Hepatitis B and the Case of the Missing Women

Emily Oster*
Harvard University
Draft: Preliminary and Incomplete

February 8, 2005

Abstract

In many Asian countries the ratio of male to female population is higher than in the West – as high as 1.07 in China and India, and even higher in Pakistan. A number of authors (most notably Sen (1992)) have suggested that this imbalance reflects excess female mortality and, as a result, have argued that as many as 100 million women are "missing." This paper proposes an explanation for much of the observed over-representation of males: the hepatitis B virus. Evidence drawn from the existing medical literature as well as new studies of recent vaccination efforts indicate that carriers of the hepatitis B virus have offspring sex ratios as high as 1.55 boys for each girl. Hepatitis B is common in many Asian countries, especially China, where some 10 to 15% of the population is infected. Using data on hepatitis B prevalence by country as well as estimates of the effect of hepatitis on sex ratio drawn from a wide range of sources, I find that hepatitis B can explain about 45% of the missing women: around 75% in China, between 20% and 50% in Bangladesh, Egypt, and West Asia, and less than 20% in India, Pakistan and Nepal.

1 Introduction

The ratio of men to women in the Western world is close to unity. Sex ratios at birth are bit higher (around 1.05 boys for each girl), but higher male mortality results in sex ratios around 1:1 by young adulthood and even lower among older adults. Although this pattern is universal in the West, population sex ratios in a number of Asian countries seem to be

*I am deeply indebted to Baruch Blumberg's book "Hepatitis B: The Hunt for a Killer Virus" for inspiring this paper. Nava Ashraf, Baruch Blumberg, Edward Glaeser, Michael Kremer, Steve Levitt, Jeffrey Miron, Yuzo Miyakawa, Karen Norberg, Nancy Qian, Jesse Shapiro, Amartya Sen, Andrei Shleifer and participants in seminars at the Fox Chase Cancer Center, the Center for Population Studies at Harvard and the Harvard Economics department labor lunch provided helpful comments. Mike Matthews was extremely helpful in providing data.

abnormally high – 1.07 in China, 1.08 in India and as high as 1.11 in Pakistan. This imbalance has caused a number of authors to argue that there is substantial excess female mortality in Asia (Sen(1992), Coale(1991), Klasen(1994)). In general, they argue that neglect of female children and poor conditions for women contribute to the higher sex ratio. These authors have estimated a large number of the "missing women" – between 60 and 100 million.

This paper connects the missing women puzzle to a separate literature on the hepatitis B virus (HBV) and offspring sex ratios. There is substantial evidence that women who are carriers of HBV give birth to a higher ratio of boys to girls than non-carriers; since many of the countries with missing women also have relatively high hepatitis B carrier prevalence this naturally occurring higher sex ratio at birth could produce a higher population sex ratio even with no excess female mortality. Indeed, I argue in the paper that after adjusting for differences in sex ratio at birth caused by HBV, the number of missing women in the countries highlighted by Sen(1992) drops to 32 million, from the 60 million calculated by Coale(1991) and the 107 million suggested by Sen(1992). There is significant variation in the share of missing women explained by country: I find that HBV can explain 75% of the missing women calculated by Coale(1991) in China, but less than 20% in India, Pakistan and Nepal. Other countries lie between these polar cases.

The issue of the missing women has become a flashpoint for development economists, demographers and policy-makers. For many, presumed excess female mortality is only the most severe example of discrimination against women in general in this region; more broadly, women have lower human capital attainment, less favorable health outcomes and less control in relationships. The missing women are emblematic of what seems to be a much larger problem and understanding why these populations have such a widely skewed sex ratio seems vital to moving forward on issues of gender equality. The results here suggest that sex ratio bias may be less attributable to differences in tastes and preferences than many have suggested. Understanding whether the biased sex ratios are a result of biology or behavior is vital to understanding what policies will be effective in decreasing them.

This paper begins with a brief discussion of how the number of missing women are calculated and the connection to sex ratio at birth and HBV. Section 3 then presents a discussion of the origin of the sex ratio bias. The hepatitis B hypothesis would suggest that the sex ratio bias arises at birth; to the extent that the bias arises later, there cannot be a role for hepatitis B. I present evidence that sex ratio at birth is substantially higher in most of the missing women countries than it is in the West. In addition, the pattern of sex ratio varies across these countries. China, for example, exhibits very high sex ratios at birth and declining sex ratios over childhood while India, in contrast, sees increasing sex ratios during childhood. The overall high sex ratio at birth suggests a role for HBV in the missing women puzzle, especially in countries like China with a high sex ratio at birth.

Section 4 presents a variety of evidence pointing to an effect of HBV on sex ratios at birth. Section 4.2 discusses individual-level evidence drawn from the existing literature. In 1972, researchers studying HBV in Greece noted that the sex ratio among 131 children ever born to women with HBV was 1.85 boys for each girl, versus only 1.13 among 542 children ever born to women without HBV. Subsequently, six additional studies in different areas confirmed these results. Maximum likelihood estimates of the effect of HBV on sex ratio at birth from the existing studies suggest a sex ratio at birth of 1.55 boys for each girl among HBV carriers.

In addition to the existing estimates, this paper presents new evidence on the effect of HBV on sex ratio at birth. First, Section 4.3 takes advantage of the availability of a vaccination for HBV for the past 20 years as a natural experiment. I consider first offspring sex ratios for Alaskan Natives (historically high HBV) and non-Natives in Alaska (low HBV). Data on births to these groups are compared before and after a massive vaccination campaign. I find that sex ratios among white Alaskans are unaffected, but sex ratios among Alaskan Natives drop dramatically over the period being considered. In addition, I consider sex ratio among young mothers in Taiwan who were covered by a universal vaccination program beginning in 1984; there is suggestive evidence of a drop in sex ratio at birth among that group, as well.

Further, Section 4.4 presents aggregate evidence on the cross-country relationship between HBV and sex ratio. This section takes advantage of both a categorical distribution of HBV and continuous estimates of HBV prevalence across countries. Both types of evidence suggest that HBV and sex ratio are strongly related, and this finding is robust to restricting to countries within a particular region, countries within the OECD, and to controlling for income levels. In fact, I find about 40% of the cross country variation in sex ratio at birth can be explained by HBV.

Combining data on HBV prevalence and estimates of the effect of HBV on sex ratio at birth enables me to estimate the HBV-adjusted number of missing women. The results suggest around 45% of the missing women estimated by Coale(1991) are explained by HBV. The HBV-adjusted number of missing women is 32 million – 8 million in China, 19 million in India and smaller numbers elsewhere.

First and foremost, the findings in this paper have implications for the puzzle of the missing women. However, the results also have implications for interpreting work on the correlates of gender preference and theories on the economics of the family. For example, drawing on Becker(1981), a number of papers have developed theories that provide a rational motivation for sex-selective mistreatment, infanticide and abortion, based on future earnings capacity of children (see Rosenzweig and Shultz(1982), Edulund(1999), Qian(2005), Kojima(2005)). One natural way to test these theories is to regress sex ratio on economic factors of interest (maternal education, earnings power, etc).[1] However, the conclusions from that type of regression rest fundamentally on the premise that there is no biological explanation for differences across space in sex ratios. In order to fully address these theories, it may be necessary to first adjust for naturally occurring differences in sex ratio, and then explore preferential correlates.

Moving a bit further afield, these results may also have implications for economists' study of marriage and labor market outcomes. In his work on the economics of the family, Becker (1981) highlights the importance of population sex ratios in the formation of polygamous societies, as well as in the determination of dowry and brideprice levels. Building on this, others have argued that high sex ratios (many men in the population) lead to higher rates of marriage among women and lower outside labor market participation (Angrist, 2002; Chiappori, Fortin and Lacroix, 2002). Population sex ratio movements during wars have also been connected to female labor market participation and wages (Goldin, 1991; Acemoglu, Autor and Lyle, 2004). The connection between hepatitis and sex ratio therefore has the

[1] For examples, see Rosenzweig and Shultz(1982), Klasen(1996), Agnihotri(1999) and Qian(2005). Others (Sen,1998) have simply used sex ratios across space to draw conclusions about level of development.

potential to provide a cleaner test for the relationship between sex ratio and market outcomes.

Finally, attributing a share of the missing women to hepatitis rate has implications for changes in the number of missing women over time. Since the early 1980s a vaccine for HBV has been available, and HBV prevalence in many of these countries has gone down and continues to decline. This suggests that sex ratios at birth will fall over time as the whole population is exposed to the vaccine. This provides yet another way to explore the effect of sex ratio on market and matching outcomes, and provides some hint about what the future holds for women in historically high HBV countries.

2 Calibrating the Missing Women

The population sex ratio in Europe and the United States – number of men divided by number of women – is around one. That is, there are roughly equal numbers of men and women in the population overall. Sex ratio at birth weakly favors men (about 1.05 men for 1 woman), but higher male mortality over the lifetime means that the ratio is equalized in the overall population. In Asia, population sex ratios are higher. In China, for example, the population sex ratio is around 1.066 men per 1 woman and it is even higher in the Indian sub-continent. It is this observation, coupled with existing evidence about son preference in areas of Asia, that led demographers to first suggest that there were "missing women" in Asia.

"Missing women" are defined as women who once lived but have died prematurely because of artificially high female mortality (presumably induced by human behavior). The number of missing women in a country is therefore calculated by assuming that the number of men reflects appropriate mortality patterns and then using the difference between the actual and expected sex ratio to calculate the level of excess female mortality. Denote the expected sex ratio based on the number of men as SR_e, the actual sex ratio observed as SR_a and the number of women observed as N_w. Then the missing women can be calculated as below:

$$Missing = \left(\frac{SR_a}{SR_e} - 1\right) N_w \tag{1}$$

Obviously, the larger the difference between the actual and the expected sex ratio, the greater the share of women who are missing.

From the perspective of this paper, the origin of the sex ratio bias is crucial. Parental HBV infection is posited to affect offspring sex ratio at birth, which will then play into later sex ratio of the population. In particular, overall sex ratio of a given cohort in a population is a function of their sex ratio at birth and mortality patterns by gender. Denoting the actual sex ratio in cohort c as $SR_{a,c}$, sex ratio at birth as $SR_{b,c}$, male mortality for the cohort as M_m and female mortality M_f, I can write:

$$SR_{a,c} = SR_{b,c}\frac{1-M_m}{1-M_f} \tag{2}$$

It is easy to see, then, that even if two populations have identical mortality by gender, if their sex ratios at birth differ then the sex ratio later in life will differ also. A higher sex ratio at birth will produce a higher population sex ratio, as will more imbalanced mortality patterns.

In Section 4 I provide a body of evidence suggesting that the offspring sex ratio is higher for HBV carriers than non-carriers. Assuming that HBV carriers have a higher sex ratio among offspring, then the prevalence of HBV in the population will influence the sex ratio at birth. More specifically, the sex ratio at birth will be the average of the sex ratio of HBV carriers (SR_{HBV}) and non-carriers (SR_N), weighted by their population shares:

$$SR_{b,c} = (SR_{HBV})(Prev_{HBV}) + SR_N(1 - Prev_{HBV}) \tag{3}$$

The final step in connecting HBV and the missing women is to calibrate the effect of HBV on sex ratio – to calculate SR_{HBV} from the above equation. For the purposes of calibration the adjusted number of missing women, I use maximum likelihood estimates of the effect of HBV on sex ratio at birth that are calculated from the individual level data. In particular, I assume that sex ratio at birth among HBV carriers is determined by draws from a binomial distribution. I assume each study has a study-specific sex ratio for non-carriers and that there is a constant additive shift parameter introduced by HBV carrier status. I then solve for the maximum likelihood estimate of the shift parameter using data from available studies; standard errors are also calculated. Details of the calculations are in Appendix A.

As a robustness check, I also calculated the SR_{HBV} suggested by aggregate population level data. Consider two populations: Population A with a high hepatitis rate H_A

and a high sex ratio SR_A and population B with a lower hepatitis rate H_B and a lower sex ratio SR_B. Assuming that the non-HBV carrier sex ratio is the same in the two areas, solving for SR_{HBV} yields the following expression:

$$SR_{HBV} = \frac{SR_A(1 - H_B) - SR_B(1 - H_A)}{H_A - H_B} \tag{4}$$

The discussion above provides a framework for the rest of the paper. I now move to discussion the origin of the sex ratio bias in the missing women countries, and then to the possible role for hepatitis.

3 The Origin of the Sex Ratio Bias

This paper argues that hepatitis B may help account for some of the missing women. This argument, however, relies crucially on the sex ratio bias arising at birth.[2] It is worth noting here that what may seem like relatively small differences in sex ratios at birth can result in sex ratio bias in the overall population similar to what is seen in Asia. For example, moving from a sex ratio at birth of 1.05 to 1.10 could move the overall population sex ratio – even with identical mortality – from 1.00 to 1.05, which is a large majority of the U.S./China difference.

To get some sense of sex ratios over the life cycle in the missing women countries, Figure 1 shows the sex ratio by age category in each country. This figure is created by calculating the sex ratio by age group from data in the Demographic Yearbook Historical Supplement (United Nations, 1997).[3] The data used are from censuses or birth registration data taken between 1970 and 1990 (inclusive). It is worth noting, of course, that this methodology combines cohort and age effects and if there are changes in population demographics over time, this will be problematic; below I address this issue in India and China by following individual age cohorts over time.

[2]It is worth noting that men are more likely to be Hepatitis B carriers and therefore may have higher HBV-related mortality from liver issues. However, this effect is unlikely to be a significant factor, given that because only a relatively small share of carriers experience liver problems.

[3]For India and Nepal, sex ratio at birth is not available from the Demographic Yearbook, so household survey data are used to construct a measure of the sex ratio at birth for these countries based on reported sex ratio among ever born children. The reliability of these data is checked using data from Pakistan, for which both household survey data and Demographic Yearbook data are available, and the results suggest that, if anything, the household survey data understate the sex ratio at birth.

The pattern in sex ratios over the life cycle appears to vary across countries. In China, for example, sex ratio is highest at birth and then generally decreases, although is slightly higher among those 35-59 than among those aged 15-34. The pattern in India is somewhat different: Sex ratio is higher than normal at birth, but not extremely high, and increases in childhood and young adulthood before falling again for the very old. Data for Pakistan and Bangladesh suggest a slightly different pattern, with above average sex ratios at birth, increases in childhood and large increases in old age. Finally, the pattern in West Asia (Turkey and Syria) and Egypt is more typical of developed countries, but with slightly higher sex ratios at birth than would be seen in most of the developed world.

Figure 1 generally suggests that the origin of the sex ratio bias is not the same in every country (although, as mentioned, this confuses age and cohort effects). Although most of these countries seem to have higher sex ratios at birth than the average in the western world, may of them also have non-Western mortality patterns, even controlling for their income level. The increase in sex ratio in childhood – particularly in India, Pakistan and Bangladesh – suggests excess childhood mortality for girls. In contrast, in China, the data indicate high sex ratios at birth but relatively normal mortality patterns.

In order to avoid the confluence of age and cohort effects, Figure 2 follows individual cohorts in China and India rather than considering age specific sex ratios (I have chosen these countries because they represent by far the most missing women and have different sex ratio patterns by age). For example, the graph includes sex ratio by age for the 1982 birth cohort in China – their sex ratio at birth and then in subsequent years. In general, these data are from less than a total census (in China, for example, these are data from the yearly 1% censuses), but it should give a good sense of the sex ratio patterns over the early part of the life cycle. Figure 2 mirrors Figures 1: The sex ratio bias in China is clearly evident at birth, and only declines through childhood, while the sex ratio in India is low at very young ages, and increases through childhood.

Figures 1 and 2 generally demonstrate that there appears to be some sex ratio bias at birth, at least in a subset of the missing women countries. In particular, in most of these countries the average sex ratio at birth is above 1.059, which is the fixed sex ratio that Coale (1991) uses to calculate the number of missing women (this is also used in Klasen (1994) and

Klasen and Wink (2002)). As noted in Section 2, increases in sex ratio at birth will lead directly into increases in the sex ratio later in life. When re-estimating the missing women, I assume exactly the same mortality patterns as in Coale (1991), and use HBV-adjusted sex ratios at birth to explore the effect on overall sex ratios.

In general, the results in this paper suggest that the number of missing women in China in particular has been overestimated. Given this, it is worth briefly exploring whether the sex ratio bias at birth in China is naturally occurring before moving on to a discussion of Hepatitis B. More specifically, it has lately been argued that the sex ratio bias in China has been rising over time and that imbalanced sex ratios are due largely to sex-selective abortion (Hull, 1990; Yi et al, 1993; Junhong, 2001). While measurement is extremely difficult (for one attempt on a small scale, see Junhong, 2001), there can be no doubt that there has been some use of sex-selective abortion technology in China. For example, the sex ratio at birth in the 2000 Census is close to 118 boys for 100 girls, much higher than that reported in 1989 or 1982. The interaction between the one-child policy and the increased availability of ultrasound technology has likely had some effect on these ratios; as a signal, authors have noted higher sex ratios among later children, although this does not seem to be strongly true earlier in the century (see, for example, Coale and Banister (1994)).

However, the fact that sex ratio at birth has increased over the last 15 years does not directly imply that sex ratios prior to this time period were normal. Given that the missing women estimates presented in Sen(1992) and Coale(1991) rely on population sex ratios estimated between 1981 and 1991 what is relevant for the most part are sex ratios at birth before the one child policy and sex-selective abortion. Here, I present a number of pieces of evidence pointing to the fact that sex ratio at birth has been high in China historically and, more generally, evidence that suggests the bias at birth prior to the latest censuses is organic.[4]

In an early study of the effect of geography on human offspring sex ratio, Chambliss (1949) notes the high sex ratio at birth in China. In five regions of China, Chen (1947) reports sex ratios at birth that range from unity to 123 men per 100 women. Most are

[4]The first thing to note is that sex-selective abortion technology (ultrasounds) were not widely available in China prior to the 1980s (Junhong, 2001) and they dispersed over the 1980s and 1990s. The preceding analysis uses sex ratio at the 1982 and 1989 birth cohorts; this first cohort is unlikely to be heavily influenced by this technology, and yet the sex ratio is abnormally high (close to 1.09).

substantially higher than normal, and the average is 108 men per 100 women. Chiao (1934) reports a high ratio of 110.4 boys for 100 girls in his study of 12,000 farm families in 11 provinces between 1929 and 1931. He additionally reports the results of three sources of hospital registration (average sex ratio of 111.6 boys for 100 girls), three other registration surveys (average sex ratio of 114.7) and three other general surveys average sex ratio of 115.2). It is notable that even the hospital registration reports a very high sex ratio in this time period in the late 1920s.

These papers, however, generally report the results of particular censuses and individual studies that focus on one time period. In contrast, Coale and Banister (1994) do an extensive study of the path of sex ratios at birth over five decades in China. Using two retrospective fertility surveys from the 1980s, they report imputed sex ratios at birth for cohorts born between 1936 and 1989, and find high sex ratios at birth from the earliest cohorts.[5]

An alternative way to explore the importance of underreporting and infanticide is to look at the offspring sex ratios of native Chinese women in the United States. There is good evidence that hepatitis B rates among Chinese immigrants to the West are similar to those in China (Kent(2000), Gjerdingen and Lor(1997), Hayes et al (1998)), and it seems unlikely that there is significant infanticide among these groups in the United States.[6] This suggests that high sex ratios among children of these immigrants would strongly support the hypothesis of naturally occurring high sex ratios at birth.

I explore two measures of sex ratios at birth for children of Chinese immigrant mothers: vital statistics and the U.S. Census Individual Public Use Microdata Sample (IPUMS). The Center for Disease Control Vital Statistics Natality Detail Files report on all

[5]Table 1 reports some of the data from their paper, showing sex ratios at birth starting with the 1936-1940 cohort and going through the 1985-1989 cohort. The sex ratio is high in all cohorts. Although it declines a bit in the 1960s, the overall average is 1.095. It is worth noting that, although retrospective data can be problematic, Coale and Banister (1994) provide evidence suggesting that the retrospectively reported sex ratios at birth line up well with actual sex ratios from later censuses, suggesting that the data are fairly reliable.

[6]Concerns about selective migration should also be somewhat alleviated both through information on high HBV rates among immigrants and by the fact that there seems to be little evidence that HBV infection is related to socioeconomic status. Lee et al (2002) find a positive association between HBV and income in Korea; Wang et al (2002) echo these findings in Taiwan; Rahman et al (1997) find no relationship in Bangladesh; Amini et al (1993) find a negative relationship in Iran and She et al(1988) find no relationship among barbers in China. Since HBV infection does not cause active illness, there is no good reason to believe carriers would be less likely to migrate than others.

births in the U.S. from 1985 onwards. Information on sex of the child, race and origin of the mother and father, as well as basic demographic information is included. The main drawback from the perspective of this exercise is relatively limited data about national origin and race. It is possible, however, to estimate sex ratios among children born to mothers' whose race is "Chinese" and who were not born in the U.S. There are two important caveats: first, the racial coding of "Chinese" includes not only those who identify themselves as Chinese, but also those from Taiwan, Macao and other places, which may have very different (and lower) hepatitis B rates. Second, since birthplace outside the U.S. is identified only as "rest of the world", it is hard to know whether women were actually born in China. Both of these factors will bias the sex ratio downwards relative to that for Chinese women immigrants from mainland China, but the sex ratio at birth is still substantially higher than normal, as 1.082. This is strongly significantly different from the sex ratio for whites.

In contrast to the vital statistics, the IPUMS data (which give detailed information on 1% or 5% of U.S. households) provides very detailed information about country of origin; the downside is a smaller sample size and the fact that it is necessary to use data on sex ratios among young children rather than at birth. With limited infant mortality sex ratio among young children should mirror sex ratios at birth, so the latter may not be a concern, particularly in later time periods. Panel A of Table 2 shows sex ratios for U.S. born children under 6 with Chinese mothers from the IPUMS. The sex ratio overall is quite high – 1.105. In the last three censuses, it is even higher, at 1.115. In all periods except 1940-1970 the sex ratio is very high; the lower sex ratio between 1940 and 1970 is driven by quite low sex ratios among children in the 1960 census. It is worth noting that an equality test of proportions cannot reject equality between this ratio and the ratio in the later censuses. As a comparison, in Panel B two comparisons are presented: the sex ratio of births to immigrant mothers from low-HBV countries, and the sex ratio of births to native whites. Both are normal – around 1.04 – and significantly different from the rate for Chinese mothers. The comparison to other immigrant mothers is particularly noteworthy, as it suggests that the high sex ratio among Chinese mothers is not driven by their immigrant status.[7]

[7] I elaborate on this in Section 4.4, which demonstrates that HBV prevalence in the home country strongly predicts sex ratios at birth for immigrant women in the IPUMS.

Overall, the historical evidence from China points to a long-lived bias in the sex ratio at birth. This seems to be true of Chinese immigrants outside of China itself, lending support to a non-cultural explanation. Having argued that the sex ratio bias overall is due at least in part to sex ratio bias at birth, I turn now to the hepatitis B virus and its possible role in determining offspring sex ratio.

4 Hepatitis B and Offspring Sex Ratio

This section discusses the hepatitis B virus and the effect of the virus on offspring sex ratio. The first subsection provides background on the virus; I then discuss the evidence for an effect of hepatitis B on offspring sex ratio. The final subsection calibrates the size of the effect of hepatitis B on sex ratio and uses data on hepatitis B prevalence to estimate the share of the sex ratio at birth that is explained by hepatitis rates.

4.1 Background on Hepatitis B

Hepatitis B virus (HBV) is a viral infection of the liver that is spread through exchange of blood or bodily fluids. The virus can be spread through sexual contact, needle exchange and from mother to child, as well as through more casual contact. Some infected individuals – even those who do not show symptoms – do not clear the virus and go on to become chronic carriers. Chronic viral infection leads to serious liver problems for around 25% of carriers. The chance of becoming a carrier decreases with age of infection – it is close to 90% for infants, but under 10% for those infected as adults (Department of Vaccines, 2001). Since the late 1970s a vaccine for HBV has been available. It has been widely used in a number of places – the U.S., China, Taiwan, Singapore and some countries in Southern Europe – although has not been universally adopted, particulary in low and medium prevalence countries.

Cross-country HBV rates vary significantly. High hepatitis regions include East Asia, Sub-Saharan Africa, parts of the former USSR, parts of Eastern Europe, northern South America, Alaska and northern Canada. There is suggestive evidence that differences across countries result from differences in the timing of viral transmission. As mentioned above, the chance of becoming a carrier of the virus once infected decreases as an individual ages. In high epidemicity countries viral transmission is most common among infants and toddlers; in those with medium epidemicity transmission is most common in childhood, either at home or school; in low epidemicity countries sexual or intravenous drug use transmission is most common and occurs in early adulthood (Dienstag, 1982). Because of the differences in the probability of becoming a carrier depending on when infection takes place this pattern may explain differences in the overall rate across the world.[8]

The origin of these differences in transmission patterns is unknown, but may be due to differences in viral or human genotypes. HBV virus has eight known genotypes (A-H) and there is some circumstantial evidence that their distribution is correlated with HBV prevalence (Miyakawa and Mizokami, 2003) and that the genotypes differ in their transmission patterns (Duong et al (2004), Kobayasi et al (2003)). Human genetic differences in susceptibility are also a possible explanation for the variation in epidemicity. Geneticists have identified a number of polymorphisms[9] that appear to affect susceptibility to HBV (for a review, see Andrade and Andrade (2004)). It is possible that this type of genetic variation may influence HBV patterns across the world.

Relevant for this paper, there is substantial evidence that women who are carriers of HBV are more likely to have male children than those who are not. The mechanism for this relationship is not precisely known. One hypothesis is that female foeti of HBV carrier parents are much more likely to spontaneously abort early in the pregnancy, resulting in more live births that are male (Drew et al, 1978; Livadas et al, 1979), but this has not been medically confirmed. It is consistent, however, with evidence that carriers of HBV have fewer daughters than non-carriers, but similar numbers of sons (London et al, 1982). An alternative theory suggests that carrier women are less likely to conceive daughters. Despite uncertainty about the biological mechanism of the effect, the effect on sex ratio at birth is quite robust and large, as discussed in more detail below.

[8]As a corollary to this, prevalence of HBV carriers is not closely related to sexual activity, even though HBV can be transmitted sexually (Department of Vaccines, 2001). This is true because infection as an adult only rarely leads to carrier status, whereas infection as an infant or young child often does; sexual transmission is therefore responsible for only a tiny fraction of HBV carriers. It is also worth noting (consistent with the above) that HBV prevalence across countries is not correlated with rates of other sexually transmitted infections; for example, the correlation with genital herpes prevalence is around -0.06.

[9]Polymorphisms are mutations in the human genetic code that appear in at least 1% of the human population. An example is human blood groups.

4.2 Individual-Level Evidence

The relationship between HBV carrier state and offspring sex ratio was first noticed in four studies on families of carriers and non-carriers (Hesser, Economidou and Blumberg, 1975; Drew, London, Blumberg and Serjeanston, 1982; Drew, Blumberg and Robert-Lamblin, 1986; Chahnazarian, Blumberg and London, 1988). The methodology in these studies is relatively simple. In each case, married women with children were identified in villages with known high HBV rates. A sample of women and (in all cases except Greenland) their husbands were tested for HBV carriage and were given a detailed survey about their reproductive history. Subsequently, two additional studies (Cazal, Lemiare and Robinet-Levy, 1976; Livadas et al, 1979) specifically sought to test this theory with the same methodology, and came to the same conclusions. All six studies tabulate the gender of all children of carrier and non-carrier parents and compare the sex ratio across the two groups. It is worth noting that there is no *a priori* reason to expect carrier and non-carrier families to be different from each other, absent HBV. In most cases women would not know about their HBV status and would have no symptoms during childbearing years.

The estimated offspring sex ratios for carrier and non-carrier parents in each of the six studies cited above are presented in Table 3. Overall, families with an HBV carrier parent had 350 male children and 243 female children (sex ratio of 1.44), while families without an HBV carrier parent had 2083 male children and 2043 female children (sex ratio of 1.02). The difference is statistically significant overall and in each individual study.

There is one further study of this issue that employed a slightly different procedure (Mazzur and Watson, 1974). In the British Solomon Islands families were tested for HBV and the sex ratio among carriers and siblings of carriers was explored. Assuming that the transmission of the virus to the carriers was parent-to-child, then the families with siblings who are carriers will be offspring of HBV-positive individuals. In this study, among siblings of carriers there were 90 men and 56 women (a sex ratio of 1.61), whereas among siblings of non-carriers the ratio was 91 women and 89 men (sex ratio of 0.97). Although this study is slightly less precise than the other evidence given the procedure employed (in particular the fact that parents were not tested directly), it is nevertheless consistent and provides an estimate of similar magnitude.

Although in many ways the type of evidence presented here is the cleanest way to test the hypothesis that HBV affects sex ratio, the sample sizes are relatively small. In addition, the missing women problem relies on aggregate sex ratios, so it is worth exploring the effect of HBV on sex ratios in the aggregate, as is done in the next two sections.

4.3 Time Series Evidence

A vaccine for HBV became available very shortly after the virus was discovered in the late 1970s. To the extent that vaccinated women are now of childbearing age, this provides a natural experiment to explore the effect of HBV on sex ratio. In general, universal vaccination programs (in Singapore, Italy, Spain, etc) did not start until the mid-1980s or early 1990s. However, there are two areas that allows us to consider variation over time: Alaska and Taiwan.

The Native Alaskan population has historically been the only population in the U.S. with high hepatitis rates – as high as 16% to 20% in some villages, while non-Native Alaskans have very low HBV rates. Further, HBV prevalence varied considerably across Alaska (McMahon et al, 1993), with particulary high prevalence in West and Central areas, and much lower in the Far North and East. The U.S. began vaccination of the entire Native population in 1982 and the vaccination campaign was complete, including not only infants but adolescents and adults (Harpaz et al, 2000).

Unlike in vaccination campaigns that focus only on infants, the overall coverage of this campaign meant that the incidence in the population overall fell immediately. The share of carriers (prevalence) among women of childbearing age decreased gradually, since anyone who was infected already would remain infected even after the vaccination campaign. However, by the end of the 1990s nearly all women who were giving birth would have been vaccinated and therefore would not be carriers. In addition, over the period from the early 1980s through the late 1990s the number of infected women giving birth would have decreased gradually as a share of the total births, as carriers were replaced by vaccinated individuals in the population of women of childbearing age. It may therefore be possible to see changes in offspring sex ratio over time, even considering only the last 20 years.

Data on the number and gender of births are available across regions in Alaska from 1980 through 2000. I compare the changes in sex ratio at birth across three groups: high-native, high-HBV regions (Yukon-Koyukuk, Aleutians East Borough, Lake and Penninsula, Bristol Bay), high-native, low-HBV regions (Prince of Wales, Skagway-Hoonah-Angoon, Wrangall Peterson) and a low-native region (Anchorage). The effect of the vaccination program on these three groups can be seen in Figure 3.

While sex ratios are relatively constant in the high-native, low-HBV and low-native regions, they decrease dramatically in the high-native, high-HBV region. The sample sizes are relatively small, but it is possible to reject equality between the normal sex ratio at birth and the ratio observed in the high-native, high-HBV regions in the earliest time periods.

It is also possible to explore this issue using the IPUMS to explore the changes in sex ratio for young children among the two groups over this time period. Specifically, I consider the sex ratio among very young children born to whites and Natives in Alaska from the 1980 through 2000 censuses. These data cover before (1980) and after the vaccination campaign. I expect a similar pattern to that seen in the vital statistics: decreases in sex ratio among natives between 1980 and 1990 and then between 1990 and 2000. Column 1 of Table 4 reports the results of a regression of child sex on whether the mother is an Alaskan Native, year dummies and interactions between Native and year, as well as other controls.

The overall effect of native is zero, but the interactions are both positive and significant. Sex ratio among natives is higher than that among non-natives in the 1990 census, and higher still in the 1980 census. In Column 2, I limit to Eskimos, the ethnic group which is generally thought to have the highest HBV rate (McMahon et al, 1993). Although sufficiently detailed racial identifiers are not available until the 1990 census it is nevertheless possible to get a sense of the pattern using 1990 and 2000. In this case, the dummy on Native is positive, as is the interaction: this ethnic group has a higher offspring sex ratio during both 1990 and 2000, but it is even higher in 1990 (when the vaccination program would have covered fewer people) than later.

In addition to Alaska, there is some suggestive evidence on HBV and sex ratio from Taiwan. Taiwan was historically among the highest HBV regions in the world and began universal vaccination of newborns in July of 1984, and HBV carriage rates have dropped dramatically (Lin et al, 2003). The first cohort of newborns vaccinated were 15 in 1999, implying that it might be possible to see effects of HBV on sex ratio by looking at changes in sex ratio by age cohort.

The data are still quite limited, but Figure 4 shows the sex ratios by age cohort over time in Taiwan, for individuals 15-35. The data are from the Taiwan-Fukien Demographic Factbooks (Republic of China, 1974-2002) and represent all births in Taiwan during this time period. There appears to be some suggestive evidence for an effect of the vaccine. In all but the youngest age cohort, sex ratio at birth is increasing secularly over time; in the youngest age cohort the effect appears to be truncated in the last time period, when the vaccination coverage should have taken effect. Future years of data will be necessary to speak concretely to this hypothesis, but the existing data are certainly suggestive.

4.4 Cross Country Evidence

The individual level and time series evidence points to an effect of HBV on offspring sex ratio. If the effect is true, it should be possible to see a cross country relationship between HBV prevalence and sex ratio at birth and evidence of such a relationship would strengthen the results above. This section explores that relationship, using both categorical measures of HBV prevalence and continuous prevalence from a created database.

The World Health Organization (WHO) categorizes countries into high (8% to 20% carriers), medium (2% to 7%) or low (<2%) HBV prevalence. Appendix B lists the countries in each group.[10] These data on hepatitis are combined with data on sex ratios at birth from the United Nations Demographic Yearbook Historical Supplement (United Nations, 1997). This covers the late 1960s through present; sex ratios are weighted by the number of births recorded in each available year.

Figure 5 shows sex ratio at birth by WHO category. The leftmost set of columns includes all countries. The sex ratio at birth in high HBV countries is 1.075, versus 1.056 in countries with medium prevalence and 1.048 in countries with low prevalence. The middle set

[10] Although these data are useful in the sense that it covers all countries, there are drawbacks. The methodology for assigning countries to groups is not made obvious. In addition, in some cases, no studies have been done on hepatitis B in a given country and yet the WHO assigns the country to a group based on its geographic location.

of columns presents the world sex ratio excluding China to ensure that the results are not driven by this one high-population outlier and the results are still very strong. The third set of columns limits the analysis to Europe, which has substantial variation in HBV rate but less racial or income level heterogeneity. The sex ratio-HBV relationship holds strongly even in this limited sample: sex ratio in high HBV regions is around 1.07, versus 1.063 for medium and 1.055 for low HBV countries.[11]

The HBV-sex ratio relationship in the WHO data is suggestive but the data are somewhat difficult to interpret. In addition, this type of evidence is limited by the fact that HBV prevalence may vary quite widely even within a category. I therefore move now to consider more continuous measures of HBV prevalence. Although no centralized database with cross-country HBV rates is available, I have created such a database by aggregating published studies that report HBV rates. To this end, HBV estimates for each country were constructed by searching "prevalence HBsAg *country name*" in PubMed and extracting all published studies that reported a prevalence rate and sample size and did not include high-risk groups.[12] Overall prevalence rate for each country is a weighted average of the prevalence rates for all studies. A list of studies used for each country appears in Appendix C.

Although data were available for 71 countries, in many cases the total number of individuals tested was very small. To avoid attenuation bias from measurement error as much as possible, the data were limited to countries in which at least 15,000 people were tested. This results in a sample of 25 countries. Figure 6 shows a graph of sex ratio on hepatitis rate for these countries. The graph is clearly upward sloping, even if China is excluded. In fact, China is not a particular outlier and lies quite close to the regression line. The major outlier is Iran, for which the sex ratio at birth is extremely low. The data from Iran are known to be incomplete (as is true for a number of countries), which may explain this somewhat. In Figure 7, the data are limited to OECD countries to avoid the suggestion that income and development level differences may be driving the results. The relationship is strongly positive, even excluding South Korea.

Table 5 reports estimates of the effect of HBV on sex ratio using these continuous

[11] All differences discussed here are significant at less than the 1% level in a test of equality of proportions.
[12] HBsAg is the HBV surface antigen that is present in HBV carriers.

prevalence data. Columns 1 and 2 show the relationship between HBV prevalence and sex ratio across the world; Columns 3 and 4 are limited to the OECD. In Columns 2 and 4 controls for log GDP per capita are included in the regression to demonstrate that the relationship is not driven by development patterns. In all regressions the coefficient on HBV is strongly statistically significant, the magnitude is similar across all four columns and including GDP controls does not seem to affect the result. Further, HBV prevalence explains about 40% of the variation in sex ratio across the world and an even larger share of the variation within the OECD. The worldwide estimates imply that a 1% increase in HBV carrier rate increases the sex ratio by around 0.005. This implies that moving from Finland to Germany increases the sex ratio from 1.049 to around 1.054. It is worth noting the that the results are still statistically significant and of similar magnitude when China and South Korea, respectively, are excluded.[13]

Despite the results above, there remains the possibility that cultural differences across countries drive differences in sex ratio and HBV prevalence. In order to alleviate this concern it is possible replicate the results using data on sex ratios among children of immigrants to the United States. This analysis also allows me to consider additional controls in a way that was not possible in the overall cross country analysis. Using the data from the IPUMS (discussed in more detail Section 3) I calculate the sex ratio of young children born in the United States to foreign-born mothers.[14] In this way it is possible to create a cross-country dataset on sex ratio that is not contaminated by cross-country differences in the circumstances of pregnancy and birth.

One possible problem with using data from the IPUMS is the possibility that immigrants are selected on, among other things, low hepatitis rates. If this is true, then the results above may reflect differences in offspring sex ratio due to something other than differences in hepatitis rates. To the extent that it is available, the data generally do not

[13]The analysis is very similar when historical sex ratios – from the first half of the century or earlier – are used rather than present day ratios. This is encouraging, as it suggests that the patterns are long-lived, which is consistent with an explanation relying on an old disease.

[14]This includes all children under 6. As long as I assume that childhood mortality in the U.S. is not different for immigrants from high-hepatitis countries and those from low-hepatitis countries, then using sex ratio among young children rather than at birth will not be a problem. Although using vital statistics would provide sex ratios at birth, the information about country of origin is extremely limited and would make this type of analysis impossible.

support this. In particular, there is good evidence that immigrants from Asia to North America have very high HBV rates.[15] In the case of Asian immigrants, these studies cite HBV rates as high as 20%, which is actually higher than what is typically seen in the home country. Further, as discussed in Section 3 there is little evidence that HBV infection and socioeconomic status are negatively, so the concern that only the rich emigrate to the U.S. and they have lower disease burden probably does not hold in the case of HBV.

The relationship between hepatitis and sex ratio holds in the IPUMS. Figure 8 replicates Figure 6 using the IPUMS data from 1990 and 2000 rather than the actual country data on sex ratios (the sample is somewhat truncated due to data imitations – there are not sufficiently large numbers of immigrants in many countries). The pattern in sex ratios is also apparent in these data. In a regression, the simple relationship is significant at the 1% level (results not shown).

In addition to the obvious advantage of eliminating cross country variation in pregnancy conditions, using the IPUMS allows me to adjust for other factors that have been argued to affect the sex ratio, including female social status, marital status and age. To do this, I extend the sample back through the 20th century, and control for year of birth, in order to allow more countries to be used in the analysis. In Table 6 I present estimates from a regression using the IPUMS sample, regressing whether or not a child is a boy on HBV prevalence rate in the maternal county of origin and a number of controls. In Column 1 I include only the hepatitis rate measure and dummies for year of birth. The coefficient on prevalence in Column 1 is highly significant.[16]

A number of researchers have noted that maternal age (or birth order) appears to affect sex ratio, with older mothers having higher sex ratios among their children (Lazarus, 2002). If the age distribution at maternity across countries is correlated with HBV prevalence, estimates could be biased. Column 2 of Table 6 includes dummies for maternal age in the regression, with virtually no change in coefficient size – if anything, the effect becomes more

[15]See, for example: Kent(2000)on Chinese immigrants to Canada, Gjerdingen and Lor(1997) on Hmong patients in Minnesota, Hayes et al(1998) on pediatric refugees to Maine, Adair and Nwaneri(1999)**?** on African immigrants to Minneapolis and Patel and Voigt(2002) on Vietnamese immigrants to the midwest.

[16]The magnitude is somewhat smaller than that estimated using data from the countries themselves, but it is somewhat difficult to compare magnitudes when this regression uses young children rather than births, and the estimates of sex ratio are less precise.

pronounced.

The Trivers-Willard hypothesis (Trivers and Willard, 1973) suggests that sex ratio of offspring may be influenced by female status; in particular, females of higher status have more male children. Although some studies have supported this, others have not (for a summary, see Lazarus (2002)). It is worth noting that the hypothesis has to do with female rank within the particular society, suggesting that it is probably not responsible for variation across countries. However, in order to control for this Column 3 in Table 6 includes a quadratic in maternal education relative to the her cohort. Neither control for status is significant and they do not affect the estimates on hepatitis.

Finally, the presence of a partner in the household has been hypothesized to play a role in offspring sex ratio. Norberg (2004) presents evidence that women who live with a spouse or partner are more likely to have male children than those who do not. If women from high hepatitis regions are more likely to be conceive children while not cohabitating then this could bias the coefficients.[17] Marital status is reported in the IPUMS, and Column 4 of Table 6 includes a dummy for married or single. While the coefficient on marital status has the expected sign, it is not significant and the coefficient on HBV prevalence is essentially unchanged.

4.5 Calibrating the Effects of HBV on Sex Ratio

As discussed in Section 2, I use a maximum likelihood technique to estimate the effect of HBV on sex ratio at birth from the individual-level data. This technique also allows estimation of standard errors (see Appendix A). In addition, I calculate the effects of HBV on sex ratio that are implied by the aggregate data using the methodology outlined in Equation (4). These estimates are presented in Table 7.

The maximum likelihood estimate of the effect of HBV on sex ratio is 1.55, implying that HBV carriers have 1.55 boys for each girl. The 95% confidence interval is relatively large, from 1.32 to 1.83, due largely to the small samples. However, it is worth noting even the lower

[17]Obviously cohabitation status at conception and after birth are not perfectly correlated. There should be some correlation, however. In addition, this would only affect the coefficient on HBV if the breakup or consummation of marriage was related to the country of origin HBV status, which seems unlikely.

bound is far above the average sex ratio for non-carriers. The estimates derived from the aggregate data are consistent with the individual-level estimates – they range from a low of 1.48 to a high of 1.96.

Using the calibration from Table 7, it is possible to calculate the predicted sex ratio at birth using data on HBV prevalence in the missing women countries. In this case, I use prevalence for pregnant women only (where available). Table 8 presents estimates of HBV prevalence, predicted sex ratio at birth and the range of empirical estimates of sex ratio at birth for each country. The predicted sex ratios at birth are all well inside the range of empirical estimates for each country.[18] These results suggest that much of the differences in sex ratio at birth displayed in these countries can be explained by variation in HBV rate.

5 Estimating the Missing Women

The HBV-adjusted number of missing women is calculated exactly as in Coale (1991), except that in the new estimates I use the sex ratio at birth predicted by HBV rate in Table 8, rather than a constant of 1.059 for each country. The mortality patterns assumed after birth are the same as those used in Coale (1991) and reflect mortality estimates from the model life tables in Coale, Demeny and Vaughan (1983).[19]

The new estimation appears in Table 9. In Panel A, I compare the HBV-adjusted estimates to those of Coale (1991). After adjusting for HBV, roughly 45% of the Coale (1991) missing women are explained – nearly 30 million women "found". The country-by-country breakdown reveals large variation in the percent of the puzzle that is explicable. In particular, the high HBV rate in China explains 75% of the missing women in that country. In contrast, in India, Pakistan and Nepal less than 20% of the puzzle is explained.[20] The numbers for

[18]For all countries but India and Nepal, the range represents the range of sex ratios across years in the United Nations Demographic Yearbook; for India and Nepal it represents the range by maternal age for women in household surveys. In the case of Nepal there are relatively few women of any given age in the surveys, which explains the unusually large range.

[19]These life tables are widely used by demographers. They use historical data on population in a number of regions to calculate population projections given any particular life expectancy and gross reproductive rate. The tables can be easily used to calculate expected sex ratios by gender, which are adjusted for sex ratio at birth to give expected sex ratio overall.

[20]The results suggest that the number of missing women in Nepal actually increases after adjusting for HBV. This is due to the fact that in the original Coale(1991) estimates the assumed sex ratio at birth is actually quite

Bangladesh, Egypt and Turkey/Syria are intermediate. The 95% confidence interval on share explained is fairly large – roughly 17% to 70% – but it is worth noting that in China even the lower bound suggests around one-third of the missing women located.

These results are very consistent with the evidence presented in Section 3. India and Pakistan present sex ratio patterns by age that are not consistent with a high sex ratio at birth and normal mortality – they clearly display excess female mortality among young children. In China, however, the pattern of mortality and the evidence on historical sex ratios supports the possibility that differences in sex ratio at birth is largely the root of the gender disparity.

In Panel B of Table 9 I compare the new estimates to the original estimates from Sen (1992). In his original work, Sen assumed that sex ratios from Sub-Saharan Africa would prevail in Asia absent intervention – in particular, he assumed the sex ratio would be 0.977. As can be seen in Table 10, relative to that estimate the HBV-adjusted estimate explains 69% of the puzzle, with a 95% confidence interval ranging from 50% to 90%. More accurately, combining both realistic mortality rates (as Coale (1991) does) *and* adjusting for the HBV-determined sex ratio at birth results in around 75 million missing women found.

Before concluding, it is interesting to note the possibility that differences in HBV levels may correlate with cross-regional variation in sex ratio in China and India. In India, data on HBV are extremely limited. However, consistent estimates of HBV rate are available from eight states (World Health Organization, 2002) and these are matched with sex ratios among children ages 0-6 from Dreze and Sen(2003). The data on HBV across China are somewhat broader; in particular, one study has done HBV testing in 60 of the roughly 2300 Chinese counties(Campbell et al, 1990), although only about 60 people were tested in each county. These HBV rates are matched with births from 1989 and 1990, as reported in the 1989 Chinese population census. There are, of course, reliability issues associated with all of these data – both with the HBV data and the sex ratio data. In addition, even in China, they cover only a small subset of the counties, implying that the analysis here should be viewed as suggestive.

high (1.059). Although the Coale(1991) ratio is lower than the HBV-adjusted sex ratio for most countries in this region, Nepal has a very low HBV rate, so the predicted sex ratio at birth is only 1.054.

Having said that, the data suggest that Hepatitis B is related to cross-regional variation in both countries. Table 10 shows a regression of sex ratio on HBV across region in India (Column 1) and China (Column 2). In both cases the relationship is positive. In the case of India the relationship is significant at the 10% level (although with only eight observations it is difficult to draw conclusions). In China, however, this relationship is significant at the 1% level (the significance of relationship is not driven by outliers – it remains strongly significant in quantile and outlier-robust regressions).

The magnitude of the coefficient in the regression on Chinese data is very similar to the magnitude of the effect seen in the overall cross country data, providing support for the relationship. Obviously more data on cross regional HBV prevalence would be necessary before drawing strong conclusions. However, the evidence presented here is consistent with the positive relationship between HBV and sex ratio, and suggests that there may well be a role for HBV in understanding cross-regional differences within these countries. It is also worth noting that in the case of China the strongly significant positive relationship provides support for the claims that HBV can explain a large share of the sex ratio bias there.

6 Discussion and Conclusion

The discussion above clearly has important impacts on the demography of the missing women puzzle. Introducing heterogeneity in hepatitis B rates into the model decreases the number of missing women by around one-half. In addition, the HBV-adjusted puzzle is focused much more heavily on India and Pakistan. This focus may help target policy interventions and better understand the dynamics of this important issue.

Having attributed such a large fraction of this problem to HBV carrier rates, it is worth acknowledging the effect of the hepatitis B vaccine on future sex ratios. Already the vaccine is widely administered across the world and, as the estimates in Section 4.3 suggest, the effects can be seen already. The estimates here indicate that as fewer and fewer carriers are giving birth, sex ratios at birth in these countries are likely to continue to decline, resulting in more female babies than in the past. With a strong tradition of son preference, one might even expect more sex-selective abortion and other interventions designed to once again increase the probability of a male child.

The introduction of the vaccine has implications outside of these countries, as well. A number of European countries are likely to also see a decrease in their sex ratio as fewer and fewer pregnancies are to HBV carriers. If one takes seriously the theory in Becker(1981) and the evidence presented in Angrist(2002) and Chiappori et al. (2002) then these changes may have serious consequences for marriage markets and labor market outcomes in these countries.

The missing women problem has historically been understood as arising from a difference in preferences – some countries value women less than others, and in these places women are mistreated to the point of excess mortality. The conclusions in this paper argue that preferences play a smaller role than previously expected, and biological differences a much larger one.

References

Acemoglu, Daron, David Autor, and David Lyle, "Women, War and Wages: The effect of Female Labor Supply on the Wage Structure at Mid-Century," *Journal of Political Economy*, 2004, *112* (3), 497–551.

Agnihotri, Satish, "Inferring Gender Bias from Mortality Data: A Discussion Note," *Journal of Development Studies*, 1999, *35* (4), 175–200.

Amini, S, MF Mahmoodi, S Andalibi, and AA Solati, "Seroepidemiology of hepatitis B, delta and human immunodeficiency virus infections in Hamadan province, Iran: a population based study," *Journal of Tropical Medicine and Hygine*, 1993, *96* (5), 277–287.

Andrade, Dahir Ramos Junior and Dahir Ramos Andrade, "The Influence of the Human Genome on Chronic Viral Hepatitis Outcome," *Review of the Institute of Tropical Medicine*, 2004.

Angrist, Josh, "How do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation," *Quarterly Journal of Economics*, 2002, *117* (3), 997–1038.

Becker, Gary, *A Treatise on the Family*, Cambridge, MA: Harvard University Press, 1981.

Blumberg, Baruch, *Hepatitis B: The Hunt for a Killer Virus*, Princeton: Princeton University Press, 2002.

Campbell, T. Colin, Junshi Chen, Chongbo Liu, Junyao Li, and Banoo Parpia, "Nonassociation of Aflatoxin with Primary Liver Cancer in a Cross-Sectional Ecological Survey in the People's Republic of China," *Cancer Research*, 1990, *50*, 6882–6893.

Cazal, P, JM Lemiare, and M Robinet-Levy, "Hepatitis B et Rapport du Masculinite," *Revue franaise de transfusion et immunohmatologie*, 1976, *19*, 577–581.

Chahnazarian, A, BS Blumberg, and WT London, "Hepatitis B and the sex ratio at birth: a comparative analysis of four populations," *Journal of Biosocial Science*, 1978, *20* (3), 357–370.

Chambliss, Rollin, "The Geographic Factor in the Human Sex Ratio at Birth," *Social Forces*, 1949, *28* (2), 190–195.

Chen, Ta, "Sex, Age, Size of Family and Density of Population," *American Journal of Sociology*, 1947, *52*, 17–24.

Chiao, Chi-Ming, "A Study of Chinese Population," *Milbank Memorial Fund Quarterly Bulletin*, 1934, *12* (2), 171–184.

Chiappori, Pierre-Andre, Bernard Fortin, and Guy Lacroix, "Marriage Market, Divorce Legislatoin and Household Labor Supply," *Journal of Political Economy*, 2002, *110* (11), 37–72.

Coale, Ansley, “Excess Female Mortality and the Balance of the Sexes: An Estimate of the Number of ‘Missing Females’,” *Population and Development Review*, 1991, *17*, 517–523.

___ **and Judith Banister**, “Five Decades of Missing Females in China,” *Demography*, 1994, *31* (3), 459–479.

___ **, Paul Demeny, and Barbara Vaughan**, *Regional Model Life Tables and Stable Populations*, Princeton: Princeton University Press, 1983.

Department of Vaccines, “Introduction of Hepatitis B Vaccine into Childhood Immunization Services,” Technical Report, World Health Organization 2001.

Dienstag, Jules, “The Epidemiology of Hepatitis B,” in Irving Millman, Toby Eisenstein, and Baruch Blumberg, eds., *Hepatitis B*, Plenum Press New York, NY 1982.

Drew, JS, BS Blumberg, and J Robert-Lamblin, “Hepatitis B virus and sex ratio of offspring in East Greenland,” *Human Biology*, 1986, *58* (1), 115–120.

___ **, WT London, BS Blumberg, and S Serjeanston**, “Hepatitis B virus and sex ratio on Kar Kar Island,” *Human Biology*, 1982, *54* (1), 123–135.

___ **,** ___ **, ET Lustbader, JE Hesser, and BS Blumberg**, “Hepatitis B Virus and the Sex Ratio of Offspring,” *Science*, 1978, *201* (4357), 687–692.

Dreze, Jean and Amartya Sen, *India: Development and Participation*, 2nd ed., Oxford: Oxford University Press, 2002.

Duong, Tran Nhu, Norio Horiike, Kojiro Michitaka, Chen Yan et al., “Comparison of Genotypes C and D of the Hepatitis B Virus in Japan: A Clinical and Molecular Biological Study,” *Journal of Medical Virology*, 2004, *72*, 551–557.

Edlund, Lena, “Son Preference, Sex Ratios and Marriage Patterns,” *Journal of Political Economy*, 1999, *107* (6), 1275–1304.

Gjerdingen, DK and V Lor, “Hepatitis B status of Hmong patients,” *Journal of the American Board of Family Practitioners*, 1997, *10* (5), 322–328.

Goldin, Claudia, “The Role of World War II in the Rise in Women’s Employment,” *American Economic Review*, 1991, *81* (4), 741–756.

Harpaz, Rafael, Brian McMahon, Harold Margolis et al., “Elimination of New Chronic Hepatitis B Virus Infections: Results of the Alaska Immunization Program,” *Journal of Infectious Disease*, 2000, *181*, 413–418.

Hayes, EB, SB Talbot, ES Matheson, HM Pressler, AB Hanna, and CA McCarthy, “Health status of pediatric refugees in Portland, ME,” *Archives of Pediatric and Adolescent Medicine*, 1998, *152* (6), 564–568.

Hesser, JE, J Economidou, and BS Blumberg, "Hepatitis B surface antigen (Australia antigen) in parents and sex ratio of offspring in a Greek population," *Human Biology*, 1975, *47* (4), 415–425.

Hull, Terence, "Recent Trends in Sex Ratio at Birth in China," *Population and Development Review*, 1990, *16* (1), 63–83.

Junhong, Chu, "Prenatal Sex Determination and Sex Selective Abortion in Rural Central China," *Population and Development Review*, 2001, *27* (2), 259–281.

Kent, Heather, "BCs Chinese migrants a healthy lot, MDs find," *Journal of the Canadian Medical Association*, 2000, *162* (2), 256.

Klasen, Stephan, "Missing Women Reconsidered," *World Development*, 1994, *22* (7), 1061–1071.

____ , "Nutrition, Health and Mortality in Sub-Saharan Africa: Is there a Gender Bias?," *Journal of Development Studies*, 1996, *32* (6), 913–932.

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

Kobayashi, Mariko, Yasuji Arase, Kenji Ikeda, Akihito Tsubota et al., "Clinical features of hepatitis B virus genotype A in Japanese patients," *Journal of Gastroenterology*, 2003, *38*, 656–662.

Kojima, Fuhito, "Economics of Infanticide," 2005. Mimeo, Harvard University.

Lazarus, John, *Sex Ratios: Concepts and Research Methods*, Cambridge University Press, 2002.

Lee, DH, JH Kim, JJ Nam, HR Kim, and HR Shin, "Epidemiological findings of hepatitis B infection based on 1998 National Health and Nutrition Survey in Korea," *Journal of Korean Medical Science*, 2002, *17* (4), 457–462.

Lin, HH, LY Wang, CT Hu, SC Huang, LC Huang, SS Lin, YM Chiang, TT Liu, and CL Chen, "Decline of hepatitis B carrier rate in vaccinated and unvaccinated subjects: sixteen years after newborn vaccination program in Taiwan," *Journal of Medical Virology*, 2003, *69* (4), 471–474.

Livadas, D, J Economidou, JE Hesser, DA Koutras, and S Hadziyannis, "Fertility and Sex Ratio of Offspring of Female HBsAg Carriers," *Journal of the Royal Society of Medicine*, 1979, *72*, 509–512.

London, W. Thomas, Richard STevens, Frances Shofer, Jean Drew, Joann Brunhofer, and Baruch Blumberg, *Viral Hepatitis*, Philadelphia, PA: The Franklin Institute Press,

Mazzur, S and TM Watson, "Excess males among siblings of Australian antigen carriers," *Nature*, 1974, *250* (461), 60–61.

McMahon, BJ, S Schoenberg, L Bulkow, RB Wainwright, MA Fitzgerald, AJ Parkinson, E Coker, and D Ritter, "Seroprevalence of hepatitis B viral markers in 52,000 Alaska Natives.," *American Journal of Epidemiology*, 1993, *138* (7), 544–549.

Miyakawa, Yuzo and Masashi Mizokami, "Classifying Hepatitis B Virus Genotypes," *Intervirology*, 2003, *46*, 329–338.

Norberg, Karen, "Partnership Status and the Human Sex Ratio at Birth," *Proceedings of the Royal Society B*, 2004, *Forthcoming.*

of China, Republic, *Taiwan-Fukien Demographic Factbook*, Minguo: Republic of China, 1974-2002,annual.

Patel, PA and MD Voigt, "Prevalence and interaction of hepatitis B and latent tuberculosis in Vietnamese immigrants to the United States," *American Journal of Gastroenterology*, 2002, *97* (5), 1198–1203.

Qian, Nancy, "Missing Women and the Price of Tea in China: The Effect of Relative Female Income on Sex Imbalance," 2005. MIT Working Paper.

Rahman, M, Amanullah, H Sattar, M Rahman, HA Rashid, and AS Mollah, "Sero-epidemiological study of hepatitis B virus infection in a village," *Bangladesh Medical Research Council Bulletin*, 1997, *23* (2), 38–41.

Rosenzweig, Mark and T. Paul Shultz, "Market Opportunities, Genetic Endowments and Intrafamily Resource Distribution: Child Survival in India," *American Economic Review*, 1982, *72* (4), 803–815.

Sen, Amartya, "Missing Women," *British Medical Journal*, 1992, *304*, 587–588.

——, "More than 100 Million Women are Missing," *New York Review of Books*, December 20, 1990.

She, SL, LY Shi, YJ Wu, ZZ Li, CZ Zheng, YP Wu, and XH Yu XH, "A seroepidemiologic study of hepatitis B virus infection among barbers in Huangshi City, Hubei, China," *Microbiology and Immunology*, 1988, *32* (2), 229–233.

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

United Nations, "Demographic Yearbook Historical Supplement," Technical Report, United Nations 1997.

Wang, CS, TT Chang, WJ Yao, and P Chou, "Comparison of hepatitis B virus and hepatitis C virus prevalence and risk factors in a community-based study," *American Journal of Tropical Medicine and Hygine*, 2002, *66* (4), 389–393.

World Health Organization, "Prevention of Hepatitis B in India: An Overview," Technical Report, World Health Organization, Regional Office for South-East Asia 2002.

Yi, Zeng, Tu Ping, Gu Baochang, Xu Yi, Li Bohua, and Li Yongping, "Causes and Implications of the Recent Increase in Reported Sex Ratio at Birth in China," *Population and Development Review*, 1993, *19* (2), 283–302.

Appendix A: Maximum Likelihood Calculations

There are six studies, detailed in Table 3, providing information on sex ratios for carriers and non-carriers of the HBV virus. For each study, we observe the share of boys born to carriers, denoted p_i and the share born to non-carriers, denoted f_i. We assume there is a constant shift parameter x that is the same across areas, so that $p_i = f_i + x$. This is equivalent to allowing for a constant shift and a location-specific underlying mean. Assume that f_i is given by the sex ratio observed for non-carriers in each study. [21] Consistent with the underlying process, assume that the number of boys born in each area is drawn from a binomial distribution with parameter p_i.

Given these parameters, the likelihood of observing b_i boys and g_i girls in study i is::

$$(1 - f_i - x)^{g_i} (f_i + x)^{b_i}$$

Taking the product over all i, we can write the overall likelihood for all studies as:

$$\mathcal{L} = \prod_{i=1}^{n} (1 - f_i - x)^{g_i} (f_i + x)^{b_i}$$

The log likelihood is therefore:

$$L = \sum_{i-1}^{n} g_i \ln(1 - f_i - x) + b_i \ln(f_i + x)$$

This is maximized over the choice of x either numerically or by solving for the value of x that satisfies:

$$\frac{dL}{dx} = \sum_{i-1}^{n} -g_i \frac{1}{1 - f_i - x} + b_i \frac{1}{f_i + x} = 0$$

The variance of the likelihood function is simply the negative expectation of the second derivative, as below:

$$\begin{aligned} \frac{d^2 L}{dx} &= \sum_{i-1}^{n} \frac{-g_i}{(1 - f_i - x)^2} - \frac{b_i}{(f_i + x)^2} \\ Var(L) &= \sum_{i-1}^{n} \frac{g_i}{(1 - f_i - x)^2} + \frac{b_i}{(f_i + x)^2} \end{aligned}$$

The variance of the estimated value of x is the inverse of this variance.

[21] The results are very similar if we assume that the overall mean for non-carriers is constant across studies.

Appendix B: Hepatitis Rate Categories

High (>8%)		**Medium** (2%-7%)		**Low** (< 2%)	
Armenia	Philippines	Albania	Romania	American Samoa	Ireland
Azerbaijan	Reunion	Algeria	Russia	Antigua	Jamaica
Bahrain	Seychelles	Bangladesh	Slovenia	Argentina	Jersey
Bolivia	Singapore	Bosnia	Spain	Australia	Latvia
Botswana	South Korea	Brazil	Sri Lanka	Austria	Lithuania
Brunei	Tadjikistan	Croatia	Suriname	Bahamas	Luxembourg
Bulgaria	Thailand	Cyprus	Syria	Belarus	Malta
Cape Verde	Togo	Dominican Rep.	Tunisia	Belgium	Martinique
China	Uzbekistan	Egypt	Turkey	Belize	Mexico
Columbia	Zimbabwe	El Salvador	UAE	Bermuda	Netherlands
Georgia		French Guiana	Venezuela	Canada	New Caledonia
Greenland		Gibraltar	Yugoslavia	Cayman Islands	New Zealand
Hong Kong		Greece		Channel Islands	Nicaragua
Israel		Guatemala		Chile	North Ireland
Jordan		Guyana		Cook Islands	Norway
Kyrgyzstan		Honduras		Costa Rica	Panama
Kzakhstan		Iran		Cuba	Paraguay
Liberia		Iraq		Czech Republic	Puerto Rico
Macaco		Italy		Czechoslovakia	Samoa
Malawi		Japan		Denmark	San Marino
Malaysia		Kuwait		Dominica	Scotland
Maldives		Lybia		Ecuador	Slovakia
Mali		Macedonia		Estonia	St Lucia
Mauritius		Morocco		Fiji	Sweden
Moldova		Pakistan		Finland	Switzerland
Mongolia		Poland		France	Tonga
North Korea		Portugal		Germany	Trinidad
Peru		Qatar		Guadeloupe	Ukraine
				Guam	United Kingdom
				Hungary	United States
				Iceland	Uruguay

Appendix C: Citations for Estimates of Hepatitis Rates

This document is separate, given the length, and can be accessed at "www.people.fas.harvard.eduẽoster"

Table 1

Historical Estimates of Sex Ratio at Birth
Data from Coale and Banister (1994)

Year of Birth	Sex Ratio at Birth (Boys/Girls)
1936-1940	1.138
1941-1945	1.114
1946-1950	1.083
1951-1955	1.094
1956-1960	1.096
1961-1965	1.065
1966-1970	1.079
1971-1975	1.073
1976-1980	1.088
1981-1985	1.086
1985-1989	1.128
Average	**1.095**

Notes: The data in the table are taken from Coale and Banister (1994), Table 3. Their data come from a retrospective fertility survey run in China in 1982.

Table 2

Offspring Sex Ratios Immigrants and Natives in the IPUMS

Panel A: Offspring Sex Ratios among Chinese Immigrants

Year	Number of Children	Sex Ratio (Boys/Girls)
1880-1930	45	1.250
1940-1970	1144	1.047
1980	1109	1.136
1990	1312	1.092
2000	2823	1.119
Average		**1.105**
Average, 1980-2000		**1.115**

Panel B: Offspring Sex Ratios among Immigrants from Low-HBV Countries and Native Whites

Other Immigrants, 1900-2000	**1.041**
Natives, 1900-2000	**1.044**

Notes: This table reports sex ratios among children ages 1-5 observed in the U.S. Census IPUMS 1% and 5% samples. Panel A is limited to children who were born in the United States to mothers who are Chinese immigrants. The first line of Panel B is limited to children whose mothers are immigrants from countries with a hepatitis B prevalence of less than 2%. The second line of Panel B is the average for children of Native whites. There is no data for 1890 or 1930.

Table 3
Hepatitis B and Sex Ratio: Individual Level Estimates

Location	HBV Status	Sons	Daughters	Sex Ratio
Greenland	Positive	64	60	1.07
Greenland	Negative	174	194	0.90
Kar Kar Island	Positive	63	54	1.17
Kar Kar Island	Negative	163	206	0.79
Greece 1	Positive	85	46	1.85
Greece 1	Negative	287	255	1.13
Philippines	Positive	66	41	1.61
Philippines	Negative	304	301	1.01
Greece 2	Positive	52	30	1.73
Greece 2	Negative	1006	955	1.05
France	Positive	20	12	1.66
France	Negative	149	122	1.22

Notes: This table shows sex ratios among the children of carrier and non-carrier parents in four regions. Data were collected by testing married women and, in all cases except for Greenland, their husbands for HBV. Detailed reproductive histories were also collected. The table represents all births to women in these samples, with generally more than one birth to each woman. The last two studies (Greece 2 and France) were designed specifically to test the hypothesis that HBV affects offspring sex ratio, and were run after the original theory was expressed.

Table 4[a]

Sex Ratio and HBV Vaccination: Alaskan Natives and Non-Natives

Dependent Variable: Child is Male

	(1)	(2)
Explanatory Variables:		
Native (0/1)	.003	.024***
	(.37)	(4.19)
Native × Year=1980	.034***	
	(6.76)	
Native × Year=1990	.019***	.006***
	(9.27)	(4.12)
Year=1980	.031***	
	(7.33)	
Year=1990	.034***	.02***
	(6.46)	(18.65)
constant	.489***	.492***
	(85.58)	(125.04)
CONTROLS: Maternal and child age dummies, birth year dummies, martial status dummy		
Number of Observations	5041	3163

Notes: The table presents results of a regression of child sex on whether the mother is an Alaskan Native and census year and interactions. The HBV vaccination program covered Natives beginning in the early 1980s. Regression is limited to children under 4. Data from the IPUMS 1980, 1990 and 2000 5% samples.

[a] t-statistics in parenthesis

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5[a]

Cross Country Hepatitis B Prevalence and Sex Ratio

Dependent Variable: Sex Ratio at Birth

	(1) All	(2) All	(3) OECD	(4) OECD
Explanatory Variables:				
Hepatitis B Prevalence (%)	.0043***	.0050***	.0070***	.0073***
	(3.96)	(3.87)	(6.43)	(5.54)
Log GDP Per Capita (1995)		0.0038		0.0020
		(.87)		(.43)
constant	1.049***	1.012***	1.052***	1.032***
	(243.83)	(23.99)	(472.32)	(22.57)
Number of Observations	26	26	16	16
R^2	.39	.43	.75	.75

Notes: HBV prevalence is calculated by aggregating published studies; studies used for each country appear in Appendix B. Only countries with more than 15,000 people tested for HBV are used in the analysis. Sex ratio is the number of boys born divided by number of girls, and data on births are from the United Nations Demographic Yearbook Historical Supplement for all available years from the 1960s through the present.

[a] t-statistics in parenthesis

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6[a]

Male Births and Hepatitis: IPUMS

Dependent Variable: Birth is Male

	(1)	(2)	(3)	(4)
Explanatory Variables:				
Hepatitis Prev. (%)	.0013***	.0013***	.0014***	.0014***
	(5.54)	(5.08)	(5.44)	(5.6)
Educ. Relative to Cohort			.00035	.00035
			(.2)	(.2)
Educ. Rel to Cohort Sq.			−.00014	−.00014
			(−.71)	(−.71)
Unmarried (0/1)				−.0035
				(−.89)
constant	.509***	.511***	.511***	.514***
	(312.95)	(10.15)	(10.15)	(10.09)
Year of Birth Dummies	YES	YES	YES	YES
Maternal Age Dummies	NO	YES	YES	YES
Number of Observations	215708	215708	197706	197706

Notes: The table presents regressions of child sex on controls, including HBV prevalence in maternal country of origin. Data used are from the IPUMS and are limited to children ages 1-6 born to immigrant mothers. Only countries with HBV prevalence estimated from at least 15,000 subjects are included, and only countries with at least 2000 births during the 20th century. Columns 3 and 4 are limited to post-1920 data.

[a] t-statistics in parenthesis

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 7

Estimates of Effect of HBV on Sex Ratio at Birth

Maximum Likelihood Estimate from Individual-Level Data		
Mean	1.55	
95% Confidence Interval	(1.32,1.84)	
Estimates from Aggregate Data		
Study	**Estimate**	**Description**
Cross Country Estimates: World	1.48	Using regression results for effect of hepatitis on sex ratios for 26 countries
Cross Country Estimates: OECD	1.75	Using regression results for effect of hepatitis on sex ratios in OECD countries
Alaskan Evidence	1.96	Using regression results from the IPUMS, assuming HBV rate of 10% among Alaskan Natives before vaccination

Notes: Maximum likelihood estimates calculated as discussed in Section 2. All other estimates calculated as in Equation (4). Discussion of the individual-level evidence is in section 4.2; discussion of the cross country evidence is in section 4.3; discussion of the Alaskan evidence is in section 4.4.

Table 8

Estimates of HBV Prevalence and Predicted Sex Ratio at Birth

Country	Estimated HBV Prevalence	Predicted Sex Ratio at Birth	Range in Empirical Sex Ratio at Birth
China	11.24%	1.102	(1.084-1.14)
India	4.33%	1.069	(1.053-1.11)
Pakistan	2.28%	1.059	(1.025-1.10)
Bangladesh	3.80%	1.066	(1.057-1.076)
Nepal	1.22%	1.054	(0.90-1.14)
West Asia	4.05%	1.068	(1.036-1.091)
Egypt	5.26%	1.073	(1.032-1.112)

Notes: HBV prevalence is calculated from published studies; studies used for each country are in Appendix B. Prevalence reported is for pregnant women in all countries but Pakistan and Nepal. Due to lack of data these are for all healthy individual. Predicted sex ratio at birth is calculated as in Equation (3) in Section 2, assuming a sex ratio of 1.53 for HBV carriers and 1.048 for non-carriers. Range in empirical sex ratio at birth is created from the Demographic Yearbook data (United Nations (1997,2001)) for countries other than India and Nepal; for those countries the range is the range of children's sex ratio across ages as reported in Demographic and Health Surveys from 1990-2000. West Asia includes Turkey and Syria.

Table 9
Re-Estimating the Missing Women

PANEL A: Comparison with Coale (1991)							
Country	Coale (1991) Population Sex Ratio	HBV-Adjusted Population Sex Ratio	Actual Population Sex Ratio	Female Population	Coale (1991) Missing	HBV-Adjusted Missing	% "Found"
China	1.010	1.051	1.066	548.7	30.42	7.75	74.53%
India	1.020	1.029	1.077	406.3	22.71	18.75	17.41%
Pakistan	1.025	1.025	1.105	40	3.12	3.12	0.00%
Bangladesh	1.025	1.032	1.064	42.2	1.61	1.31	18.68%
Nepal	1.025	1.020	1.050	7.3	0.18	0.21	0.00%
West Asia	1.030	1.038	1.060	55	1.60	1.15	28.21%
Egypt	1.020	1.034	1.047	23.5	0.62	0.30	51.80%
			TOTAL		60.26	32.59	45.91%
			95% Confidence Interval			(17.91,49.9)	(17.1%,70.3%)
PANEL B: Comparison with Sen (1992)							
Country	Sen (1992) Population Sex Ratio	HBV-Adjusted Population Sex Ratio	Actual Population Sex Ratio	Female Population	Sen (1992) Missing	HBV-Adjusted Missing	% "Found"
China	0.977	1.051	1.066	548.7	49.98	7.75	84.50%
India	0.977	1.029	1.077	406.3	41.59	18.75	54.91%
Pakistan	0.977	1.025	1.105	40	5.24	3.12	40.42%
Bangladesh	0.977	1.032	1.064	42.2	3.76	1.31	65.25%
Nepal	0.977	1.020	1.050	7.3	0.55	0.21	60.68%
West Asia	0.977	1.038	1.060	55	4.67	1.15	75.39%
Egypt	0.977	1.034	1.047	23.5	1.68	0.30	82.19%
			TOTAL		107.47	32.59	69.67%
			95% Confidence Interval			(17.91,49.9)	(53.5%,89.1%)

Notes: Coale(1991) and Sen(1992) population sex ratios taken from their papers. HBV-adjusted population sex ratio is calculated using information on sex ratio at birth from Table 9 and model life tables from Coale, Demeny and Vaughan (1983). Actual sex ratio is as reported in Coale(1991) and represents population sex ratio in the mid-1980s. The number of missing women is calculated as in Equation (1) in Section 2. Number of missing women in millions.

Table 10[a]

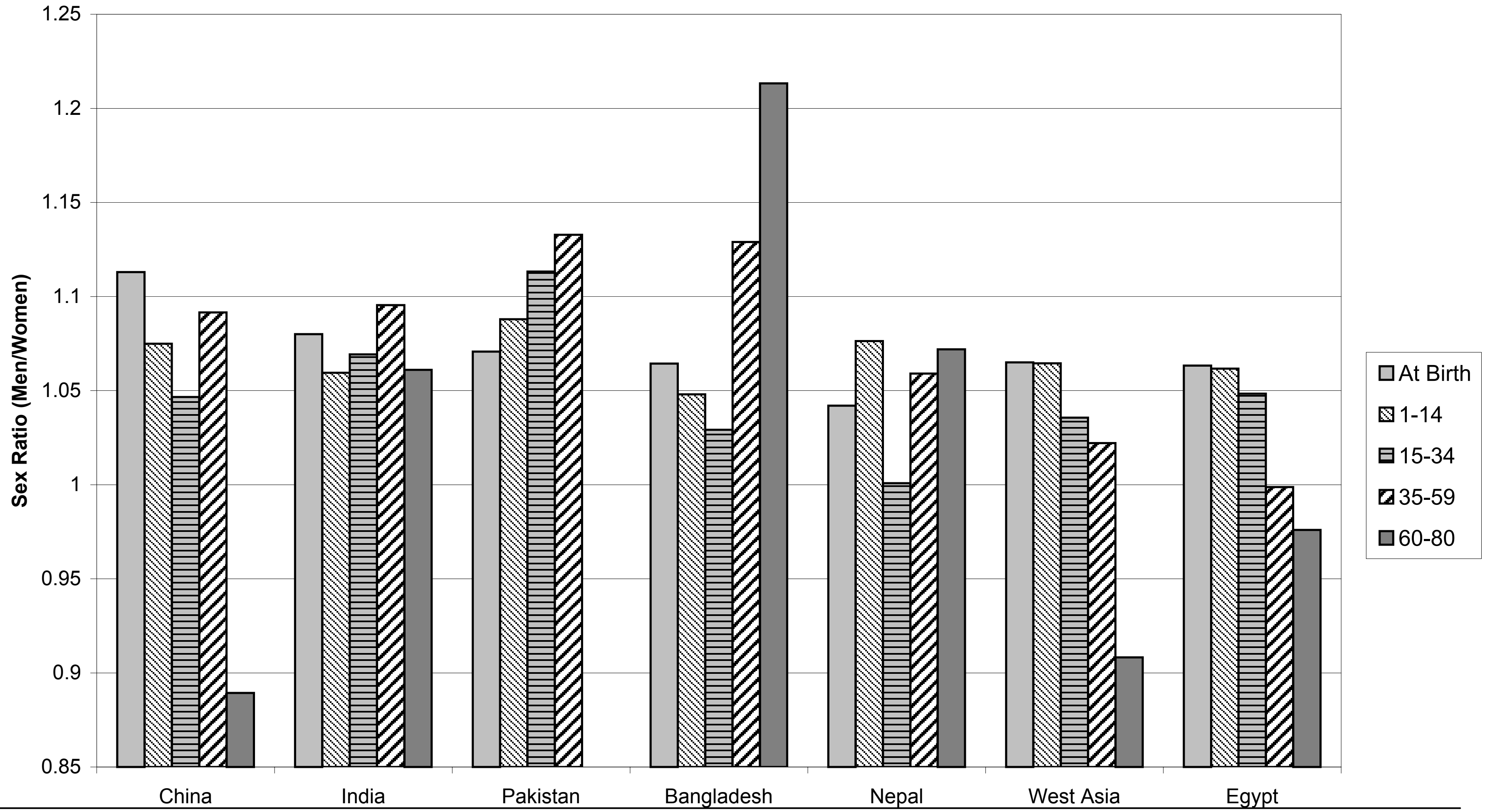
HBV and Sex Ratio Across Region: India and China

Dependent Variable	*Sex Ratio Ages 0-6*	*Sex Ratio at Birth*
	India	China
Explanatory Variables:		
Hepatitis Prevalence (%)	.016*	.005***
	(1.93)	(2.45)
constant	1.018***	1.066***
	(31.81)	(57.21)
Number of Observations	8	60

[a] t-statistics in parenthesis

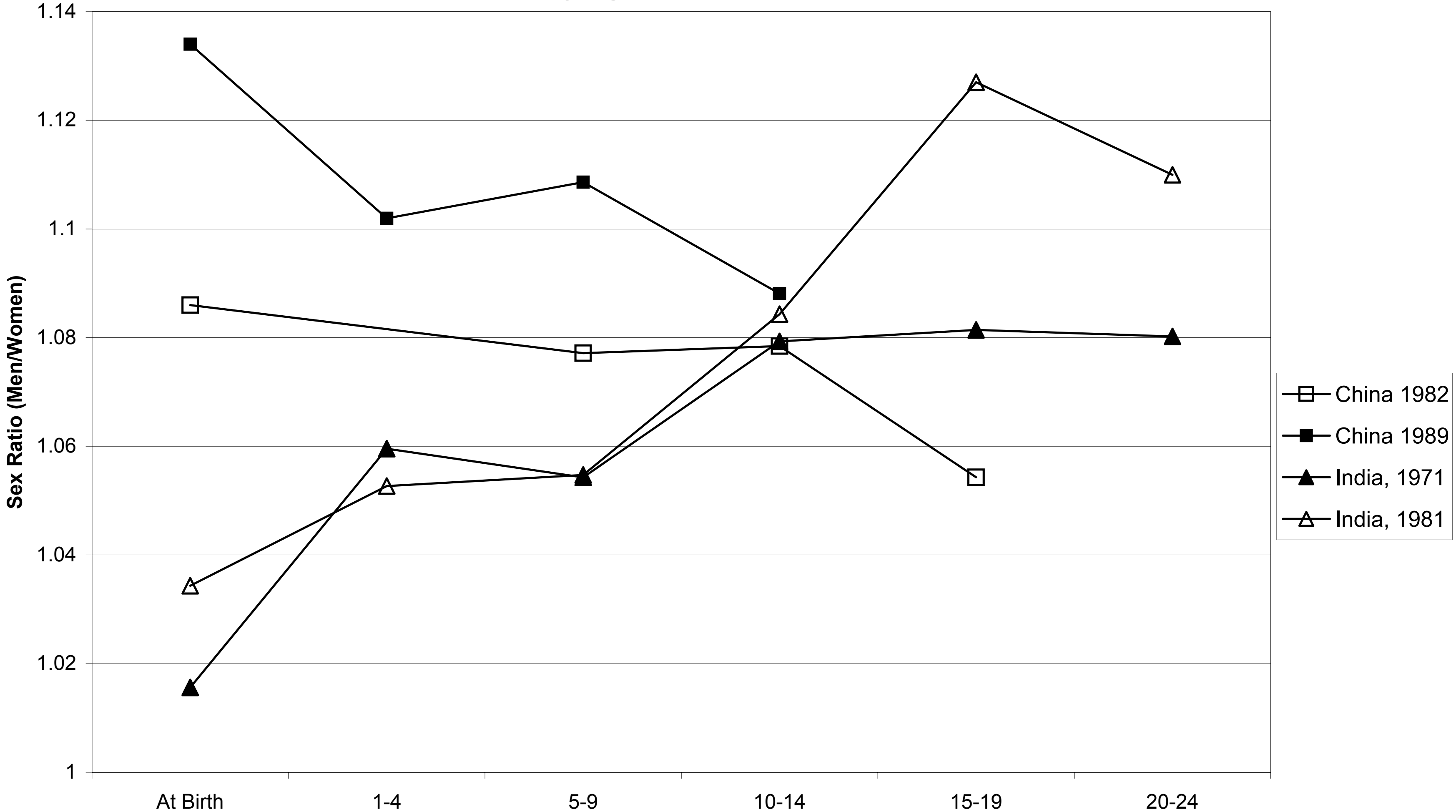
* significant at 10%; ** significant at 5%; *** significant at 1%

Figure 1
Sex Ratio by Age, Missing Women Countries



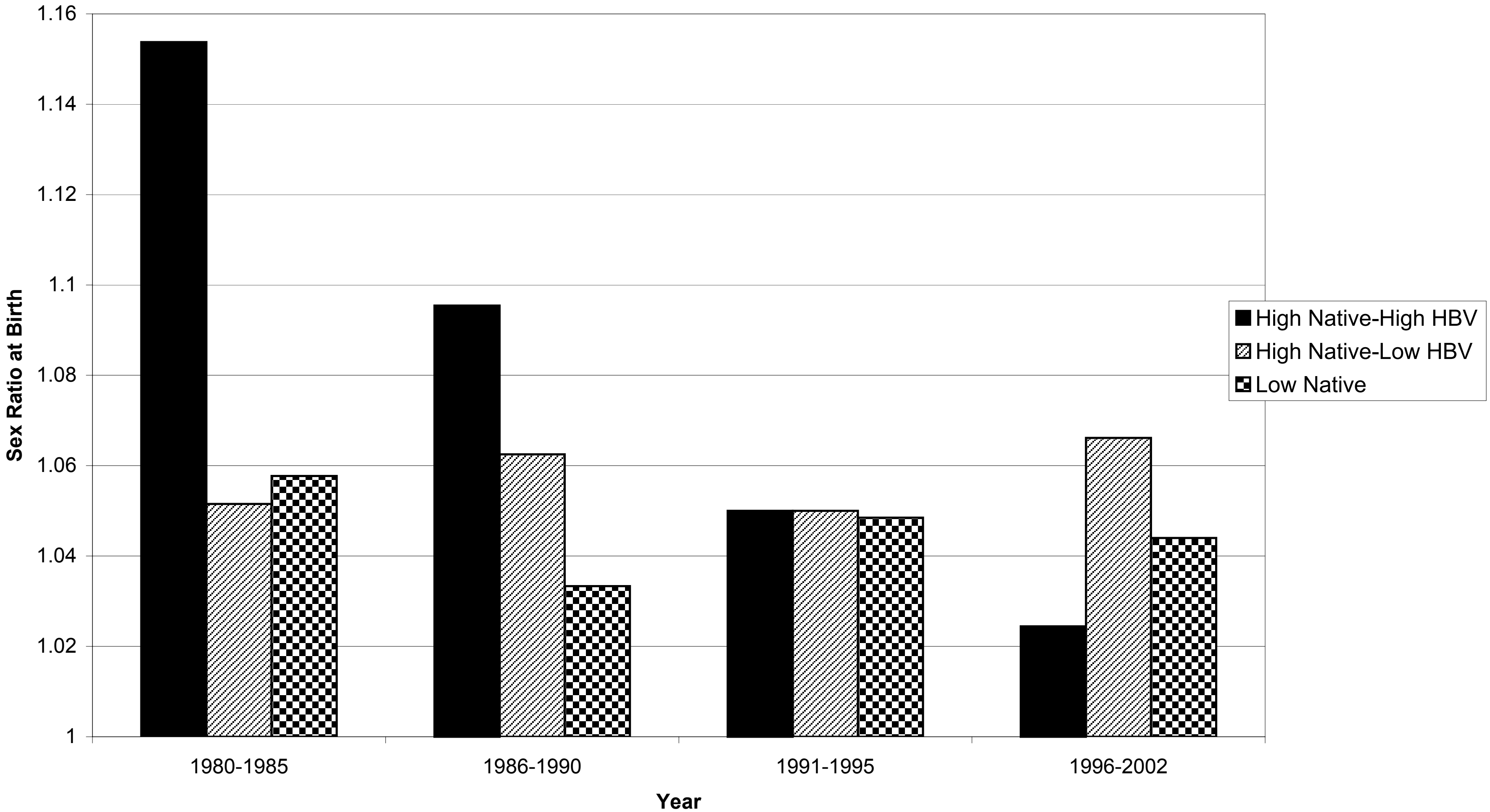
Notes: This figure presents sex ratios by age group for the missing women countries. Data on sex ratios are from the United Nations Demographic Yearbook Historical Supplement (United Nations, 1997) and represent all censuses and birth registration data between 1970 and 1990. Data on sex ratio at birth for India and Nepal is from the Demographic and Health Surveys run between 1990 and 2000 and represents reported sex ratios among chlidren ever born to survey respondents.

Figure 2
Sex Ratio by Age, Indian and Chinese Cohorts



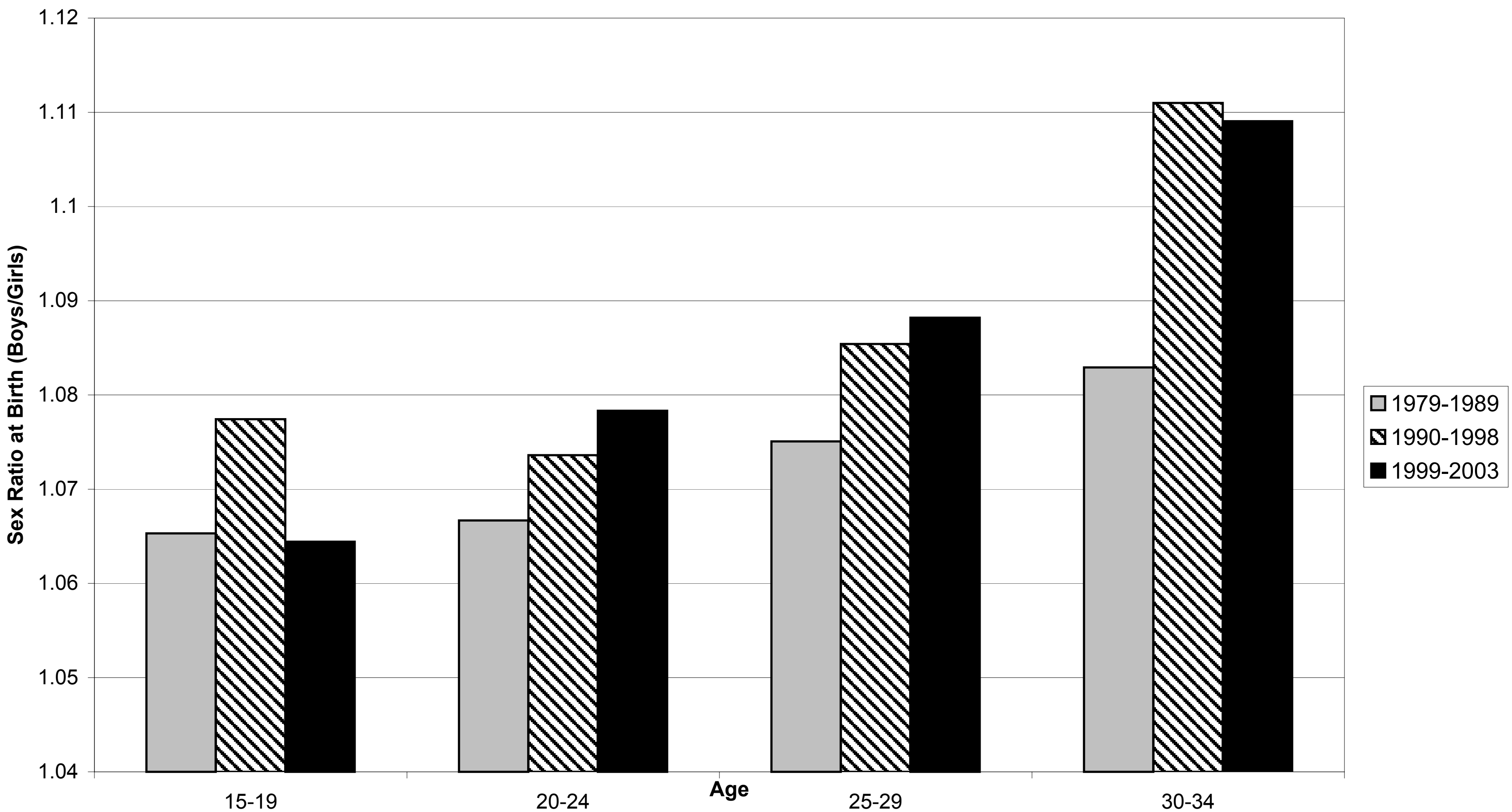
Notes: This figure presents sex ratios by age for specific birth cohorts in India and China. For example, the China 1982 line shows sex ratios at each age for the cohort born in China in 1982. Data for India is from the Demographic Yearbook Historical Supplement (United Nations, 1997) and the Demographic Yearbook (2001). Data for China is from the overall Chinese census and the Chinese 1% census.

Figure 3
Changes in Sex Ratio at birth in Alaska by Region Type



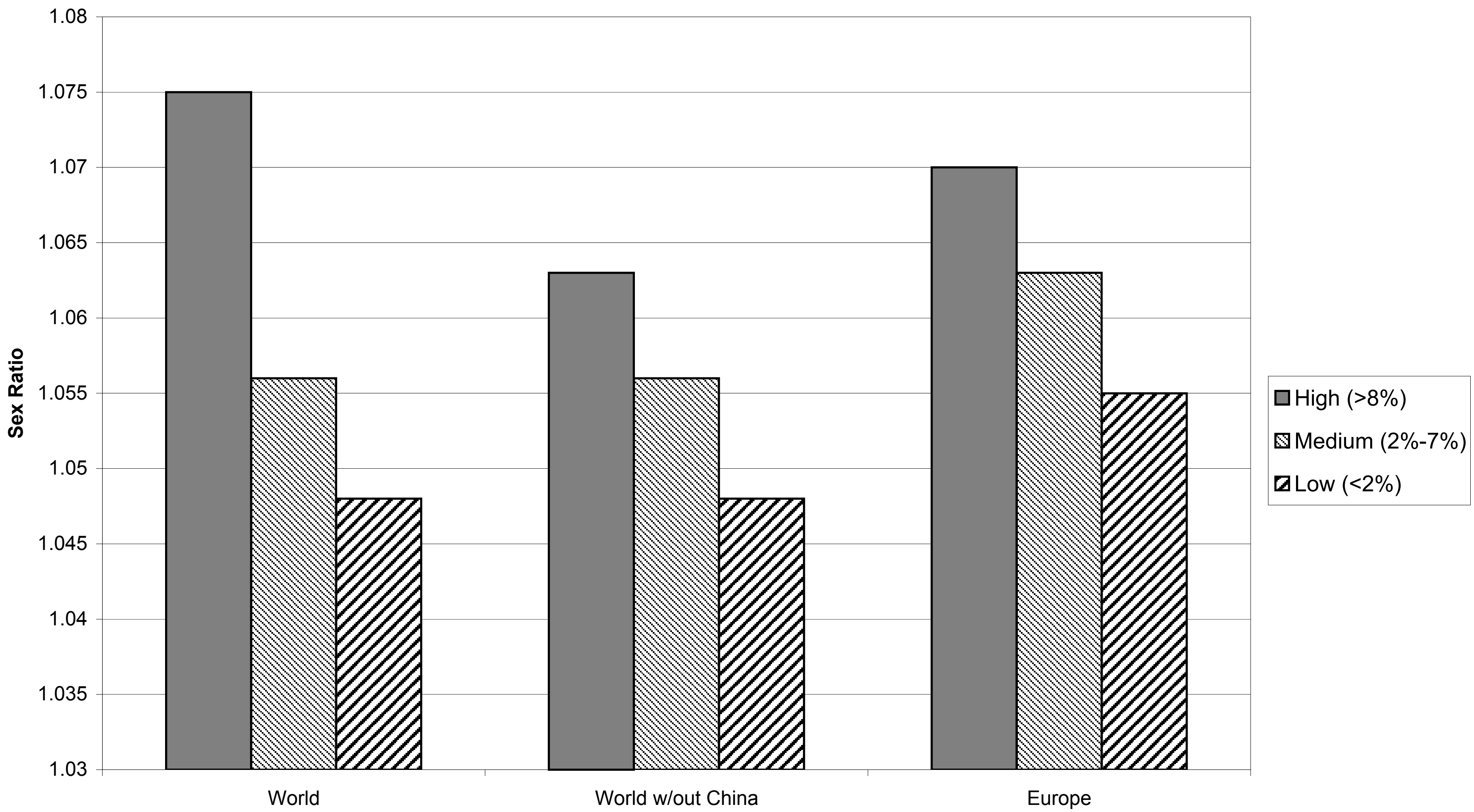
Notes: The figure compares overall sex ratios at birth across three region types: High-native, high-HBV regions (Yukon-Koyukuk, Aleutians East Borough, Lake and Penninsula, Bristol Bay), high-native, low-HBV regions (Prince of Wales, Skagway-Hoonah-Angoon, Wrangall Peterson) and a low-native region (Anchorage). Sex ratio is number of boys born divided by number of girls.

Figure 4
Changes in Sex Ratio at Birth by Cohort, Taiwan



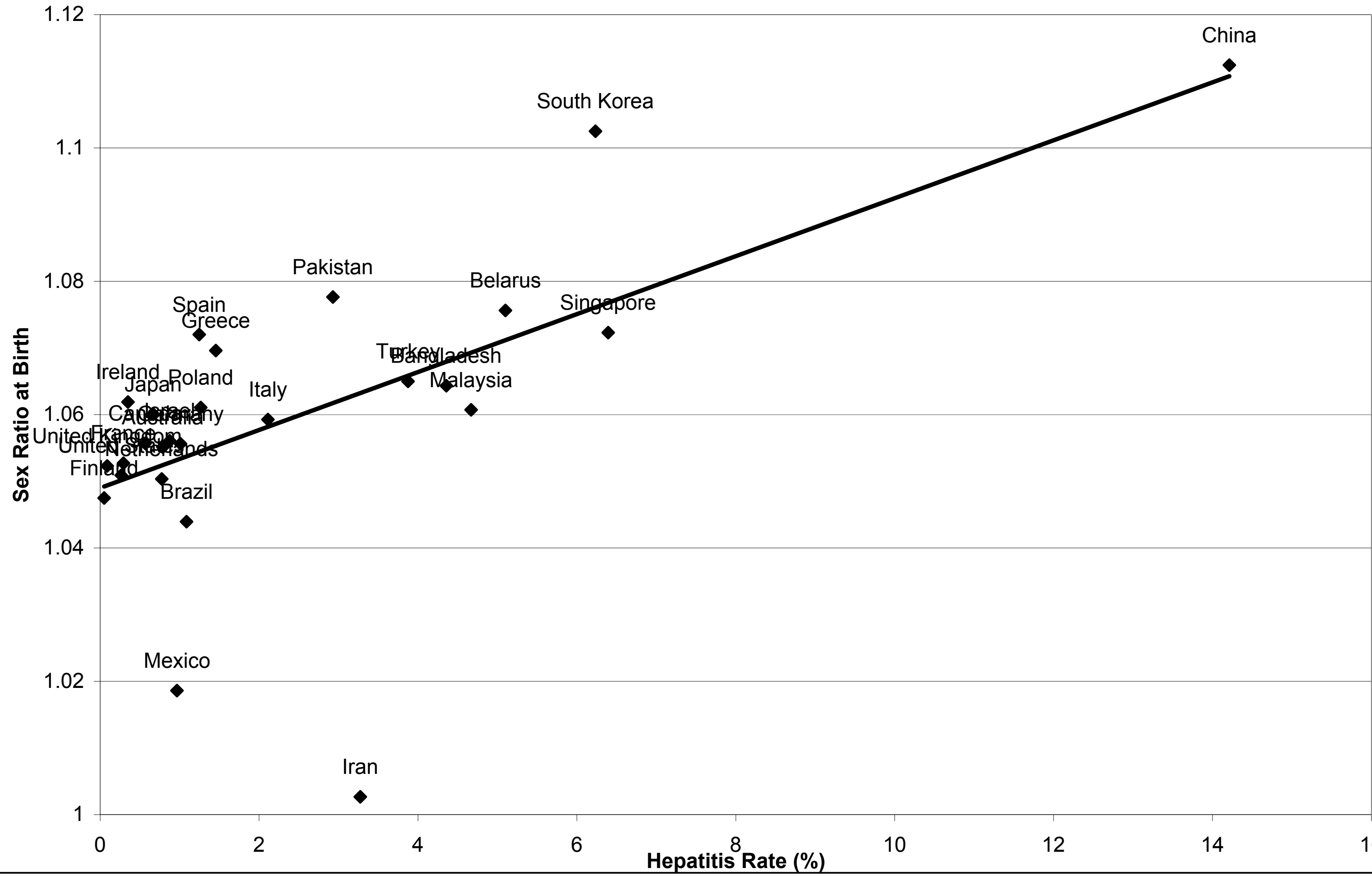
Notes: This figure represents the evolution of sex ratios at birth among four cohorts of Taiwanese women. The data a is from Taiwan-Fukien Demographic Yearbooks 1979-2003, and represents all births in Taiwan fro m1979 through 2003. Universal HBV vaccination of newborns began in July 1984, and would therefore have covered the 15-19-year old cohort after 1999.

Figure 5
Sex Ratio at Birth by WHO HBV Prevalence Categories



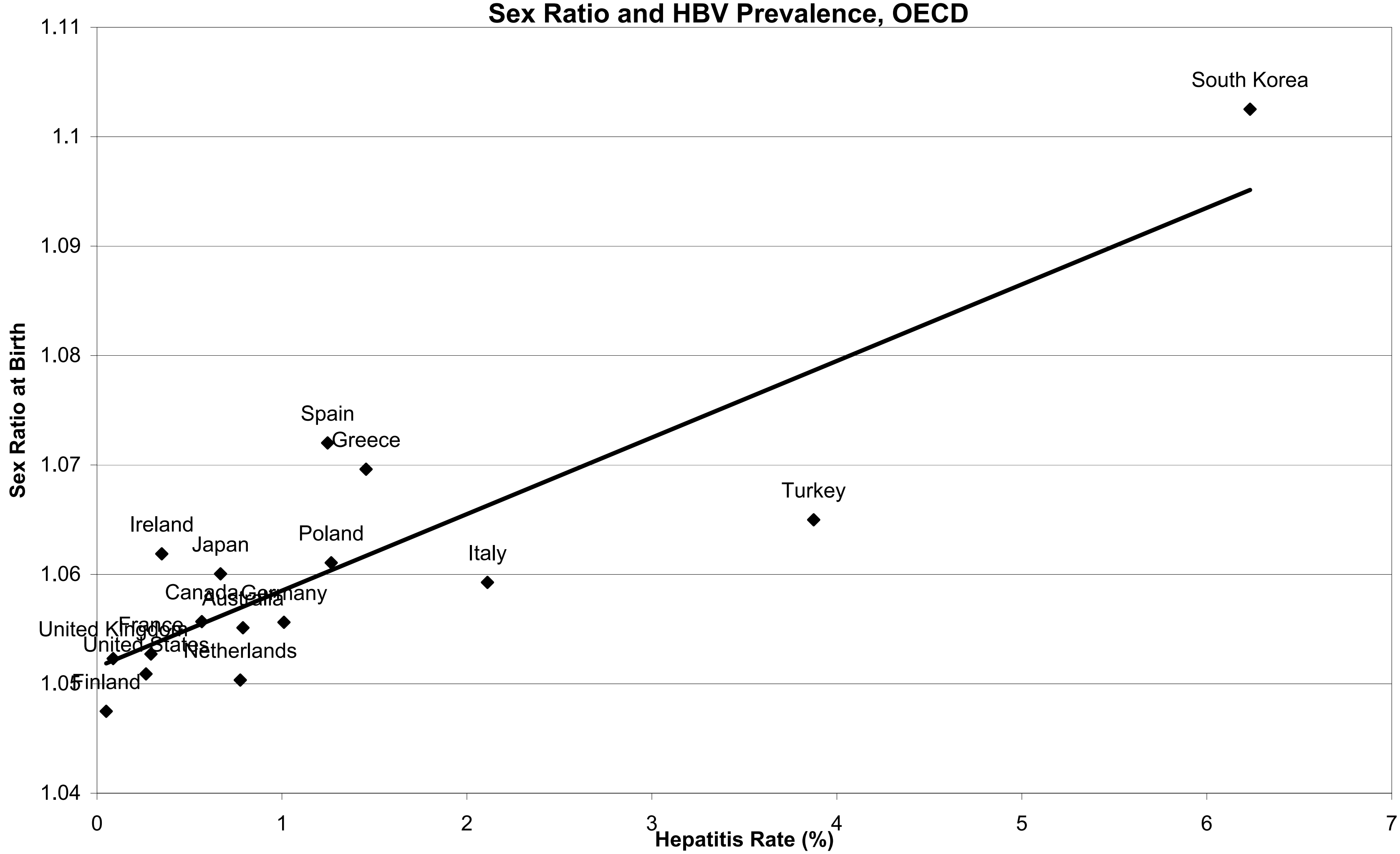
Notes: HBV prevalence information is from the WHO, and countries in each category are in Appendix A. Sex ratio is the number of boys born divided by number of girls, and data on births is from the United Nations Demographic Yearbook Historical Supplement. Sex ratio in each group is total number of boys born in that group divided by total number of girls

Figure 6
Sex Ratio and HBV Prevalence



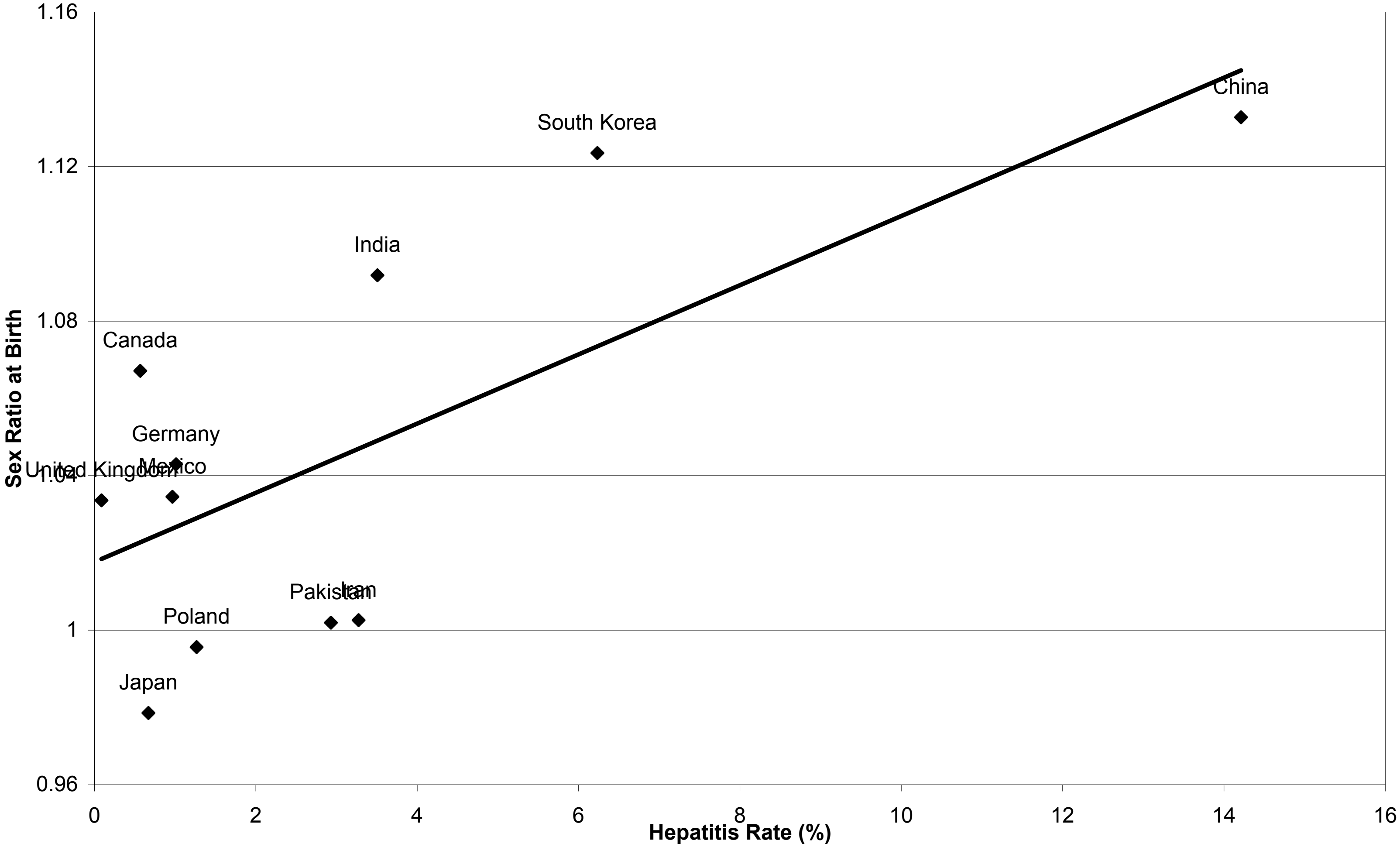
Notes: HBV prevalence is calculated by aggregating published studies; studies for each country appear in Appendix B. Only countries with more than 15,000 people used to calculate HBV prevalence are included. Sex ratio is the number of boys born divided by number of girls, and data on births is from the United Nations Demographic Yearbook Historical Supplement for all available years from the 1960s through the present.

Figure 7
Sex Ratio and HBV Prevalence, OECD



Notes: HBV prevalence is calculated by aggregating published studies; studies for each country appear in Appendix B. Only OECD countries with more than 15,000 people used to calculate HBV prevalence are included. Sex ratio is the number of boys born divided by number of girls, and data on births is from the United Nations Demographic Yearbook Historical Supplement for all available years from the 1960s through the present.

Figure 8
Sex Ratio and HBV, Data on US Immigrants from the IPUMS



Notes: HBV prevalence is calculated by aggregating published studies; studies for each country appear in Appendix B. Only OECD countries with more than 15,000 people used to calculate HBV prevalence are included. Sex ratio is number of boys for each girl and is calculated for each country using children (ages 1-6) of immigrant parents recorded in the IPUMS 5% samples for 1990 and 2000.