Hepatitis B Does Not Explain Male-Biased Sex Ratios in China

Emily Oster[*]
University of Chicago and NBER

Gang Chen
Hepatitis B Foundation

Xinsen Yu
Haimen CDC

Wenyao Lin
Haimen CDC

Draft: April 10, 2008

Abstract

In earlier work (Oster, 2005) I have argued, based on existing medical literature and new analysis of cross country data and vaccination programs, that parents who are carriers of hepatitis B have a higher offspring sex ratio (more boys) than non-carrier parents. Further, since a number of Asian countries, China in particular, have high hepatitis B carrier rates, I suggested that hepatitis B could explain a large share – approximately 50% – of Asia's "missing women". Subsequent work has questioned this conclusion. Most notably, Lin and Luoh (2008) use data from a large cohort of births in Taiwan and find only a very tiny effect of *maternal* hepatitis carrier status on offspring sex ratio. Although their work is quite compelling for the case of mothers, it leaves open the possibility that *paternal* carrier status is driving higher sex offspring sex ratios. Following up on this possibility, I collected data on the offspring gender for a cohort of 67,000 people in China who are being observed in a prospective cohort study of liver cancer; approximately 15% of these individuals are hepatitis B carriers. In this sample, I find no effect of either maternal or paternal hepatitis B carrier status on offspring sex. Carrier parents are no more likely to have male children than non-carrier parents. This finding leads me to conclude that hepatitis B cannot explain skewed sex ratios in China, and the conclusions about this in Oster (2005) were incorrect.

1 Introduction

In an earlier paper (Oster, 2005) I argued that a high rates of hepatitis B carrier status in China explained a large share of the male-biased gender ratio – the "missing women" phenomenon – in that country and, more generally, that high rates of hepatitis in a number of Asian countries explained some of the skewed gender ratios there, as well. This conclusion was based first on evidence from an existing medical literature suggesting that families where at least one parent is a carrier of the hepatitis B virus had a larger share of male children (Hesser, Economidou and Blumberg, 1975; Drew, London, Blumberg and Serjeanston, 1982; Drew, Blumberg and Robert-Lamblin, 1986; Chahnazarian, Blumberg and London, 1988; Cazal, Lemiare and Robinet-Levy, 1976; Livadas et al,

---

1979). Oster (2005) supplemented this existing evidence with evidence on vaccination campaigns, notably in Alaska, and cross country evidence suggesting that countries with higher hepatitis B carrier rates have higher male-to-female ratios. In summary, the evidence suggested that families with a hepatitis B carrier parent had children that were 60% male, versus 51% in the overall population.

Subsequent to that research, a number of paper have emerged questioning the conclusions in Oster (2005). Several papers have questioned the evidence on skewed sex ratios in China, arguing that these skewed ratios, at least in recent years, appear only at higher order births, which may be inconsistent with the hepatitis B hypothesis (Das Gupta, 2005; Ebenstein, 2007). Most notably, however, Lin and Luoh (2008) have used a new, individual-level, dataset to test the underlying premise that hepatitis B carrier status is related to sex ratio. In a sample of approximately 3 million births in Taiwan, they find that women who are carriers of hepatitis B are only very slightly more likely to have male children than non-carriers. The effect is approximately 0.25 percentage points, which is clearly a tiny fraction of the 9 percentage point effect suggested in Oster (2005).

The dataset used by Lin and Luoh (2008) clearly has many benefits relative to the data relied on in Oster (2005). The original individual-level studies reported on in the medical literature were quite small (most only a few hundred or thousand children) and were not run in East Asia. The Lin and Luoh (2008) data has the advantage of being a very large sample and being specifically applicable to the Chinese context, where the missing women conclusions were most quantitatively important. The one significant downside of the Lin and Luoh (2008) data is that they only observe *maternal* infection. The original data, as well as the population-level analysis in Oster (2005) focus on infection at the family or population level. Taken together, the two data points – the original work and the Lin and Luoh (2008) data – leave open the possibility that *paternal* hepatitis B infection could be responsible for skewed gender ratios. Indeed, in a re-analysis of the original data from Greece and the Philippines (reported in detail in Chen, Oster, Yu and Lin (2008)) we see that paternal carrier status, not maternal, is associated with higher offspring sex ratios.

In light of this remaining question, we have collected data on approximately 67,000 individuals in China, taking advantage of a cohort of individuals already taking part in a study of liver cancer and the effect of hepatitis B on health (this cohort is reported on in more detail in London et al, 1995 and Evans et al, 2002). From earlier surveys of this cohort, data was available on hepatitis B carrier status (approximately 15% of individuals are carriers). Our survey, described in more detail in Section 2, collected data on the gender of offspring. Since the sample includes both

men and women, and in some cases links spouses together, it is possible to estimate the effect of both maternal and paternal hepatitis B carrier status.

Using this data, we find no effect of either paternal or maternal hepatitis B carrier status on offspring sex ratio. Estimate of the effects on gender ratio are extremely close to zero, inconsistently signed and the standard errors are small. This is true with and without controls for age, township and child birth order. This demonstrates that hepatitis B carrier status cannot explain male-biased sex ratios in China. By extension, the conclusion in Oster(2005) that hepatitis B can explain a large share of the missing women in China is not correct. This finding is consistent with the Lin and Luoh (2008) paper, but also provides further evidence by closing off the possibility that paternal carrier status is involved. The rest of this brief note reports in more detail on these findings. Section 2 describes the survey run in China, Section 3 presents our results on sex ratios and Section 4 concludes.

2   Survey Description

This study takes advantage of an existing cohort study of individuals in China who are being followed in order to evaluate he impact of chronic hepatitis B status on health, focusing on liver cancer. Detailed data on the enrollment and study of this cohort can be found elsewhere (Evans et al, 2002; London et al, 1995). This is a large, prospective cohort study of adult residents in rural areas around Haimen City in the Jiangsu Province. Subjects were originally recruited in 1992, and have been followed since then. At enrollment, individuals completed a one-page questionnaire with details on age, behaviors (smoking, alcohol consumption) and family history (particularly, history of liver cancer). Individuals also gave blood, which was tested for hepatitis B carrier status (and in some cases other hepatitis B markers). The original cohort contained 90,836 adult residents of this area. Although follow-up has been done periodically with these subjects, they have never been asked details of their offspring.

Data on children's gender was collected by the Haimen CDC in the period between December, 2007 and February, 2008, for 67,511 of the original subjects. Re-surveying all of the individuals would have been prohibitively expensive (and additionally difficult since approximately 10% of them had died). As an alternative, we asked officials in each village to report on the gender of the offspring of cohort members who lived in their village. In each village, we invited the village head, party secretary of the village, the village doctor, the village accountant and any other village

administrative staff to come together and provide data on cohort members' children. They were given a list of cohort participants in their village and asked to report the gender and status (alive, dead) for any children born to these individuals, by birth order. Observation of these meetings suggests that the administrators made use of official records of births and families, as well as personal recollection. In addition, an extremely large share of individuals in the cohort had telephones, and in cases where they were unsure, the administrators could phone the cohort member themselves.

In addition to this, we collected quality control data on 2,590 participants. This quality control focused on cohort members with 4 or more children, since pilot surveys suggested this was the group for which administrators were most likely to make errors. The quality control data also asked details about abortions or miscarriages that the respondent was aware of, as well as gender and current status (alive, dead) of live-born offspring. Comparing the quality control to the larger dataset, the surveyors found a 2.82% error rate in the first round of data provided by the administrators.[1] In situations where there were mis-matches (i.e. the quality control did not mach with the officially provided data) the data was sent back to the village administrators and they were asked to re-check everything. If necessary, this process was repeated more than once to ensure maximum overall data quality. In the results section below, we will report results both for the overall sample and for the quality control data alone. We will also report results for various sub-samples (younger people, those with fewer children, etc) for whom we think errors are even less likely.

Table 1 reports summary statistics for the overall sample, and the quality control. Sixty-four percent of the overall sample is male (representing a purposeful over sampling of the original survey, since men are more likely to be hepatitis B carriers) and 15.4% of them are carriers of hepatitis B. The average individual has 1.0 male children, and 0.95 female children, reflecting an overall slightly skewed sex ratio in the population. Individuals in the sample range in age from 34 to 84, with an average of 58. Among the sample, approximately 10% of respondents had died by the most recent survey[2] and 1.8% of the children of sample participants have died. Finally, for 13% of the sample we also have their spouse included in the cohort. This will allow us to estimate the effect of maternal and paternal carrier status at the same time, for some share of the sample.

The quality control sample is slightly skewed towards more men (73%) and has a slightly smaller share of hepatitis carriers (13%). The sample has more children, on average, due to the

---

[1]It is worth noting the error rate would likely be lower if they had not specifically focused on quality control for the group where they thought there was most likely to be mistakes.

[2]Because of the data collection technique, it was possible to get information on offspring even of people who had died. Of course, quality control was not possible with this sub-sample.

method of selection, and is older. The older age reflects the sample selection: because of the one-child rule, younger people typically do not have 4 or more children. Death rates among cohort members in this sample are zero, since only living people could be contacted, but the death rate among their offspring is higher, at 5.7%. This also reflects the selection of larger families and older parents: among families of a similar size in the overall sample, death rates are similar.

3   Results: Effect of Hepatitis B on Offspring Sex Ratio

Figure 1 shows the primary results in the paper, focusing on the overall sample. The first set of columns focus on paternal carrier status: among men who are carriers, 51.06% of their offspring are boys. Among non-carriers, this is 51.56%. For mothers, the numbers are 51.58% and 51.43%. The final set of columns compares children for whom both parents are carriers to children for whom neither parent is a carrier; the share male among offspring are 49.45% and 50.68%, respectively. This graph suggests there is no substantial difference in child gender by parental hepatitis B carrier status.

Table 2 reports regression results focusing on the overall sample. Panel A shows the relationship between gender and parental hepatitis carrier status with no controls (corresponding exactly to Figure 1). As expected based on Figure 1, neither paternal nor maternal carrier status significantly impacts child gender, and including them jointly shows a similar result. Panel B of Table 2 includes a simple set of controls: child birth order, number of children in the family, a quadratic in parent age and fixed effects for township (to capture location or income effects)[3] . Again, we see no effect of carrier status on gender. We do see some effect of birth order – consistent with other work on parity, later birth order children are more likely to be boys. We also see a negative effect of number of children – larger families have fewer boys. This is consistent with a gender-biased stopping rule, which may well be in play here.

One concern with these data, as discussed, is the possibility of errors based on the data collection process; for example, if village leaders are unable to recall details related to residents' children. Although there is no reason to think this would be bias the results in any particular direction, we re-run the analysis with two sub-samples for whom we think this misreporting is less likely. First, in Panel A of Table 3 we limit the sample to individuals who are under 60 and who are alive, for whom we think recall by others is likely to be easier. We see no effect of hepatitis status on gender for this group. Second, in Panel B of Table 3 we limit to families with only one child, where

---

[3]We have also experimented with controlling for other behavioral and health measures – alcohol or smoking abuse, exposure to pesticides, family history of liver cancer – and these do not change the result.

the gender of the child is likely to be salient and there is no issue with mis-reporting gender by birth order. Again, we see no positive effect in this sub-sample; there appears to be some negative effect of maternal hepatitis carrier status on gender in the sample for which we observe both parents, but this is a very small sample, and this result does not seem to be generally robust.

Table 4 replicates the regressions above using the quality control sample. In this case, we focus on using all of the quality control data, since the sample size is limited to begin with. In addition, because of the limited sample size it is not possible to look at the effect of maternal and paternal infection at the same time. However, in the samples separated by gender, we see confirmation of the results in Tables 2 and 3: no effect of parent carrier status on child gender.

4 Discussion and Conclusion

Tables 2 through 4 and Figure 1 all point to a similar conclusion: in this sample we see no effect of hepatitis B carrier status on offspring gender. Given this result, we conclude that hepatitis B carrier rates cannot explain male-biased sex ratios in the China. The ultimate conclusion of my original paper (Oster, 2005) – that hepatitis B explained a large share of the "missing women" in China – is not supported. The evidence here is, of course, consistent with the finding in Lin and Luoh (2008). However, it takes a significant further step in allowing us to estimate the effect of *both* paternal and maternal carrier rates on offspring gender.

An important remaining issue is whether it is possible to reconcile the biological results in the original paper (Oster, 2005) with these results and, in particular, how the individual-level data from outside of China and the evidence from vaccination campaigns in Alaska can coexist with the new results from China. In Chen, Oster Yu and Lin (2008) we address this issue in more detail. We re-visit the original individual-level data and continue to find support for the connection between, in particular, *paternal* hepatitis B carrier status and offspring sex ratio. Moreover, in the data from China discussed here, we also find some interaction between hepatitis B, gender and fertility: women with the hepatitis B e antigen (carriers who are also replicating an additional viral antigen) seem to have fewer male children. Further, women who are carriers of the virus have fewer children overall, even with extensive controls. Together, this evidence suggests that there may still be some interaction between hepatitis B and fertility outcomes (in general) but that clearly the pathways are much more complicated that the simple carrier-male offspring connection. One possibility is that the interaction differs with different genotypes of the hepatitis B virus, but that is not something we can

provide definitively with the existing data. From the perspective of the missing women issue that is the focus of Oster (2005), however, the relevant question is whether the hepatitis B carrier rates can explain the gender imbalance in the Chinese population, which it seems clear from the evidence presented here, as well as the Lin and Luoh (2008) work, that they cannot.

References

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, Anouch, Baruch Blumberg, and W. Thomas London, "Hepatitis B and the Sex Ratio at Birth: a Comparative Analysis of Four Populations.," *Journal of Biosocial Science*, July 1988, *20* (3), 357–370.

Chen, Gang, Emily Oster, Xinsen Yu, and Wenyao Lin, "Hepatitis B and Sex Ratios at Birth: Re-visiting the Relationship," *Mimeo, University of Chicago*, 2008.

Das Gupta, Monica, "Explaining Asia's "Missing Women": A New Look at the Data," *Population and Development Review*, September 2005, *31* (3), 529–535.

Drew, Jean, Baruch Blumberg, and J Robert-Lamblin, ""Hepatitis B Virus and Sex Ratio of Offspring in East Greenland."," *Human Biology*, February 1986, *58* (1), 115–120.

⎯⎯ , **W. Thomas London, Baruch Blumberg, and Susan Serjeanston**, ""Hepatitis B Virus and Sex Ratio on Kar Kar Island."," *Human Biology*, February 1982, *54* (1), 123–135.

Ebenstein, Avraham, "Fertility Choices and Sex Selection in Asia: Analysis and Policy," *Mimeo, University of California Berkeley*, 2007.

Evans, Alison, Gang Chen, Eric Ross, Fu-Min Shen, Wen-Yao Lin, and W. Thomas London, "Eight-Year Follow-Up of the 90,000-Person Haimen City Cohort: I. Hepatocellular Carcinoma Mortality, Risk Factors and Gender Differences," *Cancer Epidemiology, Biomarkers and Prevention*, 2002, *11*, 369–376.

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*, December 1975, *47* (4), 415–425.

Lin, Ming-Jen and Ming-Ching Luoh, "Can Hepatitis B Mothers Account for the Number of Missing Women? Evidence from Three Million Newborns in Taiwan," *American Economic Review, forthcoming*, 2008.

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*, July 1979, *72*, 509–512.

London, W. Thomas, Alison Evans, K McGlynn, K Buetow, P An, L Gao, Eric Lustbader, Eric Ross, Gang Chen, and Fu-Min Shen, "Viral, Host and Enviromental Risk Factors in Hepatocellular Carcinoma: A Prospective Study in Haimen City, China," *Interviology*, 1995, *38*, 155–161.

Oster, Emily, "Hepatitis B and the Case of the Missing Women," *Journal of Political Economy*, December 2005, *113* (6), 1163–1216.

Table 1. *Summary Statistics on Haimen City Cohort*

	Entire Sample	*Quality Control Sample*
Sample Size	67,511	2,590
Respondent is Male	64.7%	72.7%
Respondent Age	58.4	70.4
Respondent Hepatitis B Carrier	15.4%	13.0%
Respondent is still Alive	90.9%	100%
Number of Male Children	1.00	2.16
Number of Female Children	0.95	2.31
Share of Children Dead	1.77%	5.70%
Share with Spouse Also Observed	13.0%	10.1%

Notes: This table reports summary statistics for the overall sample, and the quality control sample. Quality control was focused on families of 4 or more children.

Table 2. *Gender and Parental Hepatitis B: Full Sample*

	Panel A: No Controls		
	Dependent Variable: Child is Male		
Explanatory Variables:			
Father HBV Carrier	−.005		.0057
	(.005)		(.011)
Mother HBV Carrier		−.0015	−.007
		(.006)	(.01)
constant	.516***	.516***	.507***
	(.002)	(.002)	(.004)
Number of Observations	94,800	56,386	19,152
	Panel B: With Controls		
	Dependent Variable: Child is Male		
Father HBV Carrier	−.004		.0072
	(.005)		(.011)
Mother HBV Carrier		−.0019	−.0053
		(.006)	(.01)
Parent Age	−.0053***	−.0071***	−.0054
	(.002)	(.003)	(.005)
Parent Age Sq.	.0001***	.0001***	.0001
	(.000)	(.000)	(.000)
Birth Order = 2	.0252***	.0275***	.0298**
	(.004)	(.005)	(.009)
Birth Order = 3	.0349***	.0414***	.0257***
	(.006)	(.008)	(.012)
Birht Order = 4 +	.0742***	.093***	.1152**
	(.007)	(.01)	(.015)
Number of Kids	−.0435***	−.0514***	−.0569***
	(.002)	(.003)	(.004)
constant	.633***	.682***	.609***
	(.059)	(.075)	(.154)
Township Fixed Effects	YES	YES	YES
Number of Observations	94,800	56,386	19,152

standard errors in parentheses

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

Notes: This table reports effects of parental hepatitis B status on child gender in the Chinese sample. Column 1 includes all children for whom we observe paternal hepatitis; Column 2 includes all children for whom we observe maternal hepatitis. Column 3 includes only those children (a smaller sample) for whom both parents are in the cohort, so we see both of their hepatitis carrier status.

**Table 3.** *Gender and Parental Hepatitis B: Sub-Samples*

Panel A: Parents Under 60 and Living			
Dependent Variable: Child is Male			
Explanatory Variables:			
Father HBV Carrier	−.0112		−.0107
	(.008)		(.017)
Mother HBV Carrier		−.0058	−.0044
		(.009)	(.017)
Controls (in all columns): Parental Age (linear and squared), birth order (first, second, third, 4+), total number of children in family, constant, township fixed effects.			
Number of Observations	29,722	23,007	6650
Panel B: Single Children Only			
Dependent Variable: Child is Male			
Father HBV Carrier	−.0083		−.013
	(.009)		(.025)
Mother HBV Carrier		−.0162	−.0499**
		(.012)	(.024)
Controls (in all columns): Parental Age (linear and squared), constant, township fixed effects.			
Number of Observations	19,778	12,043	3148

standard errors in parentheses

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

Notes: This table reports effects of parental hepatitis B status on child gender in sub-samples of the overall surveyed population: Panel A focuses on children of parents who are under 60 and still alive and Panel B focuses on only one-child families. Column 1 includes all children for whom we observe paternal hepatitis; Column 2 includes all children for whom we observe maternal hepatitis. Column 3 includes only those children (a smaller sample) for whom both parents are in the cohort, so we see both of their hepatitis carrier status.

Table 4. *Gender and Parental Hepatitis B: Quality Control Data*

	Dependent Variable: Child is Male	
Explanatory Variables:		
Father HBV Carrier	.0227	
	(.017)	
Mother HBV Carrier		−.0192
		(.025)
Parental Age	−.0172	−.016
	(.016)	(.022)
Parental Age Sq.	.0002	.0002
	(0)	(0)
Birth Order = 2	.0077	.0368
	(.016)	(.025)
Birth Order = 3	.0206	.0603**
	(.016)	(.027)
Birth Order = 4+	.0267*	.0592**
	(.015)	(.024)
constant	.918*	.786
	(.547)	(.759)
Number of Observations	8433	3175
R^2	.01	.02

standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Notes: This table reports effects of parental hepatitis B status on child gender in the quality controls sample for which individual interviews were conducted. Column 1 includes all children for whom we observe paternal hepatitis; Column 2 includes all children for whom we observe maternal hepatitis.

Figure 1:
Hepatitis B Carrier Status and Share Male Among Offspring
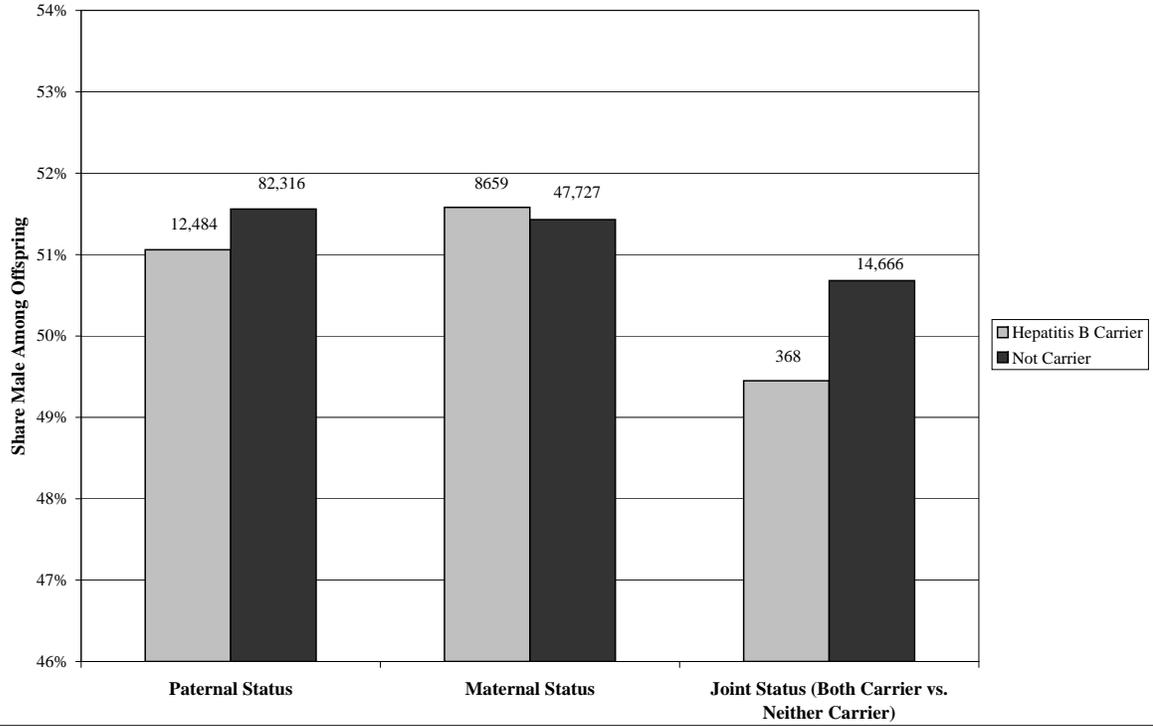
This figure shows the share male among the offspring of Hepatitis B carriers and non-carriers. The number of offspring in each cell are also reported.