ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Chew, Soo Hong; Yi, Junjian; Zhang, Junsen; Zhong, Songfa

Working Paper

Risk Aversion and Son Preference: Experimental Evidence from Chinese Twin Parents

IZA Discussion Papers, No. 10519

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Chew, Soo Hong; Yi, Junjian; Zhang, Junsen; Zhong, Songfa (2017) : Risk Aversion and Son Preference: Experimental Evidence from Chinese Twin Parents, IZA Discussion Papers, No. 10519, Institute of Labor Economics (IZA), Bonn

This Version is available at:
https://hdl.handle.net/10419/161142

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10519

Risk Aversion and Son Preference: Experimental Evidence from Chinese Twin Parents

Soo Hong Chew
Junjian Yi
Junsen Zhang
Songfa Zhong

JANUARY 2017

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10519

Risk Aversion and Son Preference: Experimental Evidence from Chinese Twin Parents

Soo Hong Chew
National University of Singapore

Junjian Yi
National University of Singapore and IZA

Junsen Zhang
Chinese University of Hong Kong and IZA

Songfa Zhong
National University of Singapore

JANUARY 2017

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Risk Aversion and Son Preference: Experimental Evidence from Chinese Twin Parents*

We study the role of risk aversion underlying son preference in patriarchal societies, where sons serve as better insurance for old-age support than daughters. The implications of an insurance motive on son preference are two-fold. First, prior to the birth of their children, more risk-averse parents have a stronger preference for sons than for daughters. Second, after the birth of their children, parents with sons are more risk seeking, compared to parents with daughters. We adopt a within-twin-pair fixed-effects estimator with a weak identification assumption, which enables us to jointly identify these two effects. We further conduct an incentivized choice experiment to assess parental risk attitude in a sample of Chinese twins with children, and follow up with a second twin sample to examine the replicability of the findings. In both samples, we find that parents with greater risk aversion before the birth of their children are more likely to have sons through sex selection than parents with less risk aversion. Additionally, having sons significantly decreases parental risk aversion. These results contribute to the literature on the sources of son preference and help shed light on the nature of gender inequality.

JEL Classification: C93, D01, D80, J13

Keywords: risk aversion, son preference, twins, experimental economics

Corresponding author:
Junjian Yi
Department of Economics
Faculty of Arts & Social Sciences
National University of Singapore
AS2 Level 6, 1 Arts Link
Singapore 117570
Singapore

E-mail: junjian@nus.edu.sg

* This research project was supported by grants from the Research Grants Council of Hong Kong (Project No. CUHK4358/01H) and from the Hong Kong University of Science and Technology (Project No. AOE-MG/H-03/06).

1. Introduction

One major form of gender inequality is son preference, in which parents favor sons over daughters. Despite modernization and rapid economic growth, this preference has been deeply rooted in human history for millennia, and persists in most modern societies (see, e.g., Gangadharan and Maitra, 2003). Son preference is present even among parents in the United States (Dahl and Moretti, 2008). Studies show that son preference is persistent among second-generation Asian migrants to the United States and Canada (Almond and Edlund, 2008; Almond, Edlund, and Milligan, 2013). With modern gender-selection technology, son preference manifests itself in terms of gender imbalance. Sen (2003) concludes that in the 2000s, about 100 million women were missing worldwide. In China alone, for instance, the sex ratio at birth rose from a biologically normal range of 1.03-1.06 to 1.11 in 1990, and further increased to 1.20 in 2000, resulting in 33 million excess men under the age of 20 (Ebenstein, 2011). Given the grave social and economic consequences, son preference has been widely explored in the literature, and the increasingly imbalanced sex ratio has raised public policy concerns at the global level.[1]

The literature has extensively investigated the socioeconomic, institutional, and cultural factors underpinning the demand for sons.[2] Studies suggest that low relative adult female earnings in traditional agrarian economies are an important factor in the persistence of son preference (Ben-Porath and Welch, 1976; Rosenzweig and Schultz, 1982; Ahn, 1995; Qian, 2008). Das Gapta et al. (2013) suggest that adult sons' support of elderly parents contributes to the persistence in son preference in China, India, and South Korea. Edlund (1999) relates son preference to institutional factors such as the caste system in India. Alesina, Giuliano, and Nunn (2013) explore historical origins of cross-cultural differences in the role of women in society, and find that traditional agricultural practices, such as plowing, influenced the historical gender division of labor and the evolution of gender norms.[3]

Our paper borrows insights from behavioral economics, and presents the first at-

[1] For the long-run socioeconomic consequences of gender imbalance, see Rao (1993), Angrist (2002), Chiappori, Fortin, and Lacroix (2002), Das Gupta (2005), Ebenstein (2010, 2011), Wei and Zhang (2011), Bethmann and Kvasnicka (2012), Das Gupta, Ebenstein, and Sharygin (2013), Edlund, Li, Yi, and Zhang (2013), and Hu and Schlosser (2015).

[2] Leung (1988, 1991) proposes the econometric method to detect son preference using fertility data.

[3] Oster (2005) suggests that the hepatitis B virus contributes to the missing women in China and India. The paper was scrutinized by Lin and Luoh (2008), and subsequently retracted by Oster et al. in 2010.

tempt in the literature to propose and test an insurance motive underpinning son preference. This approach is closely related to a broader literature linking economic preferences to economic outcomes (see, e.g., Dohmen et al., 2011; Sutter et al., 2013; Noussair, Trautmann, and Van de Kuilen, 2013; Dimmock et al., 2015). More specifically, compared to daughters, sons have higher agricultural productivity, and are additionally expected to shoulder more of the responsibility of caring for their parents. Hence, sons are perceived as a better source of insurance for old age in traditional societies in which environmental and social risks are pervasive.[4] This notion has given rise to the Chinese proverb of "raising sons to insure against old age." However, the literature has presented no rigorous analyses on the demand for sons being derived from a sense of security in the presence of risk. We link risk aversion and son preference, and investigate the following two-fold hypotheses. First, prior to the birth of the child, parents who are "endowed" with greater risk aversion (hereafter, *endowed risk attitude*)[5] may have a stronger need for security, and thus show a stronger preference for sons by engaging in gender selection. Second, after the birth of the child, parents with sons feel more secure and are consequently less risk averse, compared to parents without sons.[6]

In empirically testing our hypotheses, we are mindful that the ordinary least squares (OLS) estimate is generally biased since the endowed risk attitude is unobserved and potentially correlated with the demand for sons. To address this omitted-variable bias, we build on the well-known within-twin-pair fixed-effects (FE) estimator and generalize this method by relaxing its identification assumption. The standard twin FE method assumes that the within-twin-pair differences in children's gender composition are uncorrelated with the within-twin-pair differences in any unobserved variable, i.e., twins have exactly the same risk attitude, implying that their endowed risk attitudes are perfectly correlated. We relax this identification assumption in our causal inference and assume that endowed risk attitude is more similar between twin siblings than between two randomly chosen strangers. This could be

[4] Our paper uses the term "son preference" interchangeably with "the demand for male children," and does not distinguish among the different forms of son preference.
[5] The "endowed" risk attitude may include the genetic endowment of risk attitude and all other environmental factors affecting parental risk attitude before the births of children.
[6] We derive a simple utility-based theoretical framework based on source preference (Chew and Sagi, 2008) in Appendix I to link risk aversion and son preference.

due to either genetic or environmental factors. More specifically, a sufficient condition for our identification is for the correlation coefficient between the endowed risk attitudes of the twin siblings not less than 0.5.[7] With the weaker assumption, the FE estimates deliver a tighter lower bound on the effect of the presence of sons on parental risk attitude in terms of absolute value compared to OLS estimates. Furthermore, comparing the FE and OLS estimates helps us identify the sign of the behavioral effect of endowed risk aversion on the demand for sons. As a result, our twin-based design can be used to empirically analyze not only the effect of the gender composition of children on parental risk attitude, but also the effect of endowed risk attitude on parental demand for sons.

In addition to the methodological innovation, we also address the concern of replicability, which has received increasing attention across different scientific disciplines such as genetics, psychology, and laboratory experimental economics (Lawrence et al., 2013; Open Science Collaboration, 2015; Camerer et al., 2016; Maniadis, Tufano, and List, 2014). To address this issue, we use two separate data sets of Chinese twin parents. In the first data set—the Chinese Adult Twin Survey (CATS)—we assess the risk attitude of the twin using an incentivized experimental task. In the second data set—the Longitudinal Chinese Child Twin Survey (LCCTS)—we use a choice-based questionnaire to assess the risk attitude of twins.

We have two main findings. First, the within-twin-pair FE estimates indicate that the presence of sons significantly decreases the revealed risk aversion of parents. This finding supports our second hypothesis that sons are perceived to provide better security and decrease parental risk aversion. Second, we compare the FE and OLS estimates, and find that parents with more endowed risk aversion are more likely to have sons than parents with less endowed risk aversion. This finding supports our first hypothesis that stronger endowed risk aversion leads to greater son preference and more gender imbalance. Our two identified effects suggest a loop from preferences to outcomes, and then from outcomes to preferences, when considering behavioral dynamics.

To further test the insurance motive of son preference underlying these two hypotheses, we explore the heterogeneous effects of sons on parental risk aversion. First,

[7] See the detailed discussion in Section 2 below.

we show that the presence of sons decreases the degree of risk aversion more for mothers than for fathers. This heterogeneous effect is consistent with the notion that mothers are more likely to rely on sons' support and security in their old age. Second, we find that the negative effect of children's gender composition on parental risk attitude exists only for parents working in the private sector. Compared to employees in the private sector, those in the public sector have better public insurance and high-quality old-age pension, (Meng, 2012), and hence have less demand for sons' support in their old age. Overall, these two heterogeneous effects are both in line with our hypotheses linking risk aversion and son preference.

Our paper contributes to the understanding of the causes of sex-ratio imbalance, which has been a subject of extensive studies by researchers and policy makers worldwide (Rosenzweig and Schultz, 1982; Ahn, 1995; Edlund, 1999; Oster, 2005; Lin and Luoh, 2008; Qian, 2008; Ebenstein and Leung, 2010). We study the insurance motive as a cause of the persistent son preference from the perspective of behavioral economics. In particular, Ebenstein and Leung (2010), which is the closest study to ours, find that the old-age pension program in rural China mitigates the gender imbalance in areas with better access to the program. By contrast, our paper provides direct experimental evidence to identify the causal relationship between parental risk attitude and the demand for sons, in which the presence of sons is regarded as a proxy for insurance.

We also contribute to the literature on development economics (Rosenzweig and Schultz, 1982; Rosenzweig, 1988a, 1988b, 1993). Credit markets in developing countries are generally far from perfect (Rosenzweig, 1988a, 1988b, 1993). Therefore, families in these countries usually rely on their (extended) family for risk sharing. For instance, Rosenzweig (1988a) provides an example of the insurance motive in arranging the marriage of a daughter. We complement Rosenzweig's (1988a) analysis by demonstrating an insurance motive in the parents' demand for sons.

Finally, our research contributes to the literature on family economics (Rao, 1993; Angrist, 2002; Chiappori, Fortin, and Lacroix, 2002; Wei and Zhang, 2011). Davies and Zhang (1995) are the first to model son preference using a constrained utility maximization framework by distinguishing between two sources of son preference—the differential gender return from the budget constraint and the pure son preference from the utility function. Our paper is the first to use a utility-based model to

analyze son preference in the presence of risk, based on Chew and Sagi (2008) which models distinct parental risk attitudes depending on whether the child is a son or a daughter. This framework gives rise to an experimental and empirical basis for testing the implications of the insurance motive for the persistence of gender imbalance.

Our results have major public policy implications. We suggest that the persistence of son preference may result partly from the historically low coverage of insurance and old-age pension systems, particularly in rural areas where son preference has tended to be more prevalent. The implementation of large-scale social insurance and old-age pension programs should offer an effective way to alleviate the deterioration of gender imbalance in these areas.

The rest of the paper is organized as follows. Section 2 discusses the econometric identification strategy. Section 3 describes the data, and Section 4 presents the estimation results. Section 5 concludes.

2. Empirical Strategy

2.1 Research Questions

This paper studies two closely related research questions to understand the relationship between risk attitude and son preference. On the one hand, parents with sons, ceteris paribus, will have a reduced risk aversion than parents with daughters, because sons provide more security. On the other hand, because of the need for insurance from sons, parents who are more risk averse, ceteris paribus, may prefer to have sons rather than daughters. Appendix I presents a source-dependent expected utility model to formally investigate the relationship between son preference and risk attitude.

To study the first question, we specify the following linear structural equation relating parental risk attitude (y_{ji}) of the jth ($j=1,2$) twin in pair i to the gender composition of the subject's children (S_{ji}), the subject's observed characteristics (X_{ji}), the subject's endowed risk attitude (ϑ_{ji}), and a random shock (ε_{ji}):

$$y_{ji} = \alpha + \beta S_{ji} + X_{ji}\delta + \gamma\vartheta_{ji} + \varepsilon_{ji}\,. \tag{1}$$

We use y to measure the degree of the subject's risk aversion after the births of his or her children. Empirically, we conduct an incentivized choice experiment to assess the degree of the risk aversion (y), where a higher value of y means greater risk aversion.

We use S to denote the gender composition of the subject's children. Specifically, we construct three variables to measure the gender composition of children: (i) a dummy indicating whether there is at least one son, (ii) a dummy indicating whether the first child is a son, and (iii) the ratio of the number of sons to the number of children. We use the vector X to denote the subjects' observed characteristics, such as family income and parental education that might affect the subject's risk attitude.

We use ϑ to denote the subject's "endowed" risk attitude, which is a hypothetical variable. The endowed risk attitude is defined as the unobserved risk attitude before the births of the subject's children. Thus, it may reflect the influence of both genetic endowment and environmental factors before childbirths. Notably, the endowed risk attitude is observed by subjects but unobserved by researchers. Finally, ε is a random shock to the subject's risk attitude.

Our first research question seeks to identify β in equation (1). Because the endowed risk attitude (ϑ), a hypothetical term, captures all influences of both genetic endowment and environmental factors on parental risk attitudes before childbirths, $\beta = \frac{\partial y}{\partial s}$, which measures the causal effect of children's gender composition on the degree of parental risk aversion. In this sense, ϑ is a sufficient statistic. That is, conditional on ϑ, ε is i.i.d., and thus β provides a causal interpretation. Because sons provide old-age support and security to their parents, we expect β to be negative.

In China, the sex ratio at birth is traditionally biased toward males. Prior to the introduction of ultrasound B machines in the 1980s, infanticide was one of the main means of exercising gender selection (Coale and Banister, 1994; Scharping, 2002). Since then, the widespread use of ultrasound B machines has led to gender-abortion as a means to select the gender of a child (Ebenstein, 2010; Li, Yi, and Zhang, 2011). At the same time, parental endowed risk attitudes can affect the demand for sons, e.g., because sons provide security or insurance, parents who are more risk averse may have greater incentive to practice gender selection (Ebenstein and Leung, 2010).

This brings us to our second research question concerning the relationship between parental endowed risk attitude and the demand for sons. We specify the following behavioral equation:

$$S_{ji} = \phi + \xi \vartheta_{ji} + Z_{ji}\varphi + \nu_{ji}, \tag{2}$$

where S and ϑ are defined as in equation (1) above, and Z is a vector of observed variables affecting the parental gender-selection behavior.[8] Finally, ν is a random shock to the subject's gender selection. Without gender selection, a biological process determines the child's gender. Moreover, gender-selection technology is not perfect since gender-selective abortion cannot ensure the gender of the next birth. It follows that ν measures all random shocks in determining the child's gender, conditional on parental risk attitude (ϑ) and observed parental characteristics (Z).

In equation (2), $\xi = \partial S / \partial \vartheta$ measures the causal effect of the endowed risk aversion on the demand for sons. Therefore, our second research question concerns the identification of ξ. Because sons provide insurance for parents, parents who are more risk averse have a greater demand for sons. We expect ξ to be positive.

2.2 Empirical Challenges

Equations (1) and (2) form a recursive structure. Specifically, our analysis contains two types of parental risk attitudes. The first type is risk attitude before the birth of a child, which is unobserved to the researchers and defined as endowed risk attitude in our paper. Endowed risk attitude affects son preference due to the need for insurance from sons. The second type is risk attitude after the birth of a child, which is measured by our experiment. If the child were a son, parents would feel more insured or exhibit less risk aversion than if the child were a daughter.

One empirical challenge concerns the unobservability of ϑ. Had we observed the subject's endowed risk attitude, both β and ξ could be easily identified. However, endowed risk aversion is usually unobserved. Therefore, the OLS estimate of β in equation (1) is generally biased if the individual heterogeneity of ϑ is ignored by regressing equations, such as the following:

$$y_{ji} = \alpha^{OLS} + \beta^{OLS} S_{ji} + X_{ji}\delta^{OLS} + \varepsilon_{ji}^{OLS}. \tag{3}$$

Using equations (1) and (2),

$$\beta^{OLS} = \frac{\partial y}{\partial s} + \frac{\partial y}{\partial \theta} \cdot \frac{\partial \theta}{\partial s} = \beta + \gamma \xi^{-1}, \tag{4}$$

[8] Some elements may appear in Z but are not included in X, for example, the availability of ultrasound B machines in the nearby community hospital for the household. This variable affects the cost of gender selection, but it is plausibly exogenous to the household's risk attitude.

where $\xi^{-1} = \frac{\sigma_{\vartheta s}}{\sigma_S^{\ 2}}$.[9] $\sigma_{\vartheta s}$ is the covariance between ϑ and S, and σ_S is the variance of S. Thus, β^{OLS} is a biased estimate of β. The bias with β^{OLS} is $\beta^{OLS} - \beta$, which is equal to $\gamma \frac{\sigma_{\vartheta s}}{\sigma_S^{\ 2}}$.

The instrumental variable (IV) estimation method is generally applied to identify β in equation (1). Hypothetically, if a variable, which is included in Z (equation (2)) but excluded from X (Equation (1)), is found, it can be used as an IV to identify β. The IV should be correlated with parental demand for sons but uncorrelated with parental risk attitude. Empirically, finding this IV is difficult, if not impossible.

2.3 Empirical Identification

Our twin-based data in which parents are twins enable us to perform the identification. A major methodological innovation of our study is to relax the strong identification assumption in the standard twin method. The standard within-twin-pair FE estimations require the assumption that the within-twin-pair differences in the independent variable of interest (i.e., the children's gender composition in our context) are uncorrelated with the within-twin-pair differences in any unobserved variable. Despite using the within-twin-pair FE method, the economic inference in our paper is actually derived from a weaker identification assumption that allows the unobserved to be correlated with the children's gender composition.

We take the within-twin-pair difference of equation (1),

$$\Delta y_i = \beta \Delta S_i + \Delta X_i \delta + \gamma \Delta \vartheta_i + \Delta \varepsilon_i, \tag{5}$$

where Δ is an operator of within-twin-pair difference. Because $\Delta \vartheta_{ji}$ is still unobserved, we can only estimate the following equation:

$$\Delta y_i = \beta^{FE} \Delta S_i + \Delta X_i \delta^{FE} + \Delta \varepsilon_i. \tag{6}$$

The within-twin-pair FE estimate in equation (6) becomes:

$$\beta^{FE} = \beta + \gamma \frac{\sigma_{\Delta \vartheta \Delta s}}{\sigma_{\Delta s}^{\ 2}}, \tag{7}$$

where $\sigma_{\Delta \vartheta \Delta s}$ is the covariance between $\Delta \vartheta$ and ΔS, and $\sigma_{\Delta S}$ is the variance of

[9] We ignore the covariates of X_{ji} below to simplify our discussion.

ΔS. Generally speaking, β^{FE} is also a biased estimate of β. The bias with the within-twin-pair FE estimate is that $\beta^{FE} - \beta$, which is equal to $\gamma \frac{\sigma_{\Delta\vartheta\Delta s}}{\sigma_{\Delta s}^{2}}$. The literature on the twin method usually assumes that both twin siblings share identical endowments ($\vartheta_1 = \vartheta_2$); in this case, $\sigma_{\Delta\vartheta\Delta s} = 0$. In this extreme case, the within-twin-pair FE estimate is unbiased, such that $\beta^{FE} = \beta$.

This paper relaxes the strong assumption—common in the literature—that twin siblings share identical endowments. Conceptually, ϑ in equation (1) measures not only the genetic endowment of risk aversion, but also the prenatal factors influencing parental risk attitude. The genetic endowments are not necessarily identical even between MZ twins. Furthermore, sibling-specific shocks to risk attitude, which may affect parental gender-selection decisions if they occur prior to the births of their children, may exist. Therefore, within-twin-pair FE estimation cannot completely eliminate the bias resulting from unobserved heterogeneity in the endowed risk attitude.

We develop a weaker assumption than the conventional assumption in the literature to tighten the lower bound of β in terms of absolute value, and identify the signs of both β and ξ.

Identification Assumption. The endogenous variation of the gender composition of children, induced by the endowed risk aversion within twin pairs, is smaller than that across twin pairs; that is, $|\sigma_{\Delta\vartheta\Delta s}|/\sigma_{\Delta s}^{2} < |\sigma_{\vartheta s}|/\sigma_{s}^{2}$.

This assumption is unlikely to be violated. Specifically, we check the magnitude of the four terms, namely, $\sigma_{\Delta\vartheta\Delta s}$, $\sigma_{\vartheta s}$, $\sigma_{\Delta s}^{2}$, and σ_{s}^{2}. The latter two terms, $\sigma_{\Delta s}^{2}$ and σ_{s}^{2}, are observed, whereas the other terms, $\sigma_{\Delta\vartheta\Delta s}$ and $\sigma_{\vartheta s}$, are not. We first examine the two observed terms, namely, $\sigma_{\Delta s}^{2}$ and σ_{s}^{2}. Our variables of interest are (i) having at least one son, (ii) the first child being a son, and (iii) the fraction of sons out of all children. Based on our estimation sample described below, σ_{s}^{2} equals 0.49, 0.50, and 0.46 in terms of the three variables used to measure the gender composition of children, respectively. By contrast, $\sigma_{\Delta s}^{2}$ equals 0.72, 0.81, and 0.73 in terms of the same three variables, respectively. Therefore, $\sigma_{\Delta s}^{2} > \sigma_{s}^{2}$.

We then examine the two unobserved terms, namely, $\sigma_{\Delta\vartheta\Delta s}$ and $\sigma_{\vartheta s}$. We note that

$$\begin{aligned}\sigma_{\Delta\vartheta\Delta s} &= \sigma(\vartheta_1 - \vartheta_2, S_1 - S_2) \\ &= \sigma(\vartheta_1, S_1) + \sigma(\vartheta_2, S_2) - \sigma(\vartheta_1, S_2) - \sigma(\vartheta_2, S_1).\end{aligned}$$

If we assume that ϑ and S are drawn from the same distributions for twin siblings, then $\sigma(\vartheta_1, S_1) = \sigma(\vartheta_2, S_2) = \sigma(\vartheta, S)$ and $\sigma(\vartheta_1, S_2) = \sigma(\vartheta_2, S_1)$. The endowed risk attitude is correlated with twin siblings because of the shared genetic endowments and family background. We use ρ to denote the correlation coefficient between ϑ_1 and ϑ_2. Therefore, $|\sigma_{\Delta\vartheta\Delta s}| < |\sigma_{\vartheta s}|$ if and only if $\rho > 0.5$. Essentially, the standard within-twin-pair fixed-effects estimations in the literature assume that $\rho = 1$. Thus, $\sigma_{\Delta\vartheta\Delta s} = 0$, and $|\sigma_{\Delta\vartheta\Delta s}|$ is always less than $|\sigma_{\vartheta s}|$. Here, the identification assumption is automatically satisfied, and the within-twin-pair FE estimates are unbiased. By contrast, a sufficient condition for our identification assumption is $\rho > 0.5$ when $\sigma_{\Delta s}{}^2 > \sigma_s{}^2$, which is much weaker than the existing identification assumption in the literature.[10]

We cannot directly observe the endowed risk attitudes of twin siblings, i.e., ϑ_1 and ϑ_2. Zhong et al. (2009) show a similar risk attitude in twins. In both of our twin samples, the correlation coefficients in risk attitude between Chinese twins are above 0.75. Generally, the correlation in endowed risk attitudes should be higher than the correlation in the risk attitude measured in adulthood between identical twins, because the latter contains accumulated shocks to the risk attitude of an individual over a life cycle. Therefore, ρ should be larger than 0.5, and $|\sigma_{\Delta\vartheta\Delta s}| < |\sigma_{\vartheta s}|$.

In summary, with respect to the inequality $|\sigma_{\Delta\vartheta\Delta s}|/\sigma_{\Delta s}{}^2 < |\sigma_{\vartheta s}|/\sigma_s{}^2$ in the identification assumption, the denominator on the left-hand side is large, and the numerator on the left-hand side is small. Hence, the identification assumption, $|\sigma_{\Delta\vartheta\Delta s}|/\sigma_{\Delta s}{}^2 < |\sigma_{\vartheta s}|/\sigma_s{}^2$, is unlikely to be violated.

---

[10] It would be a necessary condition if $\sigma_{\Delta s}{}^2 = \sigma_s{}^2$. Furthermore, the larger the difference ($\sigma_{\Delta s}{}^2 - \sigma_s{}^2$), the lower the correlation coefficient (ρ) is needed for our identification. Based on adult twin data, $\sigma_{\Delta s}{}^2 = 0.72$ and $\sigma_s{}^2 = 0.49$ when we use the variable of having at least one son to measure children gender composition. In this case, the necessary and sufficient condition for our identification is $\rho > 0.23$.

We derive the following proposition based on this identification assumption:

Proposition. Under the identification assumption, the FE estimate can tighten either the lower bound or the upper bound of β in terms of absolute value, and identify the signs of both β and ξ. Specifically, when $\beta^{OLS} > \beta^{FE}$ and $\beta^{FE} < 0$, we conclude that (i) $\beta < 0$, (ii) $\xi > 0$, and (iii) β^{FE} is a less biased estimate of β than β^{OLS}, by providing a tighter lower bound of the negative effect of the gender composition of children on parental risk aversion.[11]

The proof is presented in Appendix II. We can identify the signs of both structural parameters of (i) the effect of the children's gender composition on parental risk attitude at the survey time (β), and (ii) the effect of the endowed risk attitude on parental demand for sons (ξ). We exercise caution in the indirect inference of the sign of ξ by comparing the OLS and FE estimates. In equation (1), we take the unobserved cross-individual heterogeneity to be the endowed risk attitude, which serves as a sufficient statistic.

3. Data

The data sets used in this study come from two sources. The first data set is based on the Chinese Adult Twin Survey (CATS) in 2002. We conducted an incentivized experiment on a subsample of the twins from the CATS in 2008. The second data set is based on the second wave of the Longitudinal Chinese Child Twin Survey (LCCTS) in 2013.

3.1 The Chinese Adult Twin Survey

The CATS was conducted by the Urban Survey Unit of the National Bureau of Statistics in June 2002 and July 2002 in the urban areas of five cities in China, namely, Chengdu, Chongqin, Harbin, Hefei, and Wuhan. The CATS is supported by the Research Grants Council of Hong Kong. The CATS covers a wide range of demographic, social, and economic information, and is based on existing twin questionnaires from the United States and other sources. The questionnaire was designed by one of the

[11] Other cases may be analyzed in a similar way. The case analyzed in the proposition corresponds to the empirical results presented in our paper.

authors of this paper and Chinese experts at the NBS. The local statistical bureaus identified adult twins between the ages of 18 and 65. The questionnaires were completed through face-to-face personal interviews. Both the process of conducting the survey and the inputting of data were closely monitored by one of the authors of this paper. The CATS is the first socioeconomic survey on twins in China and perhaps Asia. For a detailed description of the CATS, see Li et al. (2007), Huang et al. (2009), Li, Rosenzweig, and Zhang (2010), and Li, Liu, and Zhang (2012), who use the CATS in their empirical analyses.

We conducted an incentivized experiment in 2008 on a subsample of twins from the CATS to measure the subject's risk attitude. The experiment was financially supported by a grant from the Hong Kong University of Science and Technology. Due to budget constraints, we conducted the experiments only in Hefei and Wuhan, the capital cities of Anhui and Hubei Provinces, respectively.[12] Our sample was composed of 48 pairs (96 individuals) of monozygotic (identical) twins with children.

One approach to measuring risk attitudes is through a simple experimental elicitation known as the price-list design (Holt and Laury, 2002), in which subjects make an ordered array of binary choices. We conducted a pilot study using the price-list design, and we observed a large proportion of subjects with multiple-switching behavior, which suggests that non-student subjects may have difficulties in understanding the task. We learned from the pilot experiment, and subsequently used a simplified version of this procedure in the present study to assess the subjects' risk attitudes. More specifically, subjects chose between an even-chance lottery of receiving ¥40 and receiving zero, or receiving a guaranteed ¥20.[13] Subjects were incentivized for their choice in this comparison.[14] As their decisions reveal, a subject's valuation of the gamble is categorized as risk aversion if he chose certainty ($y = 1$ in equation (1)), and risk seeking if he chose the lottery ($y = 0$).[15] Chew et al. (2016) discuss the procedure and instructions of our experiment in detail.

[12] Seventy-seven percent of the twins in CATS are sampled from Wuhan and Hefei.
[13] The exchange rate in 2008 is USD1≈¥6.8.
[14] In these two cities, the minimum wage per hour for a full-time employee is ¥6.5, and the average wage per hour is ¥9.5 (NBS, 2009). The expected gain from this simple game is ¥20, which is much higher than their alternative hourly wage. This finding suggests that the participants have a non-negligible incentive to make decisions in the experiment.
[15] Should lotteries with smaller probabilities be used, the majority of subjects would be risk seeking. In this setting, we would need to go beyond the expected utility framework and consider the role of nonlinear probability weighting. We leave this issue for future studies.

Columns (1)-(2) in Table 1 report the means and standard deviations of the main variables based on the experimental sample. The socioeconomic variables are derived from the CATS. As the table shows, 43 percent of the subjects chose the certain amount of ¥20 over the gamble. The sex ratio of children is 52 percent, which is biased toward sons, although only 49 percent of first-born children are boys. Moreover, 60 percent of the families have at least one son. After combining the three measures of the children's gender, we find that gender-selection behavior tends to be focused on high birth parities. We use a relatively old adult-twins sample with a mean age of 52. The children were born before the one-child policy; thus, the issue of the use of fertility drugs on the occurrence of twins is not a concern. Most of the children of these adult twins were born before the one-child policy was implemented in 1979. Therefore, most of these adult twins have more than one child. Coale and Banister (1994) report on the five decades of missing women in China. They conclude that gender selection was prevalent in China even before 1979. Only one third of our subjects are males. Their educational attainments are categorized into three levels: middle school or below (omitted baseline group), high school, and technical school and college.[16] The mean annual family income in our sample is ¥19,800.

Because our experiment was conducted on a subsample of the CATS, columns (3)-(4) in Table 1 report the means and standard deviations of the main socioeconomic variables in the CATS. We include twin pairs with children as of 2002. Furthermore, for comparison, non-twin households in the five cities were taken from regular households with whom the UHS conducts regular monthly surveys. The summary statistics of the main variables for the UHS non-twin sample are reported in columns (5) and (6). Comparing the CATS sample with the UHS non-twin sample, the subjects in the experimental sample are older and more likely to be female. Due to budget constraints, we were only able to contact a small proportion of twins in the 2008. We invited these twins to participate in the experiment in hotels. The older and female twins were more likely to be contacted and to accept our invitations.

3.2 The Longitudinal Chinese Child Twin Survey

In additional to the CATS data, we use the LCCTS data. Compared to the CATS, the benefits of using the LCCTS are three-fold. First, the LCCTS sample size is larger.

[16] Because only four individuals finished technical school, we categorize technical school and college as one group.

Second, the LCCTS includes twins from both rural and urban areas, whereas twins in the CATS come only from urban areas. Third, all children of twins in the LCCTS were born after the one-child policy.

The LCCTS is the first census-type longitudinal household survey on twins conducted by the Urban Survey Unit of the National Bureau of Statistics. The first wave of the survey was conducted in late 2002 and early 2003 in Kunming, China.[17] The twins were then between the ages of 6 and 18. A follow-up survey on the households surveyed in the first wave was conducted in 2013; the twins were now between the ages of 17 and 29. Some of the twins in the LCCTS had married and had had their own children by 2013.

In the first wave in 2002, the households were initially identified by the USU based on the 2000 population census according to whether children had the same birth year and month, and whether children had the same relationship with the household head. The addresses of these households were then obtained from the census office, and a visit to the household verified the presence of twins. Therefore, the LCCTS is a census-type survey on twins. The survey covers an extensive range of information on demographic, social, and economic characteristics at both the household and individual levels. In our analysis, we restrict the LCCTS sample to both twin siblings who had married and had had their own children by 2013, for a total of 122 pairs of twins (224 individuals). For a detailed description, see Rosenzweig and Zhang (2009), Yi et al. (2015), and Yi (2016), who use the LCCTS in their empirical analyses.

We have designed a survey question to measure the subject's risk attitude in the second wave of the LCCTS: "Suppose that there is a business to run; which of the following two options do you prefer? (a) A sure profit of RMB 10,000; (b) a 50% chance with a profit of RMB 20,000 and a 50% chance that you will earn nothing." We categorize the subject as risk averse (in equation (1)) if he or she chose (a). Similar hypothetical questions are used in Kahneman and Tversky (1979), since "the method of hypothetical choices emerges as the simplest procedure by which a large number of theoretical questions can be investigated". Recently, Dohmen et al (2011) compare various measurements of risk preference and their economic consequences. In particular, they include a hypothetical lottery question on investment decision of an endowment of €100,000, and show that the risk attitude elicited from hypothetical

[17] Kunming is the capital city of Yunnan and has a total population of approximately five million. Yunnan is a less-developed province located in the far southwestern corner of China.

lottery predicts decisions in the financial domain. While we are aware that method of using hypothetical questions is not a perfect substitute for elicitation of risk preferences using incentivized experiments, we view it as a complement for our experiment with a real incentive in the CATS data.

Columns (7)-(8) in Table 1 report the means and standard deviations of the main variables in the LCCTS. As shown in the table, 53 percent of subjects chose option (a). The statistics suggest that the sex ratio of the twin subjects' children is biased toward boys. Specifically, the sex ratio of the twin subjects' children is 53 percent; 54 percent of their first-born children are boys; and 56 percent of the twin subjects have at least one son. The average age of twin subjects in the LCCTS is 25, as shown in column (7); all of their children were born after the one-child policy. Li, Yi, and Zhang (2011) discuss the cause of the gender imbalance in China since 1979. Column (7) also shows that only 41 percent of the twin subjects are male. The reason for the low percentage is that females tend to marry earlier in China. Furthermore, 5 percent of the twin subjects are high school graduates, and 9 percent are college graduates. The reason for the low education of the twins is two-fold. First, the LCCTS encompasses both rural and urban twins. Second, less-educated subjects are more likely to get married earlier, and consequently to have children at younger ages.

4. Estimation Results

This section reports and discusses our estimation results. We first present the basic results of the OLS and the FE estimates. Subsequently, we conduct two heterogeneous analyses by examining the effects of children's gender composition on mothers and fathers, respectively, and the insurance motive of the son preference. We finally discuss potential pitfalls and scientific validity.

4.1 Basic Results

The dependent variable is the measured parental risk aversion after having children. The independent variable is the gender composition of their children, for which we have three measures: (i) having at least one son, (ii) the first child being a son, and (iii) the ratio of sons to all children. These measures give rise to three regression specifications.

Using the CATS data, Table 2 reports the OLS and the FE estimates of the effects of children's gender composition on parental risk aversion. Given the small size of our

sample, we report bootstrapped standard errors throughout the paper. Our regression results remain robust when we use a permutation method to correct for potential bias with standard errors arising from the small sample. Columns (1)-(3) of Table 2 present the OLS estimates. The estimated coefficients of these three variables are consistently negative but statistically insignificant. Risk aversion initially decreases and then increases with age. Males are more risk seeking than females. However, the estimates are statistically insignificant in all three specifications.

Columns (4)-(6) of Table 2 present the FE estimates, and are parallel to columns (1)-(3). All three estimates of the effects of children's gender composition are statistically significant at the 10 percent level despite the relatively small sample size; these estimates are consistent across the three columns. We compare the three estimates and find that having at least one son has the largest negative effect on parental risk aversion. The differences between the FE and the OLS estimates are statistically significant. For example, given the null hypothesis that the FE estimate equals the OLS estimate with respect to having at least one son, the Wald cross-equation test statistic is significant at the 10 percent level.

Table 3 replicates Table 2 by using the LCCTS data. The results in Table 3 are qualitatively similar to those in Table 2. Columns (1)-(3) report the OLS estimates. The U-shaped relationship between age and risk aversion is more significant in Table 3 than in Table 2. The estimated coefficients of both age and age-squared variables are statistically significant, as shown in columns (1)-(3). Males are less risk averse than females, and the estimates are statistically significant at the 10 percent level in all three specifications. Finally, subjects with at least a college education are less risk averse than those with at most a middle school education. The estimates are statistically significant at the 5 percent level across all specifications in columns (1)-(3).

The FE estimates are reported in columns (4)-(6). Consistent with the results reported in Table 2, all three estimates of the effects of children's gender composition are statistically significant at least the 10 percent level. The differences between the FE and the OLS estimates are also statistically significant in Table 3. The LCCTS used the survey question to measure risk attitude, whereas we conducted an incentivized experiment to measure risk attitude with the CATS data. Thus, we carry out Durbin-Wu-Hausman tests between Tables 2 and 3 with respect to columns (4)-(6). We find that the FE estimates are not statistically significant. The results suggest that so-

liciting risk preference through survey questions is also valid in our study.

Following the proposition in Section 2, we can identify the signs of the structural parameters of both β and ξ by comparing the FE estimates (columns (4)-(6)) with the OLS estimates (columns (1)-(3)) in both tables. Both the FE and OLS estimates are negative, but the FE estimates are more negative. We have three main conclusions. First, β is negative; that is, having sons significantly decreases the degree of parental risk aversion. Second, ξ is positive; that is, the endowed risk aversion increases the demand for sons. These two findings are consistent with our theoretical predictions. Finally, since the FE estimates provide a lower bound of the negative effect of children's gender composition on parental risk aversion in terms of absolute value, they are thus less biased than the OLS estimates.

4.2 Heterogeneous Effects by Mothers and Fathers

This section examines the effects of children's gender composition on maternal and paternal risk attitudes. We expect the presence of sons to have a greater effect on risk seeking for mothers than fathers, because mothers are more likely to rely on sons' support and security in their old age. This reliance derives from mothers usually living longer than fathers, and from fathers generally being the main breadwinners in China. Men are more likely to be formally employed, and consequently to have pensions.

The results reported in Table 4 are consistent with our prediction. Because the CATS sample is too small to divide into subsamples, we use only the LCCTS data only. Columns (1)-(3) report the estimates for mothers, and columns (4)-(6) for fathers. Relative to columns (4)-(6), the magnitudes of the estimated coefficients on children's gender composition in columns (1)-(3) are bigger. Furthermore, the estimates in columns (1)-(3) are statistically significant at least the 5 percent level, whereas the estimates in columns (4)-(6) are statistically insignificant.[18] Therefore, the effect of children's gender composition on risk attitude is greater for mothers than fathers. The results support our hypothesis on the insurance motive for son preference.

4.3 Heterogeneous Effects by Public and Private Sectors

This section examines the insurance motive for son preference. The FE estimates in Tables 2 and 3 show that the presence of sons decreases the degree of parental risk

[18] We exercise caution in the comparison, because the sample size of the male-twin pairs is smaller than that of the female-twin pairs.

aversion. Furthermore, compared to the FE estimates, the OLS estimates are positively biased. We interpret these results as an insurance motive for son preference. On one hand, parents with sons, ceteris paribus, are more risk seeking than parents with daughters, because sons provide more security. Thus, the FE estimates are negative. On the other hand, parents who are more risk averse before the birth of their child, ceteris paribus, may prefer to have sons. Thus, children's gender composition and the endowed parental risk aversion are positively correlated in equation (1) above, leading to the OLS estimates being positively biased.

We design a test for the insurance motive for son preference.[19] Employees in the public sector in China have public insurance and high-quality old-age pension (Meng, 2012). By contrast, employees in the private sector do not have either public insurance or old-age pension. If the insurance motive is driving our results, the effect of children's gender composition on parental risk attitude should be insignificant for parents working in the public sector.

Columns (1)-(3) in Table 5 report the results using the CATS data. We conduct the FE estimation of equation (6) by adding two variables. One is a dummy variable equal to one if the subject is working in the public sector, and equal to zero otherwise. The other variable is an interaction term of the measure of children's gender composition and the dummy variable for public sector. The estimation result is consistent with our prediction. The effect on parental risk aversion for those working in the private sector is measured by the estimated coefficients of the three measures of children's gender composition. These three estimates are statistically significant. The effect for parents working in the public sector is measured by the sum of two estimated coefficients—the coefficient of the measure of children's gender composition, and the coefficient of the interaction term. The coefficients of the interaction terms are positive across the three specifications. Thus, the negative effect of children's gender composition on parental risk aversion is less severe for those working in the public sector. Furthermore, we conduct three F-tests, and find that the effect for parents working in the public sector is statistically insignificant, as shown in columns (1)-(3).

We further conduct a test to distinguish the insurance-incentive hypothesis from another competing hypothesis, namely, a motive to see their sons married. Specifically, the competitive marriage market for males might drive the positive effect of sons on

[19] We appreciate an anonymous referee's suggestion of testing the insurance motive for son preference.

parental risk seeking. As the sex ratio increases, the competition in the marriage market becomes fiercer for sons. Therefore, parents of sons become more aggressive in order to enhance their sons' chances in the marriage market.

To test the two hypotheses of insurance motive and the motive for sons' marriage, we added two variables to the FE estimation of equation (6). One is a dummy variable equal to one if all children are married, and equal to zero otherwise. We then interact this dummy variable with the measure of children's gender composition. Should our results be driven by the marriage motive, we predict a weaker negative effect of sons on parental risk aversion for parents whose children are married than for parents whose children are unmarried since parents have less worries about their children's marriages once their children are married. Therefore, the marriage-motive hypothesis predicts that the coefficient of the interaction term is positive. On the other hand, should our results be driven the insurance motive, married sons are perceived to provide more security for their parents; hence the coefficient of the interaction term is negative. Columns (4)-(6) of Table 5 show that the coefficients of the interaction terms are negative across all three specifications. These results are consistent with the insurance-motive hypothesis, and inconsistent with the marriage-motive hypothesis. However, we note that the estimates reported in Columns (4)-(6) in Table 5 are statistically insignificant.

Columns (7)-(9) in Table 5 replicate columns (1)-(3) using the LCCTS data. The results are consistent with the insurance-motive hypothesis. Specifically, the estimated coefficients of the interaction terms are positive and statistically significant, suggesting that the negative effect of children's gender composition on parental risk aversion is muted when parents work in the public sector.

We further analyze the effects of children's gender composition on parental risk aversion by rural and urban subsamples using the LCCTS data. Given that rural residents do not usually have public insurance or old-age pensions, we predict a stronger effect for the rural subsample. Our empirical results confirm this prediction. Because the rural indicator and the public sector indicator are highly correlated, we report only the results using the latter indicator in columns (7)-(9). Finally, because all children of twins are young and unmarried in the LCCTS, we are not able to replicate columns (4)-(6) using this data set.

4.4 Potential Pitfalls

The within-sibling FE estimator, especially the within-twin-pair estimator, has been used to identify causal relationships in labor economics for some time, following Behrman and Taubman's (1976) seminal study on the economic returns to education. Griliches (1979) was the first to provide a rigorous treatment of the sibling models and data in econometric research, but this method has received criticism in the literature. Behrman and Rosenzweig (1999) and Bound and Solon (1999) discuss twin-based estimators in a special issue of the *Economics of Education Review*.

The literature proposes two major concerns with the within-twin-pair FE estimator. One concern is that the assumption of the within-twin-pair differences in the independent variable being uncorrelated with any unobserved variable between the twin siblings is too strong. Our identification strategy relaxes this assumption. The other concern is the possible effect of measurement errors with respect to whether the FE estimator is more fragile than the OLS estimator (Griliches, 1979; Bound and Solon, 1999). Unlike in other studies, such as studies on estimating the returns to schooling, measurement errors may not be an issue in our analysis. First, the quality of our twin data is high, as discussed below.[20] Second, parents are less likely to misreport the number and gender of their children than to misreport schooling years. Finally, the magnitude of FE estimates is consistently larger than that of OLS estimates in terms of absolute values, as presented below. Therefore, the attenuation bias induced by measurement errors is likely not applicable to our study.

4.5 Scientific Validity

Conducting experiments on twins is scientifically valuable because twins share similar genetic endowments and family background, thereby providing a unique natural experimental setting. The within-twin comparison removes the unobserved confounders to a large degree (Behrman and Rosenzweig, 1999). However, implementing experiments on a large sample, especially with twins, is difficult. Sample size has been a concern, particularly for studies involving controlled experiments (see, e.g., List, Sadoff, and Wagner, 2011 and Maniadis, Tufano, and List, 2014). For example, List,

[20] See Li, Liu, and Zhang (2012) for a detailed description of the quality of CATS data. For example, the correlation between the self and the cross-twin sibling's reported schooling years is 0.94, which is higher than that in other major twin survey data in economic research, such as the Twinsburg Data and the Minnesota Twin Registry.

Sadoff, and Wagner (2011) suggest that the reliance on statistical significance as the sole criterion can lead to an excessive number of false positives. They further propose additional inferential information including the observed p-values, the power of the design, and the prior probability of the hypothesis. For instance, given the significance level of 0.05, the power of 0.80, and the prior change in the outcome variable of half a standard deviation, the sample size in each treatment cell should be at least 64.

Although our study relies on two relatively small samples (the CATS sample with 48 pairs of twins and the LCCTS sample with 122 pairs of twins), we believe that the inference from our data is reliable. First, the sample size of the experimental CATS data, although small, is marginally acceptable with a significance level of 0.1 because the size of our estimated effect is about half a standard deviation of the outcome variable, as reported below.[21] Second, the sample size of the LCCTS used in our analysis is sufficiently large based on the criteria in Maniadis, Tufano, and List (2014). Third, the twin design may be regarded as a within-subject design, because twin siblings share similar genetic endowments and family background. List, Sadoff, and Wagner (2011) show that fewer subjects are needed for a within-subject design if the individual fixed effects are uncorrelated with the treatment effect. Although a twin design is similar to a within-subject design, the twin design does not suffer from the disadvantages associated with the within-subject design, such as history and learning effects, and sensitization to perceived dependencies across trials.[22]

Replicability of scientific findings has been an increasing concern across multiple disciplines including the social sciences. The Open Science Collaboration (2015) conduct replications of 100 experimental studies published in three top-ranked psychology journals. While 97 percent of original studies had statistically significant results, only 36 percent of replications had statistically significant results. Camerer et al. (2016) replicated 18 studies published in in the American Economic Review and the Quarterly Journal of Economics in 2011-2014, and show a replication rate of about 67 percent to 78 percent. Relatedly, Maniadis, Tufano, and List (2014) demonstrate that the probability of false positives decreases with the number of replications (see also Levitt and List, 2009). List, Shaikh, and Xu (2016) discuss multiple hypothesis testing

[21] Using equation (5) in List, Sadoff, and Wagner (2011), the minimum number of observations in each treatment cell is about 49 if (i) the significance level is 0.10, (ii) the power is 0.80, and (iii) the prior change in the outcome variable is half a standard deviation.

[22] The concern with the twin design is that twin siblings are not identical, which we discuss above.

and offer a powerful correction procedure to address three common concerns in field experimental economics. Multiple hypothesis testing is less of a concern in our study, as we test one single hypothesis that links risk aversion to son preference. Though both the LCCTS and CATS were designed by the same principal investigators, the two surveys were conducted independently in different periods of time and for different research purposes. Thus, the results from the LCCTS can be regarded as a robust confirmation of those from the CATS. In sum, the within-twin-pair FE estimates from two independent data sets enhance the scientific reliability of our empirical results.

5. Conclusion

This paper studies the relationship between parental risk aversion and son preference. On the one hand, parents with sons may be less risk averse than parents with daughters because sons provide support and security for their elderly parents in traditional patriarchal societies. On the other hand, families with greater endowed risk aversion may have a stronger demand for sons than families with less endowed risk aversion. Therefore, the OLS estimates are generally biased because of the unobserved heterogeneity in endowed risk aversion. We elicit the risk attitudes of Chinese twin parents together with a within-twin-pair FE estimator to address the concern relating to the omitted-variable bias. Assuming endowed risk aversion is similar across twin siblings, the FE estimates will be less biased than the OLS estimates in rendering a causal interpretation of our findings. In this sense, our FE estimates indicate that having sons significantly decreases parental risk aversion. Moreover, we compare the OLS and the FE estimates, and show that families with greater endowed risk aversion indeed have a stronger demand for sons.

The identified positive effect of endowed risk aversion on the demand for sons offers an explanation for the origin of son preferences and the increased sex ratios in East Asian countries such as China, India, and Vietnam, in which parents in traditional patriarchal societies demand more for sons than for daughters to protect them against socioeconomic risks. On the other hand, Chiappori, Fortin, and Lacroix (2002) argue that the increased sex ratio may increase the bargaining power of females. This is corroborated by the observation in Bulte, Tu and List (2015) of a positive correlation between sex ratios and female bargaining power in China pointing to a potential lessening of son preference in the long run which may arise from an increased tendency

of married daughters supporting their own elderly parents. More broadly speaking, there is scope for future work on related issues with larger sample sizes, more complete social, economic, and demographic data, and more comprehensive measures of preferential traits elicited in incentivized choice experiments.

References

Abdellaoui, M., Baillon, A., Placido, L., and Wakker, P. (2011). The Rich Domain of Uncertainty: Source Functions and Their Experimental Implementation. *American Economic Review*, 101(2), 695-723.

Ahn, N. (1995). Measuring the Value of Children by Sex and Age Using a Dynamic Programming Model. *Review of Economic Studies*, 62(3), 361-379.

Alesina, A. F., Giuliano, P., and Nunn, N. (2013). On the origins of gender roles: Women and the plough. *Quarterly Journal of Economics*, 128(2), 469-530.

Almond, D. and Edlund, L. (2008). Son-biased sex ratios in the 2000 United States Census. *Proceedings of the National Academy of Sciences*, 105(15), 5681-5682.

Almond, D., Edlund, L., and Milligan, K. (2013). Son Preference and the Persistence of Culture: Evidence from Asian Immigrants to Canada. *Population and Development Review*, 39(1), 75-95.

Angrist, J. (2002). How do sex ratios affect marriage and labor markets? Evidence from America's second generation. *Quarterly Journal of Economics*, 117(3), 997-1038.

Behrman, J. and Taubman, P. (1976). Intergenerational Transmission of Income and Wealth. *American Economic Review*, 66(3), 436-440.

Behrman, J. and Rosenzweig, M. (1999). Ability Biases in Schooling Returns and Twins: A Test and New Estimates. *Economics of Education Review*, 18(2), 159-167.

Ben-Porath, Y. and Welch, F. (1976). Do sex preferences really matter? *Quarterly Journal of Economics*, 90(2), 285-307.

Bethmann, D. and Kvasnicka, M. (2012). World War II, Missing Men and Out of Wedlock Childbearing. *Economic Journal*, 123(567), 162-194.

Camerer, C. F., Dreber, A., Forsell, E., Ho, T. H., Huber, J., Johannesson, M., and Heikensten, E. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, 351(6280), 1433-1436.

Chew, S., Yi, J., Zhang, J., and Zhong, S. (2016). Education and Preferences: Experimental Evidence from Twins. *Working paper*.

Chew, S. H. and Sagi, J. S. (2008). Small Worlds: Modeling Attitudes toward Sources of Uncertainty. *Journal of Economic Theory*, 139(1), 1-24.

Chiappori, P. A., Fortin, B., and Lacroix, G. (2002). Marriage market, divorce legislation, and household labor supply. *Journal of Political Economy*, 110(1), 37-72.

Coale, A. J., and Banister, J. (1994). Five decades of missing females in China. *Demography*, 31(3), 459-479.

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

Das Gupta, M., Ebenstein, A., and Sharygin, J. (2013). China's Marriage Market and Upcoming Challenges for Elderly Men. *Population Studies*, 67(1), 39-59.

Davies, J. and Zhang, J. (1995). Gender Bias, Investments in Children and Bequests. *International Economic Review*, 36(3), 795-818.

Dahl, G. B., and Moretti, E. (2008). The demand for sons. *Review of Economic Studies*, 75(4), 1085-1120.

Dimmock, S. G., Kouwenberg, R., and Wakker, P. P. (2015). Ambiguity attitudes in a large representative sample. *Management Science*, 62(5), 1363-1380.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3), 522-550.

Ebenstein, A. (2011). Estimating a Dynamic Model of Sex Selection in China. *Demography*, 48(2), 783-811.

Ebenstein, A. (2010). The 'Missing Girls' of China and the Unintended Consequences of the One-Child Policy. *Journal of Human Resources*, 45(1), 87-115.

Ebenstein, A. and Leung, S. (2010). Son Preference and the Missing Market for Social Insurance: Evidence from China's Rural Pension Program. *Population and Development Review*, 36(1), 47-70.

Edlund, L., Li, H., Yi, J., and Zhang, J. (2013). Sex ratios and crime: Evidence from China. *Review of Economics and Statistics*, 95(5), 1520-1534.

Edlund, L. (1999). Son preference, sex ratios, and marriage patterns. *Journal of Political Economy*, 107(6), 1275-1304.

Gangadharan, L., and Maitra, P. (2003). Testing for son preference in South Africa. *Journal of African Economies*, 12(3), 371-416.

Griliches, Z. (1979). Sibling Models and Data in Economics: Beginnings of Survey. *Journal of Political Economy*, 87(5), 37-64.

Holt, A. and Laury, K. (2002). Risk Aversion and Incentive Effects. *American Economic Review*, 92(5), 1644-1655.

Hu, L. and Schlosser, A. (2015). Prenatal Sex Selection and Girls' Well-being: Evidence from India. *Economic Journal*, 125(587), 1227-1261.

Huang, C., Li, H., Liu, P., and Zhang, J. (2009). Why Does Spousal Education Matter for incomes? Assortative Mating and Cross-Productivity. *Journal of Labor Economics*, 27(4), 633-52.

Lawrence, M. S., Stojanov, P., Polak, P., Kryukov, G. V., Cibulskis, K., Sivachenko, A., and Kiezun, A. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*, 499(7457), 214-218.

Leung, S. F. (1991). A Stochastic Dynamic Analysis of Parental Sex Preferences and Fertility. *Quarterly Journal of Economics*, 106(4), 1063-1088.

Leung, S. F. (1988). On tests for sex preferences. *Journal of Population Economics*, 1(2), 95-114.

Levitt, S. D., and List, J. A. (2009). Field experiments in economics: The past, the present, and the future. *European Economic Review*, 53(1), 1-18.

Li, H., Liu, P., and Zhang, J. (2012). Estimating Returns to Education Using Twins in Urban China. *Journal of Development Economics*, 97(2), 494-504.

Li, H., Liu, P., Zhang, J., and Ma, N. (2007). Economic Returns to Communist Party Membership: Evidence from Urban Chinese Twins. *Economic Journal*, 117(523), 1504-1520.

Li, H., Rosenzweig, M., and Zhang, J. (2010). Altruism, Favoritism, and Guilt in the Allocation of Family Resources: Sophie's Choice in Mao's Mass Send−Down Movement. *Journal of Political Economy*, 118(1), 1-38.

Li, H., Yi, J., and Zhang, J. (2011). Estimating the Effect of the One-Child Policy on Sex Ratio Imbalance in China: Identification Based on the Difference-in-Differences. *Demography*, 48(4), 1535-1557.

Lin, M. J., and Luoh, M. C. (2008). Can hepatitis B mothers account for the number of missing women? Evidence from three million newborns in Taiwan. *American Economic Review*, 98(5), 2259-2273.

List, J. A., Sadoff, S., and Wagner, M. (2011). So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design. *Experimental Economics*, 14(4), 439-457.

Maniadis, Z., Tufano, F., and List, J. A. (2014). One swallow doesn't make a summer: New evidence on anchoring effects. *American Economic Review*, 104(1), 277-290.

Meng, X. (2012). Labor market outcomes and reforms in China. *Journal of Economic Perspectives*, 26(4), 75-101.

National Bureau of Statistics (NBS). (2009). *China Statistical Yearbook 2009* [2009 nian zhong guo tong ji nian jian]. Beijing: China Statistics Press.

Noussair, C. N., Trautmann, S. T., and Van de Kuilen, G. (2014). Higher order risk attitudes, demographics, and financial decisions. *Review of Economic Studies*, 81(1), 325-355.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716.

Oster, E. (2005). Hepatitis B and the Case of the Missing Women. *Journal of Political Economy*, 113(6), 1163-1216.

Oster, E., Chen, G., Yu, X., and Lin, W. (2010). Hepatitis B does not explain male-biased sex ratios in China. *Economics Letters*, 107(2), 142-144.

Qian, N. (2008). Missing Women and the Price of Tea in China. *Quarterly Journal of Economics*, 123(3), 1251-1285.

Rao, V. (1993). The Rising Price of Husbands: A Hedonic Analysis of Dowry Increases in Rural India. *Journal of Political Economics*, 101(4), 666-677.

Rosenzweig, M. (1993). Women, Insurance Capital, and Economic Development in Rural India. *Journal of Human Resources*, 28(4), 735-58.

Rosenzweig, M. (1988a). Risk, Private Information, and the Family. *American Economic Review*, 78(2), 245-50.

Rosenzweig, M. (1988b). Risk, Implicit Contracts and the Family in Rural Areas of Low-Income Countries. *Economic Journal*, 98(393), 1148-70.

Rosenzweig, M. and Schultz, T. (1982). Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India. *American Economic Review*, 72(4), 803-15.

Rosenzweig, M. R., and Zhang, J. (2009). Do population control policies induce more human capital investment? Twins, birth weight and China's "one-child" policy. *Review of Economic Studies*, 76(3), 1149-1174.

Sen, A. (2003). Missing women—revisited: Reduction in female mortality has been counterbalanced by sex selective abortions. *British Medical Journal*, 327(7427), 1297-1298.

Scharping, T. (2002). *Birth Control in China 1949-2000: Population Policy and Demographic Development*. Publisher: Routledge.

Sutter, M., Kocher, M. G., Rützler, M., and Trautmann, S. (2013). Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. *American Economic Review*, 103(1), 510-531.

Wei, S. J. and Zhang, X. (2011). The Competitive Saving Motive: Evidence from Rising Sex Ratios and Saving Rates in China. *Journal of Political Economics*, 119(3), 511-564.

Yi, J. (2016). Endogenous altruism: theory and evidence from Chinese Twins. *Working paper.*

Yi, J., Heckman, J. J., Zhang, J., and Conti, G. (2015). Early Health Shocks, Intra−household Resource Allocation and Child Outcomes. *Economic Journal*, 125(588), 347-371.

Zhong, S., Chew, S. H., Set, E., Zhang, J., Xue, H., Sham, P. C., Ebstein, R. P., and Israel, S. (2009). The Heritability of Attitude toward Economic Risk. *Twin Research and Human Genetics*, 12(1), 103-107.

Table 1: Summary Statistics of Main Variables Using the Chinese Adult Twin Survey (CATS) and Longitudinal Chinese Child Twin Survey (LCCTS)

	Experimental Sample		CATS		UHS		LCCTS	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Risk-attitude measure								
Risk aversion	0.43	0.33					0.46	0.35
Children's gender composition								
Have at least one son	0.60	0.49	0.58	0.49	0.55	0.50	0.56	0.50
The first child is a son	0.49	0.50	0.52	0.50	0.51	0.50	0.54	0.50
Sex ratio (sons / all children)	0.51	0.46	0.53	0.46	0.51	0.48	0.53	0.49
Covariates								
Age	51.92	9.07	47.52	3.95	44.39	7.81	24.87	2.91
Male	0.33	0.47	0.47	0.50	0.42	0.49	0.41	0.49
High school	0.43	0.50	0.33	0.46	0.31	0.46	0.05	0.23
College and above	0.27	0.45	0.21	0.40	0.38	0.48	0.09	0.28
Family annual income (RMB 1,000)	19.80	17.52	18.00	15.39	17.64	11.97	35.20	48.65
Public sector	0.42	0.50	0.36	0.47	0.46	0.50	0.11	0.31
All children married	0.16	0.36	0.09	0.29	0.07	0.25		
Observations	96		766		1530		244	

Data source: Columns (1)-(2): the experimental sample is based on the CATS; Columns (3)-(4): both twin siblings have children in the CATS; Columns (5)-(6): adults have children in the UHS in the same cities as in the CATS; Columns (7)-(8): both twin siblings have children in the second wave of the LCCTS.

Table 2: OLS and Within-Twin-Pair FE Estimates of the Effects of the Gender Composition of Children on Parental Risk Attitude, Using the Chinese Adult Twin Survey (CATS)

	Dependent variable: Parents are risk averse					
	OLS			FE		
	(1)	(2)	(3)	(4)	(5)	(6)
Have at least one son	-0.087			-0.203*		
	[0.104]			[0.118]		
The first child is a son		-0.084			-0.147*	
		[0.101]			[0.080]	
Sex ratio			-0.046			-0.166*
(sons / all children)			[0.127]			[0.091]
Age	-0.100	-0.097	-0.093			
	[0.063]	[0.068]	[0.070]			
Age-squared (1/100)	0.097*	0.093	0.089			
	[0.058]	[0.064]	[0.063]			
Male	-0.077	-0.080	-0.076			
	[0.126]	[0.102]	[0.113]			
High school	-0.016	-0.006	-0.017	-0.132	-0.147	-0.145
	[0.154]	[0.126]	[0.129]	[0.253]	[0.237]	[0.203]
College and above	0.019	0.025	0.020	-0.192	-0.218	-0.212
	[0.169]	[0.147]	[0.156]	[0.286]	[0.267]	[0.235]
ln (Family annual	-0.044	-0.043	-0.042	-0.116	-0.13	-0.117
income)	[0.075]	[0.078]	[0.085]	[0.074]	[0.101]	[0.124]
Observations	96	96	96	96	96	96
Twin pairs				48	48	48
R-squared	0.110	0.110	0.105	0.137	0.112	0.114

Note: Bootstrapped standard errors are in brackets; * significant at the 10 percent level; ** significant at the 5 percent level; *** significant at the 1 percent level; a city dummy is included in each OLS estimation. Data source: the experimental sample based on the CATS.

Table 3: OLS and Within-Twin-Pair FE Estimates of the Effects of the Gender Composition of Children on Parental Risk Attitude, Using the Longitudinal Chinese Child Twin Survey (LCCTS)

	Dependent variable: Parents are risk averse					
	OLS			FE		
	(1)	(2)	(3)	(4)	(5)	(6)
Have at least one son	-0.019			-0.177*		
	[0.065]			(0.0911)		
The first child is a son		-0.056			-0.235***	
		[0.065]			(0.0881)	
Sex ratio (sons / all children)			-0.041			-0.214**
			[0.066]			(0.0894)
Age	0.207***	0.202***	0.205***			
	[0.076]	[0.075]	[0.075]			
Age-squared (1/100)	-0.376***	-0.367***	-0.373***			
	[0.139]	[0.136]	[0.137]			
Male	-0.120*	-0.119*	-0.120*	-0.154	-0.150	-0.156
	[0.065]	[0.065]	[0.065]	(0.101)	(0.0995)	(0.100)
High school	0.153	0.160	0.158	-0.099	-0.107	-0.091
	[0.144]	[0.143]	[0.143]	(0.268)	(0.264)	(0.266)
College and above	-0.278**	-0.281**	-0.279**	-0.493***	-0.482***	-0.490***
	[0.108]	[0.109]	[0.109]	(0.164)	(0.161)	(0.162)
ln (Family annual income)	-0.001	-0.000	-0.001	-0.003	-0.004	-0.005
	[0.018]	[0.018]	[0.018]	(0.028)	(0.027)	(0.027)
Observations	244	244	244	244	244	244
Twin pairs				122	122	122
R-squared	0.058	0.061	0.060	0.114	0.138	0.129

Note: Robust standard errors are in brackets; * significant at the 10 percent level; ** significant at the 5 percent level; *** significant at the 1 percent level. Data source: both twin siblings have children in the second wave of the LCCTS.

Table 4: Within-Twin-Pair FE Estimates of the Effects of the Gender Composition of Children on Parental Risk Attitude, Using the Longitudinal Chinese Child Twin Survey (LCCTS) by Mothers vs. Fathers

	Dependent variable: Parents are risk averse					
	Mother			Father		
	(1)	(2)	(3)	(4)	(5)	(6)
Have at least one son	-0.337**			-0.191		
	[0.133]			(0.234)		
The first child is a son		-0.364***			-0.279	
		(0.133)			(0.206)	
Sex ratio (sons / all children)			-0.354**			-0.250
			(0.133)			(0.224)
High school	-0.293	-0.291	-0.292	-0.198	-0.198	-0.215
	[0.373]	(0.370)	(0.371)	(0.686)	(0.650)	(0.667)
College and above	-0.664***	-0.633***	-0.650***	-0.696*	-0.600*	-0.654*
	[0.218]	(0.214)	(0.216)	(0.529)	(0.519)	(0.523)
ln (Family annual income)	0.038	0.0394	0.0388	0.0537	0.0559	0.0582
	[0.043]	(0.0427)	(0.0429)	(0.0878)	(0.0798)	(0.0836)
Observations	96	96	96	60	60	60
Twin pairs	48	48	48	30	30	30
R-squared	0.245	0.260	0.254	0.195	0.236	0.216

Note: Robust standard errors are in brackets; * significant at the 10 percent level; ** significant at the 5 percent level; *** significant at the 1 percent level. Data source: both twin siblings have children in the second wave of the LCCTS.

Table 5: Within-Twin-Pair FE Estimates of the Effects of the Gender Composition of Children on Parental Risk Attitude, Using the Chinese Adult Twin Survey (CATS) and the Longitudinal Chinese Child Twin Survey (LCCTS)

	Dependent variable: Parents are risk averse								
	Sample: CATS						Sample: LCCTS		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Have at least one son	-0.341***			-0.186*			-0.226**		
	[0.123]			[0.111]			[0.099]		
The first child is a son		-0.323***			-0.100			-0.267***	
		[0.099]			[0.116]			(0.0944)	
Sex ratio (sons / all children)			-0.341***			-0.123			-0.250**
			[0.110]			[0.128]			(0.0956)
Have at least one son*public sector	0.399						0.270		
	[0.309]						[0.284]		
The first child is a son*public sector		0.491*						0.123	
		[0.283]						(0.282)	
Sex ratio (sons / all children) *public sector			0.403						0.194
			[0.310]						(0.286)
Have at least one son*children married				-0.188					
				[0.338]					
The first child is a son* children married					-0.233				
					[0.381]				
Sex ratio (sons / all children) * children married						-0.307			
						[0.474]			
Observations	96	96	96	96	96	96	244	244	244
Twin pairs	48	48	48	48	48	48	122	122	122
R-squared	0.198	0.209	0.199	0.194	0.161	0.174	0.117	0.139	0.128

Note: Bootstrapped standard errors are in brackets in columns (1)-(6), and robust standard errors are in brackets in columns (7)-(9); * significant at the 10 percent level; ** significant at the 5 percent level; *** significant at the 1 percent level. Data source: Columns (1)-(6): the experimental sample based on the CATS; Columns (7)-(9): both twin siblings have children in the second wave of the LCCTS.

Appendix I: Theoretical Framework on Son Preference and Risk Attitude

Consider a setting in which the agent has initial wealth *w* and faces an investment decision between a sure asset and a risky asset. For convenience, we assume that the sure asset has a zero rate of return, and the risky asset has a random return of *x* distributed through a cumulative distribution function *F(x)*. Let α be the dollar amount invested in the risky asset, and let $w-\alpha$ be the amount invested in the sure asset, given budget constraint *w*. The expected utility of this situation is given by the following:

$$EU = \int u(w+\alpha x)dF(x) .$$

Suppose *u* is increasing and twice differentiable. The decision maker engages in a positive level of investment, namely, $\alpha > 0$, if and only if the risky asset has a positive expected return *x*.

Sons provide major support and security to their ageing parents in traditional patriarchal societies. Parents with sons enjoy a stronger sense of security and are thus more risk tolerant than parents without sons. We apply a source-dependent expected utility model to capture this intuition (see Chew and Sagi, 2008; Abdellaoui et al., 2011). The model enables the agent to have distinct risk attitudes with different sources of risks. Specifically, the individual may have distinct utility functions, i.e., u_S with sons and u_D with daughters. If the Arrow–Pratt measure for u_D is uniformly lower than that for u_S, the agent is more risk averse with daughter-related risks than with son-related risks.

We consider how individual differences in the endowed risk attitude may affect son preference or the demand for sons. Each individual may have distinct utility functions for sons and for daughters. Therefore, we cannot compare the risk aversion between two individuals in a straightforward manner. For simplicity, we assume that the attitude toward risk with sons is identical for each individual, whereas the attitude toward risk with daughters is diverse. Consequently, individuals who are more averse to risk with daughters tend to favor risk with sons more than those who are less averse to risk with daughters. Put differently, when the agent is differentially more averse to risk with daughters, he/she exhibits a stronger son preference and is more likely to have sons, thus securing his/her future by practicing gender selection.

Appendix II: Proof of Proposition

Both the OLS and the FE estimates are biased when ϑ_{1i} is not equal to ϑ_{2i}. β^{OLS} and β^{FE} should be biased in the same direction, because $\sigma_{\Delta\vartheta\Delta s} / \sigma^2{}_{\Delta s}$ and $\sigma_{\vartheta s} / \sigma^2{}_s$ are of the same sign. If $\beta^{OLS} > \beta^{FE}$, the following two cases must be considered: $\beta^{OLS} > \beta^{FE} > \beta$ or $\beta > \beta^{OLS} > \beta^{FE}$.

Case 1: $\beta^{OLS} > \beta^{FE} > \beta$. In this case, we can derive that $|\sigma_{\Delta\vartheta\Delta s} / \sigma_{\Delta s}{}^2| < |\sigma_{\vartheta s} / \sigma_s{}^2|$ because of equations (5) and (7) and $\gamma > 0$. This derivation is consistent with the identification assumption.

Case 2: $\beta > \beta^{OLS} > \beta^{FE}$. In this case, putting equations (5) and (7) together, when $\gamma > 0$, both $\sigma_{\Delta\vartheta\Delta s} / \sigma_{\Delta s}{}^2$ and $\sigma_{\vartheta s} / \sigma_s{}^2$ are negative. Moreover, $\sigma_{\Delta\vartheta\Delta s} / \sigma^2{}_{\Delta s}$ is more negative than $\sigma_{\vartheta s} / \sigma^2{}_s$. Therefore, $|\sigma_{\Delta\vartheta\Delta s} / \sigma_{\Delta s}{}^2| > |\sigma_{\vartheta s} / \sigma_s{}^2|$. However, this finding contradicts the identification assumption. Therefore, Case 2 is impossible under the identification assumption.

In sum, only Case 1 is plausible under the identification assumption, $\beta^{OLS} > \beta^{FE}$ and $\beta^{FE} < 0$. Note that if $\vartheta_{1i} = \vartheta_{2i}$, β^{FE} will be unbiased, and hence $\beta^{OLS} > \beta^{FE}$. The proposition in section 2 directly follows when $\beta^{OLS} > \beta^{FE} > \beta$.