese	Second birth	0.115	0.149
	Third birth	2.041**	0.459
	Fourth birth	1.847**	0.572
Korean	Second birth	0.284	0.318
	Third birth	0.472	0.496
	Fourth birth	0.873	0.917

These results are quite similar to the results found by Abrevaya. One notable exception is that in our analysis, none of the coefficients on birth parity for Korean-American mothers are significant. This could be related to the decrease in son preference in South Korea noted by Chung and Das Gupta (2007). All of the other results follow the basic pattern in Abrevaya's analysis, with positive effects of later birth parities for Chinese and Indian mothers.

Conclusion

The basic fact which we learn from the global literature is that boy-birth ratios began to increase in the mid-1980s in various countries due to the introduction of technologies which allow parents to determine the gender of their unborn child. The effects of bans on gender-targeted abortion have been ineffective in India and China. In South Korea, though, boy-birth ratios have fallen in recent years, perhaps because of changes in cultural attitudes toward the necessity of having a male child.

One problem with the US data is that we have limited information on specific communities in the US, and there is no direct evidence of gender-targeted abortion (there are apparently no surveys, for example, which ask women whether they have aborted a child because the child was female). Therefore, we must make the best possible use of the limited indirect data to which we have access. Using this data, Abrevaya finds that there exists some evidence of the missing girls phenomenon, at least among Chinese-American and Indian-American mothers and perhaps also among Korean-American mothers; our analysis extends Abrevaya's analysis and finds that this trend appears to hold up until the most recent year, although the phenomenon appears to have leveled off among Korean-American women.

Some state legislatures have banned sex-selective abortion, apparently in an effort to prevent the missing girls phenomenon from developing in the United States; or perhaps as part of a larger political backlash against abortion. However, in Illinois and Pennsylvania these bans have proven to have little or no effect (Nandi, Kalantry, & Citro, 2015). Further, Nandi, Kalantry

and Citro argue that sex-selective abortion bans may serve to stigmatize and limit abortion access to Asian-American women. This information, taken with the apparent decline in the missing girls phenomenon in South Korea, suggest that perhaps only shifts in cultural attitudes may serve to limit this phenomenon.

Works Cited

Abrevaya, J. (2005). Are there missing girls in the United States? Evidence on gender preference and gender selection. *Unpublished Working Paper. Purdue University*.

Abrevaya, J. (2009). Are There Missing Girls in the United States? Evidence from Birth Data. *American Economic Journal: Applied Economics*, 1: 1-34.

Almond, D., & Edlund, L. (2008). Son-biased sex ratios in the 2000 United States Census. *Proceedings of the National Academy of Sciences*, 8.

Centers for Disease Control and Prevention. (2001). *Vital Statistics of the United States, 2018 Volume 1, Natality.* Hyatsville, MD.

Cheng, Y., Li, H., & Meng, L. (2013). Prenatal Sex Selection and Missing Girls in China: Evidence From the Diffusion of Diagnostic Ultrasound. *Journal of Human Resources*, 48: 36-70.

Chung, & Das Gupta, M. (2007). The Decline of Son Preference in South Korea: The Roles of Development and Public Policy. *Population and Development Review*, 33: 757-783.

Dahl, G. B., & Moretti, E. (2004). The Demand For Sons: Evidence from Divorce, Fertility, and Shotgun Marriage. *NBER Working Paper 10281*.

Das Gupta, M. (2005). Explaining Asia's "Missing Women": A New Look at the Data. *Population and Development Review*, 31: 529-535.

Ebenstein, A. (2010). The Consequences of the "Missing Girls" of China. *Journal of Human Resources*, 45: 87-115.

Guilmoto, C. Z., & Bongaarts, J. (2015). How Many More Missing Women? Excess Female Mortality and Prenatal Sex Selection. *Population and Development Review*, 41(2): 241-269.

Jha P, e. a. (2006).) Low female-to-male sex ratio of children born in India: National survey of 1.1 million households. *Lancet*, 367: 211–218.

Korea Institute for Health and Social Affairs. (1992). *1991 National Fertility and Family Health Survey.* Seoul: Korea Institute for Health and Social Affairs.

Lhila, A., & Simon, K. I. (2008). Prenatal Health Investment Decisions: Does the Child's Sex Matter? *Demograpy*, 45(4): 885-905.

Lundberg, S., & Rose, E. (2003). Child Gender and the Transition to Marriage. *Demography*, 40: 333-349.

Nandi, A., Kalantry, S., & Citro, B. (2015). Sex-Selective Abortion Bans are Not Associated with Changes in Sex Ratios at Birth among Asian populations in Illinois and Pennsylvania. *Forum for Health Economics and Policy*, 18(1): 41-64.

Office of the Registrar General & Census Commissioner, India. (2002). *2001 Census of India.* Mumbai: Office of the Registrar General & Census Commissioner, India.

Park, C. B., & Cho, N. H. (1995).) Consequences of son preferences in a low-fertility society: Imbalance of the sex ratio at birth in Korea. *Population and Development Review*, 21:59–84.

Scharping, T. (2003). Birth Control in China 1949–2000: Population Policy and Demographic Development. *London and New York: Routledge Curzon.*

Sen, A. (1990). More Than 100 Million Women Are Missing. *New York Review of Books*, 37: 20.

Yeung, W., Sandberg, J., Davis-Kean, P., & Hofferth, S. (2001). Children's Time with Fathers in Intact Families. *Journal of Marriage and the Family*, 63: 136-154.