

From Missing Women to Too Many Men:

Examining consequences of sex imbalance and approaches to understanding the phenomenon.

Tori Haukland Løge



Master Thesis at the Institute of Economics, Faculty of Social Sciences

UNIVERSITETET I OSLO

10.01.2013

From Missing Women to Too Many Men:

Examining consequences of sex imbalance and approaches to understanding the phenomenon.

"不孝有三，无后为大"

("There are three ways to be unfilial; having no sons is the worst")

– **孟子 (Mencius, Chinese Philosopher)**

"Historically, societies in which men substantially outnumber women are not nice places to live(...)Often they are unstable. Sometimes they are violent."

– **Mara Hvistendahl, Science Magazine**

© Tori Haukland Løge

2013

From Missing Women to Too Many Men

Tori Haukland Løge

<http://www.duo.uio.no/>

Trykk: Reprosentralen, Universitetet i Oslo

ABSTRACT

The topic of this thesis is the causes and effects of the observed male biased sex imbalance in China, with the motivation for the study being that equality between the sexes is a human right and that a surplus of bachelors may induce non-preferential outcomes in society.

First I consider a study by Nancy Qian from 2008 which claims that an exogenous shock to female-specific wages may reduce sex imbalance because females are then valued higher and/or have more bargaining power in a household. I first show that the authors' predicted effect may be biased because treatment status is not necessarily exogenous. Secondly I examine treatment heterogeneity, in particular with respect to other policy variation. This is to highlight the importance of being critical of research results, as policy making is often based on these. Thirdly I examine consequences of sex imbalance empirically using the same method as Qian (2008), and find that the populations affected by the shock to female-specific earnings also tend to have populations with better health and children with better performance in Middle School. Finally I provide a theoretical discussion of other consequences that may be anticipated in societies with a surplus of males.

PAPER OVERVIEW

This thesis is organized as follows.

In section 1 I introduce the topic of this paper, offer some background information and discuss existing literature.

In section 2 I present a summary of "Missing women and the price of tea in China: The effect of sex-specific earnings on sex imbalance", a paper published by Nancy Qian in *The Quarterly Journal of Economics*, August 2008. I also replicate part of her results.

In section 3 I provide a critical discussion of Qian's study. I then add some interesting new interactions of Qian's variables to the regressions, which may be of importance to her results.

In Section 4 I propose ideas for further study.

In Section 5 I merge agricultural variables from Qian (2008) with survey data from The Chinese Household Income Project to examine further treatment effects by the same method as Qian (2008).

In section 6 I discuss theoretical mechanisms for the possible consequences of skewed sex ratios.

In section 7 I offer conclusive remarks.

PREFACE

This thesis marks the end of a two-year master degree programme in Economics, and altogether six and a half years at the University of Oslo with various studies, ranging from Chinese History, Introductory Italian, Accounting of the firm, and finally to Economics. I am at last completing a higher degree!

I would like to give a huge thank you to Andreas Kotsadam for his great inspirational lectures in Development Economics and for being an excellent adviser. His creativity and quick responses to any technical questions along the way has been outstanding. I think it is rare for an adviser to answer an email from a student within five minutes when he is lounging on a sunbed in Malta during his vacation, but this actually happened!

Additional thanks go to Christiane Marie Ødegård for invaluable proof reading, to Eirik Okkenhaug for helping me gain access to data, as well as to my Stata study group. Together we have struggled to learn and love this programme.

I would also like to thank my parents, my sisters and my friends for having been there for me throughout my time as a student. Another special thanks goes to my excellent roommate, friend and co-student Espen Stokkereiit for all his help throughout my studies in Economics. Last but not least, I want to thank my other roommate and my boyfriend for having been patient with me during the somewhat stressful period of writing this thesis.

Any mistakes in this thesis are my own.

TABLE OF CONTENTS

SECTION 1: INTRODUCTION

1.1	<u>Introduction to topic</u>	<u>1</u>
1.2	<u>Literature on the subject</u>	<u>3</u>

SECTION 2: INTRODUCING NANCY QIAN AND HER PAPER “MISSING WOMEN AND THE PRICE OF TEA IN CHINA: THE EFFECT OF SEX-SPECIFIC EARNINGS ON SEX IMBALANCE”, 2008.

2.1	<u>Introducing Qian (2008) & her findings</u>	<u>5</u>
2.2	<u>Empirical strategy</u>	<u>6</u>
2.3	<u>Econometric method</u>	<u>7</u>
2.4	<u>Arguments backing up proxy and control regions</u>	<u>7</u>
2.5	<u>Description of data</u>	<u>8</u>
2.6	<u>Replicating Qian (2008)</u>	<u>8</u>
2.6.1	<u>Dummy for tea as dependent variable</u>	<u>9</u>
2.6.2	<u>Amount of tea as dependent variable</u>	<u>11</u>

SECTION 3: CRITIQUE OF QIAN 2008 AND FURTHER TAKING INTO ACCOUNT ETHNIC DIVERSITY IN THE RESEARCH

3.1	<u>Potential pitfalls: simultaneous shocks</u>	<u>13</u>
3.2	<u>Possible flaws in choice of treatment and control regions</u>	<u>13</u>
3.3	<u>Interactions?</u>	<u>14</u>
3.3.1	<u>Examining a possible interaction: Does the effect differ for groups depending on the One Child Policy?</u>	<u>15</u>
3.4	<u>Adding interaction terms for Han</u>	<u>17</u>
3.4.1	<u>Continuous variable for Han as independent variable</u>	<u>17</u>
3.4.2	<u>Dummy for Han as independent variable</u>	<u>19</u>
3.5	<u>Summary and remarks from section 3</u>	<u>22</u>

SECTION 4: IDEAS FOR FURTHER STUDY

4.1	<u>Introduction: Ideas for further study</u>	<u>22</u>
4.1.1	<u>Variable Group 1: Potential societal trouble associated with too many men in 'reproductive age'</u>	<u>23</u>
4.1.2	<u>Variable Group 2: Differences in people's and society's choices when the economic value of women is raised</u>	<u>24</u>

SECTION 5: EMPIRICAL INVESTIGATION WITH OTHER SURVEY DATA FROM CHINA

5.1	<u>Introduction</u>	<u>25</u>
5.2	<u>Empirical investigation of differences in outcome for tea and orchard areas</u>	<u>25</u>
5.3	<u>Description of data</u>	<u>26</u>
5.4	<u>Sex ratios</u>	<u>26</u>
5.5	<u>Performance in Middle School</u>	<u>29</u>
5.6	<u>Health status</u>	<u>32</u>
5.7	<u>Other variables in dataset</u>	<u>33</u>

SECTION 6: THEORETICAL IMPACTS OF ECONOMIC GROWTH ON GENDER ISSUES

6.1	<u>Introduction</u>	<u>34</u>
6.2	<u>Links and mechanisms from economic growth to the status of women</u>	<u>34</u>
6.3	<u>Hypothetical consequences of male-biased sex ratios</u>	<u>35</u>

SECTION 7: CONCLUSION

SECTION 8: SOURCES

LIST OF TABLES

Table 1: OLS estimates of the effect of plating tea and orchards on the fraction of males, whole sample

Table 2: OLS estimates of the effect of planting tea and orchards on the fraction of males, restricted sample

Table 3: OLS estimates of the effect og amount of tea and orchard planted on the fraction of males, whole and restricted sample

Table 4: Correlation between the fraction of Han and the crops

Table 5: OLS estimates of the fraction of minorities on tea and orchard

Table 6: OLS estimates of the effect of planting different crops on the fraction of males, including interaction effects with Han

Table 7: OLS estimates of the effect of planting different crops on the fraction of males, including interaction effects with a dummy for Han

Table 8: OLS estimates of the effect of planting tea and orchards on the fraction of males, whole and restricted sample

Table 9: OLS estimates of the effect of planting tea and orchards on the fraction of males, including control for Han, whole and restricted sample

Table 10: OLS estimates of the effect of planting tea and orchards on performance in Middle School, restricted sample

Table 11: OLS estimates of the effect of planting tea and orchards on the average health status of children, whole and restricted sample

SECTION 1: INTRODUCTION

1.1 Introduction to topic

Several Asian countries, most notably China and India, have populations with severely male-biased sex ratios, where sex ratio is defined as the number of males divided by the number of females. Amartya Sen was among the first to investigate this phenomenon, coining it as “Missing women” (Sen 1990).

In 2011, the total world sex ratio is 1.01 males for 1 female (CIA World Factbook 2012). In China the current sex ratio at birth is 1.13, and the total sex ratio is 1.06. Similarly, in India we find the respective ratios to be 1.12 and 1.08 (CIA World Factbook 2012).

At birth, boys outnumber girls everywhere in the world by around 105 male children for every 100 female children. If boys and girls receive similar treatment at birth and in childhood, girls have a higher survival rate at all ages and live longer due to biological factors that make them more resistant to disease. This is also true when controlling for environmental differences that tend to increase mortality for men, such as higher death frequency from smoking or higher probability of being involved in violence (Sen 1990). Because of this the sex ratio in most countries over time tends to one-to-one, or even slightly in favour of females.

For poorer populations in parts of Asia, notably China and India, the survival rate for girl children has long been quite different. As the culture in these societies traditionally has a strong preference for boys, less resources have been spent on girls in case of illness because they are considered to be less worth (Duflo 2011). The survival rate for girls has thus been lower than the natural rate. As these countries have experienced extensive economic growth during the last decades, we would then expect more girl children to survive. Income has risen and healthcare has become widely available, so that parents no longer face the same severe constraints at the margin, such as having to select which of their children to treat in case of illness. Simultaneously a new form of female discrimination has emerged. Although girls receive better health treatment than before when they are actually born and have more possibilities within education and work, they are now being discriminated against whilst in the womb, in form of extensive female abortion (Qian 2008).

Thus the sex ratio at birth has risen due to other important changes that have come with economic growth; The access to ultrasound and abortion technology, the trend of having less children when income goes up, in addition to the one-child-policy in China, increasing the opportunity cost of girls (Qian 2008).

The reasons for shedding light on females' opportunities and rights are several. In many parts of the world, property rights and social rights have been or are limited for women. Moreover, even in countries where both genders have equal rights by law, females traditionally have a lower participation rate in politics and decision-making bodies, as well as a lower probability of obtaining jobs in general. Further, they will on average earn less than their male counterparts in equivalent jobs even in highly developed countries. Research from the Bureau of Labor Statistics in the US for example suggests that in 2011, women who were full-time wage and salary workers had median usual weekly earnings of about 82 percent of median earnings for male full-time wage and salary workers (\$832). (US Department of Labour Statistics 2011)

There has long been an ongoing debate among scholars on whether empowering women specifically leads to development, or whether economic development in general will bring about women's empowerment, see Duflo (2011) for an overview. This debate is important, in that the advice for policy makers is quite different depending on which side is believed to be most accurate. If the latter argument is true however, it fails to account for the observed "discrimination in the womb", and does not explain the increase in the male-biased sex ratios observed in China and India and some other Asian countries as their economies have grown.

The 2012 World Development Report introduced a more nuanced view, stating that gender equality is a desirable goal in itself, and that the "business case" for women and development is secondary (World Bank 2011). Disregarding then the economic debate and accepting that equality between the sexes is inherently important, there is also an emerging concern that the extensive favouring of males at birth will in the longer run bring negative consequences to society, in a "Tragedy of the Commons" scenario; so many families in China have chosen to have sons to ensure the best possible future for their child, with the consequence that adult males are now severely in excess of females (Sen 1990). As a result many of these sons may in fact end up struggling more than their girl counterparts due to competition in many spheres of society, notably the marriage market (Yuan 2005). It thus appears that targeting girls' and womens' rights is now not just important for females themselves and possibly the economy, but also for men and the general order of society. An example can be found in the paper "Marriage and Crime: evidence from China's rising sex ratios" (2010), where Edlund et al show that the rise in the sex ratio has coincided with a rise in the crime rate in China. It is then of increasing importance to identify factors that may reverse or dampen the extensive abortion of girls.

In the main part of this thesis I consider one such study by Nancy Qian on China, which seeks to prove that increased relative wage opportunities for women in a household have a positive effect on

the fraction of girls being born. After replicating and presenting some of her results, I claim that there may be identification problems in her study, and I show that in the areas in her sample where women's wages are allegedly higher, there is also a greater fraction of minority populations. This may help explain the less male-biased sex ratios in these areas because minority populations are not constrained to the one-child-policy and therefore may have more girls. I look into the possibility of an interaction effect between the "treatment" of increased earning opportunities for women in some counties with the ethnic composition of those counties.

It is however important to note that the same male-biased sex ratio is also observed in India, and some countries of the Caucasus region, where there are no policies dictating how many children couples may have. As populations get richer, couples have historically tended to have less children, which is probably why also these populations of cultures with a strong preference for boys view the opportunity cost of having a girl as being great. Therefore I do not claim that the one-child-policy inflicted on the Han is necessarily the most important explanatory factor, I rather wish to stress the importance of considering treatment heterogeneity when treatment interacts with other population or regional traits.

Although I somewhat question the treatment identified in Qian (2008), she does manage to show a pattern in differential sex ratios across counties. In section four I therefore examine further differences in the same areas as Qian (2008), to look for other possible (causes and) effects. I wish to stress that my rationale for conducting this study is not primarily the economic case for women, which we will see is the focus of Qian (2008). It is rather in line with the view from the World Bank Report that gender equality is intrinsically important, as well as the cause to reverse or dampen the effects of the existing "Tragedy of the Commons" scenario of too many men, as described above.

I would find it most interesting and relevant to consider county-level data on variables that can indicate a measure of the impact on society of a surplus of males, such as crime rate or the size of the sex industry. I would then use the treatment and control regions identified by Qian to examine possible correlation differentials between the regions and these variables. Because I do not have access to such data I instead consider theoretical consequences of male-biased sex ratios that may be anticipated, in the last section of this paper.

1.2 Literature on the subject

In the literature on the contemporary male-biased sex ratios in parts of Asia, Amartya Sen's "Missing Women" article from 1990 is seminal. Sen (1990) discusses the possibilities that increased economic

opportunity and education for women will increase their life expectancy and may decrease female abortion. He finds some suggestive evidence of this, but concludes that the reasons why so many women are missing must be better understood before political action and public policy is implemented.

I find it useful to categorize the subsequent literature into three subject areas. One strand is research seeking to identify the cause of the skewed sex ratios and another is research proposing methods of reversing the trend, motivated by gender equality and the “business case” for women. A more recent strand is focused on extrapolating societal consequences of the *male-bias* rather than the missing women. As cohorts with large gender gaps are now becoming of adult or early adult age, consequences are starting to become more apparent. Naturally the three subject areas are interlinked so many will discuss a mixture of the three.

With regards to the first strand, there is ample agreement that the access to prenatal ultrasound and abortion technology is the reason why sex selection can occur to such a great extent. A more basic cause is the strong cultural preference for boys in some countries, but the point is that this is revealed through abnormally high sex ratios at birth, which technology makes possible. In a working paper on skewed sex ratios in Taiwan, Qian et al. (2008) present empirical evidence that the legalization of abortion accounts for almost all of the observed increase in sex ratios at birth in Taiwan. In the paper “Missing Women: Revisiting the Debate” (2003) Klasen and Wink find that “improving female education and employment opportunities has helped to reduce gender bias, while the increasing recourse to sex-selective abortions has worsened it” (Klasen & Wink 2003, p.1). Note that revealing the sex of the child has been deemed illegal in both China and India in 1994 (Marcy Darnovsky, Center for Genetics and Society, April 2009), but this law is difficult to impose.

Qian (2008) belongs to the second strand, with the focus being on identifying a potential remedy for the deficit of women, and showing why this is important for economic reasons. I again refer to Duflo (2011) for an overview of other such studies.

I now discuss the more recent third strand. French demographer Christophe Guilmoto has written several papers on gender imbalance in Asia, to highlight what he terms an alarming demographic masculinization (Guilmoto 2007). In his 2007 paper “Characteristics of Sex-Ratio Imbalances in India and Future Scenarios” he discusses potential consequences to society with a severe sex-imbalance, analogous to the “Tragedy of the Commons”, and proposes that regulations on abortions are essential (Guilmoto 2007).

Edlund et al. (2007) look into a causal link between an increase in the fraction of males and an increase in the crime rate, via the declining marriage rate for men. Marriage is an important socializing factor, and unmarried young men are arguably the most crime-prone demographic group (Steffensmeier & Allan 1996). Apart from the socializing aspect of marriage, it is possible that the necessarily fierce competition for wives may also further induce crime (Edlund et al. 2007).

Mara Hvistendahl, author of "Unnatural Selection" (2011) expresses that "Historically, societies in which men substantially outnumber women are not nice places to live(...)Often they are unstable. Sometimes they are violent."

The UN population fund (UNFPA) has issued warnings about the correlation between a scarcity of women and increases in sex trafficking and marriage migration, albeit with caveats as it may be too early to tell. Asia-Pacific director of the UN Population Fund, Nobuko Horibe has expressed that it is likely that a marriage squeeze will lead to such phenomena, but with data still being limited it is somewhat anecdotal at this stage (Horibe 2011).

Both historical evidence and theory have pointed to consequences of a male bias, and more recently also empirical investigations, although there is still little available data. In this thesis I provide an empirical analysis of variables where I have data (Section 5), and a theoretical analyses otherwise (Section 6).

SECTION 2: INTRODUCING NANCY QIAN AND HER PAPER "MISSING WOMEN AND THE PRICE OF TEA IN CHINA: THE EFFECT OF SEX-SPECIFIC EARNINGS ON SEX IMBALANCE", 2008.

2.1 Introducing Qian (2008) & her findings

Many economists argue that sex imbalance in developing countries is driven by their economic environments. Property rights and social rights have been or are limited for women in some parts of the world. Moreover, even in countries where both genders have equal rights by the law, females traditionally have a lower probability of obtaining jobs, and if they do they will on average earn less than their male counterparts in equivalent jobs. Thus many parents will prefer male offspring to ensure a better expected economic outcome for their child and the family.

Nancy Qian performed a study on China in 2008, showing effects of sex-specific earnings on sex imbalance. Qian finds that increasing female income, holding male income constant, improves survival rate for girls. Conversely, increasing male income, holding female income constant, worsens survival rate for girls. Increasing the value of "sex-neutral crops" had no effects on survival rates or

educational attainment of girls, suggesting that the change in relative income share of women in the households is important and not a change in total household income. Additionally Qian looks at educational attainment of children as an outcome, and finds that an increase in relative adult female income increases educational attainment for children of both genders. Increasing male income worsens educational attainment for girls and has no effect on boys.

2.2 Empirical strategy

Studying sex-specific income can easily be impeded by identification problems, in form of an omitted variable bias: Higher female income shares are often found in families with higher total income and higher education, both of which likely assist in driving their preference for child gender.

Qian attempts to solve the omitted variable bias by examining sex ratio of cohorts in rural counties in China before and after the Economic Reforms of 1979, between tea-planting counties and non-tea-planting counties. Sex imbalance is shown by the size of the sex ratio, defined as male births divided by female births in each county per year. The number of females in the labour force and their wages are not directly observable, so tea is used as a proxy for female income since women have a comparative advantage in picking tea.

During the reforms, the centrally planned production targets of staple crops from the Maoist era were relaxed, prices of crops were increasingly allowed to be driven by market forces and collective farms were dissolved to give individual households more freedom to adjust production to personal profit maximum. The returns to cashcrops, including tea and orchard, increased. Women have a comparative advantage in producing tea, whereas men have a comparative advantage in orchard fruit production. Thus areas suitable for growing tea experienced an increase in female-generated income, whereas areas suitable for orchard cultivation experienced an increase in male-generated income (Qian 2008).

Qian exploits this by using a Difference-in-Difference approach to examine the effects of sex-specific income on sex imbalance. Whether a cohort was born after 1979 and whether born in a tea-planting county jointly determines its' exogenous shock or treatment status. The strategy is repeated in examining differences between orchard- and non-orchard-producing counties, and finally with general cashcrops, for which production is thought to favour neither gender.

2.3 Econometric method

The paper uses panel data regression framework, in order to difference out unobservable county-fixed effects over time, and a Difference-in-Differences estimate to examine outcomes. This strategy allows the researcher to examine differences in outcome between a treatment and control group, before and after the treatment. If only looking at the outcome for the treated, the treatment that leads to their outcome status may well be confounded with other simultaneous exogenous shocks that drive gender preference. In 'natural experiments' such as the one used by Qian, it is important to argue and provide evidence that the control group is adequate so that there is no bias. Qian provides convincing evidence, in proving that tea planting counties before the reform have similar, or even lower sex-ratios than non-tea planting counties. Through an instrumental variable approach she shows that planting tea after the reform too is not correlated with girl preference but simply with geographical suitability.

2.4 Arguments backing up proxy and control regions

To back up the use of tea as proxy for female wages, Qian uses data from the 1982 Population Census where she finds that 56% of labourers in tea production are male, whereas 62% of labourers in orchard production are male. Tea production includes picking, pruning and drying. Since women's comparative advantage is in picking, the relative share of males in the picking process is likely lower than the average of 56% for the total industry. Furthermore, the main effect of post-Mao reforms on tea production was increased picking, since most tea fields were sown during a rapid expansion program during the 1960s (Qian 2008 p. 1261). These facts provide supportive evidence of female comparative advantage and the increased value of the female-specific component in tea production.

To display the exogenous price shift that occurred after agriculture was reformed, the price of tea is plotted before and after the reform and shows that it did in fact increase significantly post 1979.

In arguing that treatment and control areas have similar pre-reform trends, a simple cross-sectional comparison of the fraction of males between counties that do and do not plant tea is included, which shows that the latter have one percentage point fewer males prior to the reform (Qian page 1257).

"If the increase in the value of tea changed the reason for women to pick tea, then the pre-reform cohort is not an adequate control group" (Qian 2008 p. 1265). In order to show that tea planting counties are not otherwise different in their choice for child gender after the reform, Qian uses geographical conditions (hilliness) as an instrumental variable for whether a county planted tea after

the reform. This is done by a first stage regression of the amount of tea planted in a county after 1979 on GIS slope data from the Michigan Data Center, to prove that counties that plant tea do so because of geographical suitability and not some other reason that could potentially be correlated with relative preference for girls. Because using the IV yields results similar to OLS both in terms of size and precision, I do not include the IV estimate in my replication but refer to Qian (2008).

2.5 Description of data

Since birth data from the relevant time period is not available, Qian uses data from the 1% sample of the 1990 China Population Census to examine population and gender share of cohorts born between 1962-1990. She restricts her sample to individuals who report having lived in the same county the last five years, then presuming it to be their birth county. The 1% sample of the 1997 China Agricultural Census is the source of statistics on the amount of land area devoted to tea, orchard and general cashcrops planting in each county. Assumptions are made that the 1997 numbers on agriculture should reflect the size also at the time of this study, 1962- 1990.

The overall sample contains data from 1621 counties from all fifteen provinces of Southern China, with 51767 observations. For the DID estimate, the sample is restricted to a narrower interval, namely to cohorts born during 1970-1986, as not to confine the effect with potential history-driven changes.

2.6 Replicating Qian (2008)

In this section I replicate part of the results from Qian 2008. The dataset provided by Qian on her personal website consists of eight variables in addition to a column for county id-number. Three of the variables are data on the sizes of the agricultural industries tea, orchard and general cashcrop production gathered from the 1997 agricultural census. The remaining variables on birth year, sex ratio, cohort size, education, and fraction of ethnic Han population in a cohort are gathered from the 1990 1% sample population census.

The data provided by Qian for replication is limited, so the results will not be identical to those posted in Qian 2008.

For the differences-in-differences, it is most relevant to use the restricted sample of cohorts 1970-1986 for analysis. I however perform all the regressions both on the whole and restricted sample, for comparison.

The regressions that follow are identical to those posted on Qian's website. They include different numbers of control variables, to show that the precision of the coefficient on tea is not reduced when adding more variables that are also significant.

Throughout the paper the regressions are run both with tea as a continuous variable representing the amount of tea-planting a county has, and with tea as a dummy variable, which equals 1 for counties that plant any tea, analogous to Qian (2008).

2.6.1 Dummy for tea as dependent variable

The following regressions are run, for the whole sample, cohorts 1963-1990. Output is displayed in Table 1 below.

$$(1) \text{Sex}_{ic} = (\text{teadum}_i \times \text{post}_c)\beta + \text{post}_{cT} + \gamma_c + \varepsilon_{ic}$$

$$(2) \text{Sex}_{ic} = (\text{teadum}_i \times \text{post}_c)\beta + (\text{cashcrop}_i \times \text{post}_c)p + \text{post}_{cT} + \gamma_c + \varepsilon_{ic}$$

$$(3) \text{Sex}_{ic} = (\text{teadum}_i \times \text{post}_c)\beta + (\text{orchard}_i \times \text{post}_c)\delta + (\text{cashcrop}_i \times \text{post}_c)p + \text{post}_{cT} + \gamma_c + \varepsilon_{ic}$$

$$(4) \text{Sex}_{ic} = (\text{teadum}_i \times \text{post}_c)\beta + (\text{orchard}_i \times \text{post}_c)\delta + (\text{cashcrop}_i \times \text{post}_c)p + \text{Han}_{ic}\zeta + \text{post}_{cT} + \gamma_c + \varepsilon_{ic}$$

In equation (1), the fraction of males in county i , Sex_{ic} is a function of the interaction terms between teadum_i , a dummy variable which equals one for counties that plant tea, and post_c , a dummy variable which equals one for individuals born after the reforms in 1979, the dummy variable post_c itself, and γ_c , unobservable county specific effects differenced out by the panel data structure. In equation (2), the fraction of males is also a function of the interaction between cashcrop_i , a measure of the amount of cashcrops planted in county i , and post_c . Equation (3) adds the interaction of orchard_i , a measure of the amount of orchard planted in county i with post_c as an explanatory variable to equation (2). Finally equation (4) is similar to equation (3), the only difference being including a control for Han_{ic} , the fraction of ethnic Han in county i for cohort c .

The difference-in-differences estimate of β is the main interest, with δ and p being added for comparison and hypothesis support. These coefficients indicate the difference between regions planting different crops, before and after the reform.

For a restricted sample, cohorts 1970-1986 I perform the same regressions as above, and the results are presented in columns (5)-(8) of Table 2.

Table 1: OLS estimates of the effect of planting tea and orchards on the fraction of males, whole sample

	Dependent variables			
	Fraction of males			
	(1)	(2)	(3)	(4)
teadumpostic	-0.0096*** (0.002)	-0.0092*** (0.002)	-0.0081*** (0.003)	-0.0081*** (0.003)
orchpostic			0.0086*** (0.003)	0.0084*** (0.003)
cashcroppostic		-0.0005 (0.001)	-0.0026*** (0.001)	-0.0025** (0.001)
Hanic				0.0587*** (0.013)
postc	0.0194*** (0.001)	0.0197*** (0.001)	0.0200*** (0.001)	0.0201*** (0.001)
Observations	49,082	49,082	49,082	49,082
R-squared	0.124	0.124	0.124	0.124

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 2: OLS estimates of the effect of planting tea and orchards on the fraction of males, restricted sample

	Dependent variable			
	Fraction of males			
	(5)	(6)	(7)	(8)
teadumpostic	-0.0096*** (0.002)	-0.0102*** (0.003)	-0.0091*** (0.003)	-0.0091*** (0.003)
orchpostic			0.0083** (0.003)	0.0081** (0.004)
cashcroppostic		0.0007 (0.001)	-0.0013 (0.001)	-0.0012 (0.001)
Hanic				0.0692*** (0.017)
postc	0.0099*** (0.001)	0.0095*** (0.001)	0.0099*** (0.001)	0.0099*** (0.001)
Observations	28,349	28,349	28,349	28,349
R-squared	0.109	0.109	0.109	0.109

Discussion of results in Tables 1 and 2

I find that the fraction of males decreases after 1979 for tea-planting counties compared to general cashcrop-planting counties, for both the whole and the restricted sample. I can see that the converse is true for orchard-planting counties; the bigger the size of the orchard industry, the larger the fraction of males after 1979. Adding a control for the fraction of Han in each county does not change the significance or the size of the coefficient β on tea nor the coefficient δ on orchard, although the coefficient ζ on Han is highly significant and larger than β and δ in both the whole and the restricted samples.

I consider the restricted sample of Table 2. All coefficients on teadumpost are significant at the 1% level, whereas the coefficient on the control cashcroppost is not significant. The coefficients on the comparison group orchpost are significant at the 5% level. The controls for Han and post , which both are significant at the 1% level are of the expected positive sign, and help to disentangle the effect of the one-child policy and general economic reform.

These results are consistent with the hypothesis and findings from Qian (2008). It is the rise in the relative share of female-generated income in households that leads to a higher survival rate for girls. The policy recommendation from Qian, “One way to reduce excess female mortality (...) is to increase the relative earnings of adult women” (Qian 2008 p. 1281), is adequate if data is believed to be representable and estimation method is appropriate.

2.6.2 Amount of tea as dependent variable

I estimate the following equations for the whole sample:

$$(5) \text{Sex}_{ic} = (\text{teasown}_i \times \text{post}_c)\beta + (\text{orchardi}_i \times \text{post}_c)\delta + (\text{cashcrop}_i \times \text{post}_c)p + \text{post}_c\tau + \gamma_c + \epsilon_{ic}$$

$$(6) \text{Sex}_{ic} = (\text{teasown}_i \times \text{post}_c)\beta + (\text{orchardi}_i \times \text{post}_c)\delta + (\text{cashcrop}_i \times \text{post}_c)p + \text{post}_c\tau + \text{Han}\zeta + \gamma_c + \epsilon_{ic}$$

Output is displayed in columns (9) and (10) in Table 3 below. The difference from equations (1)-(4) above is that the dummy for tea, teadum_i has been replaced by the continuous variable tea_i , which is a measure of the size of the tea planting industry in county i .

For the restricted sample, cohorts 1970-1986, I perform the same regressions as above. Results are displayed in columns (11) and (12) in Table 3 below.

Table 3: OLS estimate of the effect of amount of tea and orchard planted on the fraction of males, whole and restricted sample

	Dependent variables			
	Fraction of males			
	(9)	(10)	(11)	(12)
teasownpostic	-0.0088 (0.006)	-0.0083 (0.006)	-0.0068 (0.008)	-0.0065 (0.008)
orchpostic	0.0089*** (0.003)	0.0088*** (0.003)	0.0093** (0.004)	0.0092** (0.004)
cashcroppostic	-0.0030*** (0.001)	-0.0029*** (0.001)	-0.0021* (0.001)	-0.0020* (0.001)
hanic		0.0585*** (0.013)		0.0688*** (0.017)
postc	0.0191*** (0.001)	0.0192*** (0.001)	0.0089*** (0.001)	0.0089*** (0.001)
Observations	49,082	49,082	28,349	28,349
R-squared	0,124	0,124	0,108	0,109

Discussion of results in Table 3

When considering the effect of the amount of tea sown, the DID-estimator of tea, i.e. the coefficient on `teasownpostic` is still negative on average, but statistically insignificant in both the whole and the restricted sample. In Qian (2008) however, the coefficient obtained on `teasownpostic` in the restricted sample (roughly equivalent to my columns (5) and (6)), is negative and significant at the 10% level. It is important to keep in mind that the data I use has some limitations compared with the data used in Qian 2008, although my data is obtained from Qian's personal website.

In the following I consider the results from all tables 1-3 above, with the limited data. The insignificant coefficient on the amount of tea sown after 1979 (`teasownpostic`) and the significant coefficient on the dummy for tea sown after 1979 (`teadumpostic`) suggest non-linear effects of tea. The coefficient on `teasownpostic` is also smaller than the coefficient on `teadumpostic`. It might then be the presence of any tea-planting in a county that is important, in making the economic possibilities for women visible in society, and thereby having an effect of how households value girls.

In reality, however, one can imagine that the size of the tea-industry in a county should matter. If "female-favourable" tea production is extensive and plays a significant part in the local economy, this will likely assist in enhancing the probability of profitable work for women. If the tea industry in a

county is small, it will probably not provide many jobs anyway, and should not play a large part in the decision to keep a girl baby, if the economic value of girls is important.

SECTION 3: CRITIQUE OF QIAN (2008) AND FURTHER TAKING INTO ACCOUNT ETHNIC DIVERSITY IN THE RESEARCH

3.1 Potential pitfalls: simultaneous shocks

A whole set of reforms were implemented with the 'opening' of China in 1979, which could potentially drive results found by Qian.

Perhaps most importantly for this study, a policy allowing couples to have one child only was introduced. This effectively raised the opportunity cost of females, since couples now have only one chance to have a boy. China's minority ethnicities, about 8,5% of the population (2000 Census), were however exempt from the rule, which then only applies to the 91,5% of the population that are of the main ethnic group, Han. Qian controls for this by adding the fraction of ethnic Han in each cohort for each county as a separate regressor, nevertheless finding a significant effect of tea. She has not however considered interaction effects between the treatment of tea and the ethnic composition of the treated counties, which I propose to do below.

The many other simultaneous shocks due to opening and reforms in 1979 could also assist in driving results. Notably the country was opened up to foreign investment and domestic entrepreneurship. Most reforms applied to the whole country, but Special Economic Zones (SEZ) were set up in some areas in 1980, with low taxes to encourage foreign investors.

3.2 Possible flaws in choice of treatment and control regions

There are several arguments why Qian's treatment and control counties for the *most part* are adequate, I refer to section 1.3 of this paper and Qian 2008. In the following I examine some possible flaws.

Qian states that all the counties in her sample are rural, and she shows similar pre-reform trends in her variable of interest, sex ratio. When matching Qian's data with a dataset from China Household Income Project (CHIP) 1988 in section three of this paper, some of her counties match data from

counties that the CHIP classifies as urban. It is not clear how Qian distinguishes out rural counties. If the classification she has used is from 1979, the 1988 data may simply indicate that some of these counties have become urban during the following decade, which sounds reasonable considering the growth and centralization China has experienced. If the classification is obtained from the 1990 population census however, then some of these counties may not be appropriate, as they may be considered urban. Urban populations may be inherently different in what drives their choice for child gender. It may thus be reasonable to exclude these counties from the sample.

Furthermore, several cities in Guangdong province and Xiamen in Fujian province were among the first SEZ's, both of which regions are included in Qian's sample. It might also be an idea to exclude these counties from the sample to better argue for homogeneity between treatment and control group.

In addition, there even exists one matriarchal ethnic group in China, the "Mosuo", who live in the north of Yunnan and Sichuan province. Its' population is not large at approximately 56 000 people (Zhou 2005), and it is unknown whether Qian's sample includes anyone from this group. However, it is possible to imagine that the presence of matriarchal culture may have spillover effects, such that counties "exposed" to them view girls differently. There are thus some counties in Yunnan and Sichuan from Qian's sample that may be inherently different in their choice for child gender, which could make them inadequate as treatment and control regions.

3.3 Interactions?

Disregarding the differential one-child-policy, it should be mentioned that China is a centralized governed country, with officials at all levels of authority appointed by The State Council in Beijing. Policy implementations are thus to a great extent homogeneous. Apart from the SEZ's and large cities on the coast, the rural counties such as the ones that make up the majority of Qian's sample should experience similar 'treatment and effects' of the economic reforms. In this sense it does not matter whether they experienced other simultaneous shocks during the study period, as long as the rise in price of 'sex-differential agricultural products' is expected to be the only reform affecting the counties differently.

If the other reforms do interact differently with the treatment effect of tea-regions however, there might be a bias in the estimates. Even though the pre-trends suggest similar characteristics, it is not unlikely that there is some interesting variation that is missed by Qian here.

3.3.1 Examining a possible interaction: Does the effect differ for groups depending on the One Child Policy?

The way the fraction of males varies with the planting of tea after the reform may depend on the ethnic composition of the county. If counties that have tea production have a higher fraction of ethnic minorities, the negative coefficient on `teasownpost` and `teasowndumpost` may be capturing the fact that many households in these counties are exempt from the one-child policy and therefore have more girls. The causal effect is then not necessarily the increase in the economic value of girls, but rather the absence of the policy, which lowers the opportunity cost of having a girl.

Correlation between Han and crops

I organize Qian's dataset so there is only one observation per cohort per county, and examine correlations between the proportion of Han in a cohort and the sex-specific crops of tea and orchard. Output is displayed in the matrix below, with standard deviations in parentheses. Both orchard and `teasown` are negatively correlated with Han, but only the correlation between Han and `teasown` is statistically significant. Moreover the correlation between `teasown` and Han is over one hundred times the size of the correlation between orchard and Han.

Table 4: Correlation between the fraction of Han and the crops

	<code>teasown</code>	<code>orchard</code>
Han	-0.0813 (0.0000)	-0.0007 (0.8806)

Linear Regression indicator

I now examine the correlation between `teasown` and the fraction of ethnic minorities in a cohort by a simple population-weighted OLS regression of `teasown` on the fraction of minorities, the variable `minority` being constructed from subtracting the fraction of Han from 1. The coefficient obtained on `minority` is of value 0.05546, and it is highly significant.

A simple regression like this does not identify a causal relationship. Minorities may prefer to live in geographical climates that happen to favour tea production, without them being involved in the industry nor necessarily considering girls to have a higher value after 1979 because of it. The regression output nevertheless confirms a statistical correlation; counties with a higher fraction of

minorities plant more tea on average. As demographic employment data is not available then, it can be hard to disentangle the effect of the proxy “tea” and the presence of minority populations on sex ratios.

Differences-in-Differences Approach

I then perform a simple population-weighted regression on teasown and orchard on the fraction of minorities for cohorts before and after 1979, yielding the output in the table below. It is apparent that both tea and orchard production is correlated with the fraction of minorities in a cohort, but the correlation with tea is of greater magnitude. I also note that after the 1979 reform, amount of tea sown is increasingly correlated with the presence of minorities, whereas orchard is less correlated with minorities. It might then be that the less biased sex ratios observed in tea areas before and after the reform are connected to the size of the minority populations there.

Table 5: OLS estimates of the fraction of minorities on tea and orchard

	Fraction of minorities born			
	Pre 1979	Post 1979	Pre 1979	Post 1979
teasown	0.1038*** (0.008)	0.1476*** (0.011)		
orchard			0.0308*** (0.004)	0.0236*** (0.005)

Since there is a difference in the correlation between tea and minorities before and after the reform, this could reflect the fact that minorities do in fact adapt to the changing economic environment and have more children, and probably give birth to more girls. Their right to have as many children as desired implies that they face a smaller opportunity cost of having a girl than the Han. On the other hand, the increased correlation could simply reflect that the Han are having less children, so the proportion of minority births goes up and their girl preference may not have changed.

Orchard production is less associated with fraction of minority births in the sample, which could possibly explain why orchard regions have a greater fraction of boys being born after the reform. Why the fraction of minority births is smaller in orchard counties after 1979 is not clear. Perhaps they do adapt to the changing economic environments in these counties, having less children and/or preferring sons. Nonetheless they do not make up a substantial part of the population in these counties.

3.4 Adding interaction terms for Han

To further look into the relationship of tea planting and size of minority populations in a county, I add interactions of Han and the different crops multiplied by the dummy for post 1979, to the DID regressions estimated by Qian. First I look at Han as a continuous variable which is simply the fraction of Han in each county, and then as a binary variable which equals one if over half the population of a cohort is ethnic Han.

I now only perform the Qian regressions with all the controls, as in equations (3), (4), (5) and (6) of section 2, and add all the interactions as separate regressors.

3.4.1 Continuous variable for Han as independent variable

I estimate the following equation for the whole and restricted sample, with independent variables including a dummy for tea and continuous variables for orchard, cashcrop and Han. All variables and subscripts are the same as those in the section 2, but now the fraction of males sex_{ic} is also a function of the interaction of each crop after 1979 with the ethnic composition Han_{ic} . This yields output columns (1) and (2) in Table 6. Standard errors are clustered at the county level and γ_c represents unobserved county-fixed effects.

$$(7) \text{ Sex}_{ic} = (\text{teadum}_{i} \times \text{post}_{c})\beta + (\text{orchard}_{i} \times \text{post}_{c})\delta + (\text{cashcrop}_{i} \times \text{post}_{c})\rho + Han_{ic}\zeta + (\text{teadum}_{i} \times Han_{ic} \times \text{post}_{c})\eta + (\text{orchard}_{i} \times Han_{ic} \times \text{post}_{c})\phi + (\text{cashcrop}_{i} \times Han_{ic} \times \text{post}_{c})\alpha + (Han_{ic} \times \text{post}_{c})\psi + \text{post}_{c}\tau + \gamma_c + \epsilon_{ic}$$

I then estimate the same equation for the whole and the restricted samples, only replacing the dummy for tea with the continuous variable $teasown$ as independent variable. This variable takes values from 0 to 3.14. This yields column output (3) and (4) in Table 6.

$$(8) \text{ Sex}_{ic} = (\text{teasown}_{i} \times \text{post}_{c})\beta + (\text{orchard}_{i} \times \text{post}_{c})\delta + (\text{cashcrop}_{i} \times \text{post}_{c})\rho + Han_{ic}\zeta + (\text{teasown}_{i} \times Han_{ic} \times \text{post}_{c})\eta + (\text{orchard}_{i} \times Han_{ic} \times \text{post}_{c})\phi + (\text{cashcrop}_{i} \times Han_{ic} \times \text{post}_{c})\alpha + (Han_{ic} \times \text{post}_{c})\psi + \text{post}_{c}\tau + \gamma_c + \epsilon_{ic}$$

Table 6: OLS estimates of the effect of planting different crops on the fraction of males, including interaction effects with Han

	Dependent variable			
	Fraction of males			
	(1)	(2)	(3)	(4)
teasowndumpostic	-0.0002 (0.006)	-0.0002 (0.008)		
teasownpostic			-0.0046 (0.010)	0.0011 (0.012)
orchpostic	-0.0112 (0.007)	-0.0130 (0.009)	-0.0140* (0.008)	-0.0137 (0.010)
cashcroppostic	0.0029* (0.002)	0.0029 (0.002)	0.0037** (0.002)	0.0030 (0.002)
Hanic	0.0454*** (0.013)	0.0647*** (0.018)	0.0446*** (0.013)	0.0644*** (0.018)
teasowndumpostHanic	-0.0081 (0.007)	-0.0098 (0.009)		
teasownpostHanic			-0.0087 (0.014)	-0.0121 (0.018)
orchpostHanic	0.0195** (0.008)	0.0224** (0.010)	0.0219** (0.009)	0.0238** (0.011)
cashcroppostHanic	-0.0049** (0.002)	-0.0043* (0.002)	-0.0059** (0.002)	-0.0051** (0.003)
Hanpostic	0.0240*** (0.004)	0.0138*** (0.004)	0.0238*** (0.004)	0.0132*** (0.005)
postc	-0.0029 (0.004)	-0.0032 (0.004)	-0.0037 (0.004)	-0.0036 (0.005)
Observations	49,082	28,349	49,082	28,349
R-squared	0.126	0.110	0.125	0.109

Discussion of results in Table 6

All the coefficients on tea and its' interactions are insignificant, so I turn my attention to the coefficients on orchpost. The coefficient on orchpost is negative in the sample for both full and restricted samples in equations (7) and (8) but only statistically significant at the 10% level for one equation (8) with full sample, displayed in column (3). Interestingly, the coefficients on the interaction terms between orchpost and Han are of higher absolute value, and are significant at the 5% level for all columns (1) – (4).

Consider the total effect of orchard in column (3)

$$d/dorch = -0.0140post + 0.0219Han.$$

This suggests that for cohorts with a lower share of Han, the fraction of males born will not augment due to increased economic opportunity for men (here in the form of a rise in the price of the claimed male-specific orchard crop after 1979), contrary to results from Qian (2008). Orchardpost will instead assist in decreasing the fraction of males for cohorts with less than 64% Han after 1979.

This is probably because minorities do not abort female fetuses to the extent that the Han do, as they are exempt from the one-child policy. They have many chances to have a son, if that is important to them. Further, although males have a comparative advantage in orchard production, women have become increasingly important in all spheres of agriculture since 1979 as the development of a market economy initially led most male labourers to go outside the traditional farm for non-farm employment (Food and Agriculture Organization of the United Nations, 2012). The proportion of males in the orchard industry of 62% as mentioned in Qian (2008), is from the 1982 Population Census, and this number may since have changed. The value of orchard may then not be a very good proxy for male-biased wages.

Furthermore, the coefficient on Hanpost is significant at the 1% level for all four equations. I continue to look at column (3), now examining the total effect of Han.

$$d/dHan = 0.0466 + 0.0219orch + 0.0238post$$

I find that for ethnic Han cohorts the fraction of males is always larger than that for minorities. This effect is stronger after 1979. The coefficient on the interaction term with orchard suggests that the fraction of males rises with the fraction of Han for cohorts in orchard-planting counties. However, as indicated above, the effect of orchard production for cohorts with zero percent Han is actually negative. In this equation, the fraction of Han would have to exceed 64 percent for the total effect of orchard to be an increased fraction of males.

It appears that the strong cultural desire of the Han to have males, along with the one-child policy may be more important in explaining skewed sex ratios than implied by Qian (2008).

3.4.2 Dummy for Han as independent variable

I then estimate equation (9), an equation with the independent variable being a dummy for Han, $Handumic$, rather than a continuous variable. $Handumic = 1$ if $Hanic > 0.5$, in other words I separate the cohorts with a Han majority from those where the Han are in minority. This yields the output displayed in columns (5) and (6) in Table 7, with a dummy for tea as independent variable

$$(9) \text{Sex}_{ic} = (\text{teadum}_i \times \text{post}_c)\beta + (\text{orchard}_i \times \text{post}_c)\delta + (\text{cashcrop}_i \times \text{post}_c)\rho + \text{Handum}_{ic}\zeta + (\text{teadum}_i \times \text{Handum}_{ic} \times \text{post}_c)\eta + (\text{orchard}_i \times \text{Handum}_{ic} \times \text{post}_c)\phi + (\text{cashcrop}_i \times \text{Handum}_{ic} \times \text{post}_c)\alpha + (\text{Handum}_{ic} \times \text{post}_c)\psi + \text{post}_{ct} + \gamma_c + \varepsilon$$

I then estimate equation (10), the same equation as above but replacing the dummy for tea with the continuous variable tea sown, yielding output in columns (7) and (8) of Table 7.

$$(10) \text{Sex}_{ic} = (\text{teasown}_i \times \text{post}_c)\beta + (\text{orchard}_i \times \text{post}_c)\delta + (\text{cashcrop}_i \times \text{post}_c)\rho + \text{Handum}_{ic}\zeta + (\text{teasown}_i \times \text{Handum}_{ic} \times \text{post}_c)\eta + (\text{orchard}_i \times \text{Handum}_{ic} \times \text{post}_c)\phi + (\text{cashcrop}_i \times \text{Handum}_{ic} \times \text{post}_c)\alpha + (\text{Handum}_{ic} \times \text{post}_c)\psi + \text{post}_{ct} + \gamma_c + \varepsilon$$

Table 7: OLS estimates of the effect of planting different crops on the fraction of males, including interaction effects with dummy for Han.

	Dependent variable			
	Fraction of males			
	(5)	(6)	(7)	(8)
teasowndumpost	-0.0033 (0.005)	-0.0038 (0.006)		
teasownpost			-0.0176** (0.009)	-0.0150* (0.008)
orchpost	-0.0100 (0.007)	-0.0093 (0.008)	-0.0168** (0.007)	-0.0145* (0.009)
cashcroppost	0.0021 (0.001)	0.0016 (0.002)	0.0042** (0.002)	0.0033* (0.002)
handum	0.0041 (0.006)	0.0131* (0.007)	0.0035 (0.005)	0.0126* (0.007)
teasowndumpostHandum	-0.0047 (0.006)	-0.0057 (0.007)		
teasownpostHandum			0.0071 (0.011)	0.0075 (0.013)
orchpostHandum	0.0188** (0.007)	0.0181** (0.009)	0.0258*** (0.008)	0.0247** (0.010)
cashcroppostHandum	-0.0044** (0.002)	-0.0028 (0.002)	-0.0069*** (0.002)	-0.0053** (0.002)
Handumpost	0.0190*** (0.003)	0.0108*** (0.004)	0.0202*** (0.003)	0.0119*** (0.004)
post	0.0018 (0.003)	-0.0005 (0.004)	-0.0003 (0.003)	-0.0024 (0.004)
Observations	49,082	28,349	49,082	28,349
R-squared	0.125	0.110	0.125	0.109

Discussion of results in Table 7

The dummy $Handumic$ is created to look at difference in effects between cohorts with a majority of Han ($Handumic = 1$) and cohorts with a minority of Han ($Handumic = 0$).

For equation (10), with the continuous variable for amount of tea sown as an independent regressor (output in columns (7) and (8)), the coefficients on $teasownpost$ are negative and significant at the 5% and 10% level respectively. This implies that for cohorts where over 50 percent are of ethnic minority, tea sown after 1979 will lower the fraction of males. This is in line with results found by Qian (2008), although my replication in part 1 did not give statistically significant coefficients. However, as mentioned earlier, this could also be attributed to the fact that the fraction of minority births is higher in tea counties after the reform. The coefficient on $teasownpostHandum$ is positive, suggesting less effect of $teasown$ for cohorts with a Han-majority. However this coefficient is not statistically significant, so I cannot rule out that it is actually zero.

Again the results for orchard production are worth noting. The coefficient on $orchpost$ is negative and statistically significant at the 5% and 10% level for equation (10) for both samples, and the coefficient on $orchpostHandum$ is positive of a higher value, and significant at the 5% level for all four columns (5) – (8).

I examine the results in column (8) which is the narrower DID estimate for cohorts 1970-1986.

$$d/dorch = -0.0145post + 0.0247Handum$$

This suggests that a rise in the price of male-biased orchard produce in itself does not worsen survival rate for girls. It only has this effect for cohorts of Han-majority. More specifically, in this equation the proportion of Han would have to exceed 50%, before the presence of orchard may increase the fraction of males. Furthermore the coefficients on $Handum$ and $Handumpost$ are both positive and significant.

$$d/dHandum = 0.0126 + 0.0247orchpost + 0.0119post$$

It appears that the Han in the sample tend to always favour boys more than do ethnic minorities, and this favourization is stronger after 1979. The rise in the price of orchard produce after 1979 will only have a treatment effect in favour of more boys being born if the majority of the population is Han. We cannot be sure then that orchard production is the important explanatory factor, rather it could be that the inherent factors of the Han population are of greater importance. This is suggestive evidence that the one-child policy and the strong paternal culture of the Han may be more important in explaining male-biased sex ratios than the economic conditions for the sexes.

3.5 Summary and remarks from section 3

The purpose of this section is not to prove that Qian is necessarily wrong in assessing her treatment population and her treatment effects. I rather wish to highlight the potential bias in results that can be obtained if interaction effects should be included and/or two explanatory variables are correlated. In these cases it is difficult to establish how the individual regressors affect the dependent variable.

SECTION 4: IDEAS FOR FURTHER STUDY

4.1 Ideas for further study

It is important to remember that the main focus of Qian's study is not to look at the effect of tea-planting on sex ratios, but rather the effect of an increase in the share of female-generated income, as proxied by tea.

There could exist more updated information or statistics of factors that raise the economic value of women, which may differ across regions so that comparing is possible. The composition of China's labour force by sector is currently at 36.7% in agriculture, 28.7 % in industry and 34.6% in services. The respective percentages in 1980 were 68.7%, 18.3% and 13%. When considering value created by each sector through GDP composition, agriculture now only makes up 10,1%, versus 30,1% in 1980 (China Statistical Yearbook 2011). It might therefore be interesting, and perhaps more relevant today to look for differences between men and women in the secondary and tertiary industries.

Whichever method employed, examining possible differences between areas where women have (or have had) better work prospects and thus have had a higher survival rate can help shed light on the potential consequences of a severely unbalanced male-female ratio, as mentioned in the introduction chapter. Equally important is the notion of better integrating women in the economy and in society in general, if not for business reasons then because equal opportunities for the sexes are a human right.

Although it is uncertain whether the treatment of tea and orchard planting after 1979 is relevant today, it is in theory possible that it is. The exogenous shocks in 1979 could have shifted people's view on gender on a more permanent basis, or ripple effects could have lead to further differentials between tea and orchard regions. Thus it is interesting to see if there are other differences in outcomes for the counties in Qian (2008) for other outcomes.

Outcomes of interest could be many depending on what the focus of the study is. I chose to divide my variables of interest into two rough categories. In the following I describe these categories and the potential difficulties that may be encountered in studying each.

4.1.1 Variable Group 1: Potential societal trouble associated with too many men in 'reproductive age'

The first group is directly associated with consequences of an excess supply of males. One might then consider looking at marriage rate, crime rate, alcohol and drug consumption, size of sex industry or a measure of trafficking activity. This is arguably one of the most important areas of gender research at the moment, because large male-biased cohorts in parts of Asia are now becoming of adult age.

This category includes variables whose effects are more relevant to study when the cohorts are of adult age. The potentially most 'troublesome' effects to society will likely be related to distortions in the marriage market. At present (year 2012) the cohorts born during the reform year are 33 years old. Official statistics on marriage are difficult to obtain, but an uncited Wikipedia article suggests that the average age of marriage is 31.1 for men and 29.1 for women in China in 2010 (Wikipedia 2012). In China, as in many Western countries, the average marriage age has tended to rise during the last decades (China Daily 2011). Men often tend to marry younger women, so "fishing out" of the females of the younger cohorts may occur for some time. In addition, some import of wives from countries nearby has started, which reduces the shortage of supply of women (Dasgupta 2011). Eventually such practices will not be enough to satisfy demand however. There is a natural limit with China being the world's most populous country and with India, the world's second biggest country having the same male biased sex ratio. Hypothesized consequences of a surplus of bachelors include an increase in crime rate, drug/alcohol consumption, increase in demand for the sex industry. In short, increased social unrest. It may be too early to get a good overview of these consequences yet, for reasons mentioned above and for lack of data, but research on the area is emerging. Edlund et al 2010 examine such consequences empirically using regional data, and documents from the UN Population Funds Seminar on Skewed Sex Ratios at Birth (2011) discuss such effects theoretically. My primary interest would be studying this category of variables empirically. However, because my focus is on the same counties as those studied by Qian and micro-data at this level is not easily available, I instead provide a general theoretical discussion of the mechanisms of such effects in section 6.

4.1.2 Variable Group 2: Differences in people's and society's choices when the economic value of women is raised

The second category includes variables we may want to examine if the aim is to identify conditions for development as well as the general equality between the sexes, the latter being claimed to be most important by the Recent World Bank Report. Although this report deemed the “business case” for women secondary, I claim that it is still relevant for policy makers dealing with allocation of scarce resources, aiming to achieve the highest possible yield of their means. Variables we might want to look at could be differential public good priorities, outcome for children and measures of productivity in various sectors. Such variables may also be important in lowering the sex ratio in the future if their values contribute to women's status being raised, as this can then possibly help dampen the consequences of the contemporary male-biased ratio, if empowering women creates more positive externalities to society.

This category of variables could be interesting to study also before the reform-cohorts are of adult age. The realization that women can contribute in the economy may plausibly have more immediate effects on the societies studied. However, finding data on such variables before and after the reform is difficult. The decade preceding the reform was characterized by severe political and social turmoil in China, and no official statistics were collected. Survey research was further probably non-existent or at least such material is seldom publically available.

If more such data exists, it is possible to use an approach like Qians', examining later population censuses or surveys. Qian looks at the 1990 Population Census, and uses reported age of survey participants to determine birth year, and thus finds the cohort size at birth in each county. She restricts her sample to persons having lived in their respective counties for the last five years, then assuming it to be their county of birth, when examining cross-sectional sex ratios and educational attainment. This assumption is probably reasonable at the time of the data collection, as migration was strictly controlled and therefore barely occurred. If we could find more survey data from the early 90's, we may be able to use similar inference. Most data collection is however done by the Official Statistics Bureau, and most variables are aggregates at a higher level than the county. The Chinese authorities are also strict on allowing research by independent persons or institutions, whether domestic or foreign. There is however one extensive population survey called The Chinese Household Income Project, which has been carried out four times by joint Chinese and International effort. This survey is publically available for years 1988, 1995 and 2002. In the following section 5 of this paper I will use data from this survey to look at outcomes for sex ratios, Middle School performance and health status of inhabitants.

The kinds of assumptions Qian makes on birth-county are perhaps not likely to hold when looking at data sets much later than 1990. China has since the 90's been experiencing mass migration from rural regions to large coastal economic giants such as Shanghai and Guangzhou. Recently there has also been an increase in within-region migration to smaller cities which are now also experiencing growth (The Economist 2012). Other variables are however still relevant to study also in later data sets. If conditions for women improved in tea-counties and this causes more women to move there and participate in society, this is in fact an important part of the results of the treatment.

SECTION 5: EMPIRICAL INVESTIGATION WITH OTHER SURVEY DATA FROM CHINA

5.1 Introduction

In this section I make an attempt at examining some consequences of skewed sex ratios empirically by using the regression framework from Qian (2008) with outcome variables obtained from survey data from The Chinese Household Income Project.

5.2 Empirical investigation of differences in outcome for tea and orchard areas

I use the same treatment and control regions as Qian (2008), and the same independent variables $teadum_{ic}$, tea_{ic} , $orchard_{ic}$ and $cashcrop_{ic}$ interacted with $post_c$. I merge these agricultural variables gathered by Qian from the 1997 Agricultural Census with data from the Chinese Household Income Project for years 1988 and 2002.

I first want to test her methodology to see if it holds with the same outcome variable, namely the fraction of males, in another dataset from around the same time. Furthermore I want to find out whether the treatment of tea after 1979 has other effects than increasing the relative survival rate of girls and increasing educational attainment for children (Qian 2008). I do this by changing the dependent variable of equations familiar from section 2, and regressing with both the continuous variable $teasown_i$ and the dummy for tea $teadum_i$, as well as the other control regressors. The format of the equations is as follows:

$$Newoutcomevar_{ic} = (teadum_i \times post_c)\beta + (orchard_i \times post_c)\delta + (cashcrop_i \times post_c)p + post_c\tau + \gamma_c + \epsilon_{ic}$$

$$Newoutcomevar_{ic} = (teasown_i \times post_c)\beta + (orchard_i \times post_c)\delta + (cashcrop_i \times post_c)p + post_c\tau + \gamma_c + \epsilon_{ic}$$

Where “Newoutcomevaric” here is the new variable I am looking at. Where possible, a control for the fraction of Han, displayed $Hanic_z$, is also added.

5.3 Description of data

The Chinese Household Income Project (CHIP) is a survey co-funded by academic institutions in the US, UK and China, through the Inter-University Consortium for Political and Social Research.

“Data were collected through a series of questionnaire-based interviews conducted in rural and urban areas in 1988, 1995, and 2002. Individual respondents reported on their economic status, employment, level of education, sources of income, household composition, and household expenditures.” (CHIP 2002).

In my analysis I only use the surveys from 1988 and 2002. The CHIP codebook from 1988 states; “The data collection consists of two distinct samples of the urban and rural population of the People's Republic of China which were selected from significantly larger samples (67,186 rural households and 34,945 urban households) drawn by the State Statistical Bureau.” (CHIP codebook 1988)

The 2002 dataset also consists of separate urban and rural data. The urban survey contains 20,632 cases (individual urban household members) and the rural survey contains 37,969 cases (individual rural household members). The variables in the urban and rural datasets from both 1988 and 2002 often differ and/or are coded differently. Where necessary in my analysis, I describe how I merge them.

The headings of the subsequent paragraphs refer to the dependent variable whose outcome I am looking at in the regressions that follow.

5.4 Sex ratios

First I perform the same regressions as in section 2, examining the effects of tea and orchard after 1979 on sex ratios. I do this for comparison with Qian (2008) as a robustness check, to see if her results hold in another sample. In this analysis, I use the survey output from the 1988 survey, for the same reasons as Qian uses 1990 data; migration was strictly controlled, so it is natural to make the assumption that the survey recipients were born in the counties they are currently residing in. Thus

birth year and cohort-size in sample per county can be deduced. Unfortunately there is no variable in the CHIP confirming how long individuals have lived in their current county, unlike in Qian's dataset, where individuals report whether or not they have been living in the county for the past five years, allowing her to restrict her sample to individuals who report having lived in their respective counties for at least the last five years. Thus my assumptions are somewhat weaker than Qian's. In addition, the data used by Qian should be regarded as more representative of demographics, as it is a population census. The data I use is primarily a household income survey created by researchers whose main focus is income and employment data rather than population composition.

The 1988 sample contains 102 131 observations altogether. When matching these with the agricultural variables from Qian (2008) and removing observations for which age and/or gender is missing, I am left with a sample of 39 941 individuals, from 222 counties, from ages 0 to 98. When I restrict my sample to individuals born 1962 and later, I am left with 20 950 observations, and the sample for the narrower DID-estimate, cohorts 1970-1988, consists of 13 786 observations. In order to control for the fraction of ethnic Han in each cohort I have to further restrict my sample as 11 798 observations on ethnicity are missing in the larger sample and 10 759 observations are missing in the restricted sample.

I use the samples of 20 950 – and 13 786 observations for the regressions where I do not control for Han. When controlling for Han the sample sizes are reduced to 9 152 and 3 027 observations respectively.

Output from the following equations are displayed in columns (1) – (4) of Table 8 below for the whole and the restricted samples, where all variables and subscripts are the same as those explained in section 2. I repeat the equations here,

$$(3) \text{ Sex}_{ic} = (\text{teadum}_i \times \text{post}_c)\beta + (\text{orchard}_i \times \text{post}_c)\delta + (\text{cashcrop}_i \times \text{post}_c)p + \text{post}_c\tau + \gamma_c + \varepsilon_{ic}$$

$$(5) \text{ Sex}_{ic} = (\text{teasown}_i \times \text{post}_c)\beta + (\text{orchard}_i \times \text{post}_c)\delta + (\text{cashcrop}_i \times \text{post}_c)p + \text{post}_c\tau + \gamma_c + \varepsilon_{ic}$$

Output from the following equations repeated from section 2 are displayed in table 9 below, the difference from Table 8 being including a control for Han on all samples.

$$(4) \text{ Sex}_{ic} = (\text{teadum}_i \times \text{post}_c)\beta + (\text{orchard}_i \times \text{post}_c)\delta + (\text{cashcrop}_i \times \text{post}_c)p + \text{Han}_{ic}\zeta + \text{post}_c\tau + \gamma_c + \varepsilon_{ic}$$

$$(6) \text{ Sex}_{ic} = (\text{teasown}_i \times \text{post}_c)\beta + (\text{orchard}_i \times \text{post}_c)\delta + (\text{cashcrop}_i \times \text{post}_c)p + \text{Han}_{ic}\zeta + \text{post}_c\tau + \gamma_c + \varepsilon_{ic}$$

Table 8: OLS estimates of the effect of planting tea and orchards on the fraction of males, whole and restricted sample.

	Dependent variable			
	Fraction of males			
	(1)	(2)	(3)	(4)
teadumpostic	-0.0018 (0.022)	-0.0130 (0.030)		
teasownpostic			-0.1246 (0.084)	-0.2266** (0.092)
orchpostic	-0.0079 (0.041)	0.0208 (0.054)	-0.0299 (0.045)	-0.0196 (0.059)
cashcroppostic	0.0093 (0.010)	0.0042 (0.012)	0.0177 (0.012)	0.0195 (0.015)
postc	0.0203** (0.010)	0.0153 (0.012)	0.0184* (0.010)	0.0108 (0.012)
Observations	20,950	13,786	20,950	13,786
R-squared	0.073	0.086	0.074	0.087

Table 9: OLS estimates of the effect of planting tea and orchards on the fraction of males, including control for Han, whole and restricted sample.

	Dependent variable			
	Fraction of males			
	(5)	(6)	(7)	(8)
teadumpostic	0.0513 (0.037)	-0.0052 (0.065)		
teasownpostic			-0.0015 (0.154)	-0.2716 (0.265)
orchpostic	-0.2115 (0.267)	0.1530 (0.238)	-0.1113 (0.250)	0.1434 (0.255)
cashcroppostic	0.0494** (0.019)	0.0600** (0.028)	0.0517** (0.022)	0.0686** (0.031)
Hanic	-0.1349* (0.077)	-0.2365* (0.122)	-0.1359* (0.077)	-0.2390* (0.122)
postc	0.0343 (0.028)	0.0312 (0.038)	0.0452* (0.024)	0.0288 (0.034)
Observations	9,152	3,027	9,152	3,027
R-squared	0.147	0.292	0.146	0.292

Discussion of results from regression output, Tables 8 and 9

In column (4) of Table 8, the DID estimate β shows a significant negative effect of planting tea after 1979 on the fraction of males in a county, consistent with findings in Qian (2008). Columns (1) – (3) also show an average negative effect of tea, but the coefficients on tea are not statistically significant. The coefficients on the variables post and cashcroppost are also of the same sign as in Qian (2008), whereas the coefficient on orchpost is ambiguous and not statistically significant.

Note that in columns (5) – (8), where I include data on ethnicity, a higher fraction of ethnic Han actually lowers the fraction of males, contrary to the results in Qian (2008) and theoretically plausible effects of the one child policy.

Tea sown is in this sample too negatively correlated with the fraction of ethnic Han, and the converse is true for orchard. When checking for interaction effects between tea and Han, this effect is positive if I use a dummy for tea sown and negative if I use the continuous teasown.

However the sample sizes are small, so it is not likely that a general effect in the population is captured. In addition, mean sex ratio in the sample with data on Han from 1962-1988 with 9 152 observations is 0.445668, and in the sample from 1970-1986 it is 0.4557023. Since the data is collected by surveys, it may be that women are more meticulous in filling out the forms as data on ethnicity is missing on a greater fraction of males than females. We also know from Qian (2008) and other papers such as Sen (1990) as well as official statistics yearbooks from China that these sex ratios are not representative. I therefore do not conclude anything from these results.

5.5 Performance in Middle School

I use the CHIP dataset from 2002 to examine the effect of tea and orchard after 1979 on performance in middle school.

One theory is that women care more about the educational outcome for their children than do men. If this is true, and the treatment effect of tea in 1979 has an effect on women's status and girls' survival rate, then we may expect children to perform better in school in tea-regions. Better performance in middle school can lead to more people obtaining higher education, which may lead to less focus on archaic patriarchal values so that fewer girls will be aborted. Alternatively, education can assist people in achieving awareness of the existing gender bias and its' potential dangers, which could also lead to less discrimination in the womb.

Performance in Middle School is measured on a scale from 1 to 5 in both the urban and the rural dataset, with 1 being the best and 5 the lowest. Results can easily be obtained from class rankings and test scores in the much standardized Chinese school system, but are essentially reported by the subjects or their parents in this survey. Having lived in China for four years myself, and having many Chinese friends and acquaintances, my impression is that the Chinese are in fact very honest in their evaluations of themselves and their children, so I assume reported performance to comply with actual observed measures of performance.

The labels on the urban and rural scales are not identical, but I interpret them as comparable in an ordinal sense, so I merge the urban and rural variables into one, named “performance”. I create a new variable with the average school performance in each cohort c in county i , “avgperformance $_{ic}$ ”.

I estimate the following equations for cohorts 1970 – 1986¹, where all other variables and subscripts are equivalent to those explained in section 2. Results are displayed in columns (9) and (10) of Table 10 below.

$$(11) \text{ avgperformance}_{ic} = (\text{teadum}_i \times \text{post}_c)\beta + (\text{orchard}_i \times \text{post}_c)\delta + (\text{cashcrop}_i \times \text{post}_c)p + \text{post}_c\tau + \gamma_c + \epsilon_{ic}$$

$$(12) \text{ avgperformance}_{ic} = (\text{tea}_i \times \text{post}_c)\beta + (\text{orchard}_i \times \text{post}_c)\delta + (\text{cashcrop}_i \times \text{post}_c)p + \text{post}_c\tau + \gamma_c + \epsilon_{ic}$$

¹ I do not run these regressions on the sample of cohorts 1962-1990 for two reasons: The Cultural Revolution may have distorted the education of the older cohorts whereas the youngest cohorts may not yet have finished Middle School.

Table 10: OLS estimates of the effect of planting tea and orchard on performance in middle school, restricted sample

	Dependent variable average performance	
	(9)	(10)
teadumpostic	-0.0286 (0.056)	
Teapostic		-0.2119** (0.103)
Orchpostic	-0.2610** (0.098)	-0.3412*** (0.102)
cashcropostic	0.1047*** (0.038)	0.1381*** (0.038)
Postc	-0.1904*** (0.034)	-0.2058*** (0.034)
Observations	5,145	5,145
R-squared	0.521	0.523

Discussion of results from regression output in Table 10

The DID estimate in column (10) is significant and in support of the hypothesized theory. Cohorts born in tea planting counties after the reforms tend to score better in school than cohorts in counties with no gender-specific crop production. However, the coefficient on the male-specific crop orchpost is also highly significant and of a larger size than that on teapost, suggesting that cohorts born in orchard-planting counties after the reform also tend to score better in school, remembering that the negative coefficient indicates better performance, as 1 is best and 5 is worst in this data. Thus the rise in the relative share of either female or male wages causes kids to perform better in school, whereas the rise in total household income has the opposite effect.

It is difficult to interpret the results in terms of a household model where both husband and wife have bargaining power depending on his or her economic value, analogous to Qian (2008). In theory, my results may imply that both men and women value the quality of their children's education, and both will devote more resources to it when their relative income share rises. The strange result is that this should not happen if total household income rises.

Instead I propose one of the following: The labour force may be predominantly male in orchard-planting counties, so that more women will be available at home to help children with school work.

Conversely for tea-planting counties. In general cashcrop-planting counties both women and men may work, so that they have less time to help out with their children's school work.

Alternatively; orchard picking requires strength and height and may therefore favour *adult* males. In tea picking *adult* women are preferred, as the tea leaves require special care. The handling of other cashcrops may not require any special skills, so children may also be involved in their production, taking time from their schoolwork and thus causing them to perform worse. Although general cashcrops are said to be gender-neutral, it is possible that child labour is extensive in areas that plant such crops.

For further study it is possible to organize the performance by gender in this dataset to test the prediction that orchard results are driven by boys and tea results by girls.

5.6 Health status

I use the CHIP dataset from 2002 to examine the effect of tea and orchard after 1979 on the health status of cohorts in the different counties.

One theory is that women care more about their children's health status than do men. If this is true we may expect to see cohorts with better health in tea-regions than in non-tea regions. Better health improves the quality of peoples' lives and often makes people happier (Arnqvist 2009). Health and happiness are desirable in themselves, and perhaps increasingly so when the current skewed sex ratio is expected to affect the happiness of men negatively.

Health status is self-reported and not directly measured, but since the respondents are of similar socioeconomic status I make the assumption that they will report similarly. Health status is measured on a scale from 1 to 5 in both the urban and the rural dataset, 1 being the best and 5 the worst. Again the labels in the two datasets are not identical, but comparable in an ordinal sense.

I create the variable for average health by taking the average of the reported health status in each cohort i in each county c .

$$(13) \text{ avghealth}_{ic} = (\text{teadum}_i \times \text{post}_c)\beta + (\text{orchard}_i \times \text{post}_c)\delta + (\text{cashcrop}_i \times \text{post}_c)\rho + \text{post}_c\tau + \gamma_c + \varepsilon_{ic}$$

$$(14) \text{ avghealth}_{ic} = (\text{tea}_i \times \text{post}_c)\beta + (\text{orchard}_i \times \text{post}_c)\delta + (\text{cashcrop}_i \times \text{post}_c)\rho + \text{post}_c\tau + \gamma_c + \varepsilon_{ic}$$

These two regressions are run on both the whole sample of cohorts 1962-1990 and on the restricted sample of cohorts 1970-1986, yielding output columns (11)-(14) of Table 11.

Table 11: OLS estimates of the effect of tea and orchard production on the average health status of children, whole and restricted sample

	Dependent variable			
	Average health status of cohort c in county i			
	(11)	(12)	(13)	(14)
teadumpostic	-0.1078** (0.047)	-0.0705 (0.051)		
teapostic			-0.0975 (0.148)	0.0399 (0.086)
orchpostic	-0.0016 (0.068)	-0.0408 (0.059)	-0.0158 (0.083)	-0.0106 (0.064)
cashcroppostic	-0.0013 (0.029)	0.0189 (0.026)	-0.0021 (0.038)	0.0012 (0.028)
postc	-0.1066*** (0.028)	-0.0558* (0.030)	-0.1227*** (0.030)	-0.0599* (0.031)
Observations		9,277	5,146	9,277
R-squared	0.545	0.542	0.541	0.540

Discussion of results from regression output, Table 11

Column (11) shows a significant negative coefficient on teadumpost, suggesting that cohorts in tea planting counties after 1979 have significantly better health. Although not statistically significant, column (12) also points in the same direction. This is in line with the theory that women are more concerned with health. One concern with such an estimate is that the older cohorts may have worse health precisely because of their age. In the widest sample, the oldest cohorts are 40, which is still a relatively young age. Even if they do become sick slightly more often than younger cohorts, the infant and youngest cohorts will also potentially have a greater probability of being sick. I thus argue that this possible slight negative bias in both treatment and control groups evens out, such that they are comparable.

5.7 Other interesting variables in CHIP

The datasets also include variables for marriage status and employment statistics, which would be interesting to examine before and after the reforms. Since this data was collected as late as 2002 however, the oldest cohort after the reform is only 23 years old. Doing a difference-in-differences

regression with the control group consisting of individuals born close to before the reform is not a good idea as this group will be much more likely to be married or to have a job due to their older age, not because of the treatment that caused skewed sex ratios.

There was also conducted a survey in 2009, so it might be fruitful to examine such statistics with this material. The data from this survey is currently (November 2012) not publically available online. However the cohorts after the reform may still be too young a “treatment group” in 2009, so perhaps it is better to use an even later survey.

SECTION 6: THEORETICAL IMPACTS OF ECONOMIC GROWTH ON GENDER ISSUES

6.1 Introduction

In this section I will briefly present a summary of theoretical links from economic growth to the survival rate of girls and the situation for women, some of which were also mentioned in the introduction. I will then discuss potential consequences to society of excess numbers of males.

6.2 Links and mechanisms from economic growth to the status of women

Economic growth → More jobs and increased education for women

An expanding economy leads to the creation of more jobs, where much of the idle new workers are necessarily female (Duflo 2011). Growth also leads to higher demand for skilled workers, so the general education level is likely to rise. With females being increasingly integrated in all spheres of the job market, their level of education will also likely increase (The Economist 2006). In addition, educated persons are more likely to see the value of both sexes having equal opportunities to obtain an education, which will further enhance female educational enrollment.

Economic growth → Expansion of health care → Higher survival rate for girls → Better opportunities

Economic growth also tends to bring about an expansion in the health care sector (Duflo 2011). As health care becomes widely available, there will be a disproportionate increase in resource use on girls, as under extreme poverty parents have tended to favour medical treatment of their male

offspring (Duflo 2011). Furthermore, there will be less risks associated with pregnancy, so life expectancy and productivity of girls will be higher, giving them a lower 'opportunity cost' in terms of boys, than earlier (Duflo 2011).

Economic growth → Drop in fertility → More resources spent on girls

Economic growth is also associated with a drop in fertility (Duflo 2011). With publically or otherwise organized pensions and health care, parents are no longer as dependent on offspring to secure their old age. They thus have less incentive to have many children. Furthermore, working women do not have as much time for childcare as women in traditional societies. Living costs also tend to rise, as well as peoples expectations of living standards, which contributes to lowering birth rates as children become more expensive. Contraception and abortion technologies that become available also make it simpler to plan and control births. More resources spent on each child raises their living standards and opportunities, often disproportionately in favour of girls as their 'care level' is initially lower (Duflo 2011).

Economic growth → Drop in fertility /contraception available → Less girls born

Despite the good news for girls described in the above paragraph, a new form of female discrimination is taking place in that less females are being born, in large part due to sex-selective abortion (Qian et al. 2008). As explained, the societies where this is evident have a strong preference for males. The male stays with the family and takes care of the parents, works for them, and through him runs the family lineage. Religious spheres, which are very important in these societies, are also often reserved for- or in favour of men. Men have always had- and still have better work opportunities. For those who are risk-averse- which most humans are to some extent- it has thus been considered more safe to have a male. Although many of the economic incentives to have males are diminishing, the tradition of preferring males is so deeply rooted in these cultures that ultrasound technology has lead to many abortions due to male sex preference (Qian et al. 2008).

6.3 Hypothetical consequences of male-biased sex ratios

I now present hypothetical consequences of the male-biased sex ratios. Data on these variables could be interesting for further study.

Bride import

Bride import is already a rising phenomenon in China (Dasgupta 2011). In itself it is not a harmful consequence if it is handled in a respectful manner, and is done voluntarily for both bride and groom. In practice it can often be more difficult to make a life together with someone from a different culture however. A study from the US finds that racial and religious differences within couples are correlated with marital instability (Heaton 2002). Moreover there is a natural limit to the practice of bride import, as there are not enough women in the world to satisfy potential demand.

Trafficking

With bride import being both expensive and in the limit impossible, human trafficking may increase. Strong sexual frustration along with the personal desire- and the cultural norm of male-female cohabitation, may lead more men to steal brides both domestically and abroad. Within China, criminal gangs are already said to kidnap women in remote rural areas to sell them into forced marriages in towns (Smil 2005). There is also evidence of Vietnamese women being trafficked into China (Le Bach et al. 2007). This is a serious threat to women's safety and the general order of society.

Sex industry

With a shortage of brides, there is likely to be an increase in demand for sex that can be bought. This can make more women engage in prostitution or other sex-related work (Kurlantzick 2007). This makes more women criminal, changes morale and probably the women's life standard and assists in enhancing the black economy. It also takes resources away from work considered more productive, probably further enhancing the black economy at the expense of tax-paying work. On the other hand, those who are in the sex industry regardless of the increased pay will now earn more money, which may lead to better outcomes for them and their families. For both men and women increased prostitution and general risk-taking behavior with regards to sex is likely to spread HIV in China (Tucker et al 2005).

General Crime rate

A higher sex ratio will trivially raise the crime rate if men are more prone to crime than women, with ample evidence suggesting that they are (Edlund et al 2007). In addition, the fact that more men will remain unmarried increases their free time, thus there will potentially be more time spent on illegal activities. Sexual frustration may further divert men to spend more time on criminal activities, such as paying for sex.

Corruption

As professional and personal attainments are important, corruption may increase. If a wife cannot directly be bought; housing, education, titles or work can be, in both legal and illegal ways. Corrupt activity potentially affects society's morale, creates unfair competition, and draws resources away from productive work in the formal sector. Edlund et al. (2007) however reject such an effect because the rich and corrupt can usually get wives either way. Corruption is often hard to measure, so it is difficult to know for sure.

Alcohol consumption

Studies from many countries show that men on average drink more alcohol, see for example Wilsnack et al. 2000. In this study from ten countries around the world, the following was true:

“In all countries, men were more likely to drink than women, and male drinkers consumed alcohol more frequently and in larger amounts, and were more likely to have alcohol-related problems than female drinkers. The consistency of this pattern across countries (and in other research literature) suggests that gender differences in drinking behavior may be biologically influenced.” (Wilsnack et al. 2000)

An increase in the fraction of males thus likely leads to a higher consumption of alcohol. In addition, if getting married and living with a person of the opposite sex is crucial to living a happy life, then the millions of men who will not be able to marry may tend to other sources of 'amusement'. We may then expect to see a further increase in the consumption of substances such as alcohol and even stronger drugs. Abuse or addiction can lead to lower happiness as well as lower productivity, as substance abuse clearly affects job performance. Moreover it potentially draws unnecessary resources from the health sector, as alcohol and drug abuse has a negative impact on health.

Productive free time

More men with more free time could also devote their efforts to fruitful causes such as striving for a more just political system, possibly bringing a positive effect to society. An increase in the pressure for democratic reforms in China could potentially speed up the process of obtaining civil liberties along the line of “the West”.

In addition, more free time may be spent on work, possibly enhancing growth. Wei & Zhang (2011) show that new domestic firms are more likely to emerge from areas with high sex ratios, that men are more likely to be entrepreneurs than women and further that men in areas with high sex ratios are more likely to be entrepreneurs (Wei & Zhang 2011). An increase in risk-taking behavior caused by competition further enhances entrepreneurship among men or families with sons (Wei & Zhang 2011). In this sense it may not be a coincidence that China’s economic growth has coincided with a growing sex ratio.

Personal health

Research has shown that the married tend to be in better health than the unmarried (Waite & Gallagher 2000). Jin & Christakis of Harvard University find that men who reach sexual maturity in an environment with few available women are more likely to die earlier than men in environments with a more even sex ratio (Jin & Christakis 2009). If women are the main caregivers in a household, the lack of women can lead to deterioration of men’s health (The Economist 2010). If a lack of brides leads to higher alcohol and/or drug consumption among men, this can further have a negative impact on their health. Prostitution is as mentioned also likely to spread the prevalence of HIV in the population (Tucker et al. 2005).

High pressure in a competitive society

Standard microeconomic theory tells us that the price of a female will rise with the relative scarcity. With supply being fixed (for now), men may experience increased pressure on their performance in life (Kurlantzick 2007). On the one hand, this may increase productivity, as competition may drive people to perform better. However, it is not certain that economic productivity will be affected for the better. We may encounter a tragedy of the commons if the sum of each man’s selfish strive does not automatically bring about the optimal outcome, without some form of cooperation.

Furthermore, in societies where competition is already fierce due to the vast populations, men may experience a reduced quality of life because of the pressure (Kurlantzick 2007).

Educational attainment

Along with the general rise in human capital, competition among men may further increase their chosen years of schooling, as job prospects and status then generally rise. Additionally, some men may have more free time to study, because of the lack of women to marry. Increased human capital can have a positive impact for individuals and for society, but we also know that the marginal benefit of education is decreasing (Trostel 2005). Resources may thus not be allocated most efficiently.

Masculinity- Cultural Change

When more men have more free time, it likely has an impact on male culture and masculinity. Recently in China there has been a surge in the production of magazines for men, quantity of gym memberships and supply and demand of male fashion clothing (Petzinger 2012). This may be a beneficial factor for self-realization and happiness of men. It also creates new business opportunities and can stimulate the economy at the micro-level. There is however almost definitely a simultaneous effect of modernity that will drive this movement, as well as the fact that a rise in income level has tended to bring people to a different stage of Maslovz hierarchy of needs. Thus the important causal link is not necessarily the higher fraction of males, though it is possible to imagine that it plays a role.

Homosexuality

In traditional societies, homosexuality is generally not a very known phenomenon and even if known, it is not generally accepted. With millions of men never finding a mate of the opposite sex, it can be speculated that the cost of being openly gay will decrease, if sexual desire and the desire to cohabit is strong. In this way, the rate of 'open' homosexuals may increase (Hesketh 2007). However there is also the effect of modernity and impressions from the West that will drive this result simultaneously, so it may be difficult to identify such an effect.

Savings rate

If wealth is important in finding a mate, the savings rate is likely to increase. Chinese parents with a son will likely raise their savings to improve their son's possibilities within the marriage market, with the pressure of savings afflicting other households (Wei & Zhang 2009). In China the savings rate is already very high, harming growth and job creation as domestic demand is low. Former chairman of the Federal Reserve, Alan Greenspan, has also claimed that Chinese savings in form of investment abroad have contributed to the housing bubble in the US and the following global financial crisis. Thus increased savings as a consequence of sex-ratio imbalance may be harmful to society.

Social unrest

The lack of women can lead more men to have more free time as well as it can lead to more sexual frustration. It is possible that men's efforts will then increasingly be devoted to other causes, possibly causing a surge of increased political or social activity. In undemocratic countries such as China, it may be a concern that this will unravel in form of demonstrations against the regime. Such demonstrations are already prevalent in China, and some can turn violent. If such demonstrations become more frequent, and an increasing proportion of men take part in these, we may expect more violence in the future, as men are more likely to engage in violent activity (Edlund et al 2010).

Happiness

As most men and women chose to marry and live together, the lack of choice to do this will probably have a negative effect on the happiness of unmarried bachelors. Even for those who do manage to find a wife, the pressure of competition can be a big strain before "sealing the deal", which can reduce quality of life. Says Wang Zheng, Professor in Women Studies at the University of Michigan "A lot of young men are very depressed(...)There are a lot of psychological problems amongst male graduates, brought on by poor job prospects and, by extension, poor marriage prospects" (Petzinger 2012). The converse can be true for women however, as they will have more choices in their search for a mate, and can potentially "marry up" in status (Hesketh 2007).

Increased choices for women

In general, the scarcity of women can lead their choice range to be extended in many contexts. As mentioned above, they will have more alternatives in the search for a husband, causing some to be able to marry more educated men from better families (Hesketh 2007). Additionally, in work places where affirmative action is used to ensure a certain ratio of women, the smaller proportion of women in the present cohorts can increase their probability of obtaining such positions.

Public goods choices

There are studies that show that men and women will prioritize different public goods, an example from India can be seen in "Chattopadhyay and Duflo 2004". In this study the authors compare investment in public goods in areas with an affirmative action policy stating that the leader must be female, to areas where there is no such policy. Their empirical investigation shows that women elected as leaders under the reservation policy invest more in the public goods more closely linked to women's concerns: drinking water and roads in West Bengal and drinking water in Rajasthan. They invest less in public goods that are more closely linked to men's concerns: education in West Bengal and roads in Rajasthan.

If no affirmative action is required in government, the larger fraction of males in a country may lead to a greater fraction of politicians being male. In this scenario we can imagine more male-type public goods being prioritized (Guilmoto 2007). On the other hand, if affirmative action is present and the status of women simultaneously is being raised due to their scarcity, then we can imagine that women's voices will be increasingly important in decision-making, causing more investment in female-type public goods.

7 CONCLUSION

Studying mechanisms that can reverse the current trend of sex-selective abortion in China and other parts of Asia is important in order to achieve equality between the sexes and to avoid potential harmful consequences of unnatural demographics. Although the possible causal link from targeting women to spurred development has been downplayed by The World Bank, research on this area should still be important in addressing the problem of a surplus of males, as well as the strive for equality between the sexes for the sake of equality in itself.

When providing policy advice based on empirical regression analysis, it is vital to consider interaction effects and treatment heterogeneity to be sure that policies will be conducted in an optimal way. My analysis from Section 3 shows that the results provided in Qian (2008) may suffer from such treatment heterogeneity, and her results may therefore be overestimated.

From my replication of Qian (2008) in Section 2 and the testing of her method on variables from the Chinese Household Income Project in Section 5, it does appear that there are differences between tea and non-tea regions, and orchard and non-orchard regions. Tea areas do tend to have less skewed sex ratios. Further, tea areas tend to have populations with better health than orchard- and cashcrop-planting areas, and children in both tea and orchard regions tend to score better in Middle School than children in cashcrop-planting regions. Although these results may not be driven entirely by tea as a proxy for women's wages and economic opportunities, it is still relevant to study these areas to identify other potential mechanisms that help drive the results, which can possibly be adapted elsewhere.

There is an emerging empirical literature on the dangers of having excess males in a society, addressing issues such as those discussed in section 6, Edlund et al. (2007) is a good example. I propose to use the Difference-in-Differences framework and counties from Qian (2008) on recent data on such variables to examine differentials between tea and non-tea regions, as well as the other crops. It may thus be possible to examine the magnitude of different effects of skewed sex ratios, which is an important area of research at the moment, as large male-biased cohorts are becoming of adult age.

8 SOURCES

Scholarly articles, reports and speeches

Qian, Nancy 2008. "Missing women and the price of tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance" *The Quarterly Journal of Economics* (2008) 123 (3): 1251-1285.

Darnovsky, Marcy 2009. "Countries with laws or policies on sex selection", Center for Genetics and Society. Web. Accessed 10.11.2012.

http://geneticsandsociety.org/downloads/200904_sex_selection_memo.pdf.

Duflo, Esther 2011. "Women's empowerment and economic development". Massachusetts Institute of Technology. NBER Working Paper No. 17702, Issued in December 2011.

Edlund, Lena; Li, Hongbin; Yi, Junjian, Zhang, Junsen 2007. "Sex Ratios and Crime: Evidence from China's One-Child Policy". Discussion paper, Institute for the Study of Labor (IZA).

Guilmoto, Christophe 2007. "Characteristics of Sex-Ratio Imbalance in India and Future Scenarios" UNFPA for presentation at the 4th Asia Pacific Conference on Sexual and Reproductive Health. Web. Accessed 02.01.2013. <http://www.unfpa.org/gender/docs/studies/india.pdf>.

Heaton, Tim B. 2002. "Factors Contributing to Increasing Marital Stability in the United States". *Journal of Family Issues* 2002 23: 392.

Hesketh, Therese 2007. "Too many men". UCL Lunch Hour Lecture Series. Web. Accessed 20.06.2012. http://www.ucl.ac.uk/lhl/lhlpub_autumn07/11_201107.

Horibe, Nobuko 2011. "Speech of UNFPA Regional Director, Ms. Nobuko Horibe at the International Workshop on Skewed Sex Ratio at Birth". UNFPA Vietnam, October 2011. Web. Accessed 30.10.2012. <http://www.un.org.vn/en/unfpa-speeches/2032-speech-of-unfpa-regional-director-ms-nobuko-horibe-at-the-international-workshop-on-skewed-sex-ratio-at-birth.html>.

Hvistendahl, Mara 2011. "Unnatural Selection". First Edition, Public Affairs New York.

Jin, Lei & Chrisatakis, Nicholas A. 2009. "Investigating the mechanism of marital mortality reduction: The transition to widowhood and quality of health care". *Demography*, Volume 46-Number 3, August 2009.

Le Bach D, Bélanger D, Khat TH. 2007. "Transnational migration, marriage and trafficking at the China-Vietnam border". *Watering the neighbour's garden: The growing demographic female deficit in Asia*, edited by Isabelle Attane and Christophe Z. Guilmoto. Committee for International Cooperation in National Research in Demography [CICRED], 2007. p. 393-425.

Qian, Nancy; Lin, Ming-Jen; Liu, Jin-Tan 2008. "More Women Missing, Fewer Girls Dying: The Impact of Abortion on Sex Ratios at Birth and Excess Female Mortality in Taiwan". NBER Working Paper No. 14541.

Sen, Amartya 1990. "More than 100 million women are missing". The New York Review of Books, December 20, 1990.

Smil, Vaclav 2005. "The Next 50 Years: Unfolding Trends". Yale University. Population and Development Review (2005) ,Volume: 31, Issue: 4. Population Council.

Steffensmeier, Darell; Allan, Emilie 1996. "GENDER AND CRIME: Toward a Gendered Theory of Female Offending". Annual Review of Sociology 1996, 22: p. 459-87.

Trostel, Philip A. 2005. "Nonlinearity in the return to education" Journal of Applied Economics vol 001, 2005. Web. Accessed 15.12.2012. <http://redalyc.uaemex.mx/pdf/103/10380109.pdf>.

Tucker et. al 2005. "Surplus men, sex work, and the spread of HIV in China". AIDS: 8 April 2005 - Volume 19 - Issue 6 - p 539-547. Editorial Review.

Waite, Linda J & Gallagher, Maggie 2000. "The case for marriage". Broadway Books 2000

Wang, Xin & Wen, Yi 2012 "Housing Prices and the High Chinese Saving Rate Puzzle". Working Paper 2012-038A. Web. Accessed 02.01.2013 <http://research.stlouisfed.org/wp/2012/2012-038.pdf>

Wei, Shang-Jin & Zhang ,Xiaobo 2011. "Sex Ratios, Entrepreneurship, and Economic Growth in the People's Republic of China" NBER Working Paper No. 16800

Wei, Shang-Jin & Zhang, Xiaobo 2009. "The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China". NBER Working Paper No. 15093.

Yuan, Xin 2005. "High Sex Ratio at Birth and Female Deficit in China ". Institute of Population & Development, Nankai University. Paper presented at International Conference on "Female Deficit in Asia:Trends and Perspectives" in Singapore 2005.

Knowledge Databases

CIA World Factbook 2012. "China". Web. Accessed 10.10.2012
<https://www.cia.gov/library/publications/the-world-factbook/>

National Bureau of Statistics (NBS) of the People's Republic of China 2012. "China Statistical Yearbook 2011". NBS. Web. Accessed 15.09.2012. <http://www.stats.gov.cn/tjsj/ndsj/2011/indexeh.htm>

United Nations Food and Agriculture Organization (UN FAO) 2012. "China". Sustainable Development Department FAO. Web. Accessed 28.11.2012_ <http://www.fao.org/sd/WPdirect/WPre0107.htm>

United Nations Populations Fund (UNFPA) 2011. "Report of the International Workshop on Skewed Sex Ratios at Birth: Addressing the Issue and the Way Forward." UNFPA 5-6. October 2011. Web. Accessed 10.11.2012
http://www.unfpa.org/webdav/site/global/shared/documents/publications/2012/Report_SexRatios_2012.pdf

United Nations Population Fund (UNFPA) 2012. "Sex Imbalances at Birth: Current trends, consequences and policy implications". UNFPA Asia and the Pacific Regional Office, August 2012.

<http://www.unfpa.org/webdav/site/global/shared/documents/publications/2012/Sex%20Imbalances%20at%20Birth.%20PDF%20UNFPA%20APRO%20publication%202012.pdf>

US Department of Labour Statistics 2011, "Highlights of Women's Earnings in 2011". Web. Accessed 30.11.2012. <http://www.bls.gov/cps/cpswom2011.pdf>

WHO 2012 , Gender and Genetics, Genomi Resource centre. Web. Accessed 15.12.2012 <http://www.who.int/genomics/gender/en/index4.html>

World Bank 2011. "World Development Report 2012: Gender Equality and Development". The World Bank Group , published September 28, 2011.

Journal and Newspaper articles

Arnqvist, Sarah. "What makes us happy". New York Times 30.06.2009. Web. Accessed 15.12.2012 <http://newoldage.blogs.nytimes.com/2009/06/30/what-makes-us-happy/>

Dasgupta, Saibal 2011. "China fears rising bride imports due to imbalance in sex ratio". The Times of India. Web. Accessed 03.01.2013. http://articles.timesofindia.indiatimes.com/2011-08-28/china/29937854_1_china-fears-females-china-academy

Haines, Renee. "Getting married later in life is a modern trend" China Daily 21.02.2011. Web. Accessed 10.10.2012 http://www.chinadaily.com.cn/life/2011-02/21/content_12048369.htm

Kurlantzick, Josh 2007. "China's future: A Nation of Single Men?". The Los Angeles Times October 2007. Carnegie Endowment.

Petzinger, Jill 2012. "Who's the man". That's PRD magazine 01.11.2012. Web. Accessed 02.01.2013. <http://thatsmags.com/prd/article/view/10635>

The Economist 2012. "Changing Migration Patterns. Welcome Home" The Economist Magazine 25.02.2012. Web. Accessed 31.10.2012. <http://www.economist.com/node/21548273>

The Economist 2010. "Health and the sex ratio. A healthy relationship". The Economist Magazine 11.08.2010. Web. Accessed 03.02.2013. <http://www.economist.com/node/16789152>

The Economist 2006. "Women in the workforce. The importance of sex". The Economist Magazine 12.04.2006. Web. Accessed 03.01.2012. <http://www.economist.com/node/6800723>

Zhou, Xiaoli 2005. "Rough Cut : The Women's Kingdom, In China, how free can a woman be?". Public Broadcasting System. Web. Accessed 10.10.2012. http://www.pbs.org/frontlineworld/rough/2005/07/introduction_to.html

Other

Wikipedia, The Free Encyclopedia 2012. "Age at first marriage". Web. Accessed 15.10.2012.
http://en.wikipedia.org/wiki/Age_at_first_marriage

Data

Interuniversity Consortium for Political and Social Research (ICPSR) 1988. "Chinese Household Income Project 1988". Web. Downloaded 17.10.2012
<http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/9836>

Interuniversity Consortium for Political and Social Research (ICPSR) 2002. "Chinese Household Income Project 2002". Web. Downloaded 17.10.2012
<http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/21741>

Qian, Nancy 2008. "Data for replicating my results". Web. Downloaded 01.09.2012.
http://www.econ.yale.edu/~nq3/NANCYS_Yale_Website/Data.html