Two Essays on Family Behavior and Human Capital

YI, Junjian

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

in

Economics

The Chinese University of Hong Kong

June 2011

Abstract

This thesis consists of two essays on family behavior and human capital.

The first essay studies how early health shocks affect the child's human capital formation. We first formulate a theoretical model to understand how early health shocks affect child outcomes through parental responses. We nest a dynamic model of human capability formation into a standard intrahousehold resource allocation framework. By introducing the multidimensionality of child endowments, we allow parents to compensate and reinforce along different dimensions. We then test our main empirical predictions using a Chinese child twins survey, which contains detailed information on child- and parent-specific expenditures. We can differentiate between investments in money and investments in time. On the one hand, we find evidence of compensating investment in child health but of reinforcing investment in education. On the other hand, we find no change in the time spent with the child. We confirm that an early health insult negatively affects the child under several different domains, ranging from later health, to cognition, and then to personality. Our findings suggest caution in interpreting reduced-form estimates of the effects of early-life shocks. In the presence of asymmetric parental responses under different dimensions of the child's human capital, they cannot even be unambiguously interpreted as upper or lower bounds of the biological effects.

The second essay empirically estimates the effects of education on two dimensions of preference – decision making under risk and uncertainty and decision making involving time. We conduct a number of incentivized choice experiments on Chinese adult twins to measure preference, and use a within-twin-pair fixed-effects estimator to sweep out unobservable family background effects. The estimation results show that a higher level of education tends to reduce the degree of risk aversion toward moderate prospects, moderate hazards, and longshot prospects. In terms of decision making anomalies under risk and uncertainty, university educated subjects exhibit significantly more Allais-type behavior compared to pre-high school subjects, while high school educated subjects also exhibit more ambiguity aversion as well as familiarity bias relative to pre-high school subjects. For decision making involving time, a higher level of education tends to reduce the degree of impatience, hyperbolic discounting, dread, and hopefulness. The experimental evidences suggest that people with a higher level of education tends to exhibit more "biased" preference in risk attitude and less "biased" preference regarding time.

摘要

本論文集收錄了兩篇有關家庭行為和人力資本的研究論文。

第一篇文章研究早期健康衝擊對兒童人力資本形成的影響。本文首先從理論上模型了早期健康衝擊通過影響家庭資源配置從而影響兒童人力資本形成的作用機制。本文將一個動態的人力資本形成模型納入到一個標準的家庭資源配置模型。通過引入人力資本的多維性，我們推導出父母在家庭資源配置過程中可能在人力資本的不同維度上補償和強化早期健康衝擊的影響。然後，本文在中國兒童雙胞胎調查資料的基礎上，運用科學的計量方法來檢驗我們的理論預測。該調查資料不但包括了詳盡的兒童及其父母的支出資訊，並且能夠區分父母對兒童的物質資本投資和時間投資。本文的經驗分析結果表明：作為對兒童早期健康衝擊的回應，一方面，父母進行了補償性的健康投資的同時，進行了強化性的教育投資；另一方面，在時間投資方面，父母的投資行為並未表現出明顯的差異。另外，本文的研究結果證實了早期健康衝擊將對兒童的健康、認知能力與非認知能力等多個維度產生一系列的負面影響。本文的研究結果表明：應該謹慎的解釋早期健康衝擊的簡化估計值。如果父母在家庭資源的配置過程中對兒童早期健康衝擊做出了調整，該簡化估計值不但反應了在早期健康衝擊醫學的效應，而且反應了家庭資源配置的效應。更重要的是，在家庭資源配置的調整過程中，如果父母在人力資本的不同維度上採取了不同的投資策略，那麼簡化估計值甚至不能給出早期健康衝擊效應的上限或者下限。

第二篇文章研究教育對偏好的影響。本文集中在兩個維度的偏好：即風險與不確定偏好與時間偏好。基於一個中國成年人雙胞胎樣本，本文展開了一系列存在現實獎懲的實驗來度量偏好，並且運用雙胞胎固定效應模型來處理不可觀測的家庭之間的異質性問題。本文的估計結果表明：（1）教育降低了風險規避的程度；（2）然而，與初中及以下教育水準的接受實驗者相比，高中教育水準的接受實驗者更多的表現出“阿萊斯”行為、模糊規避與熟悉度偏誤；（3）在時

間偏好方面，教育提高了耐心，同時降低了雙曲貼現偏好、害怕偏好與希望偏好。本文的實驗證據表明高教育程度的接受實驗者在風險與不確定性條件下的行為更加“偏離”經典的經濟學理論，而其時間偏好更加“符合”經典的經濟學理論。

Acknowledgements

I would like to express my deepest gratitude to my supervisor Prof. Junsen Zhang for his patient supervision and devotion of his precious time. He has provided detailed, thoughtful, and constructive comments on almost every aspect of the thesis. This thesis would not have been up to this standard without his continuous guidance.

I would also like to submit my thanks to all those who ever made useful comments on the thesis, including Professors Gary Becker, Jere Behrman, Soo Hong Chew, Simon Fan, James Heckman, Hongbin Li, Pak-Wai Liu, Shelly Lundberg, Robert Pollak, Mark Rosenzweig, Wing Suen, Xiangdong Wei, Yoram Weiss, and Hongliang Zhang.

Last but not least, my thanks would go to my beloved family for their loving considerations and great confidence in me. I also owe my sincere gratitude to my friends and my fellow classmates who help me work out through this struggle, especially Gabriella Conti, Jun Han, Pinto Rodrigo, Tsz Nga Wong, Maoliang Ye, and Songfa Zhong.

To my grandfather

Contents

Essay One

Early Health Shocks, Parental Responses, and Child Outcomes[1]

[1]This essay is largely based on an on-going joint research project with Gabriella Conti, James Heckman, and Junsen Zhang. I have been the main contributor to the work so far.

1 Introduction

The literature on the effects of early-life conditions on late-life circumstances is burgeoning (Case, Fertig, and Paxson, 2005; Grantham-McGregor, Cheung, Cueto, Glewwe, Richter, Strupp, and The International Child Development Steering Group, 2007). This literature has achieved a consensus on the negative effects of an early-life health insult on both short-run (Currie, Stabile, Manivong, and Roos, 2010) and long-run outcomes (Smith, 2009). However, the role played by parental behavior is still not well understood, but its importance is being increasingly recognized (Case and Paxson, 2002; Almond and Currie, 2011). The central message of this essay is that, in general, in the presence of parental investments, the reduced-form estimates of the effects of early-life shocks do not necessarily represent a biological effect. Moreover, in case parents make compensating and reinforcing investments along different dimensions of human capital, they cannot be even unambiguously interpreted as upper or lower bounds of the biological effects.

These considerations may play a crucial role in developing countries, where national health insurance, public education, and old-age pension systems are inadequate or absent (Glewwe and Miguel, 2007). First, in the absence of public health insurance and with a tight budget, a child affected by a health insult may not receive appropriate medical treatment, and thus the early shock may have long-lasting consequences. In addition to this, in the absence of a well-functioning public education system, the consequences of an early health shock may be exacerbated, and also impair human capital formation. Finally, the absence of an old-age pension system may drive parents to base their intrahousehold resource allocation decisions on efficiency rather than on equity concerns. In this case, parents are more likely to reinforce the harmful effects of an early health insult, by devoting fewer resources to the less well-endowed child (Behrman, Rosenzweig, and Taubman, 1994). Hence, unpacking parental intrahousehold resource allocation responses is crucial to understand how early health shocks affect human capital formation, especially in developing countries. The role of the family must be taken into account when designing public policies to remediate the effects of inequality at birth or in early childhood.

Understanding how parents allocate resources across children has been researched in economics since the seminal work of Becker and Tomes (1976) and Behrman, Pollak, and Taubman (1982). However, since neither the wealth model nor the separable earnings-transfer model make unequivocal predictions regarding parental investments, whether parents exhibit a reinforcing, compensating, or neutral behavior has ultimately been an empirical question. Indeed, several papers have been devoted to testing parental strategies. The literature, nonetheless, has yet to achieve a consensus: whereas some studies have found evidence of reinforcing behavior (see, e.g., Behrman, Rosenzweig, and Taubman (1994) and Rosenzweig and Zhang (2009)), others have found empirical support for a compensating strategy (see, e.g., Behrman, Pollak, and Taubman (1982) and Pitt, Rosenzweig, and Hassan (1990)).[2] One common point to be noted is that these papers usually assume the existence of only one dimension under which parents can compensate or reinforce. Moreover, they frequently use measures of children's outcomes, such as educational attainment and test scores, to infer parental investments. We overcome both limitations in our work.

In this essay, we combine two strands of literature: the recent literature on the long-lasting effects of early-life conditions, and the more consolidated literature on intrahousehold allocation of resources. We combine them using a dynamic model of human capability formation (Heckman, 2007), which links early endowments to later outcomes through both self- and cross-productivity effects and parental investment behavior. By merging the two strands of literature we are able to model the mechanisms - parental reinforcing or compensating responses - through which an early-life

[2] Behrman (1988) finds evidence in support of both hypotheses for rural India depending on food availability:, during the lean season when food is scarce, parental allocations are significantly pro-son and quite focused on efficiency, whereas there is no gender differential during the surplus season, and parental behavior is compensating.

health shock affects later-life outcomes along different dimensions.

The key insight of our model is based on the following result: in the presence of multidimensional child endowments whose evolution is governed by a dynamic production technology, an early health shock works through a third effect in addition to the classical wealth and price effects a' la Becker and Tomes (1976) on parental investment[3] – a reallocation of resources by the parents across health and cognitive skills. Since this resource reallocation process is governed by the production technology, we call it a *technological effect*: its direction is determined by the degree of substitutability or complementarity between health and cognitive skills, and between health (cognitive skills) and investment in health (cognitive skills). In this scenario, the within-family differences in investments in children are no longer uniquely determined by parental preferences towards inequality, or the price effect.[4] Rather, these differences reflect a mixture of the price effect and of the technological effect. We show that, under plausible assumptions of complementarity between health and cognitive skills, as well as substitutability between health (cognitive skills) and investment in health (cognitive skills), our theoretical model predicts that parents will unambiguously exhibit a reinforcing investment strategy in cognitive skills, and may exhibit a compensating investment strategy in health in response to an early health shock, if they do not have preferences for inequality aversion. The intuition is as follows: if parents do not avert inequality, they will reallocate resources from the insulted child to the healthy one, improving investments on both her health and cognitive skills. However, this does not necessarily imply a reduction in both types of investment in the sick child: as a consequence of the complementarity between health and cognitive skills as well as the substitutability between health (cognitive skills) and investment in health (cognitive skills), parents will unambiguously reduce

[3]The wealth effect denotes the reduction in the human capital stock of the family as consequence of the early health shock. The price effect denotes the change in the relative valuation that the parent has of the child in response to an early health shock.

[4]The wealth effect is removed by the within-family estimator that we use.

the investment in cognitive skills, but may increase the investment in health.[5]

Our result has important implications. On the one hand, in the presence of responsive investments, reduced-form estimates of the effects of early-life shocks cannot be interpreted as a purely *biological effect*. On the other hand, if behavioral adjustments in response to shocks can be compensating and/or reinforcing along different dimensions, we cannot even unequivocally determine if reduced-form estimates represent upper- or lower-bounds of the biological effects. In our application, ignoring the intrahousehold allocation process leads to an *underestimation* of the biological effect of an early health shock on late-life health, but to an *overestimation* of its effect on cognition and related domains.

The essay is organized as follow. We derive our theoretical model in Section 2 and relate it to our econometric specification in Section 3. We describe the Chinese Child Twins Survey we use to test our theoretical predictions in Section 4. Finally, Section 5 presents the results, and Section 6 concludes.

2 A Dynamic Model of Early Shocks, Parental Responses, and Child Outcomes

In this section, we extend the dynamic model of human capability formation developed in Heckman (2007) to a multiple siblings setting, and nest it into a standard model of intrahousehold resource allocation (Becker and Tomes, 1976; Behrman, Pollak, and Taubman, 1982). We show that an early health shock can affect child outcomes through two channels: a direct channel – the production of human capital – and an indirect one – the process of intrahousehold resource allocation. The latter is affected by three factors: the wealth effect, the price effect, and the technological

[5]The study spiritually closest to ours is Behrman and Lavy (1997). However, they do not explicitly model the intrahousehold resource allocation process, which becomes enacted in response to the early-life health shock.

effect. By introducing multidimensionality of child endowments, we allow parents to compensate and reinforce along different dimensions of the child's human capital.

2.1 The Production Technology

We assume that each family has two children ($\iota = i, j$) and that they are twins.[6] There are two periods of childhood ($t = 1, 2$). Each child has a bidimensional skill set: health (H) and other skills. The latter includes both cognitive and noncognitive skills, but we refer to them as cognitive skills (C) in the theoretical section for ease of notation.[7] We denote the endowments and investments in each period as $\theta^k_{\iota,t}$ and $I^k_{\iota,t}$, respectively, where $\iota = i, j$ indexes the child, $t = 0, 1, 2$ is the time period (0 is pre-birth), and $k = H, C$.[8] Following Heckman (2007), we write the production technologies and the investment functions for child i as follows:[9]

$$\theta^H_{i,1} = f^H(\theta^H_{i,0}, \theta^C_{i,0}, I^H_{i,0}, e^H_{i,1}), \tag{1}$$

$$\theta^C_{i,1} = f^C(\theta^H_{i,0}, \theta^C_{i,0}, I^C_{i,0}), \tag{2}$$

$$I^H_{i,1} = f^H(\theta^H_{i,1}, \theta^C_{i,1}, \theta^H_{j,1}, \theta^C_{j,1}), \tag{3}$$

$$I^C_{i,1} = f^C(\theta^H_{i,1}, \theta^C_{i,1}, \theta^H_{j,1}, \theta^C_{j,1}), \tag{4}$$

$$\theta^H_{i,2} = f^H(\theta^H_{i,1}, \theta^C_{i,1}, I^H_{i,1}), \tag{5}$$

$$\theta^C_{i,2} = f^C(\theta^H_{i,1}, \theta^C_{i,1}, I^C_{i,1}), \tag{6}$$

---

[6]This assumption is dictated by the data we use in our empirical analysis. It would be natural to extend the model to a general case with n children in the family. However, fertility and birth spacing may be endogenous to health conditions of existing children (Rosenzweig and Wolpin, 1988). We leave this extension to another occasion.

[7]In our empirical analysis, we distinguish between cognitive and noncognitive skills.

[8]$I^k_{\iota,0}$ indicates maternal investment (e.g. nutritional intake) during pregnancy. Given that our empirical analysis focuses on twins, we can safely assume that $I^k_{i,0} = I^k_{j,0}$ or $I^k_{\iota,0}$ is exogenous across twin siblings. In other words, even if the mother can decide how much to invest during pregnancy, she cannot differentially allocate resources across twin pairs.

[9]For simplicity, we assume no contagion effects between twins throughout the essay.

where $e^H_{i,1}$ is defined as a negative health shock affecting child i in period 1, i.e., $\frac{\partial \theta^H_{i,1}}{\partial e^H_{i,1}} < 0$. We assume that the early health shock $(e^H_{i,1})$ only has a direct effect on her own health in the first period,[10] whereas it affects second-period outcomes through two channels: parental investments (3)-(4) and the process of health and cognitive capital accumulation (5)-(6).[11] Note that in equations (1)-(2) and (5)-(6), we assume that children born in the same family share the same production technology, whereas we allow for the production technology of health to differ from that of cognitive skills. All functions are assumed to be continuously twice differentiable and quasi-concave.

We now analyze the different channels through which an early health shock to child i $(e^H_{i,1})$ operates. First, the total effect on child's i health in the second period can be decomposed as follows:

$$\frac{d\theta^H_{i,2}}{de^H_{i,1}} = \frac{\partial \theta^H_{i,2}}{\partial \theta^H_{i,1}} \cdot \frac{\partial \theta^H_{i,1}}{\partial e^H_{i,1}} + \frac{\partial \theta^H_{i,2}}{\partial I^H_{i,1}} \cdot \frac{\partial I^H_{i,1}}{\partial \theta^H_{i,1}} \cdot \frac{\partial \theta^H_{i,1}}{\partial e^H_{i,1}}, \tag{7}$$

where the first term is a biological effect (self-productivity as in Heckman (2007)). We define the second term as a resource reallocation effect (parents reallocate family resources in response to a health shock on child i). Second, the total effect of an early health shock to child i $(e^H_{i,1})$ on her own cognitive.capacity in the second period can also be decomposed into two channels:

$$\frac{d\theta^C_{i,2}}{de^H_{i,1}} = \frac{\partial \theta^C_{i,2}}{\partial \theta^H_{i,1}} \cdot \frac{\partial \theta^H_{i,1}}{\partial e^H_{i,1}} + \frac{\partial \theta^C_{i,2}}{\partial I^C_{i,1}} \cdot \frac{\partial I^C_{i,1}}{\partial \theta^H_{i,1}} \cdot \frac{\partial \theta^H_{i,1}}{\partial e^H_{i,1}}, \tag{8}$$

where the first term is once again a biological effect (cross-productivity similar to equation (6)), and the second term is an intrahousehold resource reallocation effect.

[10]The health shock may affect the child's brain development, and then has a direct effect on the cognitive skill in the first period. Whether the health shock directly affects the cognitive skills during the same period may depend on what kinds of health shocks the child suffered.

[11]A child can also be hit by a health shock in the second period. We assume that health shocks in the second period are serially uncorrelated with health shocks in the first period, conditional on health in the first period. This assumption can be easily relaxed. It is dictated by the information we have available in our data.

Finally, an early health shock on child i can also affect child j's ($j \neq i$) health and cognitive skills through the intrahousehold resource reallocation process in both cases. Specifically, the cross-effects of child i's health shock on child j's health and cognitive skills are as follows:

$$\frac{d\theta^H_{j,2}}{de^H_{i,1}} = \frac{\partial\theta^H_{j,2}}{\partial I^H_{j,1}} \cdot \frac{\partial I^H_{j,1}}{\partial\theta^H_{i,1}} \cdot \frac{\partial\theta^H_{i,1}}{\partial e^H_{i,1}}, \tag{9}$$

$$\frac{d\theta^C_{j,2}}{de^H_{i,1}} = \frac{\partial\theta^C_{j,2}}{\partial I^C_{j,1}} \cdot \frac{\partial I^C_{j,1}}{\partial\theta^H_{i,1}} \cdot \frac{\partial\theta^H_{i,1}}{\partial e^H_{i,1}}. \tag{10}$$

Combining equations (7)-(10), we derive the net effect of an early health shock affecting child i on the twins' health and cognitive capital as follows:

$$\frac{d\theta^H_{i,2}}{de^H_{i,1}} - \frac{d\theta^H_{j,2}}{de^H_{i,1}} = \frac{\partial\theta^H_{i,2}}{\partial\theta^H_{i,1}} \cdot \frac{\partial\theta^H_{i,1}}{\partial e^H_{i,1}} + \left(\frac{\partial\theta^H_{i,2}}{\partial I^H_{i,1}} \cdot \frac{\partial I^H_{i,1}}{\partial\theta^H_{i,1}} - \frac{\partial\theta^H_{j,2}}{\partial I^H_{j,1}} \cdot \frac{\partial I^H_{j,1}}{\partial\theta^H_{i,1}}\right) \cdot \frac{\partial\theta^H_{i,1}}{\partial e^H_{i,1}}, \tag{11}$$

$$\frac{d\theta^C_{i,2}}{de^H_{i,1}} - \frac{d\theta^C_{j,2}}{de^H_{i,1}} = \frac{\partial\theta^C_{i,2}}{\partial\theta^H_{i,1}} \cdot \frac{\partial\theta^H_{i,1}}{\partial e^H_{i,1}} + \left(\frac{\partial\theta^C_{i,2}}{\partial I^C_{i,1}} \cdot \frac{\partial I^C_{i,1}}{\partial\theta^H_{i,1}} - \frac{\partial\theta^C_{j,2}}{\partial I^C_{j,1}} \cdot \frac{\partial I^C_{j,1}}{\partial\theta^H_{i,1}}\right) \cdot \frac{\partial\theta^H_{i,1}}{\partial e^H_{i,1}}. \tag{12}$$

These equations clearly show the two channels through which early health shocks affect the distribution of health and cognitive capital within families.[12] The first terms on the right-hand side of equations (11) and (12) show how an early health shock $e^H_{i,1}$ affects the health and cognitive capital of child i through self- and cross-productivity: both terms are always negative by definition. The second terms of both equations show how the early health shock operates through the intrahousehold resource allocation process. As they are governed by parental preferences, we now proceed to model them.

2.2 Parental Preferences and Budget Constraint

We assume that parents are altruistic and care about both their own consumption and the quality of their children. Thus, parental preferences can be represented by

[12] Our within-twin-pair fixed-effects estimator gives us an estimate of these effects.

a utility function of the following form:[13]

$$U_P = U_P[c, V(\theta^H_{i,2}, \theta^C_{i,2}), V(\theta^H_{j,2}, \theta^C_{j,2})], \tag{13}$$

where c is parental consumption,[14] and $V(\theta^H_{\iota,2}, \theta^C_{\iota,2})$ is the child quality function ($\iota = i, j$). Note that both children have the same quality function but may have different health and cognitive skills in the second period. The budget constraint is specified as follows:[15]

$$p_c \cdot c + I^H_{i,1} + I^H_{j,1} + I^C_{i,1} + I^C_{j,1} = Y,$$

where p_c is the price of parental consumption, Y is the parents' total resources, the price of investment is normalized to one, and it is independent of the type of investment. We denote the total value of the resources allocated to children as follows:[16]

$$I = I^H_{i,1} + I^H_{j,1} + I^C_{i,1} + I^C_{j,1} \tag{14}$$

Following Behrman, Pollak, and Taubman (1982), we assume that the utility parents derive from children can be separated from parental consumption. Thus, we can rewrite the utility function (13) as follows:[17]

$$U_P = U_P\{c, U[V(\theta^H_{i,2}, \theta^C_{i,2}), V(\theta^H_{j,2}, \theta^C_{j,2})]\}, \tag{15}$$

[13]The parental utility function should also include the number of children. However, we omit this argument because the implementation of the "One-Child" policy at the time of the data collection allows us to assume away issues of endogenous fertility.

[14]We assume that children's consumption (excluding investments) is a basic need and that parents allocate resources identically across them. Thus, we can ignore this term in the parental utility function.

[15]We assume no borrowing or saving. Although this assumption can be easily relaxed along the lines of Behrman, Pollak, and Taubman (1982), on the one hand it is dictated by the information available in our data, and on the other it has empirical plausibility given the structure of the banking system in the Yunnan region for the period we consider, as described in Section 4.

[16]We assume that parents provide all investment to children; i.e., there is no public intervention. This assumption is plausible in our case, given the absence of public programs in the Yunnan region for the period we consider, as described in Section 4.

[17]We implement a test of this separability assumption in Section 5.3.

The separability assumption is very convenient because it allows us to focus on the allocation of resources across children without considering its effects on parental consumption. Thus, we can restate the problem of parental investments in children as that of maximizing the following utility function:

$$U = U[V(\theta^{H}_{i,2}, \theta^{C}_{i,2}), V(\theta^{H}_{j,2}, \theta^{C}_{j,2})], \tag{16}$$

subject to the investment budget constraint (14),[18] the production technologies of health and cognitive skills (1)-(2) and (5)-(6), and the quality function.

2.3 Early Health Shocks and Parental Resource Reallocation

To derive the comparative static results of the effects of an early health shock on parental resource reallocation, we follow Behrman, Pollak, and Taubman (1982) and specify parental preferences using a CES utility function[19]

$$U = \{[V(\theta^{H}_{i,2}, \theta^{C}_{i,2})]^{\rho} + [V(\theta^{H}_{j,2}, \theta^{C}_{j,2})]^{\rho}\}^{\frac{1}{\rho}}, \tag{17}$$

[18] Although we also analyze the effects of the early health shock on parental time investment, for simplicity we do not include a time constraint in our theoretical model. We leave this extension to another occasion.

[19] We assume that parents have equal concerns for their children. Thus, the weights in the child quality function are equal and normalized to one. Graphically, this means that the parental welfare function (equation (17)) is symmetrical around the 45° ray from the origin. However, it does not automatically imply that resources are equally distributed across children because they may have different endowments or may be differentially affected by shocks, as the current essay shows. Note that, although the optimal level of investments will be changed, the analytical results of the comparative statics remain qualitatively the same if we assume that parents put different weights on the quality of different children. For more discussions on the parental welfare function, see Behrman, Pollak, and Taubman (1982).

where $\rho < 1$.[20] An excellent feature of the CES representation of the parental utility function is that ρ measures the degree of parental inequality aversion across children. When $\rho < 0$, parents exhibit inequality aversion and allocate more resources to the sick child. However, when $0 < \rho < 1$, parents do not exhibit inequality aversion and allocate more resources to the healthy child. Conceptually, the sign of ρ is determined by the tradeoff between efficiency and equality. If the decision of investing in children is mainly motivated by efficiency, then $0 < \rho < 1$. Otherwise, the equality motive outweighs the efficiency motive, and $\rho < 0$ (Behrman, Pollak, and Taubman, 1982). In developing countries, efficiency may be the major consideration (at least, in cases when resources are constrained), and ρ would be more likely to be positive. In contrast, equality may be the major consideration in developed countries, and thus ρ would be more likely to be negative.

We then assume the following functional form for the child quality function $V(\theta^H_{\iota,2}, \theta^C_{\iota,2})$ $(\iota = i, j)$:

$$V(\theta^H_{\iota,2}, \theta^C_{\iota,2}) = (\theta^H_{\iota,2})^{\alpha_H}(\theta^C_{\iota,2})^{\alpha_C}, \tag{18}$$

where $0 < \alpha_H, \alpha_C < 1$, and $\alpha_H(\alpha_C)$ measures the importance of health (cognition) in the quality function. Finally, following Behrman, Pollak, and Taubman (1982) and Cunha and Heckman (2008), we assume substitutability between investment in health $(I^H_{\iota,1})$ and the stock of health $(\theta^H_{\iota,1})$ in the health production function $(\theta^H_{\iota,2})$,[21] as well as between investment in cognitive skills $(I^C_{\iota,1})$ and the cognitive stock $(\theta^C_{\iota,1})$ in the cognitive skills production function $(\theta^C_{\iota,2})$. Thus, we can specify the following

[20] ρ is a continuous variable, and it implies that all parents have both efficiency and equality considerations unless $\rho = 1$ or $\rho = -\infty$. $\rho = 1$ means that parents only care about efficiency, whereas $\rho = -\infty$ means that parents only care about equality. The latter is the Rawlsian case, in which case the parental utility function (16) can be rewritten as $U = U\left[min\left(V_i, V_j\right)\right]$.

[21] This assumption is also consistent with the original formulation in Grossman (1972): $H_{t+1} = I_t + (1-\delta)\,H_t$, where H is the stock of health, I is gross investment, and δ is depreciation.

functional forms for the production technologies:[22]

$$\theta^{H}_{i,2} = (\theta^{C}_{i,1})^{\gamma}[\beta_\theta \theta^{H}_{i,1} + \beta_I I^{H}_{i,1}]^{1-\gamma}, \quad (19)$$

$$\theta^{C}_{i,2} = (\theta^{H}_{i,1})^{\gamma}[\beta_\theta \theta^{C}_{i,1} + \beta_I I^{C}_{i,1}]^{1-\gamma}, \quad (20)$$

where $0 < \gamma < 1$ and $0 < \beta_\theta, \beta_I < 1$. The parameter γ can be interpreted as the importance of the first-period cognition (health) in producing health (cognition) in the second period, whereas the parameter β_θ can be interpreted as the *relative* importance of the first-period health (cognition) in producing health (cognition) in the second period, relative to investment in health (cognition) in the first period.

By solving the parental optimization problem,[23] we derive the optimal investment in the health and cognition of child i as follows:[24]

$$I^{H*}_{i,1} = \frac{\alpha_H}{\beta_I} W \pi_i - \frac{\beta_\theta}{\beta_I} \theta^{H}_{i,1}, \quad (21)$$

$$I^{C*}_{i,1} = \frac{\alpha_C}{\beta_I} W \pi_i - \frac{\beta_\theta}{\beta_I} \theta^{C}_{i,1}, \quad (22)$$

where:

$$W = \beta_\theta \left(\theta^{H}_{i,1} + \theta^{C}_{i,1} + \theta^{H}_{j,1} + \theta^{C}_{j,1}\right) + \beta_I I, \quad (23)$$

$$\pi_i = \frac{V\left(\theta^{H}_{i,2}, \theta^{C}_{i,2}\right)^{\rho}}{U^{\rho}}. \quad (24)$$

Let us first consider equation (23). W measures the *full resources* devoted to the production of health and cognitive skills in the second period, which includes the health and cognitive stock of both children in the first period and the total resources

[22]We assume a Cobb-Douglas production technology to simplify the calculations. Our basic results are unchanged if we assume a general CES production technology and relax the assumption of substitutability between investments and stocks of skills. The results with this alternative specification are reported in the appendix.

[23]The solution to the parental optimization problem is obtained by maximizing the utility function (17) subject to the investment budget constraint (equation (14)), the production technologies (equations (19)-(20)), and the quality function (equation (18)).

[24]The formal derivation is reported in the appendix.

allocated to children in the first period, weighted by their relative importance in the production function (equations (19)-(20)). Note that $dW/d\theta_{i,1}^H = \beta_\theta > 0$: a one-unit increase in child i's health in the first period increases the *full resources* by β_θ. We call this the *wealth effect* as in Becker and Tomes (1976). The wealth effect is always positive. Let us now consider equation (24): π_i measures the relative importance of child i in the parental utility function.[25] Thus, $W\pi_i$ measures the share of total resources allocated to child i. It is important to note that the sign of $d\pi_i/d\theta_{i,1}^H$ is unambiguously determined by the parental inequality aversion parameter ρ:[26] when $\rho > 0$, parents give more weight to efficiency than to equality, so they allocate more resource to child i if this child has better health in the first period. Following Becker and Tomes (1976), we interpret $d\pi_i/d\theta_{i,1}^H$ as a "*price effect*", as an increase in child i's health stock changes the child's relative importance or *shadow price* in the parental utility function.[27] Let us finally consider the equation for optimal investment in health (equation (21)). In this equation, α_H measures the relative importance of health in the child quality function (equation (18)); β_I measures the productivity of the investment in health (equation (19)); and β_θ/β_I measures the trade-off between health in the first period and investments in health in the production technology (equation (19)). An analogous interpretation applies to equation (22) for optimal investment in cognitive skills.

We now derive the comparative static results for the effect of health in the first period on investment in health and cognitive skills for child i:

$$\frac{\partial I_{i,1}^{H*}}{\partial \theta_{i,1}^H} = \frac{\alpha_H}{\beta_I}\left(\frac{\partial W}{\partial \theta_{i,1}^H}\pi_i + \frac{\partial \pi_i}{\partial \theta_{i,1}^H}W\right) - \frac{\beta_\theta}{\beta_I}, \tag{25}$$

$$\frac{\partial I_{i,1}^{C*}}{\partial \theta_{i,1}^H} = \frac{\alpha_C}{\beta_I}\left(\frac{\partial W}{\partial \theta_{i,1}^H}\pi_i + \frac{\partial \pi_i}{\partial \theta_{i,1}^H}W\right). \tag{26}$$

[25] Note that $U^\rho = V_i(\theta_{i,2}^H, \theta_{i,2}^C)^\rho + V_j(\theta_{j,2}^H, \theta_{j,2}^C)^\rho$.

[26] The mathematical derivation is shown in the appendix.

[27] The shadow price here involves not only resources but also utility.

Note that, in addition to the wealth effect and the price effect discussed above, equation (25) also includes an additional term, $(-\beta_\theta/\beta_I)$: we call this term the *technological effect*, because it stems directly from the health production technology (equation (19)). Due to the substitutability between the health stock in the first period and the investment in health (equation (19)), an increase in the health stock in the first period will reduce the amount invested in health. Thus, the technological effect is always negative. As noted above, the wealth effect is always positive, whereas the sign of the price effect depends on the parental degree of inequality aversion: $\partial\pi_i/\partial\theta^H_{i,1}$ is positive if $\rho > 0$ (efficiency outweighs equality), whereas it is negative if $\rho < 0$ (equality outweighs efficiency). In either case, the own effect of first-period health on investment in health is ambiguous. On the contrary, the own effect of first-period health on investment in cognitive skills is always positive if parents exhibit no inequality aversion, as both the wealth effect and the price effect are positive (equation (26)).

We now investigate the cross-effects of child i's health in the first period on investment in health and cognitive skills of child j:

$$\frac{\partial I^{H*}_{j,1}}{\partial\theta^H_{i,1}} = \frac{\alpha_H}{\beta_I}\left(\frac{\partial W}{\partial\theta^H_{i,1}}\pi_j + \frac{\partial\pi_j}{\partial\theta^H_{i,1}}W\right), \tag{27}$$

$$\frac{\partial I^{C*}_{j,1}}{\partial\theta^H_{i,1}} = \frac{\alpha_C}{\beta_I}\left(\frac{\partial W}{\partial\theta^H_{i,1}}\pi_j + \frac{\partial\pi_j}{\partial\theta^H_{i,1}}W\right). \tag{28}$$

Note that $\partial\pi_j/\partial\theta^H_{i,1}$ has a sign opposite to $\partial\pi_i/\partial\theta^H_{i,1}$ because $\pi_i + \pi_j = 1$.[28] Hence, the price effects on investments in health and cognition are always negative if parents exhibit no inequality aversion. Subtracting pairwise the equations (25)-(28), we

[28] For example, when efficiency outweighs equality ($\rho > 0$), $\partial\pi_i/\partial\theta^H_{i,1} > 0$, while $\partial\pi_j/\partial\theta^H_{i,1} < 0$.

obtain the following:[29]

$$\frac{\partial I_{i,1}^{H*}}{\partial \theta_{i,1}^{H}} - \frac{\partial I_{j,1}^{H*}}{\partial \theta_{i,1}^{H}} = \frac{\alpha_H}{\beta_I}\left(\frac{\partial \pi_i}{\partial \theta_{i,1}^{H}} - \frac{\partial \pi_j}{\partial \theta_{i,1}^{H}}\right) W - \frac{\beta_\theta}{\beta_I}, \quad (29)$$

$$\frac{\partial I_{i,1}^{C*}}{\partial \theta_{i,1}^{H}} - \frac{\partial I_{j,1}^{C*}}{\partial \theta_{i,1}^{H}} = \frac{\alpha_C}{\beta_I}\left(\frac{\partial \pi_i}{\partial \theta_{i,1}^{H}} - \frac{\partial \pi_j}{\partial \theta_{i,1}^{H}}\right) W. \quad (30)$$

When parents give more weight to efficiency than to equality ($\rho > 0$), $\partial I_{i,1}^{C*}/\partial \theta_{i,1}^{H} - \partial I_{j,1}^{C*}/\partial \theta_{i,1}^{H}$ is positive, whereas the sign of $\partial I_{i,1}^{H*}/\partial \theta_{i,1}^{H} - \partial I_{j,1}^{H*}/\partial \theta_{i,1}^{H}$ is undetermined because it depends on the relative magnitude of the price effect $\left(\frac{\partial \pi_i}{\partial \theta_{i,1}^{H}} - \frac{\partial \pi_j}{\partial \theta_{i,1}^{H}}\right)$, which is positive, and the technological effect $-\left(\frac{\beta_\theta}{\beta_I}\right)$, which is negative.

We now summarize the main predictions of our theoretical model that we will test empirically. The first prediction is related to the effect of an early health shock affecting child i on the difference in investment in health and cognitive skills across twins. It is obtained directly from equations (29) and (30):

$$\left(\frac{\partial I_{i,1}^{H*}}{\partial \theta_{i,1}^{H}} - \frac{\partial I_{j,1}^{H*}}{\partial \theta_{i,1}^{H}}\right)\frac{\partial \theta_{i,1}^{H}}{\partial e_{i,1}^{H}} = \left[\frac{\alpha_H}{\beta_I}\left(\frac{\partial \pi_i}{\partial \theta_{i,1}^{H}} - \frac{\partial \pi_j}{\partial \theta_{i,1}^{H}}\right) W - \frac{\beta_\theta}{\beta_I}\right]\frac{\partial \theta_{i,1}^{H}}{\partial e_{i,1}^{H}}, \quad (31)$$

$$\left(\frac{\partial I_{i,1}^{C*}}{\partial \theta_{i,1}^{H}} - \frac{\partial I_{j,1}^{C*}}{\partial \theta_{i,1}^{H}}\right)\frac{\partial \theta_{i,1}^{H}}{\partial e_{i,1}^{H}} = \left[\frac{\alpha_C}{\beta_I}\left(\frac{\partial \pi_i}{\partial \theta_{i,1}^{H}} - \frac{\partial \pi_j}{\partial \theta_{i,1}^{H}}\right) W\right]\frac{\partial \theta_{i,1}^{H}}{\partial e_{i,1}^{H}}. \quad (32)$$

When $\rho > 0$, the within-twin-pair fixed-effects estimate of the effect of an early health shock on the investment in cognitive skills is predicted to be negative (equation (32)). However, the sign of the effect on investment in health is ambiguous (equation (31)), because it depends on the relative magnitude of the price effect (which is positive) and the technological effect (which is negative). The case when $\rho < 0$ can be analyzed in a similar way.

The second prediction is related to the effect of an early health shock on health and cognition in the second period. By plugging equations (29)-(30) into equations (11)-(12) and assuming that the productivity of the investment is the same across

[29]Note we assume $\pi_j = \pi_i$, consistently with the assumption that parents have equal concerns.

twins $\left(\text{i.e.} \frac{\partial\theta^H_{i,2}}{\partial I^H_{i,1}} = \frac{\partial\theta^H_{j,2}}{\partial I^H_{j,1}} \text{ and } \frac{\partial\theta^C_{i,2}}{\partial I^C_{i,1}} = \frac{\partial\theta^C_{j,2}}{\partial I^C_{j,1}}\right)$, we obtain the following:

$$\frac{d\theta^H_{i,2}}{de^H_{i,1}} - \frac{d\theta^H_{j,2}}{de^H_{i,1}} = \frac{\partial\theta^H_{i,2}}{\partial\theta^H_{i,1}} \cdot \frac{\partial\theta^H_{i,1}}{\partial e^H_{i,1}} + \frac{\partial\theta^H_{i,2}}{\partial I^H_{i,1}} \cdot \left[\frac{\alpha_H}{\beta_I}\left(\frac{\partial\pi_i}{\partial\theta^H_{i,1}} - \frac{\partial\pi_j}{\partial\theta^H_{i,1}}\right)W - \frac{\beta_\theta}{\beta_I}\right] \cdot \frac{\partial\theta^H_{i,1}}{\partial e^H_{i,1}} \quad (33)$$

$$\frac{d\theta^C_{i,2}}{de^H_{i,1}} - \frac{d\theta^C_{j,2}}{de^H_{i,1}} = \frac{\partial\theta^C_{i,2}}{\partial\theta^H_{i,1}} \cdot \frac{\partial\theta^H_{i,1}}{\partial e^H_{i,1}} + \frac{\partial\theta^C_{i,2}}{\partial I^C_{i,1}} \cdot \left[\frac{\alpha_C}{\beta_I}\left(\frac{\partial\pi_i}{\partial\theta^H_{i,1}} - \frac{\partial\pi_j}{\partial\theta^H_{i,1}}\right)W\right] \cdot \frac{\partial\theta^H_{i,1}}{\partial e^H_{i,1}}. \quad (34)$$

When $\rho > 0$, the within-twin-pair fixed-effects estimate of the effect of the early health shock on cognitive skills in the second period is predicted to be negative (equation (34)). However, the sign of the effect of the early health shock on health in the second period is ambiguous (equation (33)), because it depends on the relative magnitude of the price effect (which is positive) and the technological effect (which is negative). The case when $\rho < 0$ can be analyzed in a similar way.

Finally, before moving on to the econometric model, we discuss the implications of our theoretical model and its relationship with the empirical analysis below. An ambitious objective is to estimate the dynamic model as we have laid it out and to identify separately the parental preferences from the technology parameters. Unfortunately, we are not able to achieve this objective in this essay because our data do not contain information on the child's health and cognitive skill stock in the first period $\left(\theta^H_{i,1} \text{ and } \theta^C_{i,1}\right)$.[30] Thus, we carry out the reduced-form estimation of equations (31)-(34) below.[31] However, although we cannot estimate the entire structural system, our theoretical model plays a key role in guiding the interpretation of our empirical results. First, the model rationalizes that parents can make compensating and reinforcing investments along different dimensions during the same time (equations (31)-(32)). This is the key insight we plan to test in the empirical part. Second, it lays down the basic framework that can be used to interpret the

[30]The data set will be discussed in detail in section 4. The data we use are essentially cross-sectional, and the early health shock variable is constructed retrospectively.

[31]One consequence of this is that, for example, we will not be able to ascertain if the negative health effects of an early health shock are derived from a change in the production technology.

within-twin-pair fixed-effects estimates by using the estimates from equations (33)-(34) as the lower or upper bound of the biological effects.[32] Because we are able to estimate the reduced-form effects of an early health shock on investment in health $\left[\frac{\alpha_H}{\beta_I}\left(\frac{\partial \pi_i}{\partial \theta^H_{i,1}} - \frac{\partial \pi_j}{\partial \theta^H_{i,1}}\right)W - \frac{\beta_\theta}{\beta_I}\right]$ and cognitive skills $\left[\frac{\alpha_C}{\beta_I}\left(\frac{\partial \pi_i}{\partial \theta^H_{i,1}} - \frac{\partial \pi_j}{\partial \theta^H_{i,1}}\right)W\right]$ from equations (31) and (32), and the signs of $\frac{\partial \theta^H_{i,2}}{\partial I^H_{i,1}}$ (equation (33)) and $\frac{\partial \theta^C_{i,2}}{\partial I^C_{i,1}}$ (equation (34)) are always positive, we are able to infer whether the reduced-form estimates of $\frac{d\theta^H_{i,2}}{de^H_{i,1}} - \frac{d\theta^H_{j,2}}{de^H_{i,1}}$ (equation (33)) and $\frac{d\theta^C_{i,2}}{de^H_{i,1}} - \frac{d\theta^C_{j,2}}{de^H_{i,1}}$ (equation (34)) are *lower* or *upper* bounds of the biological effects. Third, the theoretical model provides a framework to interpret the differences between the OLS and the within-twin-pair fixed-effects estimates, as we will discuss in a later section. Therefore, we will discuss our empirical results in light of our theoretical framework in the following sections.

3 The Econometric Analysis

In this section, we present the econometric specification we estimate, guided by our theoretical model. We first analyze how parents respond to an early health shock, by specifying the stochastic version of the parental investment equation as follows:

$$I^\kappa_{i,\tau} = a^\kappa_1 e^H_{i,\tau} + a^\kappa_2 e^H_{j,\tau} + X_{i,\tau} b^\kappa_1 + X_{j,\tau} b^\kappa_2 + \zeta_\tau \varphi^\kappa + \mu_\tau + \epsilon^\kappa_{i,\tau}, \qquad (35)$$

where $\kappa = H, C$, i and j index the two twin siblings in household τ. I^κ is the investment in κ during the first period; e^H is a health shock in the first period;[33] X is a vector of child-specific characteristics; ζ_τ is a vector of observed household characteristics affecting parental investment decisions; μ_τ is the unobservable household heterogeneity such as reporting heterogeneity which will be discussed later; and ϵ^κ

[32]The biological effects are represented by $\frac{\partial \theta^H_{i,2}}{\partial \theta^H_{i,1}} \cdot \frac{\partial \theta^H_{i,1}}{\partial e^H_{i,1}}$ (equation (33)) and $\frac{\partial \theta^C_{i,2}}{\partial \theta^H_{i,1}} \cdot \frac{\partial \theta^H_{i,1}}{\partial e^H_{i,1}}$ (equation (34)).

[33]As clarified in Section 4, the health shock, as measured in our data, occurs between the ages of 0 and 3, and parental investment refers to the year prior to the survey (the twins are between 6 and 18 years old, with mean age 11, at the time of the survey; see Table 1).

is the disturbance term. To sweep out family-level unobserved heterogeneity, we use the following within-twin-pair fixed-effects specification:

$$I^{\kappa}_{i,\tau} - I^{\kappa}_{j,\tau} = \alpha^{\kappa}\left(e^{H}_{i,\tau} - e^{H}_{j,\tau}\right) + \left(X_{i,\tau} - X_{j,\tau}\right)\beta^{\kappa} + \epsilon^{\kappa}_{i,\tau} - \epsilon^{\kappa}_{j,\tau}, \tag{36}$$

where $\alpha^{\kappa} = a^{\kappa}_{1} - a^{\kappa}_{2}$ and $\beta^{\kappa} = b^{\kappa}_{1} - b^{\kappa}_{2}$. Equation (36) is the empirical counterpart of equations (31)-(32). Our theoretical model shows that the within-twin-pair estimator, in removing the family-level unobserved heterogeneity, also sweeps out the wealth effect induced by an early health shock. Thus, when parents give more weight to efficiency than to equality, i.e., $\rho > 0$, our theoretical model unambiguously predicts α^{C} to be negative. However, the sign of α^{H} remains undetermined because it depends on the trade-off between the degree of parental inequality aversion (the price effect) and the substitutability between investment in health and the stock of health in the first period to produce health in the second period (the technological effect).

We then analyze how an early health shock affects later outcomes, using the following specification:

$$\theta^{\kappa}_{\iota,\tau} = \gamma^{\kappa} e^{H}_{\iota,\tau} + X_{\iota,\tau}\delta^{\kappa} + \zeta_{\tau}\psi^{\kappa} + \mu_{\tau} + \varepsilon^{\kappa}_{\iota,\tau}, \tag{37}$$

where $\theta^{\kappa}_{\iota,\tau}$ is the outcome κ for the twin child ι ($\iota = i, j$)in household τ in the second period,[34] and all the other terms are defined as in equation (35). The corresponding within-twin-pair fixed-effects specification is:

$$\theta^{\kappa}_{i,\tau} - \theta^{\kappa}_{j,\tau} = \gamma^{\kappa}(e^{H}_{i,\tau} - e^{H}_{j,\tau}) + (X_{i,\tau} - X_{j,\tau})\delta^{\kappa} + \varepsilon^{\kappa}_{i,\tau} - \varepsilon^{\kappa}_{j,\tau}. \tag{38}$$

Equation (38) is the empirical counterpart of equations (33)-(34). Our theoretical

[34]As clarified in Section 4, the outcomes refer to the year of the survey when the twins are between 6 and 18 years old (mean age 11; see Table 1).

model predicts the sign of γ^C to be unambiguously negative if efficiency outweighs equality when parents make investment decisions. However, the sign of γ^H is undetermined, as discussed in Section 2.

Before proceeding to the data description, we now discuss our identification strategy. On the one hand, although siblings are biologically similar to dizygotic twins, the within-twin-pair fixed-effects estimator requires much weaker identification assumptions than the within-siblings fixed-effects estimator when estimating child outcomes production functions (Todd and Wolpin, 2007). Specifically, the within-siblings fixed-effects estimator requires three additional assumptions. First, the effects of an early health shock must be either independent of age if siblings' outcomes are measured at different ages but at the same point in time, or independent of time if siblings' outcomes are measured at the same age but at different points in time. Second, parents must not make time-varying investments across siblings. Third, parents must not adjust their fertility choices and investment behavior in response to a health shock affecting their existing children, an assumption which seems untenable according to Rosenzweig and Wolpin (1988) and the suggestive evidence we provide.[35]

On the other hand, our within-twin-pair fixed-effects estimator still relies on the assumption that $\varepsilon_{i,\tau}-\varepsilon_{j,\tau}$ and $\epsilon_{i,\tau}-\epsilon_{j,\tau}$ are uncorrelated with $e^H_{i,\tau}-e^H_{j,\tau}$, conditional on the observables. In other words, our key identification assumption is that, conditional on the observed covariates, the early health shock occurs randomly within twin pairs. Of course, there is always the possibility that it can reflect unobserved health differences. Unfortunately, due to data limitations, we cannot estimate a model that also includes individual-level unobserved heterogeneity, but we try to address this

[35]Table 1 in the appendix provides suggestive evidence that the fertility decision is significantly affected by the health status of the first child: the occurrence of a health shock in the child at ages 0-1 has a significantly negative association with the probability that the mother has a second child. This table is based on the comparison group of non-twin households in our survey data (see the data description section below). Our results are consistent with the findings of Rosenzweig and Wolpin (1988), and show the usefulness of adopting the twin-fixed-effects method in the presence of the "One-Child" policy.

concern by controlling for birth weight in all our specifications. Our rationale for doing so is that birth weight can be considered a proxy for the child's stock of health capital at birth, before the occurrence of the early health shock at ages 0-3 (Behrman and Rosenzweig, 2004; Almond, Chay, and Lee, 2005).[36]

4 Data

4.1 The Chinese Child Twins Survey

The data we use for this study come from the Chinese Child Twins Survey (CCTS), which is the first census-type child twins survey of which we are aware.[37] The survey was carried out by the Urban Survey Unit (USU) of the National Bureau of Statistics (NBS) in late 2002 and early 2003 in the Kunming district of China. Kunming is the capital of Yunnan Province, which is located in the far southwestern corner of China and has a total population of about 5 million.

The CCTS includes a sample of households with twins aged between 6 and 18 years living in Kunming in 2002. The households were initially identified by the USU based on the 2000 population census according to whether the children have the same birth year and month and the same relationship with the household head. The addresses of these households were then obtained from the census office, and the presence of twins was verified with a visit to the household. Starting from 2,300 pairs of potential twins identified in the census, 1,694 households with twins were successfully interviewed; among these, 1,300 households had twins on the first birth and 394 households had twins on the second birth.[38] A comparison sample of 1,693

[36]Evidence and discussion on the randomness of early health shocks within twin pairs are shown in a later section.

[37]See Rosenzweig and Zhang (2009) for a detailed description of the CCTS.

[38]The "One-Child" policy is strictly implemented in urban areas in Kunming. In rural areas, however, households are encouraged to have one child, but are exempted from the strict "One-Child" policy (although they are allowed to have two children at most (Family Planning Commission of Yunnan Province, 2003)). This is evident in Panel H of Table 1, where the proportion of twins

households with no twins was also surveyed using the same questionnaire.[39]

The questionnaire was designed by Junsen Zhang in close consultation with Mark Rosenzweig and Chinese experts at the National Bureau of Statistics. Based on existing twins and child questionnaires in the US and elsewhere, the survey covers an extensive range of information about inputs and outcomes of children, in addition to a wide range of demographic, social, and economic information at the household level. The questionnaire is divided in two parts. The first part is answered jointly by the father, mother, and children, and collects information on the household situation, parents, schooling and health of the children, and parental investments. After completing the first part, each parent and each child are separately interviewed in different rooms. The second part covers information on home tutoring, children's schooling and academic performance, entertainment, and social activities.

We exploit two features of the Chinese institutional system in our empirical analysis. First, the existence of the "One-Child" policy serves as a natural experiment to eliminate the possibility that the fertility decision will be endogenously affected by the health condition of the twin children (this is an issue raised by Rosenzweig and Wolpin (1988)). The second feature of the Chinese system that we exploit in our empirical analysis is the strict household registration policy known as *hukou.* The *hukou* system was established in the early 1950s to consolidate socialist governance, control population flow, and administer the planned economy. Under this system, every person is required to be registered where she is born and to obtain a *hukou* certificate: all administrative activities, such as land distribution, issuance of ID cards, registration of a child in school, and registration of marriage, are based on the *hukou*

born at the second birth is much higher in rural (0.33) than in urban (0.07) areas. In our analysis, we include both first-birth (in which case parents are not allowed to have any more children) and second-birth twins because the results are qualitatively the same if we exclude the latter sample.

[39]To guarantee the comparability of the non-twin group, the fourth household on the right-hand side of the same block of the twin household was chosen as the non-twin comparison. If the fourth household had no children aged 6-18, then interviewers continued with the fifth, sixth, etc.

status.[40] Conveniently, at the time the survey was carried out, the *hukou* system was still very strict in the Kunming district. This allows us to compare rural and urban samples without worrying about selectivity concerns arising from migration into the richer urban areas. Therefore, we can interpret these results in light of the differences in the institutional backgrounds between urban and rural areas (West and Zhao, 2000). First, at the time of the survey, the medical insurance system was almost absent in rural areas,[41] whereas medical expenditures on children could be partly reimbursed by the government if the parents were affiliated with government departments or state-owned enterprises in urban areas (Liu, Rao, Wu, and Gakidou, 2008). Second, although public education was not free in both urban and rural areas at the time of the survey,[42] its quality in urban areas was much higher than that in rural areas. Finally, residents in urban areas were covered by the old-age pension system (although the amount of money provided by the government may have been insufficient to satisfy the basic needs), whereas there was no old-age pension system in rural areas at all. We will return to and take all these institutional features into account in the interpretation of our empirical results.[43]

[40]Until the early 1990s, it was also used to distribute food, cooking oil, and clothing coupons. Moreover, it imposed strong restrictions in moving across localities, both in urban and rural areas. Although the Chinese government has been gradually reforming it since the mid-1990s, the *hukou* system is still very strict in most places (Yusuf and Saich, 2008).

[41]The Chinese government began to promote the New Cooperative Medical System (NCMS) in rural areas after 2003. NCMS is a co-pay insurance system financed by the central government, local government, and individuals (Brown, de Brauw, and Du, 2009).

[42]The tuition fee for compulsory education (six years of primary school and three years of middle school) has been exempted in both rural and urban areas only since September 1, 2008.

[43]In addition to urban and rural, we also compare the results between the male and female subsamples. The results for the mixed-gender subsamples are only reported in the appendix because the sample size becomes much smaller in this case, as it only includes DZ twin pairs. Note that we do not distinguish between MZ and DZ twins in our estimation, given that the criterion to establish zigosity in our data is not based on DNA testing but on physical resemblance. Thus, it is subject to considerable error, which is likely to be correlated with parental behavior (e.g. parents may actually themselves attenuate pre-existing differences among twins). In any case, previous results in the literature do not point to the existence of marked differences in analyses based only on the MZ or DZ subsample, as, for example, in Black, Devereux, and Salvanes (2007).

4.2 The Summary Statistics

We now describe the main variables that are used in our empirical analysis. The summary statistics are reported in Table 1.

Early Health Shocks Our independent variable of interest (early health shocks, $e^H_{i,1}$) is defined as a dummy variable indicating whether the child suffered from a serious disease during ages 0-3.[44] Table 1 (Panel A) shows that the prevalence rate in our sample is 9%. The most prevalent diseases are serious diarrhoea and calcium deficiency, as is the case for children in developing countries (Strauss and Thomas, 1998).[45] Table 2 in the appendix tabulates the distribution of serious diseases suffered during ages 0-3.[46]

We now address some potential concerns regarding the measurement of early health shocks as they are based on health histories constructed retrospectively. First, retrospective data may suffer from recall error, particularly, parents may report that the child who is currently sick was also sick in the past. The fixed-effects estimator may exacerbate this problem (Strauss and Thomas, 1998).[47] There are three reasons why we believe this to be less of a concern in our

---

[44] The illness, which is used to measure a health shock, can be either an outcome of a shock or reflects an individual-specific health endowment which would be a persistent component. If we assume that the individual-specific persistent component to be identical across twin siblings conditional on observable variables, the within-twin-pair variation in the illness will only reflect the part induced by the shocks. Such heterogeneity in health endowment would be removed by the within-twin-pair fixed-effects estimator.

[45] The complete list also includes asthma, fracture, attention deficit disorder, heart disease, serious hearing difficulties, whooping cough, stammer, and serious eyesight problems. Unfortunately, we cannot distinguish between mental and physical diseases because the former have low prevalence in our sample; see Currie and Stabile (2006) for an analysis of the effect of child mental health on human capital accumulation. Another interesting extension that we leave to a future occasion is to separate the effect of life-threatening shocks. This may well be a circumstance with infinite parental inequality aversion.

[46] We also defined our main independent variable as the number of serious diseases suffered during ages 0-3. The empirical results obtained using this alternative definition are almost identical to the ones reported in the essay and are available from the authors upon request.

[47] In general, the classical measurement error will bias the fixed-effects estimates towards zero. This is not the case in our study, because, as discussed in Section 5, our fixed-effects estimates are generally of a bigger magnitude than the corresponding OLS estimates.

case: (a) the health history questions are answered together by the father, mother, and children (in the first part of the questionnaire); (b) given the young age of the twin sample, the recall period is not very long; (c) parents and children are also asked to specify the timing and duration of each disease. This contextualization has the potential to increase recollection effort and further minimize recall error.

Second, respondents may use different thresholds so that some of the differences in the reported illnesses across households may simply reflect differences in the standards (Strauss and Thomas, 1998; Smith, 2009). For example, more educated households can both keep more accurate medical records and have higher standards. This is termed as reporting heterogeneity in the literature (Strauss and Thomas, 1998).[48] The problem of reporting heterogeneity may also exist in our case. Although the medical and economic environments are much better in urban areas than in rural areas as discussed above, Table 1 shows that the prevalence rate of early health shocks is 10% in urban areas, whereas it is only 8% in rural areas. The difference is statistically significant indicating that urban families are more likely to report early health shocks (rather than children in urban areas being more likely to suffer from them). However, these differences are unlikely to exist across twin siblings in the same family. Thus, our within-twin-pair fixed-effects estimation strategy will also avoid the bias arising from reporting heterogeneity.[49]

Parental Investments Our main dependent variables are the measures of the parental investments in children in the year before the survey $\left(I^{K}_{i,1}\right)$. Due

[48]The reporting heterogeneity can be regarded as a component of μ_r in equation (35).

[49]Another interesting aspect of the twin design is that it overcomes the usual problem of the lack of an explicit reference group (or anchoring): it is natural for the parents to think of one twin as the reference point for the other. Curiously, Bago d'Uva, van Doorslaer, Lindeboom, and O'Donnell (2008) find that reporting heterogeneity does not seem to be a source of distortion for the measurement of health disparities in China.

to the richness of our data, we are able to differentiate between investment in money (i.e., medical, education and clothing expenditures) and investment in time (i.e., minutes per day the parents spend tutoring each twin). Medical expenditures include money spent on medical treatments and on the purchase of medicine or health products;[50] educational expenditures include school tuition fees, money spent on the purchase of books and stationery, home tutors, and tutoring-class expenses. The summary statistics are shown in Table 1 (Panel B). There are several things that should be noted. First, the medical and educational expenditures on children constitute a substantial fraction of the family income: educational expenditures alone amount to ¥911.58/year, out of a per capita family income of ¥3,030/year (Table 1, Panel H).[51] Second, there are significant differences between rural and urban households: not only do urban households spend, on average, twice the amount as rural households for all the three types of expenditures, but they also constitute a bigger share of the family income. This suggests that rural families may face a much tighter budget than families in urban areas. Third, parents in urban areas also spend, on average, more time tutoring their children, a statistic which can be partly rationalized by their higher level of education compared to that of the parents in rural areas. Finally, it is interesting that we do not find significant differences by gender.

Child Health As measures of child health $(\theta^{H}_{i,2})$, we use anthropometric indicators (i.e., height, weight, and body mass index (BMI)),[52] general health status, and occurrence of visits to the hospital, which are all reported by both parents.

[50]Grossman (2000) also measures medical care by personal medical expenditures on doctors, dentists, hospital care, prescribed and nonprescribed drugs, nonmedical practitioners, and medical appliances.

[51]Unfortunately, our survey does not contain information on family's the total expenditures.

[52]The height, weight, and BMI are standardized by age and gender on the basis of US growth charts.

The summary statistics in Table 1 (Panel C) show that the height and weight of Chinese child twins are about one standard deviation lower than those of US children of the same age and gender, with the differences being particularly pronounced in rural areas. In contrast, rural children appear less likely to go to the hospital than children in urban areas. This fact may be due to the higher medical costs or the tighter budgets faced by rural households, rather than being a reflection of better health conditions.

Child Academic and Schooling Performance As measures of academic performance $(\theta^C_{i,2})$, we use both objective (exam transcripts) and subjective (self-reported evaluations in comparison with the class norm) measures in two different subjects: Literature and Mathematics.[53] Table 1 (Panel D) shows that urban children, on average, perform better than rural ones, and that girls perform better than boys in Literature. We also analyze several outcomes related to schooling performance, both recorded from transcripts (i.e., grade repetition, good student awards, and awards in contests) and reported by the parents (i.e., whether the parents have been interviewed by the teacher because of the poor performance of the child and whether the child is doing minor actions in class). Note that children in urban areas and girls in general perform better (Table 1, Panel E).

Child Noncognitive Skills Different from the administrative data commonly used in twin-based analysis, our data are also rich in terms of noncognitive measures, which are categorical and reported by both parents (Table 1, Panel F). On the one hand, it is noted that children in urban areas are more likely to be reported by their parents as experiencing greater emotional instability, feeling more lonely, or anxious. On the other hand, girls are reported to have a stronger

[53]Literature and Mathematics are compulsory courses from primary school to high school (from age 6 to 18).

personality than boys.[54]

Parental Labor Supply and Expenditures Finally, we also analyze the effect of an early health shock on parental labor supply, measured as days worked per month, and on parental expenditures. We sum up the expenditures on several goods: cigarettes, alcohol, clothes, and cosmetics. Note that expenditures are separately recorded for both the mother and father. Panel G in Table 1 shows that both parents work longer hours in rural areas, whereas they have higher expenditures in urban areas.

Control Variables We include a rich set of control variables in all our empirical specifications: birth weight, gender, age, birth order,[55] number of siblings, mother's age, mother's years of schooling, per capita family income, binary indicators for household ownership of a washing machine, refrigerator, cell phone, whether the mother has a job in the public sector, and living in a rural area (of course, among these variables only birth weight and the gender dummy when required are included in the twins fixed-effects specifications). The summary statistics are reported in Table 1 (Panel H).

[54] We also analyze the effect of an early health shock on the parent-child relationship, both from the parents' (educational expectations and quality of the relationship) and the child's perspective (openness of the communication and time spent with the parents). As observed in Table 82 in the appendix, there are significant differences between the urban and the rural subsamples, that likely reflect different parenting styles. On the one hand, parents in urban areas have higher expectations regarding the educational achievement of their children. On the other hand, children in urban areas report that they spend more time with their parents.

[55] Interestingly, although we do not find any other evidence of gender discrimination, the proportion of males born at second birth (0.23) is significantly higher than that of females (0.16), and we observe that the mothers of female twins are significantly more educated (9.10) than those of males (8.70). We interpret this finding as evidence that more educated mothers are less likely to practice selective abortion.

5 Results

5.1 Effects of an Early Health Shock on Parental Investments

We first present evidence in support of our identifying assumption of the randomness of the early health shocks. Table 2 (the first two columns) presents both OLS and within-twin-pair FE estimates of the determinants of early health shocks.[56] Clearly, both across and within households, the occurrence of an early health shock is unrelated to birth weight.[57]

We now turn to our main estimation results, starting with the effects of an early health shock on parental investments. Our main finding is that parents adopt a compensatory strategy when deciding how much to invest in health but use a reinforcement strategy with respect to investment in education in response to an early health shock affecting one of the twin children. The estimates are both statistically and economically significant. As shown in Table 2 (column 4), the gap in medical expenditures on average increases by ¥305 in favor of the sick twin, but the gap in educational expenditures increases by ¥186, on average, in favor of the healthy one.

To interpret these findings, we refer to our theoretical model. The key point is that, in our framework with multidimensional child endowments, the compensating or reinforcing nature of investment in health depends on both the price effect and

[56]Table 3 in the appendix presents the full OLS and FE results. It shows that there is a positive correlation between the level of education of the mother and the probability of reporting that the child has suffered from an early health shock. As discussed above, this reporting bias is swept out by the within-twin-pair FE estimator.

[57]However, why would an early health shock uncorrelated with birth weight differentially affect only one of the twins? One plausible explanation is the epigenetic effect: according to the Developmental Origins of Health and Disease (DOHaD) theory, small variations in prenatal experiences may affect the risk of disease in the absence of any effect on birth weight (Godfrey, Gluckman, and Hanson, 2010). In our case, what is critical to our identification strategy is the assumption that the first manifestation of this latent (or epigenetic) effect occurs with the health shock recorded in the data, thereby ruling out any previous parental response. This assumption is supported by the fact that, on average, 60% of the early health shocks affect the child within the first year of life and are not short-term episodes.

the technological effect (equation (31)), whereas that of investment in education is unambiguously determined by the price effect (equation(32)). We first examine the effect of an early health shock on educational expenditures. The result of a reinforcing investment in education suggests that the price effect of an early health shock is negative. This finding implies that efficiency outweighs equality and that ρ is positive in the parental utility function.[58] We then examine the effect on health expenditures. The result of a compensating investment in health reflects the fact that the technological effect (the substitutability between health stock and investment in health) dominates the price effect.[59] Therefore, we observe that parents compensate and reinforce along different dimensions of the child's human capital at the same time.

These results have important implications. In the presence of parental responses, the reduced-form estimates of the effects of an early health shock cannot be purely interpreted as "biological" effects. They constitute either an upper or a lower bound on the true biological effect depending on whether parents adopt a reinforcing or compensatory strategy: this is something that we will not know unless we observe parental behavior. These results are also policy relevant. Parental responses should be taken into account when designing interventions aimed at remediating disadvantage, as parents can exacerbate or annihilate their effects by reallocating resources within the family. Moreover, compared with the within-twin-pair FE estimates, the OLS estimates (also reported in Table 2) systematically underestimate (in absolute value) the effects of early health shocks on parental investments.

[58] Referring to equation (32), the negative estimate in the educational expenditure equation implies that $\partial\pi_i/\partial\theta^H_{i,1}$ is positive because $\partial\theta^H_{i,1}/\partial e^H_{i,1}$ is negative and that $\partial\pi_j/\partial\theta^H_{i,1}$ is opposite to the sign of $\partial\pi_i/\partial\theta^H_{i,1}$. The positive $\partial\pi_i/\partial\theta^H_{i,1}$ implies that ρ is positive and that efficiency outweighs equality when parents make investment decisions.

[59] Referring to equation (31), $\frac{\alpha_H}{\beta_I}\left(\frac{\partial\pi_i}{\partial\theta^H_{i,1}} - \frac{\partial\pi_j}{\partial\theta^H_{i,1}}\right)W$ is positive on the basis of the estimate in the educational expenditure equation. Therefore, the positive estimate in the health expenditures equation implies that $\frac{\alpha_H}{\beta_I}\left(\frac{\partial\pi_i}{\partial\theta^H_{i,1}} - \frac{\partial\pi_j}{\partial\theta^H_{i,1}}\right)W < \frac{\beta_e}{\beta_I}$, suggesting that the price effect is dominated by the technological effect.

The richness of our data allows us to investigate the effects of an early health shock not only on investment in money but also on investment in time. Interestingly, we find that parents spend the same amount of time on both twins,[60] a finding which may reflect the fact that parental time is a non-excludable public good within the household because the parents usually tutor the twins together.[61]

Finally, we find significant differences across subsamples (Table 3). On the one hand, the increase in medical expenditures in favor of the sick twin is not accompanied by a corresponding decrease in educational expenditures in rural areas. We rationalize this finding in light of the fact that the budget is already very tight in rural areas, and thus no changes in educational expenditures are possible. Instead, in urban areas, the extra educational resources allocated in favor of the healthy twin have almost the same monetary value as the amount redistributed to pay for the medical expenses of the sick twin. On the other hand, we also find significant differences by gender. The amount of money reallocated for both medical and educational expenditures in case of female twins is almost twice the amount allocated in case of male twins.

5.2 Effects of an Early Health Shock on Child Outcomes

5.2.1 Child Health

We now examine the effects of an early health shock on child outcomes. We first examine its effects on health in Table 4. Overall, we find some evidence of a long-lasting effect on anthropometric measures. The twin child affected by the early insult is evaluated by the parents as being in worse health and is reported to have a greater occurrence of hospital visits. We now refer to the predictions of our theoretical model (equation (33)) as a guide to interpret the results. Despite the fact that parents have

[60] Note that this question was answered by each twin separately.

[61] Price (2008) shows that most of the variation in the time spent with the child is driven by birth order and maternal employment, which do not vary within twin pairs.

allocated more money as medical expenditures to the sick child, the negative effects of an early health shock are persistent. This finding implies that the direct medical damage (the first term in equation (33)) outweighs the positive intrahousehold resource allocation effect (the second term in equation (33)). Moreover, importantly, in the presence of compensatory parental responses concerning health expenditures,[62] reduced-form estimates *understate* (in absolute value) the biological effect. Given the difficulties of observing all the relevant inputs, we can only say that what we are estimating is a lower bound.

There are also substantial differences across subsamples (Table 4). An early health insult has a consistently negative effect on weight, BMI, and general health status but not on the occurrence of hospital visits in the rural sample. In contrast, an early health shock increases the occurrence of hospital visits in the urban sample and worsens the reported general health status, but it does not have a significantly negative effect on the anthropometric measures. We interpret this evidence by speculating that health shocks may have more long-lasting effects in rural areas where a tighter budget may not allow the parents to go to the hospital for the child to receive the necessary medical care every time it is required. This result has important implications. It suggests that, on the one hand, the negative health effects of an early health shock may be partly offset by compensating investments in families with adequate resources, as our theoretical model predicts. In other words, remediation is possible. On the other hand, the negative effects of an early health shock may persist throughout the life-cycle of children born in poor families because of a tight budget.[63] In the latter case, government subsidies or public health insurance might be crucial policy tools for preventing an early health shock from impairing the child's

[62]The second term on the right-hand side of equation (33) is positive on the basis of our estimates in the health expenditures equation (Table 2).

[63]This is consistent with the evidence reported in Condliffe and Link (2008) for the United States. Note that, in the richer urban areas, both the level of medical expenditures (Table 1) and the money allocated to the sick twin (Table 3) have a larger magnitude than in rural areas.

human capital formation. The gender differences are also noteworthy. Whereas an early health shock has a negative effect on the anthropometric measures only for females, it increases the occurrence of hospital visits only for males. This finding can be interpreted as evidence of greater vulnerability in terms of physical growth for females, and of greater susceptibility to disease for males, given that we do not find any gender differences in the reallocation of medical expenditures.

Finally, it is also interesting to compare the FE and the OLS estimates, which are reported in Table 4 and in Tables 10-12 in the appendix, respectively. We note that, for both the whole sample and the rural sample, OLS estimates underestimate the negative effects of an early health shock. However, for the urban sample, they overestimate them. To interpret these findings, we need to refer to our theoretical model once again. As discussed above, on the one hand, the FE estimator sweeps out the cross-household reporting heterogeneity. On the other hand, the effect of intrahousehold resource allocation is more important in driving the FE than the OLS estimates. Therefore, the difference between the OLS and the FE estimates depends on the relative importance of these factors. As our empirical evidence shows, to the extent the compensation in health via increased medical expenditures is stronger in urban than in rural areas (parents in urban areas allocate, on average, ¥130 more in medical expenditures to the insulted child than parents in rural areas, as shown in Table 3), the OLS estimates will be biased downward in the latter but upward in the former. The conceptual clarification that our theoretical model allows between OLS and FE estimates of the reduced-form effects of an early-life shock on late-life outcomes also has important implications in reconciling the contradictory empirical results present in the literature: although some studies find that, compared with within-family fixed-effects estimates, OLS estimates underestimate the negative effects of early-life health conditions, others find evidence of upward bias.

5.2.2 Other Child Outcomes

We then examine the effects of an early health shock on educational outcomes.[64] Table 5 shows that the twin affected by an early health insult has poorer academic achievement, both perceived and actual. Table 6 shows that an early health insult also negatively affects the twin's schooling performance. Whereas these results come as no surprise, the point that we want to stress here is that we find these negative effects in the presence of parental reinforcing behavior (as noted in Section 5.1). Hence, reduced-form estimates *overstate* (in absolute value) the true biological effect. Given the difficulties of observing all the relevant inputs, we can only say that what we are estimating is an *upper* bound.[65]

We also uncover a significant gender difference. In the case of female twin pairs, the difference in academic achievements between the healthy and sick sisters is only perceived, not real. A significant difference also emerges between the rural and the urban subsamples. Whereas in rural areas we see the effects mainly operating through a problematic behavior in the classroom, in urban areas the long-lasting effects of early-life insults seem to affect mainly purely educational performance. This is consistent with the evidence reported earlier of a reduction in educational expenditures in the urban areas but not in the rural ones. Lastly, Table 7 shows that an early health insult consistently and negatively affects the child's personality in several different domains, with no significant differences between the rural and urban subsamples, but with the girls significantly more affected than the boys.[66]

[64] In this case, we restrict our analysis to 95% of the sample who is still in school.

[65] Referring to equation (34), the second term on the right-hand side of this equation is negative based on our estimates in the educational expenditure equation (Table 2). Therefore, the reduced-form estimates overstate the true biological effect of an early health shock on the child's academic outcomes (the first term on the right-hand side of this equation).

[66] Finally, in Tables 76-81 in the appendix, we report the results on the effects of an early health shock on the relationship between parents and children. From the parental standpoint, parents consistently lower their expectations for the expected educational level of the child affected by the shock, and they also report a worsening relationship between them. The only exception to this pattern occurs in the rural sample, which can be explained in the context of a more traditional type of parent-child relationship, where parents have expectations and children have duties unaffected

Before moving on to the last section, we make several observations about the role that birth weight plays in our analysis, which appears to be more marginal than the pervasive and long-lasting consequences of the early health shock. First, we note that birth weight has an effect on parental investments (Tables 46-51 in the appendix) only in rural areas, where parents allocate more medical expenditures to children lighter at birth (Table 47 in the appendix). Second, among all outcomes studied (Tables 52-81 in the appendix), birth weight has only a strong and negative effect on physical growth (the anthropometric indicators; Tables 52-57 in the appendix), one of the outcomes for which the early health shock has less of an impact, especially for males and in the urban areas. This suggests that, if parents do not respond to the difference in birth weight across twins (e.g., because they do not perceive a difference), these will work only through the biological channel, and birth weight will only be an indicator of physical fitness.[67]

5.3 Effects of an Early Health Shock on Parental Labor Supply and Expenditures

Lastly, we go beyond the within-twin-pair estimation framework to understand how money is reallocated within households by exploiting the richness of our data to investigate the effects of an early health shock on parental work and consumption

by changes in circumstances. From the child standpoint, instead, there is no change in the way the relationship with the parents is perceived compared with the healthy twin under a wide variety of common activities, ranging from playing to having dinner together. This is consistent with our previous result where we find evidence of no change in the time parents spend tutoring the child.

[67]Note that we can recast our findings in light of the recent literature on gene-environment correlation (rGE) and gene-environment interactions (GXE), according to which the observed phenotypic differences among twin pairs are a function of the complex interplay between genetic and environmental factors. Under the interpretation that the early health shock is a manifestation of an epigenetic effect, the differential parental responses can be considered an instance of gene-environment correlation (rGE - genes determine the selection into certain environments; in the current context, they trigger certain parental responses), whereas the phenotypic differences in health and other outcomes can be considered an instance of gene-environment interaction (GXE - parental behavior amplifies or reduces the genetic predisposition). See Conti and Heckman (2010) for a proposed application of gene-environment interaction models to twins data.

patterns.[68] As these characteristics are invariant within twin pairs, we conduct the analysis at the household level. We analyze whether there are differences in parental labor supply and expenditures in case only one twin child is affected by the health shock compared with the case where none of them is.[69] The results are reported in Table 8. We highlight two main findings: in households where one twin child is affected by a health shock, the father is significantly less likely to spend money on goods for himself, whereas the mother is significantly more likely to work. Moreover, we note that these results are driven by different subsamples: mothers are more likely to work longer hours in households located in urban areas and in the presence of sons, whereas fathers are less likely to spend money on goods for themselves in rural areas and in the presence of daughters. These results can be explained in light of the fact that, in families with male twins, expenditures on non-children goods are already reduced to a minimum. This is due to the need for parents to save money to buy housing and stock wealth to help their sons to attract a mate, given the sex ratio imbalance occurring after the implementation of the "One-Child" policy as result of the preferences for sons (Wei and Zhang, 2009). We derive two main implications from these findings. First, we claim that they provide a direct test of the separability assumption between parental consumption and the utility they derive from their children. Although this is a standard assumption adopted in the literature, it is strongly rejected in our data. Second, they imply that the within-twin-pair FE estimates of the effects of an early health shock on parental investments provide only a partial picture because they ignore the reallocation process arising through parental consumption. As such, the within-twin-pair FE estimates understate the overall negative effect of an early health shock at the household level.

[68]See Pitt and Rosenzweig (1990) for an analysis of the effects of child health on intrafamily allocation of time.

[69]We also include the case where both twins are affected by a health shock as a separate category.

6 Conclusions

In this essay, we have studied how early health shocks affect human capital formation. We have first formulated a theoretical model to understand how early health shocks affect child outcomes through parental responses. We have nested a dynamic model of human capability formation into a standard intrahousehold resource allocation framework. By introducing multidimensionality of child endowments, we have allowed parents to compensate and reinforce along different dimensions. We have then tested our main empirical predictions using the CCTS, which contains detailed information on child- and parent-specific expenditures. We have differentiated between investments in money and investments in time. On the one hand, we have found evidence of compensating investment in child health but of reinforcing investment in education. On the other hand, we have found no change in the time spent with the child. We have confirmed that an early health insult negatively affects the child under several different domains, ranging from later health, to cognition, to personality. We have also showed that, in presence of adequate resources, partial remediation may be possible, at least with respect to the child's physical growth. Our findings emphasize the importance of accounting for behavioral responses to early health shocks: parental responses should be taken into account when designing interventions to remediate disadvantage, as parents can exacerbate or annihilate their effects by reallocating resources within the family. They also suggest caution in interpreting reduced-form estimates as purely biological effects. In the presence of asymmetric parental responses under different dimensions of the child's human capital, reduced-form estimates cannot even be unequivocally interpreted as either lower- or upper-bounds of the biological effects.

References

ALMOND, D., K. Y. CHAY, AND D. S. LEE (2005): "The Costs of Low Birth Weight," *Quarterly Journal of Economics*, 120(3), 1031–1083.

ALMOND, D., AND J. CURRIE (2011): "Human capital development before age five," *Handbook of Labor Economics*, 4B, 1315–1486.

BAGO D'UVA, T., E. VAN DOORSLAER, M. LINDEBOOM, AND O. O'DONNELL (2008): "Does reporting heterogeneity bias the measurement of health disparities?," *Health Economics*, 17, 351–375.

BECKER, G., AND N. TOMES (1976): "Child endowments and the quantity and quality of children," *The Journal of Political Economy*, 84(4), 143–162.

BEHRMAN, J. (1988): "Intrahousehold allocation of nutrients in rural India: Are boys favored? Do parents exhibit inequality aversion?," *Oxford Economic Papers*, 40(1), 32–54.

BEHRMAN, J., AND V. LAVY (1997): "Child health and schooling achievement: association, causality and household allocations," *CARESS Working Papres.*

BEHRMAN, J., R. POLLAK, AND P. TAUBMAN (1982): "Parental preferences and provision for progeny," *The Journal of Political Economy*, 90(1), 52–73.

BEHRMAN, J., AND M. ROSENZWEIG (2004): "Returns to birthweight," *Review of Economics and Statistics*, 86(2), 586–601.

BEHRMAN, J., M. ROSENZWEIG, AND P. TAUBMAN (1994): "Endowments and the allocation of schooling in the family and in the marriage market: the twins experiment," *The Journal of Political Economy*, 102(6), 1131–1174.

Black, S., P. Devereux, and K. Salvanes (2007): "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes*," *The Quarterly Journal of Economics*, 122(1), 409–439.

Brown, P., A. de Brauw, and Y. Du (2009): "Understanding Variation in the Design of China's New Co-operative Medical System," *The China Quarterly*, 198, 304–329.

Case, A., A. Fertig, and C. Paxson (2005): "The Lasting Impact of Childhood Health and Circumstance," *Journal of Health Economics*, 24(2), 365–389.

Case, A., and C. Paxson (2002): "Parental Behavior And Child Health," *Health Affairs*, 21(2), 164–178.

Condliffe, S., and C. Link (2008): "The relationship between economic status and child health: Evidence from the United States," *The American Economic Review*, 98(4), 1605–1618.

Conti, G., and J. Heckman (2010): "Understanding the Early Origins of the Education-Health Gradient: A Framework That Can Also Be Applied to Analyze Gene-Environment Interactions," *Perspectives on Psychological Science*, 5, 585–605.

Cunha, F., and J. Heckman (2008): "Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation," *Journal of Human Resources*, 43(4), 738–782.

Currie, J., and M. Stabile (2006): "Child mental health and human capital accumulation: the case of ADHD," *Journal of Health Economics*, 25(6), 1094–1118.

Currie, J., M. Stabile, P. Manivong, and L. Roos (2010): "Child Health and Young Adult Outcomes," *Journal of Human Resources*, 45(3), 517.

Family Planning Commission of Yunnan Province (2003): "Regulations on Population and Family Planning in Yunnan Province," *China Population Press.*

Glewwe, P., and E. Miguel (2007): "The impact of child health and nutrition on education in less developed countries," *Handbook of development economics*, 4, 3561–3606.

Godfrey, K., P. Gluckman, and M. Hanson (2010): "Developmental origins of metabolic disease: life course and intergenerational perspectives," *Trends in Endocrinology and Metabolism.*

Grantham-McGregor, S., Y. B. Cheung, S. Cueto, P. Glewwe, L. Richter, B. Strupp, and The International Child Development Steering Group (2007): "Developmental Potential in the First 5 Years for Children in Developing Countries," *The Lancet*, 369, 60–70, Series, Child development in developing countries.

Grossman, M. (1972): "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy*, 80(2), 223–255.

——— (2000): "The Human Capital Model," in *Handbook of Health Economics*, ed. by A. J. Culyer, and J. P. Newhouse, vol. 1, pp. 347–408. Elsevier, Amsterdam.

Heckman, J. J. (2007): "The Economics, Technology and Neuroscience of Human Capability Formation," *Proceedings of the National Academy of Sciences*, 104(3), 13250–13255.

Liu, Y., K. Rao, J. Wu, and E. Gakidou (2008): "China's health system performance," *The Lancet*, 372(9653), 1914–1923.

Pitt, M., and M. Rosenzweig (1990): "Estimating the Intrahousehold Incidence of Illness: Child Health and Gender-Inequality in the Allocation of Time," *International Economic Review*, 31, 969–989.

Pitt, M., M. Rosenzweig, and M. Hassan (1990): "Productivity, health, and inequality in the intrahousehold distribution of food in low-income countries," *The American Economic Review*, 80(5), 1139–1156.

Price, J. (2008): "Parent-Child Quality Time: Does Birth Order Matter?," *Journal of Human Resources*, 43(1), 26.

Rosenzweig, M., and K. Wolpin (1988): "Heterogeneity, intrafamily distribution, and child health," *The Journal of Human Resources*, 23(4), 437–461.

Rosenzweig, M., and J. Zhang (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.

Smith, J. (2009): "The impact of childhood health on adult labor market outcomes," *The Review of Economics and Statistics*, 91(3), 478–489.

Strauss, J., and D. Thomas (1998): "Health, nutrition, and economic development," *Journal of Economic Literature*, 36(2), 766–817.

Todd, P., and K. Wolpin (2007): "The production of cognitive achievement in children: Home, school, and racial test score gaps," *Journal of Human capital*, 1(1), 91–136.

Wei, S., and X. Zhang (2009): "The competitive saving motive: evidence from rising sex ratios and savings rates in China," *NBER working paper*.

West, L., and Y. Zhao (2000): *Rural labor flows in China*. Univ of California Inst of East.

Yusuf, S., and T. Saich (2008): *China urbanizes: consequences, strategies, and policies.* World Bank Publications.

Table 1: Summary Statistics

	Whole	Rural	Urban	Difference	Male	Female	Difference
A. Main Independent Variables							
Early health shocks (dummy)	0.09	0.08	0.10	-0.02**	0.11	0.07	0.04***
Birth weight (kg)	2.46	2.48	2.43	0.04***	2.50	2.37	0.13***
Birth weight: <2 (dummy)	0.13	0.11	0.14	-0.03**	0.10	0.18	-0.08***
Birth weight: 2-2.5 (dummy)	0.36	0.35	0.37	-0.02	0.35	0.38	-0.03
Birth weight: 2.5-3 (dummy)	0.38	0.38	0.37	0.01	0.39	0.34	0.06***
Birth weight: 3- (dummy)	0.14	0.16	0.12	0.03***	0.16	0.10	0.05***
B. Investments							
Medical expenditures (¥/year)	225.76	150.16	310.69	-160.53***	238.32	254.70	-16.38
Education expenditures (¥/year)	911.58	630.41	1227.50	-597.09***	909.44	981.11	-71.67
Clothing expenditures (¥/year)	241.95	173.24	319.16	-145.92***	242.84	259.27	-16.43
Parents home tutor (minutes/day)	19.37	14.93	24.30	-9.36***	20.52	19.41	1.12
C. Health							
Height (cm)	137.35	133.37	141.83	-8.46***	137.74	136.84	0.90
Weight (kg)	33.60	32.07	35.30	-3.23***	34.26	32.87	1.39***
BMI	17.38	17.66	17.08	0.58***	17.57	17.20	0.37***
Height-for-age z-score	-1.05	-1.47	-0.59	-0.88***	-1.07	-1.01	-0.06
Weight-for-age z-score	-0.87	-1.03	-0.69	-0.34***	-0.79	-0.92	0.13***
BMI-for-age z-score	-0.31	-0.15	-0.50	0.35***	-0.19	-0.43	0.24***
Self reported general health status (GHS) (1: worst; 4: best)	2.92	2.91	2.93	-0.02	2.91	2.94	-0.02
Hospital visits (dummy)	0.33	0.30	0.36	-0.06***	0.33	0.34	-0.01
D. Academic performance							
Relative measures [1: lowest; 5: highest one-fifth in class]							
Literature	3.53	3.47	3.61	-0.14***	3.44	3.65	-0.21***
Mathematics	3.48	3.45	3.51	-0.06*	3.48	3.51	-0.03
Marks [1-100]							
Literature	81.95	78.89	85.30	-6.41***	81.18	83.59	-2.41***
Mathematics	80.89	78.54	83.48	-4.95***	81.46	81.39	0.07

The variables in Panel A, the expenditures in Panel B and the health measures in Panel C are reported by both parents; the amount of time the parents spend tutoring the child in Panel B and the relative measures of academic performance in Panel D are reported by each twin; the marks in Panel D are recorded from examination transcripts in the term before the survey. The "Difference" columns report the results of the t-test for the difference in the means in the two previous columns. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table 1: Summary Statistics (ctd.)

	Whole	Rural	Urban	Difference	Male	Female	Difference
E. Schooling performance							
Good Student Awards (dummy)	0.24	0.18	0.31	-0.13***	0.21	0.29	-0.08***
Awards in contests (dummy)	0.07	0.03	0.11	-0.09***	0.07	0.08	-0.01
Grade repetition for poor performance (dummy)	0.04	0.06	0.03	0.02***	0.05	0.04	0.01
Parents interviewed for poor performance (dummy)	0.13	0.08	0.18	-0.10***	0.18	0.08	0.10***
Child doing minor actions in class	1.73	1.76	1.70	0.06**	1.89	1.59	0.30***
measured as 1: never; 2: seldom; 3: sometimes; 4: always							
F. Noncognitive skills							
Measured as 1: disagree; 2: agree; 3: strongly agree							
Always feel lonely	1.20	1.17	1.23	-0.05***	1.20	1.21	-0.01
Always feel anxious or fretful	1.28	1.26	1.31	-0.06***	1.38	1.23	0.15***
Easily distracted	1.59	1.52	1.66	-0.14***	1.66	1.55	0.11***
Always careless	1.94	1.96	1.91	0.05**	2.01	1.91	0.10***
Easily frightened	1.36	1.36	1.37	0.00	1.34	1.43	-0.09***
Emotionally instable	1.12	1.06	1.18	-0.12***	1.13	1.14	-0.01

The first three measures of schooling performance in Panel E are recorded from the official transcripts; the last two measures in the same Panel, and all the noncognitive skills measures in Panel F are reported by parents. The "Difference" columns report the results of the t-test for the difference in the means in the two previous columns. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table 1: Summary Statistics (ctd.)

	Whole	Rural	Urban	Difference	Male	Female	Difference
G. Parental Labor Supply and Expenditures							
Father							
Labor supply (days/month)	25.50	26.18	24.65	1.53***	25.87	24.96	0.91***
Expenditures (¥/past 6 months)	700.38	536.58	886.55	-349.97***	686.88	725.62	-38.74***
Mother							
Labor supply (days/month)	25.32	26.03	24.34	1.69***	25.40	24.94	0.46**
Expenditures (¥/past 6 months)	283.93	169.66	412.33	-242.67***	261.83	348.22	-86.39***
H. Control variables							
Age	11.19	10.97	11.43	-0.46***	11.15	11.19	-0.05
Birth order (1: born at the second birth; 0: born at the first birth)	0.20	0.33	0.07	0.26***	0.23	0.16	0.06***
# siblings	2.24	2.39	2.08	0.30***	2.26	2.21	0.04**
Mother's age	36.84	36.49	37.24	-0.75***	36.55	36.53	0.01
Mother's schooling years	8.67	7.38	10.11	-2.73***	8.70	9.10	-0.40***
Per capital family income (¥1,000/year 2002)	3.03	2.46	3.68	-1.22***	3.10	3.04	0.05
Own washing machine (dummy)	0.65	0.47	0.85	-0.38***	0.66	0.69	-0.03
Own refrigerator (dummy)	0.38	0.14	0.65	-0.51***	0.39	0.41	-0.02
Own cell phone (dummy)	0.36	0.19	0.56	-0.37***	0.38	0.38	0.00
Mother's occupation (dummy: 1= Public sector)	0.08	0.02	0.15	-0.13***	0.09	0.10	-0.01
Rural indicator	0.53	1.00	0.00	0.00.	0.52	0.50	0.01
# Observations	2922	1326	1546		1082	1120	
# Pairs of twins	1461	688	773		541	560	

The variables in Panel H are reported by parents. The "Difference" columns report the results of the t-test for the difference in the means in the two previous columns. * Significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2: OLS and Within-Twin-Pair Fixed-Effects Estimates of the Determinants of Early Health Shocks and the Effects of Early Health Shocks on Parental Investments (Whole Sample)

	Dependent variables									
	Early health shock		Health		Education		Clothing		Parents home tutor	
	OLS	FE	OLS	FE	OLS	FE	OLS	FE	OLS	FE
Early health shocks			1.054***	1.351***	0.021	-0.204***	0.030	-0.058	0.784	-1.493
			[0.153]	[0.314]	[0.056]	[0.073]	[0.086]	[0.042]	[1.644]	[1.263]
Birth weight (kg): <2	0.040	0.010	0.440**	0.521**	-0.076	-0.015	-0.079	0.006	0.004	0.255
	[0.026]	[0.026]	[0.189]	[0.210]	[0.078]	[0.047]	[0.113]	[0.036]	[1.694]	[1.411]
Birth weight (kg): 2-2.5	0.010	0.000	0.255*	0.468***	0.054	0.016	-0.016	0.027	-1.860	0.257
	[0.019]	[0.021]	[0.150]	[0.163]	[0.060]	[0.026]	[0.090]	[0.030]	[1.315]	[1.194]
Birth weight (kg): 2.5-3	-0.010	-0.020	0.199	0.421***	-0.046	-0.013	-0.052	-0.004	-0.955	-0.319
	[0.018]	[0.018]	[0.143]	[0.149]	[0.059]	[0.023]	[0.087]	[0.020]	[1.243]	[0.993]
Gender (boy=1)	0.040***	0.020**	0.233**	0.086	-0.036	-0.024	-0.023	-0.028	1.394*	0.235
	[0.013]	[0.010]	[0.094]	[0.093]	[0.038]	[0.021]	[0.054]	[0.017]	[0.833]	[0.779]
Δ Expenditure			237.951	305.002	19.143	-185.962	7.259	-14.033		
# Observations	2922	2922	2922	2922	2922	2922	2902	2902	2922	2922
# pairs of twins		1461		1461		1461		1451		1461

Notes: Each entry comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. The dependent variables of health, education, and clothing expenditures are in log form. The row "Δ Expenditure" reports the implied change in the level of expenditure.

Table 3: **Within-Twin-Pair Fixed-Effects Estimates of the Effects of Early Health Shocks on Parental Investments (Subsamples)**

	Health	Education	Clothing	Parents home tutor
Rural Sample				
Early health shocks	1.523***	-0.058	-0.120	-2.041
	[0.538]	[0.069]	[0.092]	[1.966]
Δ Expenditure	228.694	-36.564	-20.789	
# pairs of twins	773	773	773	764
Urban Sample				
Early health shocks	1.149***	-0.328***	-0.018	-0.962
	[0.374]	[0.116]	[0.021]	[1.678]
Δ Expenditure	356.983	-402.620	-5.745	
# pairs of twins	688	688	688	687
Male Sample				
Early health shocks	1.085**	-0.171**	-0.108	-2.393
	[0.426]	[0.074]	[0.091]	[2.565]
Δ Expenditure	258.577	-155.514	-26.227	
# pairs of twins	541	541	541	539
Female Sample				
Early health shocks	2.080***	-0.410**	-0.028	0.868
	[0.708]	[0.188]	[0.030]	[0.628]
Δ Expenditure	529.776	-402.255	-7.260	
# pairs of twins	560	560	560	556

Notes: Each entry comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Birth weight is controlled for in each regression; gender has been controlled for in the estimations based on the rural and urban samples. The dependent variables of health, education, and clothing expenditures are in log form. The row "Δ Expenditure" reports the implied change in the level of expenditure.

Table 4: Within-Twin-Pair Fixed-Effects Estimates of the Effects of Early Health Shocks on Health

	Height (z-score)	Weight (z-score)	BMI (z-score)	Health Status	Hospital Visits
Whole Sample					
Early health shocks	-0.005	-0.270**	-0.204*	-0.449***	0.163***
	[0.102]	[0.115]	[0.121]	[0.111]	[0.051]
# pairs of twins	1423	1435	1411	1455	1451
Rural Sample					
Early health shocks	-0.114	-0.494***	-0.418***	-0.522***	0.085
	[0.165]	[0.154]	[0.121]	[0.164]	[0.060]
# pairs of twins	740	755	745	773	771
Urban Sample					
Early health shocks	0.079	-0.091	-0.040	-0.409***	0.221***
	[0.127]	[0.158]	[0.184]	[0.151]	[0.076]
# pairs of twins	683	680	666	682	680
Male Sample					
Early health shocks	0.038	-0.073	0.102	-0.441**	0.205**
	[0.100]	[0.186]	[0.157]	[0.186]	[0.084]
# pairs of twins	524	527	519	538	535
Female Sample					
Early health shocks	0.107	-0.251**	-0.355**	-0.327*	0.115
	[0.154]	[0.122]	[0.159]	[0.191]	[0.081]
# pairs of twins	548	553	544	558	558

Notes: Each entry comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Birth weight is controlled for in each regression; gender has been controlled for in estimations based on whole, rural, and urban samples.

Table 5: **Within-Twin-Pair Fixed-Effects Estimates of the Effects of Early Health Shocks on Academic Achievement**

	Literature (self-reported)	Mathematics (self-reported)	Literature (exam record)	Mathematics (exam record)
Whole Sample				
Early health shocks	-0.345**	-0.529***	-5.158***	-5.384**
	[0.150]	[0.147]	[1.659]	[2.659]
# pairs of twins	1431	1425	1362	1343
Rural Sample				
Early health shocks	-0.186	-0.598***	-5.441**	-2.604
	[0.160]	[0.168]	[2.572]	[2.666]
# pairs of twins	759	757	711	705
Urban Sample				
Early health shocks	-0.511**	-0.510**	-4.632**	-7.330*
	[0.236]	[0.227]	[2.214]	[4.279]
# pairs of twins	672	668	651	638
Male Sample				
Early health shocks	-0.401*	-0.653***	-5.630**	-6.272*
	[0.242]	[0.229]	[2.527]	[3.718]
# pairs of twins	527	523	507	500
Female Sample				
Early health shocks	-0.461*	-0.457*	-2.020	-3.776
	[0.241]	[0.270]	[1.940]	[5.815]
# pairs of twins	551	549	519	511

Notes: Each entry comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Birth weight is controlled for in each regression; gender has been controlled for in estimations based on whole, rural, and urban samples.

Table 6: **Within-Twin-Pair Fixed-Effects Estimates of the Effects of Early Health Shocks on Schooling Performance**

	Good Student Awards	Awards in Contests	Grade repetition	Parents interviewed	Minor actions in class
Whole Sample					
Early health shocks	-0.185***	-0.087**	0.109***	0.100*	0.279**
	[0.069]	[0.039]	[0.042]	[0.061]	[0.131]
# pairs of twins	1461	1461	1461	1459	1445
Rural Sample					
Early health shocks	-0.096	0.006	0.083	-0.039	0.396*
	[0.086]	[0.006]	[0.055]	[0.042]	[0.233]
# pairs of twins	773	773	773	771	766
Urban Sample					
Early health shocks	-0.236**	-0.163**	0.125**	0.197**	0.169
	[0.105]	[0.068]	[0.062]	[0.100]	[0.143]
# pairs of twins	688	688	688	688	679
Male Sample					
Early health shocks	-0.252**	-0.131*	0.127*	0.045	0.203
	[0.103]	[0.068]	[0.068]	[0.108]	[0.182]
# pairs of twins	541	541	541	541	532
Female Sample					
Early health shocks	-0.223*	-0.054	0.109	0.111	0.282
	[0.129]	[0.056]	[0.074]	[0.075]	[0.207]
# pairs of twins	560	560	560	559	554

Notes: Each entry comes from a separate regression. Robust standard errors are in brackets: * significant at 10%; ** significant at 5%; *** significant at 1%. Birth weight is controlled for in each regression: gender has been controlled for in estimations based on whole, rural, and urban samples.

Table 7: **Within-Twin-Pair Fixed-Effects Estimates of the Effects of Early Health Shocks on Noncognitive Skills**

	Feel lonely	Anxious	Easily distracted	Careless	Easily Frightened	Emotional instability
Whole Sample						
Early health shocks	0.165***	0.121***	0.151***	0.142***	0.144***	0.112**
	[0.062]	[0.045]	[0.055]	[0.053]	[0.048]	[0.051]
# pairs of twins	1461	1461	1461	1461	1461	1455
Rural Sample						
Early health shocks	0.253***	0.071	0.148**	0.165**	0.092	0.044
	[0.089]	[0.056]	[0.071]	[0.075]	[0.057]	[0.074]
# pairs of twins	773	773	773	773	773	769
Urban Sample						
Early health shocks	0.099	0.160**	0.154*	0.125*	0.191***	0.163**
	[0.084]	[0.068]	[0.082]	[0.074]	[0.073]	[0.069]
# pairs of twins	688	688	688	688	688	686
Male Sample						
Early health shocks	0.040	0.093	0.036	0.035	0.117*	0.047
	[0.092]	[0.057]	[0.070]	[0.070]	[0.071]	[0.074]
# pairs of twins	541	541	541	541	541	539
Female Sample						
Early health shocks	0.394***	0.228**	0.279***	0.219**	0.282***	0.238**
	[0.113]	[0.099]	[0.105]	[0.098]	[0.106]	[0.103]
# pairs of twins	560	560	560	560	560	557

Notes: Each entry comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Birth weight is controlled for in each regression; gender has been controlled for in estimations based on whole, rural, and urban samples.

Table 8: OLS Estimates of the Early Health Shocks on Parental Labor Supply and Expenditures

	Father		Mother	
	Work	Expenditure	Work	Expenditure
Whole Sample				
Early health shock (only one child)	-0.044	-140.393*	0.046*	-4.554
	[0.078]	[79.028]	[0.027]	[65.095]
# Observations	1163	1423	1048	1442
R-squared	0.017	0.200	0.026	0.207
Rural Sample				
Early health shock (only one child)	0.004	-133.164*	0.006	-12.947
	[0.050]	[70.516]	[0.040]	[46.779]
# Observations	646	757	608	763
R-squared	0.004	0.206	0.002	0.109
Urban Sample				
Early health shock (only one child)	-0.089	-140.071	0.087***	-5.265
	[0.146]	[128.382]	[0.033]	[109.298]
# Observations	517	666	440	679
R-squared	0.014	0.156	0.032	0.289
Male Sample				
Early health shock (only one child)	0.005	33.393	0.072*	74.990
	[0.045]	[134.589]	[0.039]	[116.180]
# Observations	417	524	386	536
R-squared	0.017	0.218	0.051	0.332
Female Sample				
Early health shock (only one child)	-0.122	-379.628***	-0.011	-106.805
	[0.217]	[99.651]	[0.048]	[101.787]
# Observations	452	545	405	553
R-squared	0.033	0.187	0.025	0.182

Notes: Each entry comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Child's age, mother's years of schooling, and per capita family income are included as controls in each specification; rural has also been controlled for in estimations based on whole, male, and female samples.

Essay Two

Education and Preferences: Experimental Evidences from Chinese Adult Twins[1]

[1] This essay is largely based on an on-going joint research project with Soo Hong Chew, James Heckman, Songfa Zhong, and Junsen Zhang. I have been the main contributor to the work so far.

"This also takes care of the matter of whether those questioned would "correct" their behavior if it were pointed out to them that they "act" in violation of the expected utility hypothesis. That theory, as formulated by the von Neumann-Morgenstern axioms, is normative in the sense that the theory is "absolutely convincing" which implies that men will act accordingly. If they deviate from the theory, an explanation of the theory and of their deviation will cause them to readjust their behavior. This is similar to the man who tries to build a perpetuum mobile and then is shown that this will never be possible. Hence, on understanding the underlying physical theory, he will give up the vain effort."

------ Oskar Morgenstern, 1979

1 Introduction

At the heart of economic analysis is how we make decisions ranging from decision making under risk and uncertainty to decision making involving time. The behavioral economics revolution of the past decade reflects the rise in influence of psychological considerations in how economists model decision making behavior following the Allais paradox (1953) for decision making under risk and the Ellsberg paradox (1961) for decision making under uncertainty. This has led to an active literature in non-expected utility models (see, e.g., Starmer, 2000, for a review) including the highly influential prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). There is a parallel literature on decision making anomalies involving time, e.g., temporal discounting (Laibson, 1997), incidence of consumption (Lowenstein, 1987), and the timing of uncertainty resolution (Kreps and Porteus, 1989). There is a tendency in literatures to refer to departures from the classical model as representing decision making biases which, as with the opening quote from Oskar Morgenstern (1979), can potentially be rectified via human capital investment, i.e., education or perhaps re-education.

We adopt the view that a decision making bias refers to behavioral anomalies that are robust with respect to people being cognizant of them, rather than transitory ones which would generally not prevail with respect to full awareness. Savage (1954) argued that increased understanding ought to increase the frequency of the "truly" normative response; that preferences that initially contradict some normative principle may not survive thorough deliberation (what he termed "reflective equilibrium"). A related question is whether the incidence of behavioral anomalies also reflects cognitive ability or bias. The research reported in this essay contributes to understanding how preference and bias may be related through education. Among factors affecting preference formation, education appears especially important given that we learn and develop different ways of thinking and acting besides being trained to acquire professional skills.

The Study of the relationship between education and preference also directly relates to the literature on the determination or formation of preference. Stable preferences, together with maximizing behavior and market equilibrium, have once been regarded as the fundamental trinity assumptions, which establish the tractability of the analytical framework in economics. Although it is reasonable to assume that basic preferences do not change rapidly over a short period of time, they may change gradually over an extended duration. From a life-cycle perspective, it makes sense to treat preferences as being endogenous rather than exogenous.[2] Previous studies have analyzed theoretically the preference formation and preference change processes, and demonstrated their importance (e.g., Becker, 1992, 1996; Becker and Mulligan, 1997). Theories have been proposed about endogenous determination of preferences including wealth (Becker and Mulligan, 1997), market institutions (Bowles 1998), and culture (Bisin and Verdier, 2000). Recent empirical studies based on experimental data find that cognitive ability is associated with risk attitude and time discounting (Dohmen, et al., 2010; Benjamin et al., 2006; Burks et al., 2009), and non-cognitive ability and gender are related to attitude toward risk and ambiguity (Borghans et al., 2009). In a literature review by Croson and Gneezy (2009), they concluded that there are robust gender differences in preference in general. There is also a recent literature demonstrating the heritability of preference using incentivized choice (Cesarini et al., 2009; Zhong et al., 2009a).

There is almost no systematic and rigorous study addressing the causal relationship between education and preferences in the literature. An ordinary least squares estimation of the effect of education on preferences cannot address causality, because unobservable family background and individual heterogeneity may simultaneously affect educational outcomes and preference formation. In other words, education may be correlated with unobservable family background and the effects of endowment, which would render any correlation between education and preference spurious.[3] Due to the difficulty in breaking the endogeneity that results from omitted variables, the casual relationship between education and preferences remains elusive.

The primary goal in this essay is to empirically identify the causal effect of education on two dimensions of preference - decision making under risk and uncertainty and decision making involving time. We conduct a number of incentivized choice experiments on adult twins to measure preferences. We then use a within-twin-pair fixed-effects estimator to carry out the identification. As is argued

[2] Stigler and Becker (1977) assumed that preferences are treated to be fixed and exogenous across individuals. Yet the more recent work by Becker (1992, 1996) and Becker and Mulligan (1997) rejects the assumption of stable preferences.

[3] The difficulty in identifying the causal relationship between education and preferences is similar to that of the estimation of economic returns to education. See, e.g., Card (1999), for a review of the econometric issues in estimating returns to education.

in the literature (Ashenfelter and Krueger, 1994; Behrman et al., 1996; Behrman and Rosenzweig, 1999), twins have a similar family background, and monozygotic (from the same egg) twins are genetically identical.[4] The effects of unobserved family background and genetic endowment should be similar for both twins. Thus, taking the within-twin-pair difference will, to a great extent, reduces the unobservable family background and individual endowment effects that could cause bias in the ordinal least squares estimation. Intuitively, by comparing experimentally measured preference of twins with different educational attainments, we gain more confidence that the correlation that we observe between education and preference is not due to a correlation between education and family background or an individual's endowment.

Our within-twin-pair fixed-effects estimates, based on the experimental data on adult twin pairs, indicate that education affects decision making involving both risk and uncertainty and involving time. We find that a higher level of education tends to reduce the degree of risk aversion toward moderate prospects, moderate hazards, and longshot prospects. In terms of decision making anomalies, university educated subjects exhibit significantly more Allais type behavior compared to pre-high school subjects, while high school educated subjects also exhibit more ambiguity aversion as well as familiarity bias relative to pre-high school subjects. For decision making involving time, a higher level of education tends to reduce the degree of impatience, hyperbolic discounting, dread, hopefulness, except for anxiousness whose incidence is not sensitive to educational attainment. The estimation results are robust in a series of sensitivity analysis when we (i) use the instrumental variables (IV) method to take care of potential measurement errors; and (ii) control for birth weight and restrict the estimation sample to include only monozygotic (MZ) twins to take care of possible biases arising from omitted variables. In summary, our experimental evidences from Chinese adult twins suggest that people with a higher level of education tends to exhibit more "biased" preference in risk attitude and less "biased" preference regarding time.

At present, the relationship between individual's demographic, social and economic characteristics and experimental measured preferences have been

[4] Gorseline (1932) seems to be the first attempt to look at sibling data in economics. Not content with sibling data, Behrman and Taubman (1976), Taubman (1976a, 1976b), and Behrman et al. (1977) began to use twin data in the 1970s. Todd and Wolpin (2003) clarify different identification assumptions between within-sibling and within-twin-pair fixed-effects estimator. They conclude that within-twin-pair fixed-effects estimator needs much weaker identification assumptions than within-sibling fixed-effects estimator.

increasingly noticed to be important in economics (Borghans et al., 2009, Dohmen et al., 2010).[5] However, most of these studies are showing a correlation rather than a causal relationship.[6] Combining survey data, experimental data, and econometric methods, this essay appears to be the first study addressing the causal relationship between education and preference.

This essay also contributes to the literature on human capital. Integrating the recent development in neuroscience, psychology, and behavioral science, the boundary of human capital theory has been substantially outspreading in recent years (see, e.g., Cunha, et al., 2006; Rutter, 2006; Cunha and Heckman, 2007, 2008; Heckman, 2007; ter Weel, 2008; Cunha et al., 2010). Non-cognitive skills, such as personality traits in the terminology of psychologists, for example, have been widely accepted as an important dimension of human capital (Heckman and Rubinstein, 2001; Heckman, Stixrud and Urzua, 2006; Borghans et al., 2008). The within-twin-pair fixed-effects estimates of education and preference add a new piece to this strand of an emerging literature. Should preferences be shaped by education, they should be treated as endogenous in the human capital theory. Our results thus suggest future research to explore the mechanism underlying the relationship/interaction between human capital formation and preference formation.

Finally, the estimation of the return to education has been one of the major subjects in economics for several decades (Card, 1999). Yet, there have been limited studies to explore the mechanism underlying the relationship between education and socioeconomic success. Our identified relationships between educational attainments and decision making under risk and uncertainty and involving time will help our understanding of the pecuniary return to education.

The essay proceeds as follows. Section 2 reviews the behavioral concepts about the two dimensions of preference - decision making under risk and uncertainty and decision making involving time, and empirical evidences about their determinants. Section 3 describes the Chinese adult twin data and the experimental design. Section 4 specifies our empirical methodology. Section 5 presents the estimation results. Section 6 conducts a robustness analysis. Section 7 discusses and concludes.

---

[5] The next section gives a literature survey.

[6] A notable exception is Benjamin et al. (2010). They analyze the effect of social identity on both time preference and risk aversion by adopting a method from social phycology to introduce exogenous variation in identity effects.

2 Review of Behavioral Decision Models

This section reviews a number of choice models and related empirical studies on decision making involving risk and uncertainty as well as decision making involving time.

2.1 Decision Making under Risk and Uncertainty

Attitudes toward economic risk are one of the primitives in economics. They underpin a wide range of behavior, including portfolio choice and insurance purchase, which have significant economic consequences. Not surprisingly, risk is the most well studied preference in economics. Center to decision making under risk is the question how people would evaluate a gamble. The pioneering contributions of von Neumann and Morgenstern (1944) and Savage (1954) provided an axiomatic foundation of the expected utility model, where the value of a gamble equals the mean of utility of monetary outcomes. Under this framework, the expected utility (EU) of a lottery $(x_1,\ldots,x_n;p_1,\ldots,p_n)$ is given by:

$$EU(x_1,\ldots,x_n;p_1,\ldots,p_n)=\sum_{i=1}^{n}p_i u(x_i)$$

Despite the fact that the EU theory provides a convenient analytical tool, it is unsatisfactory in many aspects. First, the EU model is not able to account for the fourfold pattern of risk attitudes. In their 1979 seminal paper on prospect theory (PT) and the rejoinder in 1992, Kahneman and Tversky posited the notion of status quo relative to which gains and losses are defined. Risks are referred to as prospects (hazards) when they are oriented toward gains (losses). Risks can be further distinguished between those whose contingencies have moderate probabilities and those whose contingencies are highly unlikely or have longshot probabilities. For instance, insurance and state lotteries represent longshot hazards and longshot prospects while financial assets may be viewed as moderate prospects. Market evidence points to the prevalence of risk aversion toward longshot hazards and, to some extent, risk tolerance for longshot prospects. Kahneman and Tversky also pointed out the little reported tendency for people to be risk tolerant when it comes to moderate hazards, e.g., when people find themselves in insecure or unsafe situations prompting them to take a chance (Kunretither and Ginsberg, 1978). This is the so called fourfold pattern of risk attitudes – risk averse (preferring/tolerance) toward moderate prospects (hazards) and risk averse (preferring/tolerant) toward longshot prospects (hazards). Specifically, the fourfold pattern of risk attitudes consists of comparing a lottery of the form of a q chance of receiving an outcome x and a $1-q$ chance of receiving zero, denoted by (x,q), versus receiving its expected value, xq, for sure. PT assumes a *loss-averse* value function v that is concave over gains, convex over losses, and vanishes at the status quo, represented by zero. The implication of adopting such a loss-averse value function for an EU decision maker is immediate. For positive (negative) x, the EU of receiving the lottery (x,q) is given by $qv(x)$, which is always less (greater) than the utility $v(qx)$ of receiving its expected value qx for sure. Such an EU decision maker would be risk averse for all prospects and risk preferring for all hazards, leaving it unable to concurrently exhibit risk preference for longshot prospects and risk aversion for longshot hazards.

Second, the independence axiom of the EU theory had been challenged soon after it was propounded, in particular, by the Allais paradox (1953).[7] In Allais paradox, subjects are presented with two pairs of lotteries, consisting of bets on a random draw from 100 cards numbered from 1 to 100. In the first pair, Option A_1 is a sure thing, that is, you receive \$1M for sure; Option B_1 is: if you draw a card from 1 to 89, you receive \$1M; if you draw a card numbered 90 you receive zero; if you draw a card from 91 to 100, you receive \$5M. In the second pair, Option A_2 is: if you draw a card from 1 to 89, you receive zero; if you draw a card from 90 to 100, you receive \$1M; lottery B_2 is: if you draw a card from 1 to 90, you receive zero; if you draw a card from 91 to 100, you receive \$5M. Notice that both pairs of options share the same outcome 89% of the time. Under expected utility, these common outcomes would have no effect on the relative desirability of the A and the B options,, so that a preference of A over B in one pair implies such a preference for the other pair. Yet, it is often observed that people prefer A1 over B1 but prefer B2 over A2.

Third, there are other kinds of anomalies such as ambiguity aversion and familiarity bias, which cannot be accounted for by the EU theory. Ambiguity aversion was first suggested by Keynes (1921) in his Treatise on Probability in which he stated– "If two probabilities are equal in degree, ought we, in choosing our course of action, to prefer that one which is based on a greater body of knowledge?" He

---

[7] The independence axiom is a key characteristic of this model, which follows from the additive structure of EU. It means that for the strict preference relation f over any lotteries F and G, F f G if and only if $\alpha F + (1-\alpha)H$ f $\alpha G + (1-\alpha)H$, for any probability α in $(0, 1)$ and any lottery H.

illustrated this observation with an example of two urns, one containing fifty black balls and fifty red balls while another contains one hundred balls of either color. This example reappeared in Ellsberg (1961) which observed that people tend to be ambiguity averse in preferring to bet on the urn with known probabilities rather than one with unknown probabilities. The phenomenon of ambiguity aversion is puzzling. People tend to be indifferent between betting on red or black for either urn so that drawing either color ought to have the same subjective probability of one-half, regardless of the urn used. More recently, it has become increasingly recognized that decision making under uncertainty depends not only on probabilities, but also on how uncertainty itself arises. This has been specifically referred to as source dependence or familiarity bias (Fox and Tversky, 1995). In particular, they echoed Keynes and proposed that people have familiarity bias in tend their disposition to prefer betting on risks arising from a more familiar source of uncertainty.

Over the past several decades, these anomalies have inspired an active literature in decision theory going beyond the expected utility model, e.g., by using a nonlinear probability weighting function (Kahneman and Tversky, 1979; Quiggin, 1982; Chew, 1983), by using a non-additive capacity over events (Schmeidler, 1989; Tversky and Kahneman, 1992), by assuming that decision makers have non-unique priors (Gilboa and Schmeidler, 1989), and by allowing non-indifference over identically distributed risks arising from different sources of uncertainty (Chew and Sagi, 2008; Ergin and Gul, 2009).

Different methods to elicit risk attitude have been reported in the experimental economics literature.[8] An increasingly used elicitation procedure, known as the multiple price list design (Miller, Meyer, and Lanzetta, 1969; Holt and Laury, 2002), entails giving the subject on an ordered array of binary choices. Anderson et al. (2008) offered suggestive evidence of the stability of risk preference assessed using multiple price list design over a 17-month time span. Risk attitude assessed using this design was also shown to predict risky behaviours such as cigarette smoking, heavy drinking, being overweight or obese, and acceptance of risky food (Lusk and Coble, 2005; Anderson and Mellor, 2008).

Harrison et al. (2008) conducted a field experiment in Denmark with a representative sample of 253 subjects between 19 and 75 years of age. They found

[8] See, e.g., Harrison and Rutstrom (2008) for an excellent review

strong support for a decrease in risk aversion as the age of a person increases, but no effect of gender on risk attitudes. Dohman et al. (2006) conducted a study with a representative sample of roughly 22,000 individuals in Germany, using a question that asks about willingness to take risks on an 11-point scale. They found that willingness to take risks is negatively related to age and to being female, and positively related to height and parental education. However, using a methodology from psychology, Benjamin, Choi and Strickland find that making gender identity salient has no effect on intertemporal choices. In a sample of 660 customers of a German car manufacturer, Gachter et al. (2007) showed that older people are more loss averse than younger people. Higher education decreases loss aversion, while higher income and higher wealth are both positively correlated with loss aversion. Using a Dutch sample of 1935 subjects, Booij et al. (2009) found that older people are more risk averse in the gain domain, but other social, economic, and demographic variables such as income, age, and education did not appear to have a significant effect on their risk attitudes for risks over gains as well as over losses.

In a Chinese sample of 350 subjects, Zhong et al. (2009c) found male subjects to be more risk tolerant toward longshot prospects than female subjects, but not for longshot hazards. Older subjects tend to be significantly more risk averse toward longshot hazards, but not for longshot prospects. In a study of Allais paradox using a large sample of 1426 subjects, Huck and Müller (2009) found considerable heterogeneity in the population and that violation of expected utility tend to be prevalent among subjects who had less education, poor, or unemployed.

Moore and Eckel (2003) reported that in the gain domain women were more risk averse and more ambiguity averse than men. Schubert et al. (2004) reported women to be more ambiguity averse over gains while men are more ambiguity averse over losses. In a recent experiment with 347 high school students, Borghans et al. (2009) showed that women are more risk averse, while men are more ambiguity averse than women, and that this gender gap in ambiguity vanishes after conditioning on risk aversion.

2.2 Decision Making involving Time

The standard additively separable model for time preference was first proposed by Samuelson (1937), assuming a constant discount rate with a utility function for within-period consumption. There has been accumulating evidence showing that the lack of descriptive validity of this exponential discounting model.[9] In particular, people tend to exhibit hypobolic discounting preference in being more impatient across consumptions in the immediate future than in the more distant future. Phelps and Pollak (1968), and later on Laibson (1997) introduced a two-parameter model to account for this phenomenon.

Kirby et al. (2002) collected field data from 154 Tsimane' Amerindians 10–80 years of age and found that discount rates are positively correlated with age, negatively correlated with education and income, but not with wealth. In a study of Vietnamese villagers, Tanaka et al. (2009) combined survey information with experimentally elicited measures of preferences to study risk attitudes and time preferences. Their found lower discount rates to be associated with both higher household incomes and average village incomes but they did not find a statistically significant relationship between discount rate and education. Chabris et al. (2008) found that individual discount rates predict inter-individual variation in field behaviors including exercise, body mass index, and smoking. While the correlation between the discount rate and each field behavior is small, the discount rate has at least as much predictive power as any other variable such as gender, age, and education. In a survey study, Ameriks et al. (2007) reported that time-inconsistent behavior correlates with overconsumption and low wealth. Benjamin, Choi and Strickland (2010) find that, compared with white subjects, Asian-American subjects are more likely to make more patient choice when their ethnic identity are made to be salient. When racial identity are made more salient, non-immigrant blacks are more likely to make more patient choice than immigrant blacks.

Another dimension of decision making over time concerns the timing of consumption, specifically, the idea of anticipation and dread (Loewenstein, 1987). In his experiment, subjects indicated how much they would pay to obtain (avoid) outcomes that would occur either immediately or after one of several delays. A robust difference emerged in the comparison of timing preferences for desirable and aversive outcomes. For obtaining a kiss from the movie star of one's choosing, subjects considered it almost twice as valuable if it was set to occur in three days rather than immediately; for receiving a non-lethal electric shock, subjects were willing to pay almost twice as much to avoid it for ten years as they would pay to

[9] See, Frederick, Loewenstein, and O' Donoghue (2002) for a survey.

avoid the same shock immediately. Loewenstein attributed this to the utility that people expect to derive during the period of waiting: anticipating a pleasant outcome versus dreading an unpleasant one.

Timing of the resolution of uncertainty is another dimension of decision making over time. The timing of uncertainty resolution may matter for two reasons: planning advantage of early resolution (Kreps and Porteus, 1978) and anticipatory feelings such as hopefulness in case of late resolution (Chew and Epstein, 1989; Chew and Ho, 1994). Chew and Ho conducted an experimental test, and found that hopefulness, i.e., a preference for late resolution of uncertainty, is more prevalent when there is a small probability of receiving a sizable gain; while anxiousness, a preference for early resolution, becomes more prevalent when there is there is good chance of receiving the sizable gain. Similar results are reported subsequently in Lovallo and Kahneman (2000) and Hopfensitz, Krawczyk, and van Winden (2008).

3 Experimental Design Involving Adult Twins

The data sets that we use in this study are combined from two sources. One is derived from the Chinese Adult Twin Survey (CATS) which was conducted in 2002. The other is derived from the experiments that were conducted in 2008 on a subsample of the twins in CATS. This section describes the CCTS, the experiments, and the summary statistics.

3.1 The Chinese Adult Twins Survey

The socioeconomic variables in our analysis are derived from the CATS.[10] It was conducted by the Urban Survey Unit (USU) of the National Bureau of Statistics (NBS) in June and July 2002 in five cities in China. They are Chengdu, Chongqin, Harbin, Hefei, and Wuhan. Based on existing twin questionnaires from the United States and elsewhere, CATS covers a wide range of demographic, social, and economic information. The questionnaire was designed by one of the authors of this essay in close consultation with Mark Rosenzweig and Chinese experts at the NBS. Adult twins aged between 18 and 65 were identified by the local statistical bureaus. The questionnaires were completed through face-to-face personal interviews. One of

[10] Li et al. (2007), Huang et al. (2009), Li, Rosenzweig and Zhang (2010), and Li et al. (2010) gave a detailed description of the CATS data.

the authors made several site checks of the survey work and closely monitored the data input process. Thus, the survey was carefully conducted.

The CATS is the first socioeconomic twin survey in China and perhaps the first in Asia. There is rich socioeconomic information in the data set. We consider a pair of twins to be identical (monozygotic, MZ) or non-identical (dizygotic, DZ) based on the whether they have identical hair, color, looks, and gender. Thus, we can distinguish whether the twins in the sample are identical or non-identical. There are a total of 3,002 individuals who are twins. We have complete information on education and other variables for both twins in 964 pairs (1928 individuals). Of these, 488 pairs (976 individuals) are MZ twins.

3.2 Decision Making Experiments

The measures of preferences are derived from the experiments. In June and July 2008, one of the authors conducted a set of experiments on a subsample of twins in the CATS.[11] The subjects of the experiments are from the CATS data. However, due to the budget constraint, we conducted the experiments only in two cities. They are Hefei and Wuhan, the capitals of Anhui and Hubei provinces, respectively. Some twin individuals have changed their address and contact information during the period of 2002-2008. These individuals thus could not be reached. Furthermore, because participation in the experiments was voluntary, some individuals in the CATS refused to take part in the experiments due to time constraints or unwillingness. Eventually, we recruited 70 pairs (140 individuals) of twins for our experiments.[12]

The experiments were conducted in a hotel conference room at Wuhan, and a classroom at Hefei. Each individual who took part in the experiment was paid RMB60 as show-up fee.[13] In addition, there were various payoffs in each experiment. Most individuals finished the experiments within one hour. The maximum time spent was almost one and a half hours. The money was paid in cash after participants

[11] We also hired several experiment assistants in conducting the experiments.

[12] There is a sample selection problem because of migration and selective participation. However, any sample selection bias arising from migration should not be a major issue because the inter-urban migration rate was only 0.975% based on the 2000 census (Li, Liu, and Zhang, 2011). Regarding to sample selection arising from voluntary participation, this is a general concern with all labor experiments and some field experiments (List and Rasul, 2010). Thus, we exercise caution in interpreting our empirical results.

[13] The exchange rate is US \$1≈RMB 6.8 in 2008.

finished the experiment. On average, they earned RMB42 in addition to the basic reward. In these two cities, the minimum wage per hour for a full time employee is RMB6.5, and the average wage per hour is RMB9.5 (NBS, 2009). Thus, the amounts earned by participants were, on the average, much higher than their alternative wages, and participants should have sufficient incentive to make careful decisions in the experiments. The experimental design and instructions are presented in Appendix I.

Attitudes toward Fourfold Risk

There have been various ways to assess risk attitudes, and a simple experimental elicitation procedure is known as the multiple price list design, which entails giving the subject on an ordered array of binary choices.[14] The task is simple and relatively context free with the multiple price list design, relative to other experiment-based studies of risk aversion.

In this study, we use a simplified version of this procedure to assess subjects' risk attitudes. In assessing risk attitude toward moderate prospects (GAME ONE in Appendix I), subjects chose between an even-chance lottery between receiving RMB40 and receiving zero, versus receiving the expected outcome of RMB20 for sure. Subjects were incentivized for their choice in this comparison. Based on their decisions, subjects' valuation of the gamble is categorized as follows: risk aversion if certainty is chosen; risk seeking if lottery is chosen. Correspondingly, in assessing risk attitude toward moderate hazards (GAME TWO in Appendix I), subjects begin by choosing between a lottery which involves losing RMB10 and losing zero with equal probability versus losing RMB5 for sure. Subjects were incentivized, i.e., losses were deducted from subjects' show-up fees. Based on their decisions, subjects' valuation of the gamble is categorized as follows: risk averse over losses if certainty is chosen; risk tolerant over losses if lottery is chosen.

For longshot prospects (GAME THREE in Appendix I), subjects order the value of three items: (A) RMB2 lottery ticket which has a very small chance of winning 5 millions, (B) RMB2 lottery ticket which have small chance of winning 0.1 million, and (C) RMB2 for sure. We paid subjects their most preferable choice as incentive. Subjects are classified as exhibiting longshot preference, when A is preferred to B which is in turn preferred to C. For longshot hazards (GAME FOUR in Appendix I), subjects are classified as being disposed to insure if they prefer

[14] For a survey, see, e.g., Harrison and Rutstrom (2008)

losing RMB2 for sure than losing RMB2000 with 0.1% chance. We did not incentivize GAME FOUR given the anticipated difficulty in collecting RMB2000 should this unlikely contingency occur.

Allais-type Behavior

We adopt two of the four pairs – called the H (high) and L (low) pairs – of binary choices in Chew and Waller (1984) designed to test the independence axiom's parallelism implication on the behavior of indifference curves in a probability triangle (GAME FIVE in Appendix I). They find the highest incidence of Allais-type behavior, i.e., non-parallelism, based on subjects' choices in the H (high) and the L (low) pairs of binary choices. In our design, the H pair involves subjects choosing between receiving a high outcome of RMB100 with a 80% chance and receiving an intermediate outcome of $0 with 20% chance (Option A) versus receiving RMB100 with 90% chance and receiving a low outcome of losing RMB80 with 10% chance (Option B). The L pair involves subjects' choosing between losing RMB80 with 80% chance and receiving $0 with 20% chance (Option A) versus losing RMB80 with 90% chance and receiving RMB100 with 10% chance (Option B). We classify subjects as expected utility type if they choose either A or B in both pairs as implied by the independence axiom. Otherwise, we classify subjects as being Allais type if they exhibit the choice pattern – choose A in the H pair and choose B in the L pair – which imply that their indifference curves fan out in the probability triangle, i.e., satisfies Machina's (1982) Hypothesis II. Given the significant loss amount involved, this task was not incentivized.

Ambiguity Aversion and Familiarity Bias

Most experimental studies on the original Ellsberg paradox involve choosing between betting on the urn with known probability distribution and betting on that with unknown probability distribution. Betting correctly in either case would pay the same. It is found that people tend to bet on the urn with known probability distribution (see, e.g., Camerer and Weber, 1992). In order to generate a more even split of individual difference between those preferring to bet on the "known" urn versus those preferring to bet on the "unknown" urn, we increase the payoff associated with betting on the unknown urn. This calls for a judicious choice of a threshold difference. In the ambiguity aversion task (GAME SIX in Appendix I), subjects choose between betting on a "known" deck consisting of 10 red cards and 10 black cards, and an "unknown" deck consisting of 20 cards without knowing the composition of the red and black cards. For the known deck, a correct bet pays RMB10. For the unknown deck, a correct bet pays RMB12 with an increase of RMB2 as a result of pretests.

In the original experiment on familiarity bias in Fox and Tversky (1995), the bet is on whether the temperature in San Francisco/Istanbul is above or below a specific temperature. In our design, subjects choose between betting on whether Beijing temperature at a specific historical day would be odd or even, and similarly betting on Tokyo temperature (GAME SEVEN in Appendix I). Our design induces the same objective probability (Machina, 2004) of one half for odd versus even regardless of the city chosen. To generate an even split between those betting on Beijing and those betting on Tokyo, betting correctly on Beijing temperature pays RMB11 which is RMB2 less than betting on Tokyo temperature.

Impatience and Hyperbolic Discounting

In experimental studies, binary choice is a commonly used method to elicit discount rates in which subjects choose between a smaller and more recent reward versus a larger but more delayed reward. Other methods include matching tasks, rating tasks, and pricing tasks (see, e.g., Frederick, Loewenstein, and O' donoghue, 2002). Our study makes use of a simple hypothetical choice task (GAME EIGHT in Appendix I). In Situation 1, subjects choose between getting RMB100 today (A) and getting RMB120 seven days later (B). In Situation 2, subjects choose between getting RMB100 91 days later (A) and getting RMB120 98 days later (B). If subjects choose A in the first case, they are impatient. If in addition they choose B in Situation 2, they exhibit hyperbolic discounting behavior. For these tasks, we used a questionnaire.

Anticipation and Dread

We adopt a similar design as in Loewenstein (1987) for timing-of-consumption preference with both desirable and aversive outcomes: having dinner with the movie star of one's choice, and receiving a non-lethal electric shock (GAME NINE and TEN in Appendix I). Subjects were asked whether they prefer to have the dinner today or three days later. If they choose 3 days later, we classify them as experiencing anticipation. Subjects were asked whether they prefer to receive a non-lethal electric shock today or 6 months later. If they choose today, we classify them as exhibiting dread. For these tasks, we used a questionnaire.

Hopefulness and Anxiousness

We adopt a similar design as in Chew and Ho (1994) and Lovallo and Kahneman (2000) for the timing of uncertainty resolution (GAME ELEVEN and TWELVE in Appendix I). In one task, subjects state whether to delay the resolution of uncertainty about the gender of the baby by paying RMB2 under the supposition that one of his/her relatives is pregnant. We classify the subjects as experiencing hopefulness if they prefer to pay to delay the resolution of uncertainty. In another task, subjects state when they prefer to pay RMB2 to resolve uncertainty immediately on the prospect of receiving RMB1000 with 90% chance and receiving zero otherwise versus waiting until two weeks later to resolve this uncertainty. If they choose to resolve now, we classify them having preference of anxiousness, although there may be some value for planning. For this task, we used a questionnaire.

3.3 Definitions of Variables and Summary Statistics

Appendix II defines experimental measures of preferences. All variables are 0-1 dummy variables and are all self-explanatory. The summary statistics of variables are reported in Tables 1a and 1b. Table 1a first reports the individual's education levels and other socioeconomic variables. The educational attainments are categorized into five levels. The first three are general education. The next two are professional education. We use education levels rather than years of education because education years between high school and technical school are incomparable. In other words, education year is not a cardinal variable in the Chinese education system. Specifically, in China, the student faces two choices after graduating from middle school: technical school or high school. If the student enters into technical school, she will get four years of elementary profession education and then go to work. If the student enters into high school, she will get three years of general education and then take the college entrance examination. If she passes the examination, then she goes to college. Otherwise, she will go to work.[15] Since the high school education is examination oriented, few technical school graduates will take the college entrance

[15] Appendix III depicts the Chinese education system.

examination. The qualitative difference between technical school and high school is confirmed by Li, Liu, and Zhang (2011). Using the CATS data, they find that the economic return to high school education is much lower than the economic return to technical school. The return to high school education is bounded between 4.0-4.5 percent. In contrast, the return to technical school education is bounded between 20.6-22.5 percent. Finally, we group the primary school and middle school together and treat them as a baseline group because there are very few individuals who have only primary education.

We are using a relative old adult twin data set because the mean of age is 46. The minimum age in this sample is 28. Thus, it is safe to use the education attainments in the CATS of 2002 because all twins had graduated from colleges in 2002 if they attended colleges. Only 47% of the individuals are males.[16] In the OLS regression, we also include parental education levels to check how the family background affects individual's preferences. By comparing the OLS estimates without parental education levels, OLS estimates with them, and within-twin-pair fixed-effects estimates, we are able to infer how the family background and individual heterogeneity affect adult preferences.

As a robustness check, we include birth weight to control for pre-birth differences between the twins.[17] To detect the channels by which education affects preference, we include family annual income and health into the regression equation.[18] It has concluded that education attainments increase incomes and demonstrated that higher educated people have better health (Grossman, 1975). Since incomes and health may also affect preference, we try to find whether education affects preferences through incomes or health.

Table 1b reports the summary statistics for the experimental measures of preferences. The last column of Table 1b gives the percentage of within-twin-pair variation to the total variation for each preference variable. We find that the within-twin-pair variations account for about one third to one half of the total variations for all preference variables, ranging from the lowest 34.50% to the highest 54.39%.

---

[16]All twin siblings in the data are of same gender.

[17]Recently, it is found that birth weight affects a series of short- and long-run individual outcomes, including health, academic performance, education attainments and earnings (Behrman and Rosenzweig, 2004; Almond, Chay, and Lee, 2005).

[18]In CATS, the self-reported health status is rated into 5 levels. They are poor, fair, good, very good, and excellent. We categorize health into a dummy variable. It equals one if the individual reported the health status good, very good, or excellent. Otherwise, it equals zero.

4 Empirical Strategy

Our empirical analysis focuses on the estimation of the following equations that are given as

$$y_{1i} = \alpha + E_{1i}\beta + X_i\gamma + Z_{1i}\delta + \mu_i + \varepsilon_{1i}, \tag{1}$$

$$y_{2i} = \alpha + E_{2i}\beta + X_i\gamma + Z_{2i}\delta + \mu_i + \varepsilon_{2i}, \tag{2}$$

where y_{ji} ($j = 1,2$) is the experimental measure of preference of the first and second twin in family i.[19] E_{ji} ($j = 1,2$) is a vector containing dummies of education levels for twin j in family i; X_i is the set of family background variables that are observable and varying across families but not across twins; Z_{ji} ($j = 1,2$) is a set of observed variables that vary across the twins. μ_i represents a set of unobservable variables at the family level that may also affect preferences.

The OLS estimate of the educational effect on preference in Equation (1) (or Equation (2)), β, is generally biased. The bias arises because normally we do not have a perfect measure of μ_i, which is very likely to correlated with E_{ji} and y_{ji} simultaneously. Thus, we apply a within-twin-pair fixed-effects estimator for twins which is based on the first difference between (1) and (2),

$$y_{1i} - y_{2i} = (E_{1i} - E_{2i})\beta + (Z_{1i} - Z_{2i})\delta + \varepsilon_{1i} - \varepsilon_{2i}. \tag{3}$$

Both observed and unobserved family effects, i.e., X_i and μ_i are differenced out in Equation (3). Because μ_i has been removed, we can apply the OLS method to Equation (3) without worrying about bias being caused by the omitted variable of μ_i.

It is notable that the identification assumption of Equation (3) is that the differences in within-twin-pairs educational attainments are resulted from random deviations from the optimum schooling level in the same family. In other words, the within-twin-pair differences in schooling levels are uncorrelated with any omitted variables, which may affect the preferences formation in the future. The assumption of within-twin-pair random deviation from optimal schooling with the within-twin-pair fixed-effects estimator has been extensively examined and discussed in

[19] Since all experimentally measured preferences (dependent variables) are dummies, a logit model would seem to be a natural choice. However, a linear probability model facilitates our within-twin-pair fixed-effects estimation and the interpretation the estimated coefficients.

Ashenfelter and Rouse (1998). We will systematically examine the robustness of the within-twin-pair fixed-estimates in Section 6 below.

There is a concern about the small sample size of our adult Chinese twins. The small sample size containing only a total of 140 subjects casts doubts on the validity of classic tests. Classic statistical tests such as those based on the $t-$ and F statistics hinge on central limit theorems. These statistics are only asymptotically valid. Micceri (1989) makes an extensive survey concluding that classic testing statistics may be unreliable when the sample size is small. We address the potential problem of small sample size using the permutation-based inference procedure which is valid in small samples (Freeman and Lane, 1983; Heckman et al., 2010). Specifically, the permutation tests are based on Monte Carlo simulations. All reported $t-$ statistics below are computed using 3,000 draws under the random permutation procedure.

5 Education and Preferences

This section reports and discusses our estimation results. We present successively the estimated effects of education on decision making under risk and uncertainty and on decision making involving time.

5.1 Education and Decision Making under Risk and Uncertainty

Tables 2 reports both the OLS estimates and within-twin-pair fixed-effects estimates of education and fourfold pattern of risk attitudes.. As discussed above, we categorize educational attainments into four groups. The group of middle-school-and-below is the baseline group which is omitted from the regression equation. The estimated coefficients on the three educational groups remaining in the regression equation are relative to the omitted group of middle-school-and-below, respectively. Columns (1)-(3) report the estimation results with risk attitude toward moderate prospects. From the OLS estimates in Column (1), we find that higher education increases risk tolerance marginally, although the estimates are statistically insignificant. Controlling for father and mother's educational levels, Column (2) reveals an increase in the magnitude of the estimated coefficients on technical school and college level experience. In Columns (3) of the within-twin-pair fixed-effects estimation, both observed and unobserved family characteristics are swept out. In this case, we find that the education level at college-and-above substantially increases risk tolerance toward moderate prospects, and the estimate is statistically significant.

Column 2 of Table A1 in Appendix IV shows that the effects of father's education are opposite that of mother's education on risk attitude toward moderate prospect.[20] Father's education increases risk tolerance while mother's education decreases it. This finding may help rationalize a bargaining or a collective household model rather than a unitary household model involving a dictator or a dominant preference in the family.[21] From the magnitudes of the estimated coefficients, the negative effect of mother's education seems to dominate the positive effect of father's education on child's risk attitude toward moderate prospects..This result is echoing Behrman and Rosenzweig (2002) who found that mother's education is more important in determining the education of the second generation than father's education. We find that mother's education seems to be more important in shaping children's preference formation.

Columns (4)-(6) of Table 2 report the estimation results with risk attitude toward moderate hazards. Similar to the results with moderate prospect, the within-twin-pair fixed-effects estimates in Column (6) indicate that the education level of college-and-above significantly increases subjects' risk tolerance for moderate hazards. It is noted that, in Columns (4)-(5) of Table A1 in Appendix IV, the age effect on risk attitude toward moderate hazards seems nonlinear. Age increases risk tolerance in the beginning, and then declines with old age. This nonlinear pattern of age effect exists also in the estimation of risk attitude toward moderate prospects (Columns (1)-(2) of Table A1), although the estimated coefficients are only marginally significant. The estimated nonlinear relationship between age and attitude toward moderate risks may help reconcile a controversy in the literature. For example, Harrison et al. (2008) found that risk aversion decreases as age increases, while Dohman et al. (2009) observed that the willingness to take risks is negative related to age.

Columns (7)-(12) report the estimates relating to attitudes toward longshot prospects and longshot hazards. Similar to the effects on moderate prospects and moderate hazards, the fixed-effects estimation results in Column (9) show that the

[20] To save space, we have only reported the estimated coefficients on education levels in the paper, while the estimated coefficients on other variables are reported in Appendices IV and V.

[21] For a survey, see, e.g., Behrman (1997).

education level of college-and-above significantly increases risk tolerance for longshot prospects. On the contrary, Column (12) does not show a significant effect of education on people's risk attitude toward longshot hazards.

Table 3 reports both the OLS and within-twin-pair fixed-effects estimates of education and decision making anomalies under risk and uncertainty. We first look at the estimation results with Allais-type behavior in Columns (1)-(3). It is interesting to find that more educated persons are more likely to exhibit Allais-type behavior. The within-twin-pair fixed-effects estimate is statistically significant for the group with college-and-above education. Columns (4)-(9) present the estimated results with ambiguity aversion and familiarity bias. Comparing with the baseline group with middle-school-and-below education, the fixed-effects estimates in Columns (6) and (9) show that more educated people are consistently more ambiguity averse as well as more biased toward familiarity. Interestingly, the effect of education on ambiguity aversion and familiarity bias seems to be nonlinear. The group with high school education is estimated to exhibit the strongest ambiguity aversion and familiarity aversion. It is noted that, in Column (17) of Table A1 in Appendix IV, the effect of father's education on the child's familiarity bias is opposite to that of mother's education. Father's education decreases the child's familiarity bias, while mother's education increases it. The finding further supports a bargaining or collective household model rather than a unitary household model.

Summarizing Tables 2-3, we first conclude that education increases subjects' risk tolerance toward moderate prospects, moderate hazards, and longshot prospect. Second, more educated people are more likely to deviate from the prediction of the expected utility theory. They are more likely to display Allais-type behavior, ambiguity aversion, and familiarity bias, suggesting that people with higher levels of education seem more anomalous in risk attitudes. Finally, the big differences between the OLS estimates and within-twin-pair fixed-effects estimates for each preference measure in Tables 2-3 confirm that it is important to control for the cross-family heterogeneity.[22] The results suggest caution in interpreting the OLS estimates of education and decision making under risk and uncertainty as causal.

[22] However, the small sample size limits our further effort to explore the statistical significance with respect to the differences between OLS and fixed-effects estimates.

5.2 Education and Decision Making Involving Time

The OLS and within-twin-pair fixed-effects estimates of education on decision making involving time are shown in Table 4. From Column (3), we first find that subjects with college-and-above education are significantly more patient than other groups. In addition, Column (6) shows that the group with college-and-above education is less likely to exhibit hyperbolic discounting. In other words, their decisions are more likely to be time consistent.[23] It is interesting to note, from Columns (1)-(2) and (4)-(5) in Table A2 of Appendix V, that males are more patient and less disposed to exhibiting hyperbolic discounting than females. The estimated coefficients on gender (a male indicator) are highly significant. On anticipation, Column (9) in Table 4 shows that education does not have a statistically significant effect. By contrast, Columns (12) and (15) in Table 4 indicate that having more education significantly decreases both dread and hopefulness. Finally, subjects with college-and-above educational attainments are less likely to be anxious, though the estimates are statistically insignificant (Column (19)).

In summary (see Table 4), in terms of the signs of the estimated coefficients, the within-twin-pair fixed-effects estimates show that education decreases impatience, hyperbolic discounting, dread, and hopefulness, as well as anticipation, with anxiousness being the only exception. While education seems to increase anticipation, its fixed-effects estimate is statistically insignificant. In contrast to decision making under risk and uncertainty, people with higher level of education tend to exhibit less "bias" preference regarding time. Table 4 also indicates that there are big differences between OLS estimates and within-twin-pair fixed-effects estimates regarding decision making involving time. This finding corroborates the importance in addressing the causality between education and preferences.

6 Robustness

This section reports the results from several robustness tests. To deal with potential measurement errors with educational attainments, we first conduct an instrumental variable estimation. Second, to examine possible biases with our within-twin-pair fixed-effects estimates induced by omitted variables, we conduct the estimation controlling for pre-birth endowment of birth weight and restricting our sample to MZ

[23] There are 95% of the subjects who exhibit time inconsistency are practicing hyperbolic discounting in our experiment.

twins. Third, we estimate the effects of educational attainments on preferences controlling for income and health. Fourth, we check the possibility of reversal causality. Finally, we examine any remaining potential biases of our within-twin-pair fixed-effects estimates.

6.1 Measurement Errors

The measurement error problem is a primary concern with the within-twin-pair fixed-effects estimator (Ashenfelter and Krueger, 1994). The classical measurement error in education leads to a downward biased (in terms of absolute value) estimate. The fixed-effects model exacerbates such measurement error bias. This essay follows Ashenfelter and Krueger (1994) to obtain good instrumental variables to deal with possible measurement error problem. Specifically, in the CATS we asked each twin to report both their own education and their co-twin's education. If there is a risk of measurement error in the self-reported education, the cross-reported education is potentially a good instrument. The reason is that the cross-reported education should be correlated with the true education of a twin but uncorrelated with any measurement error that might be contained in the self-reported one.

The instrumental variable approach is applied as follows. Denote E_j^k for twin $k's$ report of twin $j's$ education level. We can then use $E_1^2 - E_2^1$ to instrument $E_1^1 \quad E_2^2$ in Equation (3). This approach is valid in the presence of common family-specific measurement error because family effects are eliminated in the within-twin-pair difference. However, as Ashenfelter and Krueger (1994) demonstrated, the measurement error term in $E_1^2 - E_2^1$ and that in $E_1^1 - E_2^2$ may be correlated. In this case, the instrumental variable estimate using $E_1^2 - E_2^1$ is also biased. This consideration motives us to use $E_1^1 - E_2^1$ as the regressor and $E_1^2 - E_2^2$ as the instrumental variable. This method is valid even in the presence of correlated measurement errors because the individual-specific component of the measurement error in the estimation is swept out.

Before directly going to the IV estimates, we have compared the individual's self-reported education with co-twin's reported. It is found that there are only six individuals among the whole sample whose self-reported educations are different from those reported by co-twins. It means that the potential rate of misreports of education is only 4.29%, which is lower than the misreport rate in Ashenfelter and Krueger (1994). Considering the low misreport rate, we expect that the IV estimates will not be much different from the within-twin-pair fixed-effects estimates reported in the previous section. This prediction is confirmed by Table 5, which reports the instrumental variables within-twin-pair fixed-effects estimates of education and preferences. From this table, we find that the pattern of the education effects on various preferences still remains. In summary, our within-twin-pair fixed-effects estimates of the education effects on preferences are robust to the measurement errors problem.

6.2 Omitted Variables

Although twin siblings share similar family environment, there may still exist unobservable heterogeneity across them. For example, twin siblings may be different in womb nutrition intakes and thus there are different birth weights across them. Recent studies find that birth weight affects a series of short- and long-run outcomes such as health, education, and income (Behrman and Rosenzweig, 2004). Therefore, we include birth weight to control for pre-birth endowment as a robustness test. It is also interesting to examine the effects of birth weight on preferences. Table 6 reports the within-twin-pair fixed-effects estimates controlling for the variable of birth weight. It is found that the estimated effects of educational attainments on preferences are very similar to those in Tables 2-4, indicating that our results are robust to the inclusion of birth weight as a control variable. In Table 6, we also find that birth weight does not have a statistically significant effect on preferences, which may be due to the small sample size.

Because of the small sample size, we have included both the MZ and DZ twins in our estimation above. Although DZ twins share identical family environment, they only share half of the genetic endowments. Thus, it may be argued that the within-twin-pair fixed-effects estimation is not so clean. Table 7 reports the within-twin-pair fixed-effects estimates when we restrict the estimation sample to MZ twins only. Although there are only thirty six pairs of MZ twins, the pattern of the effects of educational attainments on preferences in the basic estimation remains similar to that reported in the previous section. Despite the small sample size, the education level at college-and-above significantly increases the risk tolerance toward moderate prospect and decreases impatience.

6.3 Estimates Controlling for Income and Health Status

It has been argued that socioeconomic variables such as incomes and health affect preferences. During the same time, education affects income and health. Thus, we have also included income and health into the regression equation to check whether the effects of education on preferences are through income or health. Tables 8-9 report the within-twin-pair fixed-effects estimates of educational attainments on preferences by controlling for income and health, respectively. We find that the pattern of the estimates of education attainments has changed little after controlling for these two variables.

6.4 Reversal Causality

Another potential problem with our within-twin-pair fixed-effects estimates is reversal causality bias. Although the unobservable family factors and individual heterogeneity may have been well taken care of by using the within-twin-pair fixed-effects estimator, there may be a reversal causality problem running from preference to educational attainments. However, our within-twin-pair fixed-effects estimates are less likely biased by the reversal causality problem. Chronologically, on the one hand, preferences are experimentally measured at, on the average, 45 years old, while education was normally finished before age 22 for all subjects in our data set. On the other hand, because twin siblings, in particular MZ twin siblings, share common family background and genetic endowments, they may be unlikely to have differences in preferences in the early stage of their life that may have affected their educational attainments. The small sample size restricts our effort to further address the possible reversal causality problem.[24]

6.5 Potential Biases of Within-Twin-Pair Fixed-Effects Estimates

Ashenfelter and Rouse (1998) emphasized that there are no genetic differences between identical twins except measurement errors. They argue that different schooling levels of identical twins are due to random deviations that are not related to the determinants of schooling choices. However, within-twin-pair estimation may not completely eliminate the bias of conventional cross-sectional estimation,

[24] Re-schooling may be a potential threat to our within-twin-pair fixed-effects estimator because preference at the adult stage may affect the subject's re-schooling choice and education level. However, we find that there are only 4 subjects in our sample (140 subjects) who had received education after age 25.

although our within-twin-pair fixed-effects estimates are consistently robust in a series of sensitive analyses above. The reason is that the within-twin-pair difference in ability may remain in $\varepsilon_{1i} - \varepsilon_{2i}$ in Equation (3), which may be correlated with $E_{1i} - E_{2i}$. If endogenous variation in education comprises as a large proportion of the remaining within-twin-pair variation as it does of the cross-sectional variation, then within-twin-pair estimation is subject to as large an endogeneity bias as cross-sectional estimation. The potential endogeneity of schooling differences between twins corresponding to remaining unobserved differences in ability or personality may exist in our fixed-effects estimates despite the common genetics or they may result from with-MZ twins different experiences. Thus, the major concern of the within-twin-pair estimate is whether it is less biased than the cross-sectional estimate, and is therefore a better estimate.

Note that the bias in the cross-sectional estimator depends on the fraction of variance in education that is accounted for by variance in unobserved ability that may also affect earnings, that is, $\frac{\text{cov}(E_i, \mu_i + \varepsilon_i)}{\text{var}(E_i)}$. Similarly, the ability bias of the fixed effects estimator depends on the fraction of within-twin-pair variance in education that is accounted for by within-twin-pair variance in unobserved ability that also affects earnings, that is, $\frac{\text{cov}(\Delta E_i, \Delta\mu_i + \Delta\varepsilon_i)}{\text{var}(\Delta E_i)}$. If the endogenous variation within a family is smaller than the endogenous variation between families, the fixed effects estimator is less biased than the cross-sectional estimator. In that case, we can credit that the within-twin-pair estimates are better than OLS estimates.

Ashenfelter and Rouse (1998) suggested a correlation analysis to examine whether the within-twin-pair estimate is less biased than the cross-sectional estimate. Using the CATS data, Li et al. (2011) conducted a correlation analysis similar to that of Ashenfelter and Rouse. They use the correlations of average family education over each twin pair with the average family characteristics that may be correlated with individual heterogeneity to indicate the expected omitted bias in a cross-sectional OLS regression. They then use the correlations of the within-twin-pair differences in education with the within-twin-pair differences in these characteristics to indicate the expected omitted bias in a within-twin-pair regression. If the correlations in the cross-sectional case are larger than those in the within-twin-pair case, then the bias in the cross-sectional regression is likely to be larger than the bias in the within-twin-pair regression. Li et al. found that the between-family correlations are all larger in magnitude than the within-twin-pair correlations, suggesting that the within-twin-pair estimation of the return to education may indeed be less affected by omitted individual heterogeneity than the cross-sectional OLS estimation.[25] Given that we also use the CATS data, Li et al. provided suggestive evidences that our within-twin-pair fixed-effects estimates are less biased than cross-sectional OLS estimates.

7 Discussion and Concluding Remarks

The essay utilizes both survey and experimental data and provides a systematic empirical study of the effect of education on two dimensions of preference – risk and time. To control for unobserved family environment and to minimize individual endowment heterogeneity, we conduct a number of economic experiments on adult twin pairs and use within-twin-pair fixed-effects estimators to carry out the identification of the effect of education on preferences. Our fixed-effects estimates indicate that people with higher level of education are less risk averse toward moderate prospects, moderate hazards, and longshot prospects. These findings are in line with previous findings about risk attitude and cognitive ability (Dohmen et al., 2010; Benjamin et al., 2006; Burks et al., 2009) and extend the findings to moderate hazards, and longshot prospects. In relation to decision making anomalies under risk and uncertainty, we find that more educated people tend to be more disposed to exhibit Allais-type behavior and longshot bias. The within-twin-pair fixed effects estimates of education and preference in decision making under risk and uncertainty may contribute to our understanding of the relationship between education and socioeconomic well-being. One reason is that risk attitudes toward uncertainty underpin a wide range of economic behavior, such as portfolio choice and insurance purchase, which have long-term economic consequences for individuals.

In terms of preference involving time, our findings suggest that people with higher education will tend to be more patient and exhibit less hyperbolic discounting. Time preferences, in particular time consistency, are essential in dynamic economic decision making. Our estimated effects of educational attainments on time preferences have important implications in economics. They can help enhance our

[25] Table A3 in Appendix VI cites the between-families and within-twin-pair correlations of education and other variables in Li et al. (2010).

understanding of the relationship between education and socioeconomic well-beings. It is found that time preferences are correlated with economic behaviors such as investment, consumption, and saving, and with health related behavior such as physical exercise and smoking. In addition, the within-twin-pair fixed-effects estimates of education on time preferences may suggest a bridge between two strands of macroeconomic models, namely, the endogenous growth model stressing human capital accumulation and the dynamic macroeconomic model stressing time consistency. Furthermore, these findings accord well with intuition which underpins the so-called Save More Tomorrow prescriptive savings program (Thaler and Benartzi, 2003) to increase the saving rate.

The contrast between the more anomalous decision making behavior under risk and uncertainty and the less anomalous decision making behavior over time suggests that they may have distinct underlying mechanisms. If anomalies reflect the limitation of cognitive ability, increased understanding of anomalies ought to increase the frequency of the "truly" normative response (Savage, 1954), and decision making may even call for prescriptive policies to correct their own behavior.

We find substantial differences between the OLS estimates and within-twin-pair fixed-effects estimates for each preference in our study. The results suggest that the OLS estimated correlations between education and preferences are far from causal effects. Although the relationships between demographic and socioeconomic variables, cognitive and non-cognitive skills, and risk attitudes have been extensively studied in the literature,[26] the issue of causality has rarely been addressed.

There remains a general question about determinants of preference: nature versus nurture. The emergence of neuroeconomics over the past decade has contributed to a further advance in behavioral economics as well as in experimental economics in going beyond psychological considerations in modeling and beyond revealed choice in testing implications of different models (Camerer et al., 2005). Recent twin studies suggest that genetics may contribute significantly to economic risk attitude (Cesarini et al., 2009; Zhong et al., 2009a) as well as altruistic giving in a dictator game (Cesarini et al., 2009). At the same time, association studies have been reported between well-characterized functional genes and risk attitude in Carpenter et al. (2009), Crisan et al. (2009), Dreber & Apicella (2009), Kuhnen &

---

[26]Refer to Section 2 of the literature review.

Chiao (2009), Roe et al. (2009), and Zhong et al. (2009b,c), as well as altruistic giving in the dictator game (Knafo et al., 2008; Israel et al., 2009), and reciprocal fairness in the ultimatum game (Zhong et al., 2010). Our essay reveals additional and direct evidence about an important factor, namely education, in preference formation. Naturally, it would be of interest to further explore the interaction between genetic and education or the interplay between nature and nurture in future studies. Given our findings about the malleability of preferences by education, the individual's preference may be treated as a kind of human capital which could be shaped by various inputs through a production technology.

While this essay seems to be, to our knowledge, the first study exploring the causal effect education on preferences, it has its own limitations. A main limitation of our research is the small sample size of 140 subjects (70 pairs of twins). As a field experiment on adult twins, 140 subjects seem moderate in terms of size. Follow up research using larger sample ought to help us in uncovering more robust findings. We plan to include more adult twins in designing subsequent field experiments in the future.

Finally, this essay, as an empirical study, does not attempt to model the mechanism through which education affects individuals' decision making under risk and uncertainty and involving time. We envisage that the empirical findings in our essay will inspire further research to explore the theoretical ground of the interaction between education and preference formation.

References

Allais, M. "Le comportement de lomme rationnel devant le risque: Critique des postulats et axiomes de lcole Americaine" [Rational Man Behavior in the Presence of Risk: Critique of the Postulates and Axioms of the American School]. *Econometrica*, 1953, 21, pp. 503-546.

Almond, D.; Chay, K. Y. and Lee, D. S. "The Costs of Low Birth Weight." *Quarterly Journal of Economics*, 2005, 120(3), pp. 1031-1083.

Ameriks, J.; Caplin, A.; Leahy, J. and Tyler, T. "Measuring Self-Control Problems." *American Economic Review*, 2007, 97(3), pp. 966-972.

Andersen, S.; Harrison, G. W.; Lau, M. I. and Elisabet Rutstrom, E. "Lost in State Space: Are Preferences Stable?" *International Economic Review*, 2008, 49(3), pp. 1091-1112.

Anderson, L. R. and Mellor, J. M. "Predicting Health Behaviors with an Experimental Measure of Risk Preference." *Journal of Health Economics*, 2008, 27(5), pp. 1260-1274.

Ashenfelter, O. and Rouse, C. "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins." *Quarterly Journal of Economics*, 1998, 113(1), pp. 253-284.

Ashenfelter, O. and Krueger, A. "Estimates of the Economic Return to Schooling for a New Sample of Twins." *American Economic Review*, 1994, 84(5), pp. 1157-1173.

Becker, Gary S. "Habits, Addictions, and Traditions." *Kyklos*, 1992, 45(3), pp. 327-345.

Becker, Gary S. "*Accounting for Tastes*". Cambridge and London: Harvard University Press, 1996.

Becker, G. S. and Mulligan, C. B. "The Endogenous Determination of Time Preference." *Quarterly Journal of Economics*, 1997, 112(3), pp. 729-758.

Behrman, Jere R. "Intrahousehold Distribution and the Family," *Handbook of Population and Family Economics*. Volume 1A. Eds. Rosenzweig, Mark R. and Stark, Oded. Amsterdam; New York and Oxford: Elsevier Science, North-Holland, 1997, pp. 125-87.

Behrman, Jere R. and Rosenzweig, Mark R. "Returns to Birthweight." *Review of Economics and Statistics*, 2004, 86(2), pp. 586-601.

Behrman, Jere R., and Mark R. Rosenzweig. "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?" *American Economic Review*, 2002, 92(1), pp. 323-34.

Behrman, Jere R. and Mark R. Rosenzweig. "'Ability' Biases in Schooling Returns and Twins: a Test and New Estimates." *Economics of Education Review*, 1999, 18(2):159-167.

Behrman, Jere R., Mark R. Rosenzweig and P. Taubman. "College Choice and Wages: Estimates Using Data on Female Twins." *Review of Economics and Statistics*, 1996, 73(4): 672-685.

Behrman, Jere R. and P. Taubman. "Intergenerational Transmission of Income and Wealth." *American Economic Review*, 1976, 66(2): 436-440.

Behrman, Jere R., P. Taubman, T. Wales, and Z. Hrubec. "Inter- and Intragenerational Determination of Socioeconomic Success with Special Reference to Genetic Endowment and Family and Other Environment." 1977, Mimeo. University of Pennsylvania.

Benjamin, D. J.; Choi, J. J. and Strickland, A. J. "Social Identity and Preference." *American Economic Review*, 2010, 100, pp. 1913-1928.

Benjamin, D. J.; Brown, S. A. and Shapiro, J. M. "Who Is 'Behavioral'? Cognitive Ability and Anomalous Preferences." 2006, Working Paper.

Bisin, A. and Verdier, T. "A Model of Cultural Transmission, Voting and Political Ideology." *European Journal of Political Economy*, 2000, 16, pp. 5-29.

Booij, A. S. and Van de Kuilen, G. "A Parameter-Free Analysis of the Utility of Money for the General Population under Prospect Theory." *Journal of Economic Psychology*, 2009, 30(4), pp. 651-666.

Borghans, L.; Golsteyn, B. H. H.; Heckman, J. J. and Meijers, H. "Gender, Personality Traits and the Ellsberg Paradox." *Journal of the European Economic Association*, 2009, forthcoming.

Borghans, Lex; Duckworth, Angela Lee; Heckman, James J. and ter Weel, Bas. "The Economics and Psychology of Personality Traits," NBER Working Papers No. 13810, 2008.

Bowles, S. "Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions." *Journal of Economic Literature*, 1998, 36(1), pp. 75-111.

Burks, S. V.; Carpenter, J. P.; Goette, L. and Rustichini, A. "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment." *Proceedings of the National Academy of Sciences*, 2009, 106(19), pp. 1-6.

Camerer C, Loewenstein G, Prelec D (2005) "Neuroeconomics: How Neuroscience Can Inform Economics." *Journal of Economic Literature*, 43(1), pp.9-64.

Camerer, C. and Weber, M. "Recent Developments in Modeling Preferences: Uncertainty and Ambiguity." *Journal of Risk and Uncertainty*, 1992, 5(4), pp. 325-370.

Card, D. "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics*, Volume 3A. Eds. Ashenfelter, O. and Card, D. Amsterdam; New York and Oxford: Elsevier Science, North-Holland, 1999, pp. 1801-63.

Cesarini, D.; Dawes, C.T.; Johannesson, M.; Lichtenstein, P. and Wallace, B. " Genetic Variation in Preferences for Giving and Risk-Taking." *Quarterly Journal of Economics*, 2009, 124, pp. 809-842.

Chabris, C. F.; Laibson, D.; Morris, C. L.; Schuldt, J. P. and Taubinsky, D. "Individual Laboratory-Measured Discount Rates Predict Field Behavior." *Journal of Risk and Uncertainty*, 2008, 37(2), pp. 237-269.

Chew, S. H. "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox." *Econometrica*, 1983, 51(4), pp. 1065-1092.

Chew, S. H., and Epstein, L. G. "The Structure of Preferences and Attitudes towards the Timing of the Resolution of Uncertainty." *International Economic Review*, 1989, 30(1), pp. 103-117.

Chew, S. H., and Ho, J. L. Hope: "An Empirical Study of Attitude toward the Timing of Uncertainty Resolution." *Journal of Risk and Uncertainty*, 1994, 8(3), pp. 267-288.

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

Chew S. H., and Waller W. S. "Empirical Tests of Weighted Utility Theory." *Journal of Mathematical Psychology*, 1986. 30(1), 55-72.

Crisan, L. G., Pana, S., Vulturar, R., Heilman, R. M., Szekely, R., Druga, B. "Genetic Contributions of the Serotonin Transporter to Social Learning of Fear and Economic Decision Making." *Social Cognitive and Affective Neuroscience*, 2009, forthcoming.

Croson, R. and Gneezy, U. "Gender Differences in Preferences." *Journal of Economic Literature*, 2009, 47(2), pp. 448-474..

Cunha, F. and Heckman, J. J. "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Journal of Human Resources*, 2008, 43(4), pp. 738-782.

Cunha, F. and Heckman, J. "The Technology of Skill Formation." *American Economic Review*, 2007, 97(2), pp. 31-47.

Cunha, F.; Heckman, J. J.; Schennach, S. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." 2010, *Econometrica*, forthcoming.

Cunha, F.; Heckman, J. J.; Lochner, L. and Masterov, D. V. "Interpreting the Evidence on Life Cycle Skill Formation." In Hanushek F. and Welch F. (eds.) *Handbook of the Economics of Education*, North Holland, 2006.

Dohmen, T. J.; Falk, A.; Huffman, D. and Sunde, U. "Are Risk Aversion and Impatience Related to Cognitive Ability?" *American Economic Review*, 2010, 100(3), 1238-1260.

Dreber, A., Apicella, C. L., Eisenberg, D. T. A., Garcia, J. R., Zamore, R. S., Lum, J. K. "The 7R Polymorphism in the Dopamine Receptor D4 Gene (DRD4) is Associated with Financial Risk Taking in Men." *Evolution and Human Behavior*, 2009, 30, pp. 85-92.

Ellsberg, D. "Risk, Ambiguity, and the Savage Axioms." *Quarterly Journal of Economics*, 1961, 75(4), pp. 643-669.

Ergin, H. and F. Gul "A Subjective Theory of Compound Lotteries." *Journal of Economic Theory*, 2009, 144, pp. 899-929.

Fox, C. R. and Tversky, A. "Ambiguity Aversion and Comparative Ignorance." *Quarterly Journal of Economics*, 1995, 110(3), pp. 585-603.

Frederick, S. "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives*, 2005, 19(4), pp. 25-42.

Frederick, S.; Loewenstein, G. and O' Donoghue, T. "Time Discounting and Time Preference: A Critical Review." *Journal of Economic Literature*, 2002, 40(2), pp. 351-401.

Freeman, D. and Lane, D. "A Nonstochastic Interpretation of Reported Significance Level." *Journal of Business and Economic Statistics*, 1983, 1, pp. 292-298.

Gachter, S.; Johnson, E. J.; Herrmann, A. and Johnson, E. "Individual-Level Loss Aversion in Riskless and Risky Choices." Working Paper, 2007.

Gilboa, I. and Schmeidler, D. "Maximin Expected Utility with a Non-Unique Prior." *Journal of Mathematical. Economics*, 1989, 18, pp. 141-153.

Gorseline, D. W. "*The Effect of Schooling upon Income.*" Bloomington: Indiana Univ. Press, 1932.

Grossman, M. (1975). The Correlation between Health and Schooling. In Household Production and Consumption, edited by Nestor E. Terleckyj, Studies in Income and Wealth, Vol. 40, Conference on Research in Income and Wealth. New York: Columbia University Press for National Bureau of Economic Research, 147-211.

Harrison, G. W. and Rutström, E. E. "Risk Aversion in the Laboratory." In J. C. Cox and G. W. Harrison (Ed.), *Risk Aversion in Experiments*, 2008, Emerald Group Publishing Limit.

Heckman, J. J. "The Technology and Neuroscience of Capacity Formation." *Proceedings of the National Academy of Sciences*, 2007, 104(33), pp. 13250-55.

Heckman, J. J.; Moon, S. H.; Pinto, R.; Savelyev, P. and Yavitz, A. "Analyzing Social Experiments as Implemented: A Reexamination of the Evidence from the HighScope Perry Preschool Program." *Quantitative Economics*, 2010, 1, 1-46.

Heckman, J. J.; Stixrud, J. and Urzua, S. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics*, 2006, 24(3), pp. 411-482.

Heckman, James J. and Rubinstein, Yona. "The Importance of Noncognitive Skills: Lessons from the Ged Testing Program." *American Economic Review*, 2001, 91(2), pp. 145-149.

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

Hopfensitz, A., Krawczyk, M., and van Winden, F. "Investment, Resolution of Risk, and the Role of Affect." 2008, working paper.

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

Huck, S., and Müller, W. "Allais for All: Revisiting the Paradox." Open Access publications from Tilburg University, 2009.

Israel, S., E. Lerer, et al. "The Oxytocin Receptor (OXTR) Contributes to Prosocial Fund Allocations in the Dictator Game and the Social Value Orientations Task." 2009, *PLoS ONE* 4(5): e5535

Kahneman, D. and Tversky, A. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 1979, 47(2), pp. 2632-91.

Keynes, J. M. "*A Treatise on Probability*." Cosimo Classics, 1921.

Kirby, K. N.; Godoy, R.; Reyes-Garc a, V.; Byron, E.; Apaza, L.; Leonard, W.; Perez, E.; Vadez, V. and Wilkie, D. "Correlates of Delay-Discount Rates: Evidence from Tsimane'amerindians of the Bolivian Rain Forest." *Journal of Economic Psychology*, 2002, 23(3), pp. 291-316.

Knafo, A., S. Israel, et al. "Individual Differences in Allocation of Funds in the Dictator Game Associated with Length of the Arginine Vasopressin 1a Receptor RS3 Promoter Region and Correlation between RS3 Length and Hippocampal mRNA." *Genes Brain and Behavior*, 2008, 7(3), pp266-275.

Kreps, D. M., and Porteus, E. L. "Temporal Resolution of Uncertainty and Dynamic Choice Theory." *Econometrica*, 1978, 46(1), pp. 185-200.

Kuhnen, C. M., and Chiao, J. Y. "Genetic Determinants of Financial Risk Taking." *PLoS One*, 2009, 4(2), pp. 43-62.

Kunretither, H., and Ginsberg, R. "Disaster Insurance Protection." *Public Policy Lessons*, 1978.

Laibson, D. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics*, 1997, 112(2), pp. 443-477.

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

Li, H.; Liu, P. and Zhang, J. "Estimating Returns to Education Using Twins in Urban China." 2011, *Journal of Development Economics*, forthcoming.

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

List, J. and Rasul, I. "Field Experiments in Labor Economics." NBER Working Paper No. 16062.

Loewenstein, G. "Anticipation and the Valuation of Delayed Consumption." *Economic Journal*, 1987, 3, pp. 666-684.

Lovallo, D., and Kahneman, D. "Living with Uncertainty: Attractiveness and Resolution Timing." *Journal of Behavioral Decision Making*, 2000, 13(2), 179-190.

Lusk, J. L. and Coble, K. H. "Risk Perceptions, Risk Preference, and Acceptance of Risky Food." *American Journal of Agricultural Economics*, 2005, 87(2), pp. 393-405.

Machina, M. J. "Expected Utility Analysis without the Independence Axiom." *Econometrica*, 1982, 277-323.

Machina, M. J. "Almost-Objective Uncertainty." *Economic Theory*, 2004, 24(1), pp. 1-54.

Micceri, T. "The Unicorn, the Normal Curve, and other Improbable Creatures." *Psychological Bulletin*, 1989, 105, pp. 156-166.

Miller, L.; Meyer, D. E. and Lanzetta, J. T. "Choice among Equal Expected Value Alternatives: Sequential Effects of Winning Probability Level on Risk Preferences." *Journal of Experimental Psychology*, 1969, 79(3), pp. 419-423.

Morgenstern, O. "Some Reflections on Utility." In: M. Allais and O. Hagen, Editors, *Expected Utility and the Allais Paradox*, D. Reidel, Dodrecht (1979).

Moore, E. and Eckel, C. "Measuring Ambiguity Aversion." Working Paper, 2003.

National Bureau of Statistics (NBS). "*China Statistical Yearbook*." Beijing, China Statistic Press, 2009.

Phelps, E. S., & Pollak, R. A. (1968). On second-best national saving and game-equilibrium growth. *Review of Economic Studies*, 35(2), 185-199.

Quiggin, J. "A Theory of Anticipated Utility." *Journal of Economic Behavior and Organization*, 1982, 3(4), pp. 323-343.

Roe, B. E., Tilley, M. R., Gu, H. H., Beversdorf, D. Q., Sadee, W., Haab, T. C. "Financial and Psychological Risk Attitudes Associated with Two Single Nucleotide Polymorphisms in the Nicotine Receptor (CHRNA4) Gene." *PLoS One*, 2009, e6704.

Rutter, M. "*Genes and Behavior: Nature-Nurture Interplay Explained.*" Oxford: Blackwell Publishing, 2006.

Samuelson, P.A. "A Note on Measurement of Utility." *Review of Economic Studies*, 4, 155-161.

Savage, L. J. "*The Foundations of Statistics.*" New York, Wiley, 1954.

Schmeidler, D. "Subjective Probability and Expected Utility without Additivity." *Econometrica*, 1989, 57(3), pp. 571-587.

Schubert, R.; Gysler, M.; Brown, M.; Brachinger, H. W. and Version, T. "Gender Specific Attitudes towards Risk and Ambiguity: An Experimental Investigation." Center for Economic Research, Swiss Federal Institute of Technology, working paper, 2004.

Starmer, C. "Developments in non-expected utility theory: The Hunt for a Descriptive Theory of Choice under Risk." *Journal of Economic Literature*, 2000, 38(2), pp. 332-382.

Stigler, George J. and Becker, Gary S. "De Gustibus Non Est Disputandum." *American Economic Review*, 1977, 67(2), pp. 76-90.

Tanaka, T.; Camerer, C. and Nguyen, Q. "Risk and Time Preferences: Experimental and Household Survey Data from Vietnam." *American Economic Review*, 2009, forthcoming.

Taubman, Paul. "The Determinates of Earnings: Genetics, Family and Other Environments, a Study of White Male Twins." *American Economic Review*. 1976, 66(5), pp. 858-870.

Taubman, Paul. "Earnings, Education, Genetics, and Environment." *Journal of Human Resources*, 1976, 11(4), pp. 447-461.

ter Weel, B. "The Noncognitive Determinants of Labor Market and Behavioral Outcomes: Introduction to the Symposium." *Journal of Human Resources*, 2008, 43(4), pp. 729-737.

Thaler, RH, and Benartzi, S. "Save More Tomorrow (TM): Using Behavioral Economics to Increase Employee Saving." *Journal of Political Economy*, 2004, 112(1), pp164-187.

Todd, P. E. and Wolpin, K. I. "On the Specification and Estimation of the Production Function for Cognitive Achievement." *Economic Journal*, 2003, 113(485), pp. 3-33.

Tversky, A. and Kahneman, D. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty*, 1992, 5(4), pp. 297-323.

von Neumann, J. and Morgenstern, O. "*Theory of Games and Economic Behavior*." New York: Princeton University Press, 1944.

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

Zhong, S., Israel, S., Xue, H., Ebstein, R. P., and Chew, S. H. "Monoamine Oxidase a Gene (MAOA) Associated with Attitude towards Longshot Risks." *PLoS One*, 2009b, 4(12), e8516.

Zhong, S., Israel, S., Xue, H., Sham, P. C., Ebstein, R. P., and Chew, S. H. "A Neurochemical Approach to Valuation Sensitivity over Gains and Losses." *Proceedings of the Royal Society B: Biological Sciences*, 2009c, 276, pp. 4181-4188.

Zhong, S., Israel, S., Shalev, I., Xue, H., Sham, P. C., Ebstein, R. P., and Chew, S. H. "Dopamine D4 Receptor Gene Associated with Fairness Preference in Ultimatum Game" *PLoS ONE*, 2010 5(11): e13765.

Table 1a: Summary Statistics of Educational Attainments and other Variables

Variables	# Obs.	Mean	S.D.
Education level			
Primary school	140	0.06	0.25
Middle school	140	0.19	0.40
High school	140	0.35	0.48
Technical school	140	0.16	0.37
College-and-above	140	0.23	0.42
Parental education level			
Father primary school	140	0.46	0.50
Father middle school	140	0.22	0.42
Father high school	140	0.09	0.28
Father technical school	140	0.04	0.19
Father college-and-above	140	0.19	0.40
Mother primary school	140	0.64	0.48
Mother middle school	140	0.16	0.37
Mother high school	140	0.09	0.28
Mother technical school	140	0.07	0.26
Mother college-and-above	140	0.04	0.20
Control variables			
Age	140	45.74	11.93
Male	140	0.47	0.50
Birth weight (kg)	140	2.47	0.73
Family annual income (RMB1000)	140	22.51	18.04
Health indicator (good=1)	140	0.56	0.50

Table 1b: Summary Statistics of Experimental Measures of Preferences

Variables	# Obs.	Mean	S.D.	Min	Max	% within-twin variation
Decision Making under Risk and Uncertainty						
Moderate prospect	140	0.64	0.48	0	1	37.33%
Moderate hazard	140	0.73	0.45	0	1	44.42%
Longshot prospect	140	0.57	0.50	0	1	40.83%
Longshot hazard	128	0.57	0.50	0	1	36.66%
Allais-type Behavior	122	0.18	0.38	0	1	44.37%
Ambiguity aversion	140	0.62	0.49	0	1	44.02%
Familiarity bias	140	0.79	0.41	0	1	50.91%
Decision Making Involving Time						
Impatience	140	0.54	0.50	0	1	47.39%
Hyperbolic discounting	140	0.23	0.42	0	1	52.65%
Anticipation	130	0.43	0.50	0	1	34.50%
Dread	136	0.73	0.45	0	1	42.69%
Hopefulness	128	0.41	0.49	0	1	35.62%
Anxiousness	140	0.76	0.43	0	1	54.39%

Note: Appendix I and II give experimental instruction and variable construction. The last column gives the percentage of within-twin-pair variation to the total variation for each preference variable.

Table 2: OLS and Within-Twin-Pair Fixed-Effects Estimates of Education and Fourfold Pattern of Risk Attitudes

	Dependent variables					
	Moderate prospect			Moderate hazard		
	OLS		FE	OLS		FE
	(1)	(2)	(3)	(4)	(5)	(6)
High school	-0.087	-0.099	0.081	-0.130	-0.073	-0.057
	(0.80)	(0.89)	(0.42)	(1.28)	(0.70)	(0.29)
Technical school	0.002	0.074	0.199	-0.090	0.042	0.183
	(0.02)	(0.53)	(0.87)	(0.74)	(0.32)	(0.79)
College-and-above	0.004	0.060	0.438**	-0.060	0.016	0.386*
	(0.034)	(0.43)	(2.03)	(0.54)	(0.12)	(1.77)
Parental education	No	Yes		No	Yes	
Observations	140	140	140	140	140	140
Twin pairs			70			70
R-squared	0.11	0.19	0.08	0.11	0.17	0.09

	Longshot prospect			Longshot hazard		
	OLS		FE	OLS		FE
	(7)	(8)	(9)	(10)	(11)	(12)
High school	0.094	0.129	0.202	-0.187	-0.167	-0.331
	(0.85)	(1.13)	(0.96)	(1.60)	(1.43)	(1.55)
Technical school	0.180	0.234	0.338	0.0186	0.081	-0.165
	(1.36)	(1.62)	(1.36)	(0.14)	(0.57)	(0.67)
College-and-above	0.061	0.158	0.482**	-0.085	-0.064	-0.061
	(0.50)	(1.10)	(2.06)	(0.69)	(0.44)	(0.26)
Parental education	No	Yes		No	Yes	
Observations	140	140	140	128	128	128
Twin pairs			70			64
R-squared	0.14	0.21	0.03	0.14	0.25	0.05

Note: Absolute values of t-statistics computed based on the permutation procedure (Freedman and Lane, 1983) are in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$; the omitted educational group for the individual is middle-school-and-below; age, age square, gender, and a city dummy are included in each OLS estimation. The four dependent variables are dummy ones. They equal one if the subject is risk tolerant toward moderate prospects, moderate hazard, longshot prospect, and longshot hazard; they equal zero if the subject exhibits risk aversion. See Appendix II for experimental variable construction.

Table 3: OLS and Within-Twin-Pair Fixed-Effects Estimates of Education and Decision Making Anomalies under Risk and Uncertainty

	Dependent variables								
	Allais-type behavior			Ambiguity aversion			Familiarity bias		
	OLS		FE	OLS		FE	OLS		FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
High school	0.092	0.094	0.295	0.176	0.077	0.372*	0.080	0.0923	0.319*
	(0.97)	(0.96)	(1.64)	(1.56)	(0.67)	(1.72)	(0.85)	(0.98)	(1.65)
Technical school	0.064	0.056	0.274	0.369***	0.240*	0.242	-0.160	-0.171	0.065
	(0.59)	(0.47)	(1.33)	(2.73)	(1.66)	(0.95)	(1.43)	(1.43)	(0.28)
College-and-above	0.177*	0.206*	0.417**	0.213*	0.230	0.175	-0.125	-0.093	0.007
	(1.77)	(1.73)	(2.15)	(1.73)	(1.60)	(0.73)	(1.22)	(0.78)	(0.03)
Parental education	No	Yes		No	Yes		No	Yes	
Observations	122	122	122	140	140	140	140	140	140
Twin pairs			61			70			70
R-squared	0.15	0.21	0.08	0.07	0.17	0.05	0.10	0.21	0.06

Note: Absolute values of t-statistics computed based on the permutation procedure (Freedman and Lane, 1983) are in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$; the omitted educational group for the individual is middle-school-and-below; age, age square, gender, and a city dummy are included in each OLS estimation. The three dependent variables are dummy ones. The dependent variable of Allais equals one if the subject exhibit Allais-type behavior in the experiment. It equals zero if the subject exhibit expected utility behavior. The dependent variable of ambiguity aversion equals one if the subject exhibits ambiguity aversion in the experiment; otherwise, it equals zero. The dependent variable of familiarity bias equals one if the subject exhibit familiarity bias in the experiment; otherwise, it equals zero. See Appendix II for experimental variable construction.

Table 4: OLS and Within-Twin-Pair Fixed-Effects Estimates of Education and Decision Making Involving Time

	Dependent variables								
	Impatience			Hyperbolic discounting			Anticipation		
	OLS		FE	OLS		FE	OLS		FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
High school	0.047	0.017	-0.004	-0.064	-0.055	-0.133	0.068	0.058	0.048
	(0.42)	(0.15)	(0.02)	(0.66)	(0.53)	(0.66)	(0.61)	(0.51)	(0.23)
Technical school	-0.133	-0.177	-0.413	0.010	-0.004	-0.057	0.127	0.058	0.136
	(1.00)	(1.19)	(1.56)	(0.09)	(0.03)	(0.24)	(0.98)	(0.41)	(0.56)
College-and-above	-0.211*	-0.214	-0.520**	-0.116	-0.188	-0.382*	0.290**	0.152	0.208
	(1.73)	(1.44)	(2.09)	(1.09)	(1.45)	(1.70)	(2.49)	(1.09)	(0.92)
Parental education	No	Yes		No	Yes		No	Yes	
Observations	140	140	140	140	140	140	130	130	130
Twin pairs			70			70			65
R-squared	0.13	0.16	0.10	0.07	0.10	0.07	0.23	0.31	0.02

	Dread			Hopefulness			Anxiousness		
	OLS		FE	OLS		FE	OLS		FE
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
High school	-0.301***	-0.284**	-0.220	-0.201*	-0.186	-0.225	-0.062	-0.069	-0.130
	(2.84)	(2.55)	(1.14)	(1.71)	(1.51)	(1.08)	(0.64)	(0.67)	(0.60)
Technical school	-0.215*	-0.233	-0.515**	-0.367***	-0.336**	-0.440*	0.011	-0.013	0.036
	(1.70)	(1.65)	(2.22)	(2.64)	(2.15)	(1.80)	(0.09)	(0.10)	(0.14)
College-and-above	-0.132	-0.214	-0.389*	-0.069	-0.145	-0.478**	0.079	0.044	-0.085
	(1.17)	(1.56)	(1.81)	(0.56)	(0.99)	(2.14)	(0.74)	(0.34)	(0.35)
Parental education	No	Yes		No	Yes		No	Yes	
Observations	136	136	136	128	128	128	140	140	140
Twin pairs			68			64			70
R-squared	0.08	0.11	0.07	0.14	0.19	0.07	0.11	0.14	0.01

Note: Absolute values of t-statistics computed based on the permutation procedure (Freedman and Lane, 1983) are in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$; the omitted educational group for the individual is middle-school-and-below; age, age square, gender, and a city dummy are included in each OLS estimation. The six dependent variables are dummy ones. If the subject exhibits impatience, hyperbolic discounting behavior, anticipation, dread, hopefulness, and anxiousness in the experiments, all six variables equal one. See Appendix II for experimental variable construction.

Table 5: Instrumental Variables Within-Twin-Pair Fixed-Effects Estimates of Education and Preferences

	Dependent variables						
	Decision making under uncertainty						
	Moderate prospect	Moderate hazard	Longshot prospect	Longshot hazard	Allais-type behavior	Ambiguity aversion	Familiarity bias
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
High school	0.066	0.040	0.177	-0.308	0.430**	0.450*	0.390*
	(0.32)	(0.19)	(0.80)	(1.37)	(2.21)	(1 87)	(1.94)
Technical school	0.153	0.287	0.231	-0.079	0.286	0 110	0.131
	(0.60)	(1.10)	(0.85)	(0.30)	(1.26)	(0.37)	(0.53)
College-and-above	0.461**	0.493**	0.509**	-0.128	0.521**	0.056	-0.098
	(2.00)	(2.08)	(2.07)	(0.54)	(2.52)	(0.21)	(0.44)
Observations	140	140	140	128	122	140	140
Twin pairs	70	70	70	64	61	70	70
R-squared	0.09	0.08	0.07	0.05	0.07	0.01	0.10

	Decision making involving time					
	Impatience	Hyperbolic discounting	Anticipation	Dread	Hopefulness	Anxiousness
	(8)	(9)	(10)	(11)	(12)	(13)
High school	-0.169	-0.220	0.084	-0.238	-0.205	-0.021
	(0.69)	(1.00)	(0.38)	(1.11)	(0.89)	(0.089)
Technical school	-0.511*	-0.043	0.042	-0.530*	-0.500*	0.205
	(1.71)	(0.16)	(0.15)	(2.00)	(1.80)	(0.71)
College-and-above	-0.504*	-0.244	0.186	-0.347	-0.492**	-0.028
	(1.86)	(1.00)	(0.76)	(1.47)	(2.03)	(0.11)
Observations	140	140	130	136	128	140
Twin pairs	70	70	65	68	64	70
R-squared	0.08	0.04	0.01	0.05	0.06	0.01

Note: Absolute values of t-statistics computed based on the permutation procedure (Freedman and Lane, 1983) are in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$; the omitted educational group for the individual is middle-school-and-below. The notes under Tables 2-3 give the definitions of each dependent variable.

Table 6: Within-Twin-Pair Fixed-Effects Estimates of Education and Preferences Controlling for Birth Weight

	Dependent variables						
	Decision making under uncertainty						
	Moderate prospect	Moderate hazard	Longshot prospect	Longshot hazard	Allais-type behavior	Ambiguity aversion	Familiarity bias
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
High school	0.065	-0.021	0.209	-0.284	0.293	0.369*	0.332*
	(0.33)	(0.11)	(0.98)	(1.33)	(1.60)	(1.68)	(1.68)
Technical school	0.186	0.212	0.343	-0.127	0.272	0.240	0.075
	(0.81)	(0.91)	(1.36)	(0.52)	(1.30)	(0.93)	(0.32)
College-and-above	0.415*	0.437*	0.491**	0.003	0.414**	0.171	0.025
	(1.88)	(1.97)	(2.04)	(0.013)	(2.08)	(0.69)	(0.11)
Birth weight	-0.076	0.171	0.030	0.206	-0.010	-0.013	0.061
	(0.55)	(1.24)	(0.20)	(1.46)	(0.08)	(0.09)	(0.44)
Observations	140	140	140	128	122	140	140
Twin pairs	70	70	70	64	61	70	70
R-squared	0.08	0.11	0.06	0.08	0.08	0.05	0.07

	Decision making involving time					
	Impatience	Hyperbolic discounting	Anticipation	Dread	Hopefulness	Anxiousness
	(8)	(9)	(10)	(11)	(12)	(13)
High school	0.021	-0.136	0.058	-0.228	-0.189	-0.116
	(0.092)	(0.66)	(0.27)	(1.16)	(0.90)	(0.53)
Technical school	-0.392	-0.060	0.143	-0.521**	-0.423*	0.048
	(1.47)	(0.25)	(0.58)	(2.22)	(1.73)	(0.18)
College-and-above	-0.485*	-0.387*	0.221	-0.400*	-0.437*	-0.065
	(1.91)	(1.68)	(0.95)	(1.82)	(1.93)	(0.26)
Birth weight	0.119	-0.016	0.042	-0.039	0.142	0.068
	(0.76)	(0.11)	(0.30)	(0.28)	(1.03)	(0.45)
Observations	140	140	130	136	128	140
Twin pairs	70	70	65	68	64	70
R-squared	0.11	0.07	0.02	0.08	0.09	0.02

Note: Absolute values of t-statistics computed based on the permutation procedure (Freedman and Lane, 1983) are in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$; the omitted educational group for the individual is middle-school-and-below. The notes under Tables 2-3 give the definitions of each dependent variable.

Table 7: Within-Twin-Pair Fixed-Effects Estimates of Education and Preferences Using MZ Twins Pairs

	Dependent variables						
	Decision making under uncertainty						
	Moderate prospect	Moderate hazard	Longshot prospect	Longshot hazard	Allais-type behavior	Ambiguity aversion	Familiarity bias
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
High school	0.115	-0.141	0.128	-0.769***	0.308	0.359	0.641**
	(0.48)	(0.58)	(0.45)	(3.10)	(1.40)	(1.18)	(2.55)
Technical school	0.269	0.449	0.410	-0.462	0.385	-0.051	0.051
	(0.96)	(1.57)	(1.22)	(1.57)	(1.48)	(0.14)	(0.17)
College-and-above	0.462*	0.325	0.402	-0.077	0.231	-0.009	0.009
	(1.96)	(1.36)	(1.43)	(0.31)	(1.06)	(0.03)	(0.03)
Observations	72	72	72	68	64	72	72
Twin pairs	36	36	36	34	32	36	36
R-squared	0.12	0.16	0.07	0.29	0.09	0.07	0.22

	Decision making involving time					
	Impatience	Hyperbolic discounting	Anticipation	Dread	Hopefulness	Anxiousness
	(8)	(9)	(10)	(11)	(12)	(13)
High school	-0.026	-0.282	0.115	-0.064	-0.100	-0.179
	(0.08)	(1.22)	(0.43)	(0.24)	(0.37)	(0.64)
Technical school	-0.282	-0.103	0.269	-0.705**	-0.467	0.026
	(0.74)	(0.38)	(0.86)	(2.25)	(1.42)	(0.08)
College-and-above	-0.547*	-0.350	0.128	-0.368	-0.411	0.171
	(1.71)	(1.54)	(0.49)	(1.40)	(1.58)	(0.62)
Observations	72	72	70	70	66	72
Twin pairs	36	36	35	35	33	36
R-squared	0.11	0.09	0.02	0.17	0.10	0.05

Note: Absolute values of t-statistics computed based on the permutation procedure (Freedman and Lane, 1983) are in parentheses; *** p<0.01, ** p<0.05, * p<0.1; the omitted educational group for the individual is middle-school-and-below. The notes under Tables 2-3 give the definitions of each dependent variable.

Table 8: Within-Twin-Pair Fixed-Effects Estimates of Education and Preferences Controlling for Income

	Dependent variables: Decision making under uncertainty						
	Moderate prospect (1)	Moderate hazard (2)	Longshot prospect (3)	Longshot hazard (4)	Allais-type behavior (5)	Ambiguity aversion (6)	Familiarity bias (7)
High school	· 0.046	-0.056	0.219	-0.309	0.240	0.372*	0.342*
	(0.24)	(0.28)	(1.03)	(1.42)	(1.33)	(1.70)	(1.75)
Technical school	0.135	0.185	0.368	-0.133	0.191	0.242	0.107
	(0.59)	(0.77)	(1.44)	(0.53)	(0.92)	(0.92)	(0.46)
College-and-above	0.404*	0.387*	0.498**	-0.042	0.368*	0.175	0.029
	(1.88)	(1.75)	(2.10)	(0.18)	(1.91)	(0.72)	(0.13)
Family annual income	0.150	-0.003	-0.069	-0.062	0.161	0.001	-0.099
	(1.49)	(0.03)	(0.62)	(0.53)	(1.67)	(0.00)	(0.97)
Observations	140	140	140	128	122	140	140
Twin pairs	70	70	70	64	61	70	70
R-squared	0.11	0.09	0.07	0.06	0.12	0.05	0.08

	Decision making involving time					
	Impatience (8)	Hyperbolic discounting (9)	Anticipation (10)	Dread (11)	Hopefulness (12)	Anxiousness (13)
High school	0.005	-0.145	0.023	-0.222	-0.166	-0.091
	(0.021)	(0.71)	(0.11)	(1.13)	(0.79)	(0.42)
Technical school	-0.396	-0.080	0.110	-0.518**	-0.366	0.106
	(1.46)	(0.33)	(0.44)	(2.20)	(1.48)	(0.41)
College-and-above	-0.511**	-0.395*	0.191	-0.390*	-0.432*	-0.048
	(2.03)	(1.73)	(0.83)	(1.79)	(1.94)	(0.20)
Family annual income	-0.039	0.054	0.064	0.009	-0.162	-0.165
	(0.34)	(0.51)	(0.59)	(0.09)	(1.45)	(1.47)
Observations	140	140	130	136	128	140
Twin pairs	70	70	65	68	64	70
R-squared	0.10	0.07	0.02	0.07	0.11	0.05

Note: Absolute values of t-statistics computed based on the permutation procedure (Freedman and Lane, 1983) are in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$; the omitted educational group for the individual is middle-school-and-below. The notes under Tables 2-3 give the definitions of each dependent variable.

Table 9: Within-Twin-Pair Fixed-Effects Estimates of Education and Preferences Controlling for Health

	Dependent variables Decision making under uncertainty						
	Moderate prospect	Moderate hazard	Longshot prospect	Longshot hazard	Allais-type behavior	Ambiguity aversion	Familiarity bias
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
High school	0.083	-0.060	0.202	-0.333	0.292	0.375*	0.318
	(0.43)	(0.31)	(0.95)	(1.55)	(1.62)	(1.73)	(1.62)
Technical school	0.196	0.193	0.339	-0.160	0.282	0.233	0.068
	(0.85)	(0.83)	(1.35)	(0.65)	(1.37)	(0.91)	(0.29)
College-and-above	0.436**	0.390*	0.483**	-0.059	0.420**	0.171	0.008
	(2.01)	(1.78)	(2.04)	(0.25)	(2.17)	(0.71)	(0.038)
Health indicator	-0.052	0.133	0.008	0.073	0.124	-0.126	0.044
	(0.37)	(0.93)	(0.052)	(0.49)	(0.98)	(0.80)	(0.31)
Observations	140	140	140	128	122	140	140
Twin pairs	70	70	70	64	61	70	70
R-squared	0.08	0.10	0.06	0.06	0.09	0.05	0.06

	Decision making involving time					
	Impatience	Hyperbolic discounting	Anticipation	Dread	Hopefulness	Anxiousness
	(8)	(9)	(10)	(11)	(12)	(13)
High school	-0.007	-0.136	0.054	-0.214	-0.225	-0.133
	(0.031)	(0.67)	(0.26)	(1.12)	(1.09)	(0.62)
Technical school	-0.406	-0.048	0.122	-0.530**	-0.440*	0.045
	(1.53)	(0.20)	(0.50)	(2.30)	(1.82)	(0.18)
College-and-above	-0.518**	-0.379*	0.203	-0.395*	-0.478**	-0.082
	(2.07)	(1.68)	(0.90)	(1.85)	(2.16)	(0.34)
Health indicator	0.094	0.123	-0.183	-0.195	0.214	0.128
	(0.58)	(0.84)	(1.24)	(1.41)	(1.40)	(0.82)
Observations	140	140	130	136	128	140
Twin pairs	70	70	65	68	64	70
R-squared	0.11	0.08	0.04	0.10	0.10	0.02

Note: Absolute values of t-statistics computed based on the permutation procedure (Freedman and Lane, 1983) are in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$; the omitted educational group for the individual is middle-school-and-below. The notes under Tables 2-3 give the definitions of each dependent variable.

Appendix for Essay One

Early Health Shocks, Parental Responses, and Child Outcomes

1 Comparative Statics of the Effects of an Early Health Shock on Parental Investments with CES Production Technology

In the paper, we assume a Cobb-Douglas functional form for the technology of production of health and cognitive skills in the second period (Equations (19)-(20)). Here we derive the comparative static results assuming the following more general CES functional form:

$$\theta^H_{i,2} = \{\gamma(\theta^C_{i,1})^{\psi} + (1-\gamma)\,[\beta_\theta\theta^H_{i,1} + \beta_I I^H_{i,1}]^{\psi}\}^{\frac{1}{\psi}}, \quad (1)$$

$$\theta^C_{i,2} = \{\gamma(\theta^H_{i,1})^{\psi} + (1-\gamma)\,[\beta_\theta\theta^C_{i,1} + \beta_I I^C_{i,1}]^{\psi}\}^{\frac{1}{\psi}}, \quad (2)$$

where $\psi < 1$, and it measures the complementarity or substitutability between $\theta^C_{i,1}$ and $\theta^H_{i,1}$ ($I^H_{i,1}$) in producing $\theta^H_{i,2}$, and between $\theta^H_{i,1}$ and $\theta^C_{i,1}$ ($I^C_{i,1}$) in producing $\theta^C_{i,2}$. When $\psi \to 0$, we are back to Equations (19)-(20) in the paper. The rest of the maximization problem is the same as before, and we derive the optimal investment in the health and cognitive skills of child i as:

$$I^{H*}_{i,1} = \frac{\alpha_H}{\beta_I} W\pi_i\phi^H_i - \frac{\beta_\theta}{\beta_I}\theta^H_{i,1}, \quad (3)$$

$$I^{C*}_{i,1} = \frac{\alpha_C}{\beta_I} W\pi_i\phi^C_i - \frac{\beta_\theta}{\beta_I}\theta^C_{i,1}, \quad (4)$$

where:

$$\phi^H_i = \frac{(1-\gamma)\,[\beta_\theta\theta^H_{i,1} + (\beta_I)I^H_{i,1}]^{\psi}}{\gamma(\theta^C_{i,1})^{\psi} + (1-\gamma)\,[\beta_\theta\theta^H_{i,1} + (\beta_I)I^H_{i,1}]^{\psi}},$$

$$\phi^C_i = \frac{(1-\gamma)\,[\beta_\theta\theta^C_{i,1} + (\beta_I)I^C_{i,1}]^{\psi}}{\gamma(\theta^H_{i,1})^{\psi} + (1-\gamma)\,[\beta_\theta\theta^C_{i,1} + (\beta_I)I^C_{i,1}]^{\psi}}.$$

and W and π_i are defined as before. We notice that there are two additional terms, ϕ^H_i and ϕ^C_i, in Equations (3)-(4), in comparison to Equations (21)-(22) in the paper. The first term, ϕ^H_i, can be interpreted as the importance of $I^H_{i,1}$ in producing $\theta^H_{i,2}$, relative to $\theta^C_{i,1}$; similarly, ϕ^C_i measures the importance of $I^C_{i,1}$ in producing $\theta^C_{i,2}$, relative to $\theta^H_{i,1}$. We notice that the signs of $\partial\phi^H_i/\partial\theta^H_{i,1}$ and $\partial\phi^C_i/\partial\theta^H_{i,1}$ are determined by ψ. Heckman (2007) argues that $\theta^C_{i,1}$ and $\theta^H_{i,1}$ (and $I^H_{i,1}$) are complements in producing $\theta^H_{i,2}$, and $\theta^H_{i,1}$ and $\theta^C_{i,1}$ (and $I^C_{i,1}$) are complements in producing $\theta^C_{i,2}$. In this case, $\psi < 0$, $\partial\phi^H_i/\partial\theta^H_{i,1} < 0$, and $\partial\phi^C_i/\partial\theta^H_{i,1} > 0$. This means that an increase in the health stock in the first period reduces investment in health and increases investment in cognitive skills in the second period, by changing the relative importance of health (ϕ_i^H) and cognitive skills (ϕ_i^C), respectively. We also notice that $\partial\phi_j^H/\partial\theta_{i,1}^H = 0$ and $\partial\phi_j^C/\partial\theta_{i,1}^H = 0$ ($j \neq i$). This means that, although an increase in the health stock of child i in the first period changes the share of resources allocated to child j through $W\pi_j$, it does not change the relative importance of health and cognitive skills in the production function, i.e., ϕ_j^H and ϕ_j^C are unaffected. The new comparative static results for the effect of health in the first period on the investment in health and cognitive skills in the second period are:

$$\frac{\partial I_{i,1}^{H*}}{\partial\theta_{i,1}^H} = \frac{\alpha_H}{\beta_I}\left(\frac{\partial W}{\partial\theta_{i,1}^H}\pi_i\phi_i^H + \frac{\partial\pi_i}{\partial\theta_{i,1}^H}W\phi_i^H + \frac{\partial\phi_i^H}{\partial\theta_{i,1}^H}W\pi_i\right) - \frac{\beta_\theta}{\beta_I}, \tag{5}$$

$$\frac{\partial I_{i,1}^{C*}}{\partial\theta_{i,1}^H} = \frac{\alpha_C}{\beta_I}\left(\frac{\partial W}{\partial\theta_{i,1}^H}\pi_i\phi_i^C + \frac{\partial\pi_i}{\partial\theta_{i,1}^H}W\phi_i^C + \frac{\partial\phi_i^C}{\partial\theta_{i,1}^H}W\pi_i\right). \tag{6}$$

We notice that there are four channels through which $\theta_{i,1}^H$ affects $I_{i,2}^{H*}$. As discussed in the paper, $\partial W/\partial\theta_{i,1}^H$ is a *wealth effect*, which is always positive, $\partial\pi_i/\partial\theta_{i,1}^H$ is a *price effect*, whose sign is determined by parental inequality aversion, and $-\beta_\theta/\beta_I$ is what we defined a *technological effect*; the new term $\partial\phi_i^H/\partial\theta_{i,1}^H$ is instead a by-product of the CES specification of the production function (Equations (1)-(2)). If parents weight efficiency more than equality ($\rho > 0$) and $\theta_{i,1}^C$ and $\theta_{i,1}^H$ (and $I_{i,1}^H$) are complements ($\psi < 0$), then $\partial\pi_i/\partial\theta_{i,1}^H$ is positive and $\partial\phi_i^H/\partial\theta_{i,1}^H$ is negative. Thus, the total effect of $\theta_{i,1}^H$ on $I_{i,1}^{H*}$ is indetermined. In contrast, the total effect of $\theta_{i,1}^H$ on $I_{i,1}^{C*}$ is positive when $\rho > 0$ and $\psi < 0$. This is because both $\partial W/\partial\theta_{i,1}^H$, $\partial\pi_i/\partial\theta_{i,1}^H$ and $\partial\phi_i^C/\partial\theta_{i,1}^H$ are positive.

Then, the cross effects of child i's health in the first period on investment in child j's ($j \neq i$) health and cognitive skills in the second period are:

$$\frac{\partial I_{j,1}^{H*}}{\partial\theta_{i,1}^H} = \frac{\alpha_H}{\beta_I}\left(\frac{\partial W}{\partial\theta_{i,1}^H}\pi_j\phi_j^H + \frac{\partial\pi_j}{\partial\theta_{i,1}^H}W\phi_j^H\right), \tag{7}$$

$$\frac{\partial I_{j,1}^{C*}}{\partial\theta_{i,1}^H} = \frac{\alpha_C}{\beta_I}\left(\frac{\partial W}{\partial\theta_{i,1}^H}\pi_j\phi_j^C + \frac{\partial\pi_j}{\partial\theta_{i,1}^H}W\phi_j^C\right). \tag{8}$$

Recall that $\partial\pi_j/\partial\theta_{i,1}^H$ has the opposite sign as $\partial\pi_i/\partial\theta_{i,1}^H$, and that we assume $\pi_j = \pi_i$, $\phi_i^H = \phi_j^H$, and $\phi_i^C = \phi_j^C$ before the occurrence of the health shock.

Subtracting pairwise equations (5)-(8), we obtain:

$$\frac{\partial I_{i,1}^{H*}}{\partial \theta_{i,1}^{H}} - \frac{\partial I_{j,1}^{H*}}{\partial \theta_{i,1}^{H}} = \frac{\alpha_H}{\beta_I}\left[\left(\frac{\partial \pi_i}{\partial \theta_{i,1}^{H}} - \frac{\partial \pi_j}{\partial \theta_{i,1}^{H}}\right) W\phi_i^H + \frac{\partial \phi_i^H}{\partial \theta_{i,1}^{H}} W\pi_i\right] - \frac{\beta_\theta}{\beta_I}, \tag{9}$$

$$\frac{\partial I_{i,1}^{C*}}{\partial \theta_{i,1}^{H}} - \frac{\partial I_{j,1}^{C*}}{\partial \theta_{i,1}^{H}} = \frac{\alpha_C}{\beta_I}\left[\left(\frac{\partial \pi_i}{\partial \theta_{i,1}^{H}} - \frac{\partial \pi_j}{\partial \theta_{i,1}^{H}}\right) W\phi_i^C + \frac{\partial \phi_i^C}{\partial \theta_{i,1}^{H}} W\pi_i\right]. \tag{10}$$

When efficiency out-weights equality when parents make investment decisions ($\rho > 0$) and $\theta_{i,1}^C$ and $\theta_{i,1}^H$ (and $I_{i,1}^H$) are complements ($\psi < 0$), $\partial I_{i,1}^{C*}/\partial \theta_{i,1}^H - \partial I_{j,1}^{C*}/\partial \theta_{i,1}^H$ is positive, while the sign of $\partial I_{i,1}^{H*}/\partial \theta_{i,1}^H - \partial I_{j,1}^{H*}/\partial \theta_{i,1}^H$ is undetermined. Since an early health shock negatively affects health ($\partial \theta_{i,1}^H/\partial e_{i,1}^H < 0$), the within-twin-pair fixed-effects estimate of the effect of an early health shock on investment in cognitive skills is predicted to be negative, while that of its effect on investment in health is undetermined, as determined by the trade-off between the degree of inequality aversion of the parents, the complementarity between investment in cognitive skills and health, and the substitutability between health in the first period and investment in health in the second period.

2 Derivation of the First Order Conditions

Under the assumption of a binding budget constraint, from equation (14) we have:

$$I^C_{j,1} = I - I^C_{i,1} - I^H_{i,1} - I^H_{j,1}.$$

Substituting the specific functional forms for the child quality function, the production technologies and $I^C_{j,1}$ into Equation (17) in the paper, and taking the first-order derivative with respect to $I^H_{i,1}$, we obtain:

$$\alpha_H(1-\gamma)(\beta_I)(\theta^C_{i,1})^{\gamma}[\beta_\theta\theta^H_{i,1} + \beta_I I^H_{i,1}]^{-\gamma}(\theta^H_{i,2})^{\alpha_H\rho-1}(\theta^C_{i,2})^{\alpha_C\rho}\left[\sum_{k=1}^{2}(\theta^H_{k,2})^{\alpha_H\rho}(\theta^C_{k,2})^{\alpha_C\rho}\right]^{\frac{1}{\rho}-1} =$$

$$\alpha_C(1-\gamma)\beta_I(\theta^H_{j,1})^{\gamma}[\beta_\theta\theta^C_{j,1} + \beta_I I^C_{j,1}]^{-\gamma}(\theta^C_{j,2})^{\alpha_C\rho-1}(\theta^H_{j,2})^{\alpha_H\rho}\left[\sum_{k=1}^{2}(\theta^H_{k,2})^{\alpha_H\rho}(\theta^C_{k,2})^{\alpha_C\rho}\right]^{\frac{1}{\rho}-1}.$$

This can be simplified to:

$$\frac{\alpha_H\left[\beta_\theta\theta^C_{j,1} + \beta_I I^{C*}_{j,1}\right]\pi_i}{\alpha_C\pi_j} = [\beta_\theta\theta^H_{i,1} + \beta_I I^{H*}_{i,1}], \tag{11}$$

where π_i is defined as in Equation (24) in the paper. Similarly, the first order derivative with respect to $I^C_{i,1}$ is:

$$\frac{\left[\beta_\theta\theta^C_{j,1} + \beta_I I^{C*}_{j,1}\right]\pi_i}{\pi_j} = \left[\beta_\theta\theta^C_{i,1} + \beta_I I^{C*}_{i,1}\right]. \tag{12}$$

Summing up Equations (11)-(12) and using the fact that $\pi_i + \pi_j = 1$, we get:

$$\frac{\left[\beta_\theta\theta^C_{j,1} + \beta_I I^{C*}_{j,1}\right]}{\alpha_C\pi_j} = W,$$

where W is defined as in Equation (22) in the paper. Rearranging yields:

$$I^{C*}_{j,1} = \frac{\alpha_C}{\beta_I}W\pi_j - \frac{\beta_\theta}{\beta_I}\theta^C_{j,1}.$$

By symmetry, rewriting:

$$I^C_{i,1} = I - I^C_{j,1} - I^H_{i,1} - I^H_{j,1},$$

and repeating the above steps, we obtain:

$$I_{i,1}^{C*} = \frac{\alpha_C}{\beta_I} W \pi_i - \frac{\beta_\theta}{\beta_I} \theta_{i,1}^C.$$

Along the same lines, we obtain:

$$\begin{aligned} I_{i,1}^{H*} &= \frac{\alpha_H}{\beta_I} W \pi_i - \frac{\beta_\theta}{\beta_I} \theta_{i,1}^H, \\ I_{j,1}^{H*} &= \frac{\alpha_H}{\beta_I} W \pi_j - \frac{\beta_\theta}{\beta_I} \theta_{j,1}^H. \end{aligned}$$

3 The Price Effect of an Early Health Shock

The first order derivative of π_i with respect to $\theta^H_{i,1}$ in Equation (24) is:

$$\frac{\partial \pi_i}{\partial \theta^H_{i,1}} = \rho \frac{V_i^{\rho-1} \frac{\partial V_i}{\partial \theta^H_{i,1}} V_j^\rho - V_j^{\rho-1} \frac{\partial V_j}{\partial \theta^H_{i,1}} V_i^\rho}{U^{2\rho}},$$

where $V_k = V(\theta^H_{k,2}, \theta^C_{k,2})$ $(k = i, j)$ is defined in Equation (18). By simple manipulation, we obtain:

$$\frac{\partial \pi_i}{\partial \theta^H_{i,1}} = \rho \frac{(V_i V_j)^\rho}{\theta^H_{i,1}} \cdot \frac{(\varepsilon_{V_i,\theta^H_{i,1}} - \varepsilon_{V_j,\theta^H_{i,1}})}{U^{2\rho}},$$

where $\varepsilon_{V_i,\theta^H_{i,1}}$ is the own elasticity of child quality with respect to health in period one, and $\varepsilon_{V_j,\theta^H_{i,1}}$ is the cross-elasticity of child quality with respect to health in period one. It is reasonable to assume that $\varepsilon_{V_i,\theta^H_{i,1}} - \varepsilon_{V_j,\theta^H_{i,1}} > 0$.[1] In this case, the sign of the own price effect, $\partial \pi_i / \partial \theta^H_{i,1}$, is determined by the parental inequality aversion parameter ρ. Since $\pi_i + \pi_j = 1$, the cross-price effect, $\partial \pi_j / \partial \theta^H_{i,1}$, has the opposite sign as the own price effect.

[1]This assumption derives from the parental utility function: if parents have symmetric preference, i.e., $U(V_1, V_2) = U(V_2, V_1)$, then it is automatically satisfied. The assumption of symmetric preference is also invoked in Behrman, Pollak, and Taubman (1982).

4 The Fertility Effects of Early Health Shocks

Table 1: OLS Estimates of the Effects of an Early Health Shock at Ages 0-1 on Fertility

	Dependent variable: Whether the mother has a second child					
Early health shock	-0.175***	-0.126**	-0.128**	-0.126**	-0.123**	-0.121*
(age 0-1)	[0.057]	[0.054]	[0.057]	[0.055]	[0.054]	[0.055]
Age		0.053***	0.075***	0.071***	0.071***	0.070***
	[0.005]	[0.007]	[0.007]	[0.007]	[0.007]	[0.007]
Mother's age			-0.024***	-0.020***	-0.021***	-0.020***
			[0.005]	[0.005]	[0.005]	[0.005]
Mother's education				-0.018***	-0.018***	-0.018***
				[0.006]	[0.006]	[0.006]
Gender (boy=1)					-0.027	-0.023
					[0.032]	[0.032]
Birth weight (kg): <2						-0.053
						[0.069]
Birth weight (kg): 2-2.5						0.112
						[0.083]
Birth weight (kg): 2.5-3						-0.022
						[0.034]
# Observations	665	665	665	665	665	665

Source: The sample is the comparison group (non-twin households) in CCTS. It is restricted to rural households with one or two children because the one-child policy is strictly implemented in urban areas (Family Planning Commission of Yunnan Province, 2003). Notes: Robust standard errors are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

5 The Distribution of Early Health Shocks

Table 2: Distribution of Health Shocks at Ages 0-3

# Early Health Shocks	Freq.	Percent
0	2670	91.41
1	210	7.19
2	38	1.30
3	3	0.1

6 The Determinants of Early Health Shocks

Table 3: OLS and Within-Twin-Pair FE Results of the Determinants of Early Health Shocks

	Dependent variable: Early health shocks				
	OLS				Within-twin-pair FE
Birth weight(kg): <2	0.04	0.046*	0.04	0.04	0.01
	[0.026]	[0.027]	[0.026]	[0.026]	[0.026]
Birth weight(kg): 2-2.5	0.01	0.01	0.01	0.01	0.00
	[0.019]	[0.019]	[0.019]	[0.019]	[0.021]
Birth weight(kg): >3	-0.01	-0.01	-0.01	-0.01	-0.02
	[0.018]	[0.018]	[0.018]	[0.018]	[0.018]
Gender (boy=1)		0.039***	0.040***	0.040***	0.020**
		[0.013]	[0.013]	[0.013]	[0.010]
Age		0.00	0.00	0.00	
		[0.002]	[0.003]	[0.003]	
Birth order		0.02	0.02	0.03	
		[0.031]	[0.031]	[0.031]	
# Siblings		-0.050*	-0.04	-0.04	
		[0.025]	[0.026]	[0.026]	
Mother's age			0.00	0.00	
			[0.002]	[0.002]	
Mother's education			0.005**	0.005*	
			[0.003]	[0.003]	
Per capital family income			0.00	0.00	
			[0.003]	[0.003]	
Own washing machine			-0.01	-0.01	
			[0.016]	[0.016]	
Own refridgerator			0.01	0.00	
			[0.018]	[0.019]	
Own cell phone			-0.01	-0.01	
			[0.017]	[0.018]	
Mother working in public sector				0.01	
				[0.032]	
Rural				-0.01	
				[0.017]	
# Observations	2922	2922	2922	2922	1461
R-squared	0.00	0.01	0.02	0.02	0.01

Source: CCTS. Notes: Robust standard errors, clustered at the household level, are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%.

7 OLS Estimation Results

Table 4: OLS Estimates of the Determinants of Family Investments (Whole Sample)

	Health	Education	Clothing	Parents home tutor
Early health shocks	1.054***	0.021	0.030	0.784
	[0.153]	[0.056]	[0.086]	[1.644]
Birth weight(kg): <2	0.440**	-0.076	-0.079	0.004
	[0.189]	[0.078]	[0.113]	[1.694]
Birth weight(kg): 2-2.5	0.255*	0.054	-0.016	-1.860
	[0.150]	[0.060]	[0.090]	[1.315]
Birth weight (kg): 2.5-3	0.199	-0.046	-0.052	-0.955
	[0.143]	[0.059]	[0.087]	[1.243]
Gender (boy=1)	0.233**	-0.036	-0.023	1.394*
	[0.094]	[0.038]	[0.054]	[0.833]
Age	-0.037*	0.088***	0.049***	-2.062***
	[0.021]	[0.009]	[0.014]	[0.192]
Birth order	0.182	0.040	-0.045	-0.706
	[0.264]	[0.096]	[0.118]	[2.420]
# Siblings	-0.553**	-0.088	-0.083	-4.084*
	[0.244]	[0.089]	[0.100]	[2.142]
Mother's age	0.001	0.015***	-0.020**	-0.209*
	[0.014]	[0.006]	[0.008]	[0.127]
Mother's education	0.081***	0.026***	0.016	0.874***
	[0.022]	[0.009]	[0.014]	[0.195]
Per capital family income	0.045**	0.056***	0.087***	-0.609***
	[0.022]	[0.012]	[0.013]	[0.198]
Own washing machine	0.244*	0.100*	0.303***	2.704**
	[0.130]	[0.055]	[0.086]	[1.110]
Own refrigerator	0.165	0.097	0.064	1.834
	[0.153]	[0.065]	[0.091]	[1.309]
Own cell phone	0.024	0.249***	0.390***	2.762**
	[0.135]	[0.059]	[0.079]	[1.211]
Mother working in public sector	-0.356	0.045	0.153	-1.006
	[0.241]	[0.084]	[0.108]	[2.013]
Rural	0.201	-0.212***	-0.009	-4.558***
	[0.140]	[0.056]	[0.081]	[1.191]
# Observations	2922	2922	2922	2902
R-squared	0.071	0.296	0.182	0.202

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 5: OLS Estimates of the Determinants of Family Investments (Rural Sample)

	Health	Education	Clothing	Parents home tutor
Early health shocks	0.861***	0.032	-0.068	1.112
	[0.209]	[0.075]	[0.147]	[2.218]
Birth weight(kg): <2	0.214	-0.054	-0.022	-0.889
	[0.253]	[0.082]	[0.138]	[2.234]
Birth weight(kg): 2-2.5	0.420**	0.040	-0.010	-2.522
	[0.185]	[0.065]	[0.118]	[1.644]
Birth weight (kg): 2.5-3	0.337*	-0.054	0.033	-0.948
	[0.173]	[0.063]	[0.112]	[1.529]
Gender (boy=1)	0.208*	-0.054	-0.032	1.252
	[0.115]	[0.043]	[0.063]	[1.043]
Age	-0.032	0.104***	0.059***	-1.947***
	[0.026]	[0.011]	[0.016]	[0.248]
Birth order	0.241	0.063	-0.096	-2.134
	[0.283]	[0.108]	[0.124]	[2.518]
# Siblings	-0.563**	-0.123	-0.060	-2.524
	[0.263]	[0.104]	[0.110]	[2.222]
Mother's age	-0.001	0.013**	-0.005	-0.115
	[0.017]	[0.006]	[0.009]	[0.165]
Mother's education	0.059**	0.031**	0.035**	0.945***
	[0.028]	[0.013]	[0.017]	[0.278]
Per capital family income	0.067**	0.067***	0.100***	-0.180
	[0.026]	[0.015]	[0.017]	[0.246]
Own washing machine	0.119	0.105**	0.184**	2.060
	[0.145]	[0.050]	[0.081]	[1.312]
Own refrigerator	-0.253	0.107	0.153	4.624**
	[0.231]	[0.077]	[0.101]	[2.073]
Own cell phone	0.048	0.261***	0.312***	3.344*
	[0.183]	[0.069]	[0.088]	[1.889]
Mother working in public sector	0.217	0.082	0.131	-0.510
	[0.412]	[0.146]	[0.158]	[4.733]
Rural				
# Observations	1546	1546	1546	1528
R-squared	0.058	0.336	0.193	0.154

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 6: OLS Estimates of the Determinants of Family Investments (Urban Sample)

	Health	Education	Clothing	Parents home tutor
Early health shocks	1.163***	0.013	0.123	0.784
	[0.217]	[0.081]	[0.094]	[2.359]
Birth weight(kg): <2	0.620**	-0.081	-0.137	1.096
	[0.282]	[0.142]	[0.181]	[2.600]
Birth weight(kg): 2-2.5	0.084	0.082	0.004	-0.875
	[0.246]	[0.115]	[0.140]	[2.152]
Birth weight (kg): 2.5-3	0.008	-0.024	-0.137	-0.788
	[0.239]	[0.115]	[0.139]	[2.077]
Gender (boy=1)	0.258*	-0.008	0.001	1.565
	[0.148]	[0.064]	[0.089]	[1.310]
Age	-0.040	0.071***	0.037	-2.252***
	[0.034]	[0.016]	[0.023]	[0.302]
Birth order	0.025	-0.018	0.108	4.403
	[0.756]	[0.189]	[0.353]	[6.895]
# Siblings	-0.734	0.089	-0.232	-8.902
	[0.693]	[0.167]	[0.279]	[5.937]
Mother's age	0.004	0.018*	-0.035**	-0.267
	[0.023]	[0.010]	[0.015]	[0.203]
Mother's education	0.091***	0.021	-0.006	0.751***
	[0.033]	[0.013]	[0.021]	[0.274]
Per capital family income	0.021	0.051***	0.081***	-0.794***
	[0.030]	[0.016]	[0.017]	[0.268]
Own washing machine	0.656**	0.047	0.507**	2.731
	[0.268]	[0.143]	[0.213]	[2.105]
Own refrigerator	0.354*	0.103	0.047	0.624
	[0.208]	[0.097]	[0.139]	[1.749]
Own cell phone	0.039	0.230***	0.407***	2.089
	[0.189]	[0.088]	[0.115]	[1.579]
Mother working in public sector	-0.510*	0.073	0.212	-0.354
	[0.284]	[0.100]	[0.129]	[2.293]
Rural				
# Observations	1376	1376	1376	1374
R-squared	0.090	0.159	0.138	0.180

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 7: OLS Estimates of the Determinants of Family Investments (Male Sample)

	Health	Education	Clothing	Parents home tutor
Early health shocks	0.838***	-0.051	-0.024	4.975*
	[0.209]	[0.087]	[0.127]	[2.612]
Birth weight(kg): <2	0.468	-0.235*	-0.210	-2.006
	[0.310]	[0.130]	[0.212]	[3.097]
Birth weight(kg): 2-2.5	0.105	-0.145*	-0.093	-2.571
	[0.223]	[0.080]	[0.150]	[2.239]
Birth weight (kg): 2.5-3	0.004	-0.138*	-0.051	-0.656
	[0.218]	[0.078]	[0.146]	[2.084]
Gender (boy=1)				
Age	-0.042	0.085***	0.073***	-1.969***
	[0.033]	[0.014]	[0.025]	[0.342]
Birth order	0.698	0.047	-0.071	3.090
	[0.458]	[0.151]	[0.203]	[3.721]
# Siblings	-1.033**	-0.113	-0.068	-5.654*
	[0.440]	[0.141]	[0.168]	[3.332]
Mother's age	-0.004	0.019**	-0.037**	-0.172
	[0.022]	[0.009]	[0.019]	[0.238]
Mother's education	0.058*	0.036***	-0.009	1.019***
	[0.035]	[0.014]	[0.024]	[0.344]
Per capital family income	0.019	0.035**	0.078***	-0.658**
	[0.030]	[0.017]	[0.019]	[0.286]
Own washing machine	0.413*	0.081	0.310**	2.735
	[0.215]	[0.081]	[0.145]	[1.933]
Own refrigerator	0.299	0.208**	0.226	-0.740
	[0.247]	[0.102]	[0.152]	[2.343]
Own cell phone	-0.151	0.180**	0.192	2.708
	[0.223]	[0.085]	[0.119]	[2.167]
Mother working in public sector	-0.403	-0.309***	0.209	-8.016**
	[0.350]	[0.118]	[0.157]	[3.373]
Rural	0.043	-0.293***	-0.125	-6.930***
	[0.228]	[0.098]	[0.123]	[2.073]
# Observations	1082	1082	1082	1078
R-squared	0.072	0.324	0.172	0.184

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 8: OLS Estimates of the Determinants of Family Investments (Female Sample)

	Health	Education	Clothing	Parents home tutor
Early health shocks	1.337***	0.139	0.018	-2.647
	[0.296]	[0.102]	[0.136]	[2.669]
Birth weight(kg): <2	0.358	0.039	-0.123	2.904
	[0.323]	[0.116]	[0.179]	[2.524]
Birth weight(kg): 2-2.5	0.343	0.188*	-0.024	0.393
	[0.288]	[0.097]	[0.163]	[2.285]
Birth weight (kg): 2.5-3	0.217	0.044	-0.086	0.009
	[0.270]	[0.106]	[0.153]	[2.184]
Gender (boy=1)				
Age	-0.009	0.082***	0.010	-2.134***
	[0.037]	[0.016]	[0.021]	[0.288]
Birth order	-0.088	0.066	0.094	-1.178
	[0.384]	[0.161]	[0.184]	[3.654]
# Siblings	-0.403	-0.128	-0.099	-3.629
	[0.335]	[0.146]	[0.161]	[3.227]
Mother's age	-0.006	0.022**	-0.009	-0.347*
	[0.027]	[0.011]	[0.012]	[0.207]
Mother's education	0.077**	0.020	0.048***	1.081***
	[0.039]	[0.013]	[0.018]	[0.288]
Per capital family income	0.147***	0.057***	0.105***	-0.737*
	[0.040]	[0.016]	[0.024]	[0.382]
Own washing machine	0.118	0.097	0.253*	1.215
	[0.215]	[0.105]	[0.149]	[1.746]
Own refrigerator	0.185	0.081	0.068	2.391
	[0.249]	[0.112]	[0.150]	[2.013]
Own cell phone	0.019	0.378***	0.577***	2.914*
	[0.216]	[0.089]	[0.131]	[1.736]
Mother working in public sector	-0.367	0.300***	-0.077	1.864
	[0.381]	[0.110]	[0.169]	[2.746]
Rural	0.221	-0.119	0.135	-4.160**
	[0.231]	[0.085]	[0.143]	[1.889]
# Observations	1120	1120	1120	1112
R-squared	0.105	0.333	0.225	0.269

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 9: OLS Estimates of Determinants of Family Investments (Mixed-Gender Sample)

	Health	Education	Clothing	Parents home tutor
Early health shocks	1.208***	0.102	0.128	-1.857
	[0.375]	[0.120]	[0.210]	[3.281]
Birth weight(kg): <2	0.485	0.070	0.267	-0.480
	[0.373]	[0.158]	[0.184]	[3.954]
Birth weight(kg): 2-2.5	0.192	0.137	0.117	-2.782
	[0.291]	[0.136]	[0.165]	[2.296]
Birth weight (kg): 2.5-3	0.387	-0.007	0.001	-1.211
	[0.271]	[0.134]	[0.156]	[2.163]
Gender (boy=1)	0.088	-0.026	-0.022	0.165
	[0.096]	[0.024]	[0.021]	[0.811]
Age	-0.052	0.095***	0.082***	-2.073***
	[0.046]	[0.022]	[0.026]	[0.389]
Birth order	-0.044	-0.004	0.052	-4.484
	[0.658]	[0.173]	[0.270]	[6.031]
# Siblings	-0.289	-0.003	-0.282	-2.719
	[0.605]	[0.161]	[0.205]	[5.399]
Mother's age	0.020	0.007	-0.020	-0.087
	[0.024]	[0.009]	[0.013]	[0.219]
Mother's education	0.092**	0.017	-0.007	0.331
	[0.043]	[0.024]	[0.032]	[0.390]
Per capital family income	-0.019	0.088***	0.081***	-0.380
	[0.046]	[0.025]	[0.021]	[0.346]
Own washing machine	0.149	0.129	0.355**	4.969**
	[0.255]	[0.106]	[0.151]	[2.182]
Own refrigerator	-0.139	-0.042	-0.244	5.176**
	[0.313]	[0.134]	[0.174]	[2.507]
Own cell phone	0.196	0.130	0.468***	3.041
	[0.282]	[0.144]	[0.171]	[2.501]
Mother working in public sector	-0.460	0.185	0.468**	7.498
	[0.654]	[0.271]	[0.233]	[4.822]
Rural	0.396	-0.205*	-0.044	-1.189
	[0.278]	[0.115]	[0.152]	[2.254]
# Observations	720	720	720	712
R-squared	0.054	0.265	0.198	0.185

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 10: OLS Estimates of the Determinants of Health (Whole Sample)

	Height (z-score)	Weight (z-score)	BMI (z-score)	Health Status	Hospital Visits
Early health shocks	-0.132	-0.206**	-0.142	-0.297***	0.196***
	[0.115]	[0.090]	[0.107]	[0.063]	[0.041]
Birth weight(kg): <2	-0.453***	-0.393***	-0.271**	-0.148***	0.063
	[0.113]	[0.089]	[0.109]	[0.057]	[0.038]
Birth weight(kg): 2-2.5	-0.221**	-0.178**	-0.062	-0.104**	0.042
	[0.094]	[0.073]	[0.085]	[0.044]	[0.030]
Birth weight (kg): 2.5-3	-0.142	-0.032	0.007	-0.039	0.036
	[0.088]	[0.067]	[0.077]	[0.041]	[0.028]
Gender (boy=1)	-0.030	0.099**	0.158***	-0.027	0.015
	[0.060]	[0.048]	[0.058]	[0.028]	[0.020]
Age	0.005	-0.035***	-0.073***	0.015**	-0.010**
	[0.014]	[0.011]	[0.013]	[0.006]	[0.005]
Birth order	0.182	0.286*	0.076	-0.044	0.018
	[0.214]	[0.151]	[0.174]	[0.089]	[0.053]
# Siblings	-0.441**	-0.272**	0.121	0.036	-0.123***
	[0.189]	[0.135]	[0.146]	[0.080]	[0.047]
Mother's age	0.032***	0.007	-0.010	-0.002	0.001
	[0.009]	[0.006]	[0.009]	[0.004]	[0.003]
Mother's education	0.053***	0.028***	-0.001	-0.003	0.019***
	[0.013]	[0.010]	[0.013]	[0.006]	[0.005]
Per capital family income	0.039***	0.027***	-0.004	0.020***	-0.010**
	[0.011]	[0.009]	[0.011]	[0.005]	[0.004]
Own washing machine	0.205**	0.054	-0.082	-0.011	0.024
	[0.088]	[0.067]	[0.080]	[0.038]	[0.027]
Own refrigerator	0.130	0.021	-0.063	0.006	0.081***
	[0.087]	[0.070]	[0.084]	[0.043]	[0.031]
Own cell phone	0.082	0.044	-0.070	0.026	-0.013
	[0.082]	[0.065]	[0.080]	[0.040]	[0.028]
Mother working in public sector	-0.038	-0.042	-0.077	0.051	0.008
	[0.128]	[0.100]	[0.123]	[0.063]	[0.049]
Rural	-0.431***	-0.204***	0.128*	0.003	0.060**
	[0.082]	[0.066]	[0.078]	[0.038]	[0.028]
# Observations	2846	2870	2822	2910	2902
R-squared	0.173	0.073	0.079	0.040	0.063

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 11: OLS Estimates of the Determinants of Health (Rural Sample)

	Height (*z*-score)	Weight (*z*-score)	BMI (*z*-score)	Health Status	Hospital Visits
Early health shocks	-0.081	-0.217*	-0.082	-0.202**	0.078
	[0.173]	[0.123]	[0.145]	[0.083]	[0.060]
Birth weight(kg): <2	-0.457***	-0.271**	-0.115	-0.167**	0.069
	[0.158]	[0.126]	[0.155]	[0.079]	[0.055]
Birth weight(kg): 2-2.5	-0.415***	-0.179*	0.093	-0.103*	0.057
	[0.123]	[0.102]	[0.115]	[0.057]	[0.039]
Birth weight (kg): 2.5-3	-0.261**	-0.008	0.108	-0.016	0.022
	[0.114]	[0.093]	[0.106]	[0.053]	[0.035]
Gender (boy=1)	0.028	0.170**	0.220***	-0.034	0.061**
	[0.084]	[0.068]	[0.082]	[0.036]	[0.027]
Age	0.051***	-0.010	-0.087***	0.014*	-0.009
	[0.019]	[0.015]	[0.019]	[0.008]	[0.006]
Birth order	0.303	0.276	-0.087	-0.065	-0.027
	[0.243]	[0.168]	[0.186]	[0.104]	[0.063]
# Siblings	-0.644***	-0.319**	0.245	0.009	-0.099*
	[0.215]	[0.152]	[0.152]	[0.093]	[0.056]
Mother's age	0.030**	0.012	0.003	0.005	0.001
	[0.013]	[0.009]	[0.012]	[0.006]	[0.004]
Mother's education	0.032	0.017	-0.009	-0.004	0.018***
	[0.021]	[0.017]	[0.018]	[0.009]	[0.007]
Per capital family income	0.046**	0.041**	0.006	0.025***	-0.008
	[0.019]	[0.016]	[0.017]	[0.009]	[0.007]
Own washing machine	0.243**	0.087	-0.048	-0.030	-0.008
	[0.111]	[0.082]	[0.098]	[0.045]	[0.034]
Own refrigerator	0.175	-0.069	-0.151	0.000	0.075
	[0.137]	[0.116]	[0.133]	[0.067]	[0.054]
Own cell phone	0.300**	0.113	-0.142	0.111*	-0.021
	[0.126]	[0.103]	[0.130]	[0.065]	[0.044]
Mother working in public sector	0.367	0.283	0.118	-0.114	-0.028
	[0.304]	[0.253]	[0.249]	[0.104]	[0.119]
Rural					
# Observations	1480	1510	1490	1546	1542
R-squared	0.144	0.053	0.067	0.047	0.047

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 12: OLS Estimates of the Determinants of Health (Urban Sample)

	Height (*z*-score)	Weight (*z*-score)	BMI (*z*-score)	Health Status	Hospital Visits
Early health shocks	-0.191	-0.212*	-0.202	-0.382***	0.302***
	[0.141]	[0.123]	[0.155]	[0.091]	[0.051]
Birth weight(kg): <2	-0.333**	-0.476***	-0.452***	-0.112	0.066
	[0.164]	[0.124]	[0.151]	[0.082]	[0.055]
Birth weight(kg): 2-2.5	0.082	-0.137	-0.256**	-0.086	0.031
	[0.145]	[0.104]	[0.124]	[0.068]	[0.047]
Birth weight (kg): 2.5-3	0.101	-0.018	-0.132	-0.047	0.062
	[0.137]	[0.092]	[0.111]	[0.064]	[0.046]
Gender (boy=1)	-0.071	0.033	0.086	-0.016	-0.034
	[0.084]	[0.066]	[0.082]	[0.042]	[0.030]
Age	-0.036*	-0.055***	-0.053***	0.017**	-0.010
	[0.020]	[0.015]	[0.020]	[0.008]	[0.006]
Birth order	-0.252	0.349	0.654	-0.086	0.262***
	[0.415]	[0.403]	[0.452]	[0.136]	[0.076]
# Siblings	0.421	-0.099	-0.386	0.224**	-0.313***
	[0.340]	[0.351]	[0.386]	[0.114]	[0.045]
Mother's age	0.033**	0.001	-0.024*	-0.010*	0.001
	[0.014]	[0.009]	[0.013]	[0.006]	[0.004]
Mother's education	0.062***	0.032**	0.005	-0.002	0.021***
	[0.016]	[0.013]	[0.017]	[0.009]	[0.006]
Per capital family income	0.040***	0.021*	-0.009	0.019***	-0.012**
	[0.014]	[0.011]	[0.014]	[0.007]	[0.005]
Own washing machine	0.052	-0.032	-0.120	-0.004	0.108**
	[0.147]	[0.124]	[0.143]	[0.071]	[0.046]
Own refrigerator	0.110	0.081	-0.001	0.033	0.062
	[0.108]	[0.087]	[0.110]	[0.057]	[0.039]
Own cell phone	-0.098	-0.016	-0.010	-0.048	-0.002
	[0.106]	[0.082]	[0.102]	[0.052]	[0.037]
Mother working in public sector	-0.080	-0.084	-0.131	0.097	0.005
	[0.138]	[0.110]	[0.141]	[0.074]	[0.056]
Rural					
# Observations	1366	1360	1332	1364	1360
R-squared	0.078	0.079	0.071	0.053	0.096

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 13: OLS Estimates of the Determinants of Health (Male Sample)

	Height (*z*-score)	Weight (*z*-score)	BMI (*z*-score)	Health Status	Hospital Visits
Early health shocks	-0.181	-0.394***	-0.353**	-0.290***	0.125**
	[0.169]	[0.133]	[0.172]	[0.091]	[0.058]
Birth weight(kg): <2	-0.381**	-0.500***	-0.366**	-0.131	0.103
	[0.194]	[0.150]	[0.185]	[0.097]	[0.068]
Birth weight(kg): 2-2.5	-0.162	-0.348***	-0.263*	-0.168**	0.051
	[0.148]	[0.113]	[0.134]	[0.070]	[0.051]
Birth weight (kg): 2.5-3	-0.095	-0.021	-0.031	-0.081	0.013
	[0.141]	[0.099]	[0.119]	[0.068]	[0.049]
Gender (boy=1)					
Age	-0.035	-0.055***	-0.084***	0.008	-0.011
	[0.023]	[0.020]	[0.024]	[0.011]	[0.008]
Birth order	-0.361	-0.233	-0.269	-0.207	0.049
	[0.353]	[0.271]	[0.268]	[0.136]	[0.099]
# Siblings	-0.043	0.073	0.446**	0.196	-0.125
	[0.315]	[0.243]	[0.218]	[0.127]	[0.088]
Mother's age	0.076***	0.026**	-0.014	0.002	-0.005
	[0.016]	[0.013]	[0.016]	[0.008]	[0.005]
Mother's education	0.055***	0.030*	-0.009	-0.009	0.012
	[0.021]	[0.017]	[0.022]	[0.011]	[0.008]
Per capital family income	0.055***	0.046***	0.003	0.022***	-0.002
	[0.017]	[0.015]	[0.018]	[0.008]	[0.007]
Own washing machine	0.067	0.017	-0.036	-0.007	0.022
	[0.150]	[0.113]	[0.134]	[0.067]	[0.047]
Own refrigerator	-0.021	-0.120	0.003	0.101	0.146***
	[0.149]	[0.117]	[0.136]	[0.073]	[0.052]
Own cell phone	0.181	0.036	-0.123	-0.008	0.023
	[0.144]	[0.114]	[0.141]	[0.069]	[0.048]
Mother working in public sector	0.058	0.092	-0.031	0.091	0.007
	[0.196]	[0.160]	[0.200]	[0.094]	[0.082]
Rural	-0.379***	-0.129	0.185	0.057	0.135***
	[0.139]	[0.114]	[0.132]	[0.064]	[0.048]
# Observations	1048	1054	1038	1076	1070
R-squared	0.196	0.101	0.107	0.052	0.064

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 14: **OLS Estimates of the Determinants of Health (Female Sample)**

	Height (z-score)	Weight (z-score)	BMI (z-score)	Health Status	Hospital Visits
Early health shocks	-0.122	-0.170	-0.027	-0.251**	0.254***
	[0.224]	[0.149]	[0.190]	[0.123]	[0.073]
Birth weight(kg): <2	-0.377**	-0.305**	-0.180	-0.137	0.123**
	[0.185]	[0.134]	[0.168]	[0.092]	[0.057]
Birth weight(kg): 2-2.5	-0.049	0.007	0.063	-0.069	0.080*
	[0.164]	[0.121]	[0.147]	[0.079]	[0.048]
Birth weight (kg): 2.5-3	0.000	0.075	0.072	-0.046	0.098**
	[0.155]	[0.113]	[0.140]	[0.070]	[0.045]
Gender (boy=1)					
Age	0.048**	-0.027	-0.074***	0.019**	-0.001
	[0.022]	[0.017]	[0.022]	[0.009]	[0.007]
Birth order	0.415	0.651***	0.353	0.139	-0.009
	[0.328]	[0.235]	[0.302]	[0.132]	[0.079]
# Siblings	-0.568**	-0.423**	-0.001	-0.116	-0.106
	[0.282]	[0.205]	[0.257]	[0.113]	[0.066]
Mother's age	-0.003	0.002	0.007	-0.006	0.005
	[0.016]	[0.011]	[0.016]	[0.007]	[0.005]
Mother's education	0.074***	0.038**	0.018	0.014	0.022***
	[0.021]	[0.017]	[0.021]	[0.011]	[0.008]
Per capital family income	0.014	-0.010	-0.030	0.004	-0.009
	[0.021]	[0.018]	[0.022]	[0.010]	[0.008]
Own washing machine	0.113	-0.062	-0.226	-0.049	0.005
	[0.143]	[0.117]	[0.140]	[0.063]	[0.046]
Own refrigerator	0.235*	0.145	-0.062	-0.020	0.008
	[0.142]	[0.114]	[0.145]	[0.067]	[0.050]
Own cell phone	0.037	0.073	-0.042	0.025	-0.025
	[0.123]	[0.104]	[0.131]	[0.065]	[0.045]
Mother working in public sector	-0.180	-0.168	-0.104	0.014	0.067
	[0.209]	[0.148]	[0.181]	[0.104]	[0.075]
Rural	-0.520***	-0.315***	0.039	0.039	-0.022
	[0.134]	[0.102]	[0.132]	[0.059]	[0.046]
# Observations	1096	1106	1088	1116	1116
R-squared	0.194	0.095	0.074	0.032	0.082

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 15: OLS Estimates of the Determinants of Health (Mixed-Gender Sample)

	Height (*z*-score)	Weight (*z*-score)	BMI (*z*-score)	Health Status	Hospital Visits
Early health shocks	-0.008	0.017	0.025	-0.374***	0.283***
	[0.199]	[0.191]	[0.189]	[0.123]	[0.088]
Birth weight(kg): <2	-0.431**	-0.185	-0.138	-0.202*	-0.095
	[0.204]	[0.201]	[0.248]	[0.107]	[0.077]
Birth weight(kg): 2-2.5	-0.435**	-0.063	0.159	-0.067	0.001
	[0.179]	[0.144]	[0.160]	[0.081]	[0.056]
Birth weight (kg): 2.5-3	-0.252	-0.074	0.042	0.021	-0.014
	[0.158]	[0.136]	[0.148]	[0.073]	[0.049]
Gender (boy=1)	0.025	0.070	0.039	-0.036	0.039*
	[0.054]	[0.044]	[0.055]	[0.027]	[0.021]
Age	-0.003	-0.020	-0.056**	0.010	-0.018*
	[0.026]	[0.021]	[0.026]	[0.011]	[0.009]
Birth order	0.433	0.424	0.063	-0.167	-0.129
	[0.414]	[0.274]	[0.286]	[0.236]	[0.124]
# Siblings	-0.607*	-0.432*	-0.049	0.129	-0.021
	[0.364]	[0.228]	[0.208]	[0.210]	[0.120]
Mother's age	0.020	-0.006	-0.023	0.001	0.002
	[0.015]	[0.011]	[0.015]	[0.007]	[0.005]
Mother's education	0.020	0.012	-0.011	-0.013	0.027***
	[0.024]	[0.019]	[0.024]	[0.011]	[0.009]
Per capital family income	0.050***	0.034***	0.006	0.035***	-0.025***
	[0.017]	[0.012]	[0.014]	[0.010]	[0.006]
Own washing machine	0.470***	0.271**	0.075	-0.004	0.048
	[0.167]	[0.119]	[0.145]	[0.068]	[0.051]
Own refrigerator	0.234	0.050	-0.165	-0.102	0.100*
	[0.172]	[0.135]	[0.159]	[0.084]	[0.060]
Own cell phone	-0.005	0.079	0.012	0.082	-0.064
	[0.165]	[0.122]	[0.144]	[0.077]	[0.054]
Mother working in public sector	-0.063	-0.118	-0.052	0.076	-0.168
	[0.253]	[0.260]	[0.294]	[0.137]	[0.116]
Rural	-0.393**	-0.113	0.220	-0.127*	0.056
	[0.155]	[0.131]	[0.145]	[0.072]	[0.053]
# Observations	702	710	696	718	716
R-squared	0.169	0.074	0.075	0.091	0.117

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 16: OLS Estimates of the Determinants of Academic Performance (Whole Sample)

	Literature (self-reported)	Mathematics (self-reported)	Literature (exam record)	Mathematics (exam record)
Early health shocks	-0.291***	-0.268***	-2.319**	-3.071**
	[0.065]	[0.067]	[1.093]	[1.299]
Birth weight(kg): <2	-0.127*	-0.177**	-1.328	-0.831
	[0.069]	[0.070]	[1.074]	[1.316]
Birth weight(kg): 2-2.5	-0.093*	-0.125**	-1.171	-0.540
	[0.053]	[0.057]	[0.823]	[0.995]
Birth weight (kg): 2.5-3	-0.059	-0.055	-1.128	-0.264
	[0.051]	[0.055]	[0.786]	[0.934]
Gender (boy=1)	-0.183***	-0.009	-1.931***	0.533
	[0.035]	[0.038]	[0.534]	[0.658]
Age	-0.037***	-0.063***	-1.595***	-1.914***
	[0.007]	[0.008]	[0.127]	[0.156]
Birth order	-0.188*	-0.186	-1.097	-0.979
	[0.108]	[0.115]	[1.938]	[2.162]
# Siblings	0.121	0.116	0.770	1.070
	[0.099]	[0.105]	[1.741]	[1.955]
Mother's age	0.000	0.004	0.072	-0.018
	[0.005]	[0.005]	[0.088]	[0.103]
Mother's education	0.028***	0.037***	0.530***	0.431***
	[0.008]	[0.008]	[0.122]	[0.159]
Per capital family income	0.018**	0.019**	0.118	0.267**
	[0.008]	[0.009]	[0.098]	[0.129]
Own washing machine	-0.030	-0.107**	1.970***	1.325
	[0.047]	[0.049]	[0.762]	[0.904]
Own refrigerator	0.109**	0.059	1.345*	0.892
	[0.053]	[0.056]	[0.774]	[1.001]
Own cell phone	-0.018	0.099*	1.045	2.062**
	[0.047]	[0.051]	[0.710]	[0.901]
Mother working in public sector	0.145*	0.242***	0.452	3.339***
	[0.075]	[0.084]	[0.918]	[1.173]
Rural	-0.001	0.096*	-3.409***	-2.065**
	[0.048]	[0.051]	[0.705]	[0.903]
# Observations	2862	2850	2724	2686
R-squared	0.078	0.091	0.202	0.177

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 17: OLS Estimates of the Determinants of Academic Performance (Rural Sample)

	Literature (self-reported)	Mathematics (self-reported)	Literature (exam record)	Mathematics (exam record)
Early health shocks	-0.222**	-0.231**	-4.180**	-3.635*
	[0.090]	[0.090]	[1.914]	[1.966]
Birth weight(kg): <2	-0.104	-0.228**	-1.579	-1.161
	[0.093]	[0.096]	[1.605]	[1.789]
Birth weight(kg): 2-2.5	-0.091	-0.104	-1.265	-1.208
	[0.068]	[0.073]	[1.142]	[1.314]
Birth weight (kg): 2.5-3	-0.024	0.004	-1.222	-1.290
	[0.067]	[0.072]	[1.091]	[1.255]
Gender (boy=1)	-0.172***	-0.076	-2.422***	-0.444
	[0.048]	[0.050]	[0.803]	[0.916]
Age	-0.025**	-0.035***	-1.273***	-1.315***
	[0.010]	[0.011]	[0.196]	[0.229]
Birth order	-0.203*	-0.166	-1.497	-0.633
	[0.121]	[0.129]	[2.119]	[2.305]
# Siblings	0.148	0.115	0.688	0.876
	[0.112]	[0.119]	[1.879]	[2.055]
Mother's age	-0.006	-0.004	0.126	-0.105
	[0.006]	[0.007]	[0.124]	[0.134]
Mother's education	0.019	0.014	0.428**	0.352
	[0.012]	[0.012]	[0.204]	[0.225]
Per capital family income	0.011	0.016	0.131	0.249
	[0.011]	[0.013]	[0.208]	[0.223]
Own washing machine	0.014	-0.095	3.357***	2.692**
	[0.057]	[0.059]	[1.001]	[1.123]
Own refrigerator	0.032	-0.014	2.061	1.173
	[0.090]	[0.089]	[1.263]	[1.465]
Own cell phone	0.004	0.154**	2.078*	4.009***
	[0.071]	[0.075]	[1.203]	[1.297]
Mother working in public sector	0.308*	0.450**	-0.822	1.870
	[0.167]	[0.204]	[2.705]	[3.932]
Rural				
# Observations	1518	1514	1422	1410
R-squared	0.043	0.048	0.108	0.097

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 18: OLS Estimates of the Determinants of Academic Performance (Urban Sample)

	Literature (self-reported)	Mathematics (self-reported)	Literature (exam record)	Mathematics (exam record)
Early health shocks	-0.367***	-0.321***	-0.704	-2.502
	[0.094]	[0.096]	[1.177]	[1.627]
Birth weight(kg): <2	-0.154	-0.127	-0.989	-0.070
	[0.106]	[0.104]	[1.422]	[1.903]
Birth weight(kg): 2-2.5	-0.105	-0.149*	-1.010	0.298
	[0.084]	[0.089]	[1.183]	[1.500]
Birth weight (kg): 2.5-3	-0.111	-0.133	-0.850	1.028
	[0.081]	[0.086]	[1.129]	[1.389]
Gender (boy=1)	-0.192***	0.072	-1.245*	1.885**
	[0.053]	[0.057]	[0.675]	[0.932]
Age	-0.049***	-0.093***	-1.861***	-2.474***
	[0.011]	[0.012]	[0.165]	[0.212]
Birth order	0.034	0.014	0.741	0.299
	[0.217]	[0.225]	[4.542]	[5.446]
# Siblings	-0.103	-0.089	0.804	0.817
	[0.180]	[0.188]	[4.229]	[5.023]
Mother's age	0.007	0.017**	0.027	0.088
	[0.007]	[0.008]	[0.120]	[0.158]
Mother's education	0.034***	0.051***	0.552***	0.418*
	[0.011]	[0.011]	[0.147]	[0.219]
Per capital family income	0.021**	0.017	0.137	0.314**
	[0.010]	[0.011]	[0.099]	[0.154]
Own washing machine	-0.075	-0.055	-1.564	-2.161
	[0.089]	[0.094]	[1.137]	[1.509]
Own refrigerator	0.133*	0.060	1.346	0.886
	[0.069]	[0.074]	[0.979]	[1.351]
Own cell phone	-0.020	0.077	0.238	0.681
	[0.062]	[0.068]	[0.864]	[1.200]
Mother working in public sector	0.095	0.174*	1.381	4.318***
	[0.086]	[0.093]	[0.949]	[1.240]
Rural				
# Observations	1344	1336	1302	1276
R-squared	0.110	0.154	0.263	0.261

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 19: **OLS Estimates of the Determinants of Academic Performance (Male Sample)**

	Literature (self-reported)	Mathematics (self-reported)	Literature (exam record)	Mathematics (exam record)
Early health shocks	-0.393***	-0.250**	-3.814**	-3.771*
	[0.105]	[0.107]	[1.842]	[1.951]
Birth weight(kg): <2	-0.047	-0.284**	0.393	-2.499
	[0.132]	[0.125]	[1.573]	[1.999]
Birth weight(kg): 2-2.5	-0.164*	-0.216**	-1.376	-2.035
	[0.087]	[0.090]	[1.220]	[1.469]
Birth weight (kg): 2.5-3	-0.096	-0.105	-0.962	-1.765
	[0.085]	[0.090]	[1.123]	[1.356]
Gender (boy=1)				
Age	-0.038***	-0.054***	-[illegible].914***	-1.987***
	[0.013]	[0.014]	[0.191]	[0.232]
Birth order	-0.415**	-0.093	-0.856	-2.059
	[0.183]	[0.190]	[3.194]	[3.101]
# Siblings	0.287*	0.019	0.025	1.769
	[0.161]	[0.167]	[2.869]	[2.628]
Mother's age	-0.001	0.004	0.265**	0.006
	[0.009]	[0.009]	[0.124]	[0.149]
Mother's education	0.022	0.031**	0.565***	0.349
	[0.014]	[0.015]	[0.210]	[0.257]
Per capital family income	0.023*	0.032**	0.006	0.140
	[0.013]	[0.015]	[0.152]	[0.173]
Own washing machine	-0.093	-0.171*	0.828	0.742
	[0.085]	[0.087]	[1.208]	[1.389]
Own refrigerator	0.202**	0.154	1.781	0.840
	[0.093]	[0.096]	[1.308]	[1.527]
Own cell phone	-0.067	0.004	0.505	-0.261
	[0.081]	[0.088]	[1.193]	[1.378]
Mother working in public sector	0.077	0.219	-0.746	4.094**
	[0.126]	[0.150]	[1.635]	[2.032]
Rural	0.040	0.063	-4.419***	-4.030***
	[0.081]	[0.091]	[1.165]	[1.332]
# Observations	1054	1046	1014	1000
R-squared	0.075	0.087	0.231	0.198

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 20: OLS Estimates of the Determinants of Academic Performance (Female Sample)

	Literature (self-reported)	Mathematics (self-reported)	Literature (exam record)	Mathematics (exam record)
Early health shocks	-0.292***	-0.366***	-0.140	-2.800
	[0.104]	[0.111]	[1.617]	[2.534]
Birth weight(kg): <2	-0.146	-0.153	-3.160*	0.087
	[0.111]	[0.115]	[1.801]	[2.284]
Birth weight(kg): 2-2.5	-0.069	-0.158	-2.467*	-0.651
	[0.098]	[0.104]	[1.456]	[1.896]
Birth weight (kg): 2.5-3	-0.048	-0.108	-2.402*	-1.012
	[0.096]	[0.101]	[1.427]	[1.812]
Gender (boy=1)				
Age	-0.050***	-0.077***	-1.588***	-2.240***
	[0.011]	[0.013]	[0.195]	[0.249]
Birth order	-0.079	-0.164	-2.787	-3.743
	[0.167]	[0.185]	[3.229]	[3.510]
# Siblings	0.009	0.067	0.108	1.020
	[0.157]	[0.174]	[2.844]	[3.095]
Mother's age	0.006	0.007	0.076	0.140
	[0.008]	[0.009]	[0.140]	[0.161]
Mother's education	0.022*	0.041***	0.295	0.363
	[0.012]	[0.014]	[0.196]	[0.291]
Per capital family income	0.028**	0.025**	0.413**	0.676***
	[0.012]	[0.012]	[0.166]	[0.242]
Own washing machine	-0.049	-0.120	2.576**	0.847
	[0.077]	[0.084]	[1.293]	[1.633]
Own refrigerator	0.114	-0.022	1.593	0.900
	[0.086]	[0.094]	[1.213]	[1.833]
Own cell phone	0.031	0.155*	0.406	3.504**
	[0.073]	[0.082]	[1.116]	[1.580]
Mother working in public sector	0.171	0.166	1.878	1.974
	[0.115]	[0.117]	[1.327]	[1.767]
Rural	0.037	0.147*	-2.168*	0.650
	[0.082]	[0.086]	[1.154]	[1.687]
# Observations	1102	1098	1038	1022
R-squared	0.091	0.124	0.221	0.204

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 21: OLS Estimates of the Determinants of Academic Performance (Mixed-Gender Sample)

	Literature (self-reported)	Mathematics (self-reported)	Literature (exam record)	Mathematics (exam record)
Early health shocks	-0.075	-0.200*	-1.472	-2.125
	[0.122]	[0.121]	[1.954]	[2.392]
Birth weight(kg): <2	-0.193	-0.153	-1.635	-1.480
	[0.143]	[0.158]	[2.529]	[3.127]
Birth weight(kg): 2-2.5	-0.053	0.047	-0.582	0.856
	[0.097]	[0.110]	[1.724]	[1.922]
Birth weight (kg): 2.5-3	-0.062	0.049	-1.004	2.308
	[0.093]	[0.106]	[1.651]	[1.841]
Gender (boy=1)	-0.185***	0.009	-2.700***	-0.631
	[0.048]	[0.055]	[0.729]	[0.860]
Age	-0.020	-0.063***	-1.217***	-1.451***
	[0.015]	[0.017]	[0.299]	[0.336]
Birth order	-0.070	-0.395*	0.861	7.047
	[0.205]	[0.208]	[3.423]	[5.273]
# Siblings	0.061	0.334*	1.367	-3.822
	[0.177]	[0.186]	[3.198]	[4.899]
Mother's age	-0.002	0.002	-0.055	-0.163
	[0.008]	[0.009]	[0.179]	[0.200]
Mother's education	0.042***	0.032**	0.643***	0.343
	[0.015]	[0.015]	[0.243]	[0.270]
Per capital family income	0.001	-0.006	-0.036	0.050
	[0.015]	[0.015]	[0.186]	[0.288]
Own washing machine	0.052	0.012	2.706*	2.675
	[0.082]	[0.085]	[1.503]	[1.718]
Own refrigerator	-0.042	0.047	0.392	1.037
	[0.100]	[0.104]	[1.645]	[1.877]
Own cell phone	-0.010	0.166*	2.691*	3.365*
	[0.094]	[0.090]	[1.533]	[1.744]
Mother working in public sector	0.176	0.415**	0.103	5.996**
	[0.161]	[0.190]	[2.042]	[2.365]
Rural	-0.106	0.077	-3.191**	-2.602
	[0.090]	[0.094]	[1.391]	[1.652]
# Observations	706	706	672	664
R-squared	0.059	0.082	0.151	0.152

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 22: OLS Estimates of the Determinants of Schooling Performance (Whole Sample)

	Good Student Awards	Awards in Contests	Grade repetition	Parents interviewed	Minor actions in class
Early health shocks	-0.034	-0.037**	0.042**	0.091***	0.099*
	[0.030]	[0.016]	[0.020]	[0.032]	[0.058]
Birth weight(kg): <2	-0.077**	-0.020	0.048**	0.028	0.092
	[0.034]	[0.019]	[0.021]	[0.028]	[0.062]
Birth weight(kg): 2-2.5	-0.036	-0.008	0.006	0.001	0.061
	[0.027]	[0.016]	[0.012]	[0.021]	[0.050]
Birth weight (kg): 2.5-3	-0.029	0.004	0.004	0.008	0.089*
	[0.027]	[0.015]	[0.012]	[0.020]	[0.049]
Gender (boy=1)	-0.078***	-0.015	0.009	0.098***	0.297***
	[0.017]	[0.010]	[0.009]	[0.014]	[0.033]
Age	0.032***	0.008***	0.010***	0.001	-0.016**
	[0.004]	[0.002]	[0.002]	[0.003]	[0.007]
Birth order	-0.068	0.005	0.033	-0.047	0.081
	[0.057]	[0.016]	[0.025]	[0.038]	[0.095]
# Siblings	0.045	-0.010	-0.017	0.002	-0.149*
	[0.052]	[0.015]	[0.020]	[0.037]	[0.083]
Mother's age	0.000	0.002*	0.000	-0.001	0.000
	[0.002]	[0.001]	[0.001]	[0.002]	[0.005]
Mother's education	0.011***	0.005**	-0.003*	0.000	-0.010
	[0.004]	[0.002]	[0.002]	[0.003]	[0.007]
Per capital family income	0.007*	0.007***	-0.001	-0.003	-0.001
	[0.004]	[0.003]	[0.002]	[0.003]	[0.006]
Own washing machine	-0.021	-0.004	-0.002	-0.006	0.050
	[0.022]	[0.010]	[0.011]	[0.018]	[0.043]
Own refrigerator	-0.008	0.018	0.000	-0.004	0.011
	[0.026]	[0.015]	[0.014]	[0.021]	[0.048]
Own cell phone	0.045*	0.019	0.007	0.004	-0.012
	[0.025]	[0.014]	[0.012]	[0.019]	[0.043]
Mother working in public sector	0.089**	0.084**	0.000	-0.029	-0.090
	[0.043]	[0.036]	[0.018]	[0.034]	[0.065]
Rural	-0.059**	-0.033***	0.020	-0.092***	0.057
	[0.024]	[0.012]	[0.013]	[0.020]	[0.043]
# Observations	2922	2922	2922	2918	2890
R-squared	0.112	0.096	0.035	0.052	0.053

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 23: OLS Estimates of the Determinants of Schooling Performance (Rural Sample)

	Good Student Awards	Awards in Contests	Grade repetition	Parents interviewed	Minor actions in class
Early health shocks	-0.024	-0.018	0.058*	0.036	0.212**
	[0.035]	[0.012]	[0.034]	[0.041]	[0.085]
Birth weight(kg): <2	-0.083**	-0.007	0.062*	0.029	0.107
	[0.042]	[0.019]	[0.032]	[0.034]	[0.081]
Birth weight(kg): 2-2.5	-0.038	-0.015	0.014	0.003	0.094
	[0.033]	[0.015]	[0.018]	[0.023]	[0.063]
Birth weight (kg): 2.5-3	-0.031	-0.013	0.010	-0.002	0.127**
	[0.031]	[0.014]	[0.018]	[0.020]	[0.062]
Gender (boy=1)	-0.095***	-0.019**	0.003	0.075***	0.321***
	[0.021]	[0.009]	[0.014]	[0.016]	[0.045]
Age	0.026***	0.006***	0.015***	0.001	-0.024**
	[0.005]	[0.002]	[0.003]	[0.004]	[0.010]
Birth order	-0.036	0.016	0.038	0.034	0.071
	[0.057]	[0.012]	[0.027]	[0.026]	[0.099]
# Siblings	0.011	-0.016	-0.024	-0.063**	-0.173**
	[0.052]	[0.012]	[0.022]	[0.026]	[0.085]
Mother's age	0.000	0.001	0.001	-0.002	0.003
	[0.003]	[0.001]	[0.002]	[0.002]	[0.006]
Mother's education	0.006	0.000	-0.002	0.004	0.003
	[0.005]	[0.002]	[0.003]	[0.003]	[0.011]
Per capital family income	0.001	0.005	-0.002	-0.007**	-0.006
	[0.006]	[0.003]	[0.003]	[0.003]	[0.012]
Own washing machine	-0.004	-0.004	-0.013	0.000	0.031
	[0.025]	[0.010]	[0.015]	[0.019]	[0.053]
Own refrigerator	-0.029	0.021	0.019	0.004	0.033
	[0.041]	[0.019]	[0.026]	[0.027]	[0.079]
Own cell phone	0.116***	0.023	0.028	0.010	-0.029
	[0.038]	[0.019]	[0.022]	[0.026]	[0.065]
Mother working in public sector	0.259***	0.077	-0.051	0.024	-0.256*
	[0.094]	[0.074]	[0.042]	[0.084]	[0.131]
Rural					
# Observations	1546	1546	1546	1542	1532
R-squared	0.098	0.051	0.050	0.031	0.069

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 24: OLS Estimates of the Determinants of Schooling Performance (Urban Sample)

	Good Student Awards	Awards in Contests	Grade repetition	Parents interviewed	Minor actions in class
Early health shocks	-0.053	-0.054*	0.031	0.138***	-0.006
	[0.046]	[0.029]	[0.021]	[0.048]	[0.080]
Birth weight(kg): <2	-0.070	-0.028	0.035	0.019	0.065
	[0.053]	[0.034]	[0.026]	[0.048]	[0.098]
Birth weight(kg): 2-2.5	-0.028	-0.003	-0.001	0.004	0.016
	[0.045]	[0.030]	[0.016]	[0.039]	[0.084]
Birth weight (kg): 2.5-3	-0.018	0.024	-0.001	0.023	0.041
	[0.045]	[0.031]	[0.016]	[0.037]	[0.081]
Gender (boy=1)	-0.063**	-0.013	0.018*	0.123***	0.271***
	[0.028]	[0.020]	[0.010]	[0.024]	[0.048]
Age	0.038***	0.011***	0.005**	0.001	-0.008
	[0.006]	[0.004]	[0.002]	[0.005]	[0.010]
Birth order	-0.255	-0.002	0.024	-0.433***	-0.019
	[0.172]	[0.069]	[0.067]	[0.145]	[0.275]
# Siblings	0.248	-0.012	0.006	0.329**	0.046
	[0.160]	[0.063]	[0.058]	[0.136]	[0.243]
Mother's age	-0.001	0.004*	0.000	-0.002	-0.004
	[0.004]	[0.002]	[0.002]	[0.003]	[0.007]
Mother's education	0.015***	0.010***	-0.005**	-0.003	-0.017*
	[0.005]	[0.003]	[0.002]	[0.005]	[0.010]
Per capital family income	0.012**	0.009**	0.001	0.000	0.003
	[0.006]	[0.004]	[0.002]	[0.004]	[0.008]
Own washing machine	-0.071*	0.016	0.011	-0.040	0.057
	[0.041]	[0.021]	[0.016]	[0.042]	[0.080]
Own refrigerator	0.005	0.004	-0.009	0.001	0.019
	[0.034]	[0.022]	[0.016]	[0.032]	[0.063]
Own cell phone	-0.002	0.021	-0.011	-0.008	-0.011
	[0.033]	[0.021]	[0.013]	[0.028]	[0.057]
Mother working in public sector	0.055	0.064	0.018	-0.024	-0.043
	[0.049]	[0.042]	[0.019]	[0.038]	[0.074]
Rural					
# Observations	1376	1376	1376	1376	1358
R-squared	0.104	0.086	0.030	0.056	0.043

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 25: OLS Estimates of the Determinants of Schooling Performance (Male Sample)

	Good Student Awards	Awards in Contests	Grade repetition	Parents interviewed	Minor actions in class
Early health shocks	-0.016	-0.042	0.045	0.175***	0.152*
	[0.044]	[0.026]	[0.032]	[0.057]	[0.084]
Birth weight(kg): <2	-0.168***	-0.035	0.019	-0.030	0.109
	[0.051]	[0.034]	[0.039]	[0.052]	[0.106]
Birth weight(kg): 2-2.5	-0.040	-0.001	-0.018	0.043	0.136
	[0.041]	[0.026]	[0.023]	[0.039]	[0.084]
Birth weight (kg): 2.5-3	-0.049	0.008	-0.003	0.032	0.123
	[0.040]	[0.025]	[0.025]	[0.036]	[0.081]
Gender (boy=1)					
Age	0.031***	0.013***	0.010***	0.003	-0.016
	[0.006]	[0.004]	[0.003]	[0.006]	[0.013]
Birth order	-0.057	0.032	0.071**	0.012	0.054
	[0.098]	[0.022]	[0.030]	[0.061]	[0.192]
# Siblings	0.059	-0.026	-0.068***	-0.058	-0.046
	[0.090]	[0.017]	[0.016]	[0.055]	[0.171]
Mother's age	0.001	0.002	0.000	0.002	-0.004
	[0.004]	[0.002]	[0.002]	[0.004]	[0.008]
Mother's education	-0.001	0.003	-0.007**	-0.002	-0.012
	[0.005]	[0.004]	[0.004]	[0.006]	[0.013]
Per capital family income	0.013**	0.006	0.000	-0.006	0.006
	[0.006]	[0.004]	[0.003]	[0.005]	[0.010]
Own washing machine	-0.004	0.008	-0.014	-0.039	0.146*
	[0.034]	[0.015]	[0.020]	[0.036]	[0.080]
Own refrigerator	-0.018	0.024	0.008	0.007	0.035
	[0.042]	[0.024]	[0.022]	[0.040]	[0.088]
Own cell phone	0.022	0.017	0.016	0.000	-0.039
	[0.040]	[0.023]	[0.022]	[0.038]	[0.077]
Mother working in public sector	0.078	0.074	0.012	0.012	-0.110
	[0.065]	[0.057]	[0.041]	[0.071]	[0.108]
Rural	-0.130***	-0.043**	0.011	-0.119***	0.051
	[0.039]	[0.018]	[0.024]	[0.037]	[0.077]
# Observations	1082	1082	1082	1082	1064
R-squared	0.119	0.108	0.039	0.054	0.023

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 26: OLS Estimates of the Determinants of Schooling Performance (Female Sample)

	Good Student Awards	Awards in Contests	Grade repetition	Parents interviewed	Minor actions in class
Early health shocks	-0.035	-0.017	0.054	0.017	0.043
	[0.056]	[0.033]	[0.035]	[0.045]	[0.109]
Birth weight(kg): <2	0.017	-0.001	0.065**	0.060	0.153
	[0.060]	[0.035]	[0.030]	[0.039]	[0.098]
Birth weight(kg): 2-2.5	0.026	-0.010	0.010	-0.006	0.111
	[0.053]	[0.032]	[0.021]	[0.030]	[0.085]
Birth weight (kg): 2.5-3	0.021	0.014	-0.003	-0.023	0.148*
	[0.051]	[0.033]	[0.020]	[0.027]	[0.086]
Gender (boy=1)					
Age	0.029***	0.004	0.009***	0.002	-0.020*
	[0.007]	[0.004]	[0.003]	[0.004]	[0.011]
Birth order	-0.211***	-0.041	0.030	-0.050	0.065
	[0.079]	[0.028]	[0.038]	[0.051]	[0.129]
# Siblings	0.074	0.010	0.011	0.022	-0.193*
	[0.075]	[0.028]	[0.031]	[0.051]	[0.112]
Mother's age	0.008*	0.005*	-0.001	-0.006**	0.012
	[0.004]	[0.003]	[0.002]	[0.002]	[0.009]
Mother's education	0.017**	0.008**	0.000	0.000	-0.003
	[0.007]	[0.004]	[0.002]	[0.004]	[0.011]
Per capital family income	0.008	0.010	-0.001	-0.001	-0.021*
	[0.008]	[0.006]	[0.003]	[0.005]	[0.011]
Own washing machine	-0.063	-0.016	0.027	0.019	0.032
	[0.040]	[0.018]	[0.020]	[0.025]	[0.069]
Own refrigerator	-0.018	-0.014	-0.001	-0.050	0.022
	[0.043]	[0.024]	[0.023]	[0.031]	[0.075]
Own cell phone	0.063	0.044*	0.016	0.005	0.026
	[0.040]	[0.026]	[0.017]	[0.025]	[0.068]
Mother working in public sector	0.113	0.106*	-0.010	-0.058*	-0.044
	[0.069]	[0.057]	[0.015]	[0.031]	[0.097]
Rural	-0.021	-0.021	0.047**	-0.107***	0.076
	[0.041]	[0.022]	[0.020]	[0.029]	[0.068]
# Observations	1120	1120	1120	1118	1108
R-squared	0.137	0.107	0.057	0.058	0.021

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 27: OLS Estimates of the Determinants of Schooling Performance (Mixed-Gender Sample)

	Good Student Awards	Awards in Contests	Grade repetition	Parents interviewed	Minor actions in class
Early health shocks	-0.060	-0.046***	0.020	0.016	0.056
	[0.058]	[0.016]	[0.030]	[0.055]	[0.115]
Birth weight(kg): <2	-0.093	-0.028	0.051	0.024	-0.018
	[0.068]	[0.030]	[0.044]	[0.062]	[0.130]
Birth weight(kg): 2-2.5	-0.089*	-0.011	0.038*	-0.055	-0.090
	[0.052]	[0.023]	[0.021]	[0.038]	[0.091]
Birth weight (kg): 2.5-3	-0.041	-0.010	0.016	0.000	0.003
	[0.050]	[0.023]	[0.016]	[0.036]	[0.086]
Gender (boy=1)	-0.089***	-0.034**	0.007	0.084***	0.296***
	[0.027]	[0.014]	[0.015]	[0.022]	[0.046]
Age	0.036***	0.009**	0.011***	-0.009	-0.012
	[0.007]	[0.004]	[0.003]	[0.006]	[0.013]
Birth order	0.083	0.057**	-0.004	-0.146	0.065
	[0.136]	[0.023]	[0.069]	[0.099]	[0.188]
# Siblings	-0.033	-0.046***	0.012	0.078	-0.166
	[0.122]	[0.017]	[0.063]	[0.100]	[0.162]
Mother's age	-0.007**	0.001	0.000	-0.001	-0.006
	[0.003]	[0.002]	[0.002]	[0.003]	[0.007]
Mother's education	0.015**	0.002	-0.002	0.006	-0.012
	[0.007]	[0.004]	[0.004]	[0.005]	[0.014]
Per capital family income	-0.002	0.007	-0.001	0.001	0.006
	[0.007]	[0.005]	[0.002]	[0.004]	[0.013]
Own washing machine	0.018	0.005	-0.025	0.008	-0.055
	[0.037]	[0.019]	[0.019]	[0.033]	[0.078]
Own refrigerator	0.014	0.055**	-0.018	0.044	-0.043
	[0.052]	[0.028]	[0.024]	[0.039]	[0.090]
Own cell phone	0.055	-0.014	-0.017	-0.004	-0.031
	[0.048]	[0.025]	[0.019]	[0.039]	[0.076]
Mother working in public sector	0.080	0.015	-0.002	-0.026	-0.198
	[0.106]	[0.085]	[0.036]	[0.083]	[0.154]
Rural	-0.009	-0.032	-0.013	-0.052	0.045
	[0.046]	[0.022]	[0.023]	[0.039]	[0.076]
# Observations	720	720	720	718	718
R-squared	0.106	0.092	0.041	0.058	0.070

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 28: OLS Estimates of the Determinants of Noncognitive Skills (Whole Sample)

	Feel lonely	Anxious	Easily distracted	Careless	Easily Frightened	Emotional instability
Early health shocks	0.063*	0.046	0.094**	0.088***	0.074*	0.108***
	[0.033]	[0.037]	[0.041]	[0.034]	[0.042]	[0.034]
Birth weight(kg): <2	-0.021	0.028	0.064	0.042	0.003	0.030
	[0.031]	[0.036]	[0.043]	[0.039]	[0.040]	[0.029]
Birth weight(kg): 2-2.5	-0.023	-0.011	0.023	0.013	-0.001	-0.008
	[0.025]	[0.028]	[0.035]	[0.032]	[0.033]	[0.022]
Birth weight (kg): 2.5-3	0.008	-0.008	0.015	-0.020	-0.015	0.002
	[0.024]	[0.026]	[0.033]	[0.029]	[0.030]	[0.020]
Gender (boy=1)	-0.007	0.097***	0.087***	0.042**	-0.063***	-0.001
	[0.016]	[0.019]	[0.022]	[0.020]	[0.021]	[0.014]
Age	0.013***	-0.005	-0.019***	-0.030***	-0.004	-0.002
	[0.004]	[0.004]	[0.005]	[0.005]	[0.005]	[0.003]
Birth order	-0.033	-0.015	0.012	0.059	0.006	-0.005
	[0.050]	[0.058]	[0.071]	[0.067]	[0.067]	[0.030]
# Siblings	-0.009	-0.001	-0.058	-0.006	-0.061	-0.010
	[0.047]	[0.053]	[0.064]	[0.063]	[0.060]	[0.029]
Mother's age	-0.001	-0.005*	-0.001	0.005*	-0.006**	0.001
	[0.002]	[0.003]	[0.003]	[0.003]	[0.003]	[0.002]
Mother's education	0.006*	-0.006	0.005	-0.002	-0.006	0.012***
	[0.004]	[0.004]	[0.005]	[0.004]	[0.005]	[0.003]
Per capital family income	-0.004	-0.004	-0.003	-0.001	-0.009*	-0.006*
	[0.004]	[0.004]	[0.005]	[0.005]	[0.005]	[0.003]
Own washing machine	0.011	0.026	0.040	0.027	0.002	0.006
	[0.021]	[0.025]	[0.030]	[0.029]	[0.029]	[0.018]
Own refrigerator	-0.031	0.039	-0.005	0.030	-0.020	0.053**
	[0.025]	[0.029]	[0.034]	[0.033]	[0.032]	[0.023]
Own cell phone	-0.010	0.016	0.047	-0.009	0.013	-0.006
	[0.023]	[0.026]	[0.031]	[0.030]	[0.029]	[0.022]
Mother working in public sector	-0.034	-0.048	-0.067	-0.007	-0.032	-0.031
	[0.038]	[0.041]	[0.050]	[0.049]	[0.045]	[0.041]
Rural	-0.020	-0.039	-0.040	0.037	-0.010	-0.064***
	[0.023]	[0.027]	[0.031]	[0.029]	[0.029]	[0.021]
# Observations	2922	2922	2922	2922	2922	2911
R-squared	0.023	0.032	0.040	0.045	0.023	0.066

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 29: OLS Estimates of the Determinants of Noncognitive Skills (Rural Sample)

	Feel lonely	Anxious	Easily distracted	Careless	Easily Frightened	Emotional instability
Early health shocks	0.116**	0.000	0.064	0.007	0.056	0.029
	[0.051]	[0.050]	[0.061]	[0.053]	[0.061]	[0.032]
Birth weight(kg): <2	-0.012	-0.009	0.062	0.021	-0.022	0.001
	[0.038]	[0.046]	[0.059]	[0.054]	[0.057]	[0.031]
Birth weight(kg): 2-2.5	-0.006	-0.020	0.036	0.030	-0.045	-0.011
	[0.029]	[0.035]	[0.046]	[0.041]	[0.045]	[0.023]
Birth weight (kg): 2.5-3	0.040	0.012	0.012	0.016	-0.049	-0.006
	[0.027]	[0.034]	[0.043]	[0.038]	[0.041]	[0.021]
Gender (boy=1)	-0.007	0.130***	0.077**	0.044*	-0.054*	0.008
	[0.021]	[0.024]	[0.030]	[0.026]	[0.029]	[0.014]
Age	0.015***	-0.004	-0.025***	-0.039***	-0.002	0.001
	[0.005]	[0.006]	[0.007]	[0.006]	[0.006]	[0.003]
Birth order	-0.033	-0.010	0.034	0.037	-0.005	0.010
	[0.056]	[0.063]	[0.080]	[0.072]	[0.076]	[0.026]
# Siblings	0.011	0.006	-0.061	0.007	-0.031	-0.021
	[0.052]	[0.057]	[0.071]	[0.067]	[0.068]	[0.026]
Mother's age	0.002	-0.001	0.001	0.006*	-0.009**	-0.001
	[0.003]	[0.004]	[0.005]	[0.004]	[0.004]	[0.002]
Mother's education	0.008	-0.008	0.006	-0.007	-0.003	0.010***
	[0.005]	[0.006]	[0.007]	[0.007]	[0.007]	[0.004]
Per capital family income	-0.014***	-0.009	-0.018**	-0.010	-0.021**	-0.007**
	[0.004]	[0.006]	[0.007]	[0.009]	[0.009]	[0.003]
Own washing machine	0.030	0.060**	0.090**	0.036	0.035	0.045**
	[0.025]	[0.029]	[0.037]	[0.034]	[0.035]	[0.018]
Own refrigerator	-0.051	0.043	0.041	0.081	-0.062	0.061*
	[0.038]	[0.048]	[0.055]	[0.053]	[0.050]	[0.035]
Own cell phone	0.000	0.011	0.009	-0.050	0.055	-0.039
	[0.035]	[0.040]	[0.048]	[0.046]	[0.045]	[0.026]
Mother working in public sector	0.106	0.144	-0.087	0.104	0.009	-0.010
	[0.101]	[0.109]	[0.110]	[0.116]	[0.101]	[0.083]
Rural						
# Observations	1546	1546	1546	1546	1546	1538
R-squared	0.046	0.041	0.044	0.073	0.032	0.045

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 30: OLS Estimates of the Determinants of Noncognitive Skills (Urban Sample)

	Feel lonely	Anxious	Easily distracted	Careless	Easily Frightened	Emotional instability
Early health shocks	0.012	0.079	0.127**	0.152***	0.088	0.178***
	[0.040]	[0.052]	[0.055]	[0.043]	[0.057]	[0.056]
Birth weight(kg): <2	-0.048	0.054	0.044	0.044	0.030	0.053
	[0.050]	[0.056]	[0.063]	[0.058]	[0.056]	[0.050]
Birth weight(kg): 2-2.5	-0.047	0.003	-0.006	-0.019	0.050	-0.009
	[0.043]	[0.044]	[0.054]	[0.049]	[0.047]	[0.040]
Birth weight (kg): 2.5-3	-0.036	-0.027	0.011	-0.063	0.025	0.014
	[0.042]	[0.041]	[0.051]	[0.046]	[0.043]	[0.039]
Gender (boy=1)	-0.008	0.061**	0.093***	0.033	-0.070**	-0.012
	[0.025]	[0.028]	[0.032]	[0.031]	[0.030]	[0.026]
Age	0.012**	-0.004	-0.012	-0.020***	-0.004	-0.005
	[0.006]	[0.005]	[0.007]	[0.007]	[0.007]	[0.005]
Birth order	-0.066	-0.110	-0.120	0.145	0.087	-0.050
	[0.110]	[0.145]	[0.179]	[0.165]	[0.129]	[0.115]
# Siblings	-0.065	0.016	-0.018	-0.096	-0.197*	0.030
	[0.106]	[0.131]	[0.167]	[0.153]	[0.109]	[0.100]
Mother's age	-0.004	-0.010**	-0.004	0.004	-0.004	0.003
	[0.003]	[0.004]	[0.005]	[0.004]	[0.004]	[0.003]
Mother's education	0.002	-0.007	0.005	0.003	-0.009	0.012**
	[0.005]	[0.005]	[0.007]	[0.006]	[0.006]	[0.006]
Per capital family income	0.001	-0.001	0.006	0.003	-0.002	-0.005
	[0.006]	[0.005]	[0.006]	[0.006]	[0.005]	[0.004]
Own washing machine	-0.023	-0.048	-0.062	0.025	-0.054	-0.074*
	[0.042]	[0.050]	[0.056]	[0.054]	[0.055]	[0.044]
Own refrigerator	-0.011	0.049	-0.022	-0.009	-0.001	0.048
	[0.035]	[0.037]	[0.044]	[0.043]	[0.043]	[0.032]
Own cell phone	-0.020	0.015	0.071*	0.025	-0.008	0.019
	[0.032]	[0.035]	[0.040]	[0.039]	[0.037]	[0.032]
Mother working in public sector	-0.054	-0.076*	-0.063	-0.047	-0.037	-0.030
	[0.042]	[0.044]	[0.056]	[0.055]	[0.051]	[0.048]
Rural						
# Observations	1376	1376	1376	1376	1376	1373
R-squared	0.021	0.043	0.044	0.036	0.029	0.044

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 31: OLS Estimates of the Determinants of Noncognitive Skills (Male Sample)

	Feel lonely	Anxious	Easily distracted	Careless	Easily Frightened	Emotional instability
Early health shocks	0.081*	0.050	0.118*	0.019	0.108*	0.119**
	[0.049]	[0.058]	[0.061]	[0.056]	[0.063]	[0.050]
Birth weight(kg): <2	-0.074	-0.058	-0.011	0.048	-0.107*	-0.001
	[0.052]	[0.070]	[0.071]	[0.064]	[0.064]	[0.047]
Birth weight(kg): 2-2.5	-0.044	-0.080	0.068	-0.013	-0.029	-0.004
	[0.041]	[0.051]	[0.057]	[0.048]	[0.052]	[0.036]
Birth weight (kg): 2.5-3	-0.010	-0.036	0.029	0.008	-0.070	0.011
	[0.040]	[0.049]	[0.054]	[0.046]	[0.049]	[0.036]
Gender (boy=1)						
Age	0.014**	-0.008	-0.012	-0.024***	-0.005	-0.002
	[0.006]	[0.008]	[0.008]	[0.008]	[0.008]	[0.006]
Birth order	0.073	-0.107	-0.184	0.043	0.050	-0.059
	[0.083]	[0.129]	[0.129]	[0.117]	[0.112]	[0.076]
# Siblings	-0.069	0.099	0.113	-0.018	-0.107	0.058
	[0.071]	[0.119]	[0.120]	[0.116]	[0.099]	[0.077]
Mother's age	0.001	-0.004	-0.008	0.001	0.003	0.003
	[0.004]	[0.005]	[0.006]	[0.005]	[0.005]	[0.004]
Mother's education	0.007	-0.004	0.014*	0.008	-0.007	0.012**
	[0.006]	[0.007]	[0.009]	[0.007]	[0.008]	[0.006]
Per capital family income	-0.004	-0.002	-0.003	-0.012	-0.005	-0.007
	[0.006]	[0.007]	[0.007]	[0.007]	[0.007]	[0.005]
Own washing machine	0.030	0.004	0.015	0.100**	-0.002	-0.029
	[0.036]	[0.048]	[0.051]	[0.047]	[0.050]	[0.035]
Own refrigerator	-0.057	0.061	-0.050	0.026	-0.030	0.050
	[0.039]	[0.053]	[0.057]	[0.050]	[0.050]	[0.036]
Own cell phone	-0.030	0.090*	0.026	0.011	-0.067	0.012
	[0.038]	[0.049]	[0.052]	[0.048]	[0.047]	[0.037]
Mother working in public sector	0.036	-0.019	-0.021	-0.102	-0.053	-0.014
	[0.064]	[0.076]	[0.081]	[0.081]	[0.066]	[0.067]
Rural	-0.037	0.038	-0.071	0.096**	-0.027	-0.069**
	[0.036]	[0.049]	[0.052]	[0.048]	[0.048]	[0.032]
# Observations	1082	1082	1082	1082	1082	1078
R-squared	0.039	0.026	0.051	0.050	0.031	0.064

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 32: OLS Estimates of the Determinants of Noncognitive Skills (Female Sample)

	Feel lonely	Anxious	Easily distracted	Careless	Easily Frightened	Emotional instability
Early health shocks	0.073	0.093	0.091	0.213***	0.100	0.090
	[0.060]	[0.065]	[0.075]	[0.045]	[0.079]	[0.067]
Birth weight(kg): <2	-0.017	0.045	0.060	-0.030	0.024	0.016
	[0.054]	[0.052]	[0.070]	[0.066]	[0.070]	[0.053]
Birth weight(kg): 2-2.5	-0.040	0.002	-0.011	-0.007	-0.024	-0.034
	[0.048]	[0.041]	[0.062]	[0.059]	[0.062]	[0.044]
Birth weight (kg): 2.5-3	0.052	-0.029	-0.007	-0.067	-0.044	-0.020
	[0.046]	[0.039]	[0.059]	[0.056]	[0.058]	[0.041]
Gender (boy=1)						
Age	0.014**	-0.001	-0.024***	-0.030***	-0.003	-0.008
	[0.006]	[0.005]	[0.008]	[0.007]	[0.007]	[0.005]
Birth order	-0.015	0.047	0.105	0.049	-0.004	0.026
	[0.070]	[0.068]	[0.104]	[0.098]	[0.107]	[0.031]
# Siblings	-0.059	-0.064	-0.120	-0.016	-0.059	-0.071***
	[0.067]	[0.060]	[0.088]	[0.089]	[0.093]	[0.023]
Mother's age	0.000	-0.007*	0.001	0.009*	-0.011**	0.006
	[0.004]	[0.004]	[0.006]	[0.005]	[0.005]	[0.004]
Mother's education	0.002	-0.009	-0.007	-0.007	-0.006	0.009
	[0.006]	[0.006]	[0.008]	[0.007]	[0.008]	[0.006]
Per capital family income	-0.013*	-0.011*	-0.001	0.002	-0.011	-0.011*
	[0.008]	[0.006]	[0.009]	[0.008]	[0.008]	[0.006]
Own washing machine	0.039	0.053	0.011	0.005	0.018	0.023
	[0.036]	[0.038]	[0.052]	[0.049]	[0.050]	[0.031]
Own refrigerator	-0.021	-0.002	0.072	0.041	-0.024	0.043
	[0.045]	[0.045]	[0.055]	[0.053]	[0.056]	[0.042]
Own cell phone	0.046	-0.014	0.115**	-0.039	0.090*	-0.002
	[0.040]	[0.038]	[0.049]	[0.048]	[0.048]	[0.036]
Mother working in public sector	-0.086	-0.043	-0.090	0.034	0.014	-0.015
	[0.055]	[0.054]	[0.074]	[0.074]	[0.074]	[0.061]
Rural	-0.007	-0.126***	-0.007	0.024	0.027	-0.073*
	[0.041]	[0.042]	[0.051]	[0.047]	[0.050]	[0.039]
# Observations	1120	1120	1120	1120	1120	1115
R-squared	0.046	0.059	0.049	0.057	0.033	0.074

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 33: OLS Estimates of the Determinants of Noncognitive Skills (Mixed-Gender Sample)

	Feel lonely	Anxious	Easily distracted	Careless	Easily Frightened	Emotional instability
Early health shocks	-0.001	-0.048	0.050	0.052	0.003	0.096
	[0.063]	[0.064]	[0.079]	[0.072]	[0.074]	[0.064]
Birth weight(kg): <2	0.032	0.054	0.119	0.151*	0.065	0.069
	[0.061]	[0.066]	[0.086]	[0.079]	[0.073]	[0.046]
Birth weight(kg): 2-2.5	0.022	0.041	-0.010	0.038	0.043	0.008
	[0.042]	[0.048]	[0.066]	[0.061]	[0.053]	[0.031]
Birth weight (kg): 2.5-3	-0.017	0.034	0.027	-0.007	0.095**	0.006
	[0.035]	[0.042]	[0.061]	[0.053]	[0.048]	[0.028]
Gender (boy=1)	0.004	0.076***	0.081***	0.026*	0.001	0.012
	[0.016]	[0.016]	[0.021]	[0.014]	[0.016]	[0.008]
Age	0.000	-0.007	-0.020**	-0.033***	-0.009	0.000
	[0.007]	[0.008]	[0.010]	[0.010]	[0.009]	[0.005]
Birth order	-0.332**	-0.067	0.246**	0.154	0.025	-0.050
	[0.151]	[0.134]	[0.120]	[0.153]	[0.117]	[0.054]
# Siblings	0.246*	0.042	-0.290***	-0.023	-0.093	0.027
	[0.149]	[0.129]	[0.108]	[0.138]	[0.109]	[0.047]
Mother's age	0.000	-0.003	0.004	0.006	-0.006	-0.005**
	[0.004]	[0.004]	[0.006]	[0.005]	[0.005]	[0.002]
Mother's education	0.012	-0.003	0.008	-0.011	-0.010	0.011*
	[0.008]	[0.007]	[0.010]	[0.010]	[0.008]	[0.006]
Per capital family income	0.007	0.000	-0.003	0.014	-0.006	0.000
	[0.007]	[0.006]	[0.009]	[0.009]	[0.010]	[0.006]
Own washing machine	-0.038	0.008	0.099*	-0.040	-0.006	0.038
	[0.038]	[0.044]	[0.056]	[0.054]	[0.051]	[0.028]
Own refrigerator	0.007	0.051	-0.046	-0.004	-0.013	0.075*
	[0.047]	[0.056]	[0.066]	[0.068]	[0.060]	[0.043]
Own cell phone	-0.056	-0.035	-0.014	-0.001	0.015	-0.024
	[0.042]	[0.047]	[0.062]	[0.063]	[0.056]	[0.035]
Mother working in public sector	-0.056	-0.104	-0.101	0.150	-0.141	-0.104
	[0.106]	[0.082]	[0.122]	[0.120]	[0.086]	[0.092]
Rural	-0.018	-0.025	-0.038	-0.005	-0.039	-0.030
	[0.041]	[0.047]	[0.060]	[0.060]	[0.054]	[0.036]
# Observations	720	720	720	720	720	718
R-squared	0.043	0.024	0.045	0.084	0.040	0.091

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 34: OLS Estimates of the Early Health Shocks on Parental Labor Supply and Expenditure (Whole Sample)

	Father		Mother	
	Work	Expenditure	Work	Expenditure
Early health shock (only one child)	-0.044	-140.393*	0.046*	-4.554
	[0.078]	[79.028]	[0.027]	[65.095]
Early health shock (both children)	-0.008	-28.158	-0.052**	-56.055
	[0.025]	[98.141]	[0.024]	[75.360]
Age	-0.003	-17.651**	-0.001	-13.791**
	[0.003]	[6.981]	[0.002]	[6.915]
Mother's education	-0.007**	31.667***	-0.004	41.077***
	[0.004]	[9.263]	[0.004]	[11.259]
Per capital family income	0.001	108.823***	-0.003	88.422***
	[0.004]	[19.523]	[0.004]	[17.171]
Rural	0.038*	-143.895***	0.046**	-28.021
	[0.020]	[48.859]	[0.019]	[62.074]
# Observations	1163	1423	1048	1442
R-squared	0.017	0.200	0.026	0.207

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 35: OLS Estimates of the Early Health Shocks on Parental Labor Supply and Expenditure (Rural Sample)

	Father		Mother	
	Work	Expenditure	Work	Expenditure
Early health shock (only one child)	0.004	-133.164*	0.006	-12.947
	[0.050]	[70.516]	[0.040]	[46.779]
Early health shock (both children)	-0.028	-75.429	-0.040	-82.095
	[0.041]	[81.705]	[0.036]	[69.365]
Age	-0.004	-6.158	-0.001	-18.439
	[0.004]	[8.727]	[0.003]	[16.410]
Mother's education	-0.005	43.075***	-0.001	46.589*
	[0.005]	[14.811]	[0.004]	[24.519]
Per capital family income	0.003	97.484***	0.003	79.979**
	[0.004]	[27.488]	[0.004]	[33.518]
Rural				
# Observations	646	757	608	763
R-squared	0.004	0.206	0.002	0.109

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 36: **OLS Estimates of the Early Health Shocks on Parental Labor Supply and Expenditure (Urban Sample)**

	Father		Mother	
	Work	Expenditure	Work	Expenditure
Early health shock (only one child)	-0.089	-140.071	0.087***	-5.265
	[0.146]	[128.382]	[0.033]	[109.298]
Early health shock (both children)	0.013	12.607	-0.059*	-33.029
	[0.029]	[173.687]	[0.033]	[123.667]
Age	-0.003	-28.756***	-0.001	-9.698
	[0.004]	[10.921]	[0.004]	[6.043]
Mother's education	-0.009*	21.804*	-0.007	37.745***
	[0.005]	[11.882]	[0.006]	[9.074]
Per capital family income	0.001	115.218***	-0.005	93.231***
	[0.005]	[26.280]	[0.005]	[19.620]
Rural				
# Observations	517	666	440	679
R-squared	0.014	0.156	0.032	0.289

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 37: OLS Estimates of the Early Health Shocks on Parental Labor Supply and Expenditure (Male Sample)

	Father		Mother	
	Work	Expenditure	Work	Expenditure
Early health shock (only one child)	0.005	33.393	0.072*	74.990
	[0.045]	[134.589]	[0.039]	[116.180]
Early health shock (both children)	-0.027	-112.350	-0.035	37.919
	[0.041]	[106.573]	[0.030]	[119.717]
Age	-0.006	-25.475***	0.001	-9.872
	[0.006]	[9.320]	[0.005]	[6.252]
Mother's education	-0.002	14.344	-0.004	27.496**
	[0.007]	[11.816]	[0.008]	[11.486]
Per capital family income	0.000	112.468***	-0.007	86.125***
	[0.008]	[32.602]	[0.007]	[28.153]
Rural	0.061	-138.062**	0.059	-93.551***
	[0.043]	[57.227]	[0.044]	[27.363]
# Observations	417	524	386	536
R-squared	0.017	0.218	0.051	0.332

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 38: OLS Estimates of the Early Health Shocks on Parental Labor Supply and Expenditure (Female Sample)

	Father		Mother	
	Work	Expenditure	Work	Expenditure
Early health shock (only one child)	-0.122	-379.628***	-0.011	-106.805
	[0.217]	[99.651]	[0.048]	[101.787]
Early health shock (both children)	-0.010	18.052	-0.113**	-184.288***
	[0.055]	[184.022]	[0.056]	[61.327]
Age	-0.003	-14.235	-0.002	-22.942
	[0.004]	[12.585]	[0.003]	[15.588]
Mother's education	-0.009*	42.860***	-0.006	56.842***
	[0.005]	[15.952]	[0.005]	[20.803]
Per capital family income	0.005	109.204***	0.006*	122.770***
	[0.004]	[31.324]	[0.004]	[37.732]
Rural	0.057**	-59.049	0.040*	131.364
	[0.023]	[104.305]	[0.024]	[158.482]
# Observations	452	545	405	553
R-squared	0.033	0.187	0.025	0.182

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 39: OLS Estimates of the Early Health Shocks on Parental Labor Supply and Expenditure (Mix-Gender Sample)

	Father		Mother	
	Work	Expenditure	Work	Expenditure
Early health shock (only one child)	-0.024	-104.099	0.072	-12.351
	[0.095]	[161.995]	[0.046]	[58.623]
Early health shock (both children)	0.042	173.699	-0.022	50.398
	[0.050]	[281.145]	[0.054]	[119.871]
Age	0.001	-11.741	-0.002	-2.123
	[0.005]	[15.711]	[0.004]	[5.555]
Mother's education	-0.010	42.064**	-0.001	19.248**
	[0.008]	[19.217]	[0.004]	[7.453]
Per capital family income	-0.002	104.908***	-0.005	54.090***
	[0.004]	[36.301]	[0.005]	[12.474]
Rural	-0.017	-295.028***	0.027	-163.858***
	[0.038]	[79.320]	[0.031]	[33.161]
# Observations	294	354	257	353
R-squared	0.013	0.232	0.018	0.350

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 40: OLS Estimates of the Determinants of the Parent-Child Relationship (Whole Sample)

	Parents perception			Childs perception				
	Expected education level	Expected college level	Parents-children relationship	Discuss an issue	Tell them school life	Tell them what you think	Play games together	Watch TV together
Early health shocks	-0.197*	-0.066*	-0.024	-0.017	-0.110	0.005	0.134*	-0.010
	[0.106]	[0.035]	[0.049]	[0.078]	[0.072]	[0.079]	[0.079]	[0.084]
Birth weight(kg): <2	-0.303***	-0.127***	-0.037	0.136	0.097	0.012	0.126	-0.106
	[0.112]	[0.039]	[0.043]	[0.085]	[0.081]	[0.087]	[0.092]	[0.092]
Birth weight(kg): 2-2.5	-0.265***	-0.096***	0.025	0.088	0.092	0.027	0.134*	-0.099
	[0.089]	[0.031]	[0.035]	[0.066]	[0.066]	[0.071]	[0.071]	[0.077]
Birth weight (kg): 2.5-3	-0.165*	-0.064**	-0.022	-0.023	0.050	0.020	0.153**	-0.044
	[0.085]	[0.029]	[0.032]	[0.063]	[0.062]	[0.066]	[0.066]	[0.072]
Gender (boy=1)	-0.076	-0.042**	-0.046**	0.086**	0.218***	0.135***	0.103**	0.021
	[0.057]	[0.019]	[0.022]	[0.044]	[0.042]	[0.046]	[0.045]	[0.046]
Age	0.012	0.003	-0.001	-0.057***	0.005	0.002	0.035***	0.019*
	[0.012]	[0.004]	[0.005]	[0.010]	[0.010]	[0.010]	[0.010]	[0.011]
Birth order	-0.261	-0.036	-0.059	0.079	0.169	0.197	0.044	-0.044
	[0.195]	[0.066]	[0.082]	[0.135]	[0.131]	[0.128]	[0.149]	[0.158]
# Siblings	0.118	0.020	-0.033	0.090	-0.098	-0.145	0.094	0.157
	[0.177]	[0.060]	[0.075]	[0.123]	[0.121]	[0.111]	[0.136]	[0.143]
Mother's age	0.011	0.004	-0.001	0.014**	0.010	0.014**	0.015**	0.004
	[0.008]	[0.003]	[0.003]	[0.006]	[0.006]	[0.007]	[0.006]	[0.007]
Mother's education	0.119***	0.027***	0.010**	-0.045***	-0.040***	-0.037***	-0.043***	-0.019*
	[0.013]	[0.004]	[0.005]	[0.010]	[0.009]	[0.010]	[0.010]	[0.011]
Per capital family income	0.053***	0.011***	0.013**	-0.006	-0.009	-0.011	-0.001	-0.003
	[0.010]	[0.003]	[0.006]	[0.010]	[0.011]	[0.010]	[0.008]	[0.010]
Own washing machine	-0.004	-0.001	-0.021	-0.034	-0.047	0.000	-0.092	-0.277***
	[0.083]	[0.029]	[0.030]	[0.058]	[0.056]	[0.061]	[0.061]	[0.065]
Own refrigerator:	0.502***	0.163***	0.054	-0.199***	-0.082	-0.178***	-0.247***	0.045
	[0.086]	[0.032]	[0.036]	[0.066]	[0.064]	[0.068]	[0.071]	[0.065]
Own cell phone	0.302***	0.112***	-0.031	0.049	0.036	0.058	-0.093	0.016
	[0.080]	[0.028]	[0.033]	[0.060]	[0.058]	[0.062]	[0.065]	[0.060]
Mother working in public sector	0.355***	0.068**	0.080	-0.180*	-0.047	0.020	0.002	0.157
	[0.104]	[0.030]	[0.055]	[0.096]	[0.095]	[0.106]	[0.100]	[0.102]
Rural	-0.520***	-0.139***	-0.023	0.186***	0.029	-0.054	0.251***	0.057
	[0.083]	[0.029]	[0.032]	[0.059]	[0.059]	[0.063]	[0.063]	[0.060]
# Observations	2894	2894	2906	2898	2918	2902	2900	2910
R-squared	0.369	0.258	0.046	0.131	0.053	0.034	0.136	0.035

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 41: OLS Estimates of the Determinants of the Parent-Child Relationship (Rural Sample)

	Parents perception			Childs perception				
	Expected education level	Expected college level	Parents-children relationship	Discuss an issue	Tell them school life	Tell them what you think	Play games together	Watch TV together
Early health shocks	-0.193	-0.060	-0.007	-0.083	-0.151	0.021	0.066	0.105
	[0.167]	[0.054]	[0.055]	[0.094]	[0.100]	[0.109]	[0.108]	[0.128]
Birth weight(kg): <2	-0.190	-0.102*	-0.001	0.164	0.002	0.052	0.182	-0.043
	[0.162]	[0.057]	[0.058]	[0.113]	[0.112]	[0.116]	[0.121]	[0.133]
Birth weight(kg): 2-2.5	-0.234*	-0.107**	0.069	0.083	0.030	0.041	0.137	0.032
	[0.122]	[0.044]	[0.044]	[0.082]	[0.087]	[0.090]	[0.094]	[0.105]
Birth weight (kg): 2.5-3	-0.192	-0.074*	-0.001	0.010	-0.009	0.051	0.172**	0.072
	[0.117]	[0.042]	[0.040]	[0.078]	[0.080]	[0.083]	[0.087]	[0.095]
Gender (boy=1)	-0.023	-0.030	-0.038	0.027	0.114**	0.118**	0.104*	-0.011
	[0.082]	[0.028]	[0.028]	[0.056]	[0.055]	[0.059]	[0.059]	[0.063]
Age	0.063***	0.015**	0.008	-0.076***	-0.010	-0.015	0.014	0.027*
	[0.019]	[0.007]	[0.007]	[0.013]	[0.013]	[0.013]	[0.014]	[0.016]
Birth order	-0.177	-0.035	-0.020	0.040	0.172	0.242*	0.044	-0.046
	[0.221]	[0.073]	[0.093]	[0.150]	[0.138]	[0.137]	[0.175]	[0.177]
# Siblings	0.087	0.019	-0.053	0.105	-0.088	-0.220*	0.055	0.135
	[0.199]	[0.066]	[0.083]	[0.136]	[0.126]	[0.119]	[0.162]	[0.157]
Mother's age	-0.008	-0.002	0.000	0.013	0.014*	0.011	0.008	0.009
	[0.012]	[0.004]	[0.004]	[0.008]	[0.008]	[0.008]	[0.009]	[0.010]
Mother's education	0.106***	0.023***	0.014*	-0.030**	-0.028**	-0.032**	-0.051***	-0.068***
	[0.021]	[0.007]	[0.007]	[0.014]	[0.013]	[0.015]	[0.015]	[0.017]
Per capital family income	0.024	0.004	0.011	-0.006	0.000	-0.019	0.015	-0.009
	[0.019]	[0.008]	[0.008]	[0.014]	[0.017]	[0.017]	[0.013]	[0.020]
Own washing machine	-0.049	-0.008	-0.012	0.003	-0.081	0.023	-0.025	-0.227***
	[0.102]	[0.035]	[0.036]	[0.071]	[0.068]	[0.075]	[0.075]	[0.081]
Own refrigerator:	0.513***	0.187***	0.029	-0.144	-0.122	-0.220**	-0.326***	-0.053
	[0.140]	[0.054]	[0.062]	[0.101]	[0.099]	[0.105]	[0.113]	[0.103]
Own cell phone	0.508***	0.161***	0.037	-0.068	0.012	0.042	-0.212**	0.028
	[0.132]	[0.048]	[0.053]	[0.089]	[0.087]	[0.090]	[0.096]	[0.096]
Mother working in public sector	0.687**	0.070	-0.149	-0.474**	0.048	0.331	0.018	0.384
	[0.284]	[0.097]	[0.107]	[0.211]	[0.254]	[0.255]	[0.187]	[0.282]
Rural								
# Observations	1536	1536	1542	1536	1544	1538	1532	1542
R-squared	0.182	0.116	0.036	0.093	0.030	0.026	0.070	0.058

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 42: OLS Estimates of the Determinants of the Parent-Child Relationship (Urban Sample)

	Parents perception			Childs perception				
	Expected education level	Expected college level	Parents-children relationship	Discuss an issue	Tell them school life	Tell them what you think	Play games together	Watch TV together
Early health shocks	-0.209	-0.072	-0.029	0.055	-0.073	-0.034	0.182	-0.137
	[0.136]	[0.047]	[0.076]	[0.121]	[0.102]	[0.112]	[0.114]	[0.109]
Birth weight(kg): <2	-0.397**	-0.141***	-0.085	0.081	0.204*	-0.038	0.068	-0.187
	[0.156]	[0.053]	[0.064]	[0.129]	[0.116]	[0.132]	[0.136]	[0.124]
Birth weight(kg): 2-2.5	-0.288**	-0.082*	-0.031	0.057	0.186*	0.006	0.104	-0.258**
	[0.129]	[0.043]	[0.057]	[0.108]	[0.101]	[0.114]	[0.106]	[0.108]
Birth weight (kg): 2.5-3	-0.119	-0.045	-0.059	-0.084	0.126	-0.010	0.119	-0.189*
	[0.123]	[0.040]	[0.052]	[0.102]	[0.096]	[0.107]	[0.101]	[0.104]
Gender (boy=1)	-0.124	-0.050*	-0.049	0.151**	0.331***	0.148**	0.092	0.049
	[0.076]	[0.026]	[0.034]	[0.067]	[0.065]	[0.071]	[0.069]	[0.067]
Age	-0.043***	-0.012**	-0.008	-0.041***	0.018	0.019	0.056***	0.018
	[0.016]	[0.006]	[0.007]	[0.015]	[0.015]	[0.016]	[0.015]	[0.014]
Birth order	-0.146	0.096	-0.250*	0.195	0.082	-0.110	0.007	0.106
	[0.398]	[0.155]	[0.140]	[0.315]	[0.375]	[0.314]	[0.209]	[0.338]
# Siblings	-0.016	-0.061	0.088	0.051	-0.086	0.275	0.277*	0.046
	[0.357]	[0.142]	[0.130]	[0.279]	[0.354]	[0.265]	[0.143]	[0.308]
Mother's age	0.037***	0.011***	-0.003	0.015	0.003	0.014	0.021**	0.001
	[0.011]	[0.004]	[0.005]	[0.011]	[0.010]	[0.011]	[0.009]	[0.010]
Mother's education	0.124***	0.030***	0.006	-0.052***	-0.051***	-0.037**	-0.031**	0.018
	[0.016]	[0.005]	[0.007]	[0.014]	[0.013]	[0.014]	[0.015]	[0.014]
Per capital family income	0.069***	0.014***	0.015*	-0.006	-0.014	-0.006	-0.012	-0.002
	[0.013]	[0.003]	[0.008]	[0.013]	[0.014]	[0.012]	[0.011]	[0.011]
Own washing machine	0.133	0.024	-0.062	-0.140	0.001	-0.071	-0.208*	-0.250**
	[0.142]	[0.053]	[0.059]	[0.105]	[0.099]	[0.108]	[0.113]	[0.110]
Own refrigerator:	0.397***	0.123***	0.084*	-0.207**	-0.035	-0.147	-0.202**	0.067
	[0.107]	[0.040]	[0.045]	[0.091]	[0.088]	[0.091]	[0.092]	[0.084]
Own cell phone	0.175*	0.083**	-0.084**	0.136*	0.043	0.072	0.012	0.024
	[0.101]	[0.035]	[0.042]	[0.081]	[0.078]	[0.086]	[0.086]	[0.078]
Mother working in public sector	0.275**	0.059*	0.150**	-0.104	-0.042	-0.038	-0.034	0.016
	[0.114]	[0.031]	[0.063]	[0.108]	[0.104]	[0.118]	[0.118]	[0.112]
Rural								
# Observations	1358	1358	1364	1362	1374	1364	1368	1368
R-squared	0.315	0.213	0.053	0.086	0.069	0.047	0.101	0.021

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 43: OLS Estimates of the Determinants of the Parent-Child Relationship (Male Sample)

	Parents perception			Childs perception				
	Expected education level	Expected college level	Parents-children relationship	Discuss an issue	Tell them school life	Tell them what you think	Play games together	Watch TV together
Early health shocks	-0.305**	-0.108**	-0.062	-0.118	-0.147	-0.008	0.167	-0.039
	[0.138]	[0.044]	[0.071]	[0.118]	[0.115]	[0.123]	[0.123]	[0.120]
Birth weight(kg): <2	-0.410**	-0.125*	-0.035	0.036	0.135	0.077	0.187	-0.117
	[0.188]	[0.069]	[0.072]	[0.143]	[0.130]	[0.150]	[0.151]	[0.151]
Birth weight(kg): 2-2.5	-0.338**	-0.066	0.018	0.100	0.057	0.023	0.123	-0.143
	[0.133]	[0.049]	[0.057]	[0.101]	[0.104]	[0.115]	[0.110]	[0.116]
Birth weight (kg): 2.5-3	-0.220*	-0.051	-0.004	0.040	-0.049	0.000	0.184*	-0.144
	[0.130]	[0.047]	[0.054]	[0.096]	[0.097]	[0.109]	[0.099]	[0.108]
Gender (boy=1)								
Age	0.002	-0.001	0.005	-0.045***	0.018	0.020	0.038**	0.017
	[0.021]	[0.007]	[0.009]	[0.016]	[0.017]	[0.018]	[0.017]	[0.018]
Birth order	-0.520	-0.150	0.101	0.229	0.465**	0.325	0.212	0.321
	[0.376]	[0.126]	[0.139]	[0.237]	[0.193]	[0.215]	[0.324]	[0.267]
# Siblings	0.260	0.081	-0.191	0.086	-0.435**	-0.209	-0.081	-0.187
	[0.343]	[0.118]	[0.134]	[0.217]	[0.178]	[0.188]	[0.307]	[0.245]
Mother's age	0.021	0.008	-0.005	0.015	0.007	0.002	0.025**	0.002
	[0.015]	[0.005]	[0.006]	[0.011]	[0.011]	[0.012]	[0.011]	[0.013]
Mother's education	0.089***	0.016**	0.001	-0.043**	-0.025	-0.025	-0.036**	-0.031*
	[0.022]	[0.008]	[0.008]	[0.017]	[0.016]	[0.018]	[0.017]	[0.017]
Per capital family income	0.069***	0.020***	0.010	0.015	0.004	-0.008	0.001	0.001
	[0.016]	[0.005]	[0.010]	[0.015]	[0.018]	[0.015]	[0.011]	[0.014]
Own washing machine	-0.223	-0.029	-0.072	-0.104	-0.152	-0.160	-0.052	-0.304***
	[0.140]	[0.050]	[0.050]	[0.099]	[0.093]	[0.106]	[0.096]	[0.102]
Own refrigerator:	0.678***	0.206***	0.063	-0.172	-0.097	-0.179	-0.236**	0.065
	[0.142]	[0.053]	[0.058]	[0.111]	[0.108]	[0.117]	[0.115]	[0.110]
Own cell phone	0.327**	0.131***	0.028	0.207**	0.211**	0.180*	-0.043	-0.002
	[0.135]	[0.048]	[0.048]	[0.102]	[0.104]	[0.108]	[0.099]	[0.097]
Mother working in public sector	0.462**	0.069	0.069	-0.177	-0.009	0.084	-0.151	0.215
	[0.200]	[0.052]	[0.083]	[0.146]	[0.169]	[0.167]	[0.153]	[0.161]
Rural	-0.444***	-0.115**	-0.029	0.088	0.023	-0.071	0.324***	-0.066
	[0.135]	[0.048]	[0.052]	[0.097]	[0.101]	[0.107]	[0.096]	[0.098]
# Observations	1074	1074	1076	1074	1082	1076	1076	1082
R-squared	0.373	0.277	0.040	0.105	0.030	0.026	0.140	0.038

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 44: OLS Estimates of the Determinants of the Parent-Child Relationship (Female Sample)

	Parents perception			Childs perception				
	Expected education level	Expected college level	Parents-children relationship	Discuss an issue	Tell them school life	Tell them what you think	Play games together	Watch TV together
Early health shocks	-0.015	-0.017	0.063	0.103	-0.188	0.096	0.265*	0.007
	[0.220]	[0.074]	[0.089]	[0.145]	[0.120]	[0.133]	[0.158]	[0.148]
Birth weight(kg): <2	-0.210	-0.159***	0.009	0.316**	0.004	-0.017	0.101	-0.191
	[0.180]	[0.061]	[0.074]	[0.138]	[0.133]	[0.139]	[0.160]	[0.152]
Birth weight(kg): 2-2.5	-0.151	-0.131**	0.093	0.169	-0.029	0.067	0.113	-0.105
	[0.156]	[0.052]	[0.065]	[0.124]	[0.118]	[0.126]	[0.143]	[0.145]
Birth weight (kg): 2.5-3	-0.145	-0.116**	0.023	0.034	-0.038	0.037	0.116	0.080
	[0.149]	[0.051]	[0.057]	[0.119]	[0.113]	[0.118]	[0.135]	[0.139]
Gender (boy=1)								
Age	0.030	0.011	-0.004	-0.049***	-0.001	0.000	0.065***	0.040**
	[0.020]	[0.007]	[0.008]	[0.016]	[0.015]	[0.017]	[0.017]	[0.017]
Birth order	-0.174	0.011	-0.010	0.163	0.285	0.405**	0.082	-0.220
	[0.280]	[0.093]	[0.116]	[0.197]	[0.201]	[0.192]	[0.208]	[0.234]
# Siblings	0.121	0.014	-0.020	0.098	-0.031	-0.170	0.227	0.323
	[0.252]	[0.082]	[0.104]	[0.179]	[0.182]	[0.173]	[0.183]	[0.203]
Mother's age	0.008	-0.002	0.002	-0.002	0.017	0.012	-0.010	-0.004
	[0.015]	[0.005]	[0.006]	[0.011]	[0.011]	[0.012]	[0.011]	[0.011]
Mother's education	0.146***	0.036***	0.017**	-0.033**	-0.060***	-0.039**	-0.047**	-0.005
	[0.020]	[0.007]	[0.008]	[0.015]	[0.015]	[0.017]	[0.019]	[0.018]
Per capital family income	0.050***	0.011*	0.015	-0.023	-0.023	-0.019	0.007	-0.018
	[0.018]	[0.006]	[0.010]	[0.015]	[0.018]	[0.021]	[0.018]	[0.018]
Own washing machine	0.085	0.013	-0.012	-0.016	0.093	0.055	-0.125	-0.194*
	[0.135]	[0.047]	[0.050]	[0.096]	[0.088]	[0.101]	[0.111]	[0.108]
Own refrigerator:	0.431***	0.111**	0.075	-0.260**	-0.017	-0.195*	-0.279**	0.113
	[0.134]	[0.050]	[0.057]	[0.111]	[0.103]	[0.106]	[0.123]	[0.103]
Own cell phone	0.258**	0.119***	-0.101*	-0.136	-0.112	0.049	-0.070	-0.020
	[0.122]	[0.044]	[0.054]	[0.099]	[0.088]	[0.099]	[0.112]	[0.100]
Mother working in public sector	0.168	0.019	0.083	-0.291**	-0.034	-0.093	0.106	0.103
	[0.134]	[0.045]	[0.088]	[0.145]	[0.128]	[0.160]	[0.161]	[0.159]
Rural	-0.550***	-0.173***	-0.019	0.140	0.031	-0.106	0.177	0.065
	[0.129]	[0.047]	[0.051]	[0.098]	[0.098]	[0.101]	[0.112]	[0.101]
# Observations	1106	1106	1114	1106	1116	1110	1112	1116
R-squared	0.409	0.299	0.059	0.182	0.101	0.051	0.145	0.041

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 45: OLS Estimates of the Determinants of the Parent-Child Relationship (Mixed-Gender Sample)

	Parents perception			Childs perception				
	Expected education level	Expected college level	Parents-children relationship	Discuss an issue	Tell them school life	Tell them what you think	Play games together	Watch TV together
Early health shocks	-0.281	-0.043	-0.098	-0.089	-0.089	-0.204	-0.083	0.052
	[0.239]	[0.077]	[0.093]	[0.115]	[0.127]	[0.162]	[0.137]	[0.182]
Birth weight(kg): <2	-0.285	-0.113	-0.090	-0.100	0.091	0.065	0.204	0.232
	[0.241]	[0.082]	[0.082]	[0.170]	[0.163]	[0.181]	[0.190]	[0.207]
Birth weight(kg): 2-2.5	-0.290	-0.109*	-0.032	0.047	0.343***	0.029	0.213*	0.033
	[0.189]	[0.062]	[0.062]	[0.120]	[0.121]	[0.130]	[0.127]	[0.153]
Birth weight (kg): 2.5-3	-0.111	-0.031	-0.092*	-0.132	0.283***	0.053	0.177	-0.018
	[0.178]	[0.057]	[0.055]	[0.110]	[0.105]	[0.115]	[0.120]	[0.140]
Gender (boy=1)	-0.113**	-0.030*	-0.037**	0.080*	0.162***	0.007	-0.015	0.091**
	[0.044]	[0.017]	[0.015]	[0.047]	[0.050]	[0.052]	[0.044]	[0.045]
Age	-0.005	-0.001	-0.002	-0.070***	-0.015	-0.023	-0.014	-0.002
	[0.028]	[0.009]	[0.010]	[0.020]	[0.019]	[0.019]	[0.019]	[0.023]
Birth order	-0.315	0.044	-0.462**	-0.335	-0.598**	-0.506**	-0.222	-0.404
	[0.397]	[0.139]	[0.187]	[0.283]	[0.281]	[0.253]	[0.216]	[0.336]
# Siblings	0.145	-0.042	0.287*	0.276	0.479*	0.263	0.188	0.492*
	[0.349]	[0.124]	[0.165]	[0.255]	[0.258]	[0.207]	[0.172]	[0.292]
Mother's age	0.004	0.004	0.000	0.024**	0.003	0.019*	0.024**	0.013
	[0.015]	[0.005]	[0.005]	[0.011]	[0.010]	[0.010]	[0.010]	[0.013]
Mother's education	0.124***	0.027***	0.013	-0.055***	-0.030*	-0.039**	-0.027	-0.022
	[0.030]	[0.009]	[0.010]	[0.019]	[0.017]	[0.019]	[0.019]	[0.023]
Per capital family income	0.039*	0.001	0.015*	-0.023	-0.021	-0.011	-0.008	0.007
	[0.023]	[0.007]	[0.008]	[0.016]	[0.015]	[0.013]	[0.015]	[0.018]
Own washing machine	0.131	0.007	0.018	0.025	-0.049	0.147	-0.065	-0.323**
	[0.166]	[0.056]	[0.057]	[0.108]	[0.108]	[0.115]	[0.113]	[0.133]
Own refrigerator:	0.283	0.144**	-0.008	-0.181	-0.118	-0.087	-0.231*	-0.049
	[0.186]	[0.068]	[0.073]	[0.126]	[0.117]	[0.137]	[0.135]	[0.128]
Own cell phone	0.332*	0.064	0.017	0.186*	0.002	-0.096	-0.274**	0.083
	[0.185]	[0.062]	[0.069]	[0.110]	[0.108]	[0.126]	[0.130]	[0.123]
Mother working in public sector	0.694**	0.191***	0.146	0.265	-0.125	0.240	0.002	0.092
	[0.270]	[0.058]	[0.139]	[0.200]	[0.193]	[0.286]	[0.254]	[0.231]
Rural	-0.629***	-0.149**	-0.011	0.388***	0.050	0.050	0.198	0.200*
	[0.177]	[0.059]	[0.066]	[0.113]	[0.107]	[0.123]	[0.121]	[0.120]
# Observations	714	714	716	718	720	716	712	712
R-squared	0.324	0.191	0.082	0.158	0.070	0.041	0.128	0.065

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors, clustered at the household level, are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

8 Within-Twin-Pair FE Estimation Results

Table 46: Within-Twin-Pair FE Estimates of the Determinants of Family Investments (Whole Sample)

	Health	Education	Clothing	Parents home tutor
Early health shocks	1.351***	-0.204***	-0.058	-1.493
	[0.314]	[0.073]	[0.042]	[1.263]
Birth weight(kg): <2	0.521**	-0.015	0.006	0.255
	[0.210]	[0.047]	[0.036]	[1.411]
Birth weight(kg): 2-2.5	0.468***	0.016	0.027	0.257
	[0.163]	[0.026]	[0.030]	[1.194]
Birth weight (kg): 2.5-3	0.421***	-0.013	-0.004	-0.319
	[0.149]	[0.023]	[0.020]	[0.993]
Gender (boy=1)	0.086	-0.024	-0.028	0.235
	[0.093]	[0.021]	[0.017]	[0.779]
# Observations	1461	1461	1461	1451
R-squared	0.027	0.016	0.005	0.001

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 47: Within-Twin-Pair FE Estimates of the Determinants of Family Investments (Rural Sample)

	Health	Education	Clothing	Parents home tutor
Early health shocks	1.523***	-0.058	-0.120	-2.041
	[0.538]	[0.069]	[0.092]	[1.966]
Birth weight(kg): <2	0.483*	0.076	0.015	-0.753
	[0.272]	[0.048]	[0.040]	[1.976]
Birth weight(kg): 2-2.5	0.572***	0.028	0.004	0.364
	[0.196]	[0.020]	[0.040]	[1.478]
Birth weight (kg): 2.5-3	0.488***	-0.004	-0.024	-0.225
	[0.170]	[0.008]	[0.023]	[1.090]
Gender (boy=1)	0.139	-0.008	-0.018	0.113
	[0.122]	[0.013]	[0.011]	[0.863]
# Observations	773	773	773	764
R-squared	0.037	0.014	0.014	0.002

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 48: Within-Twin-Pair FE Estimates of the Determinants of Family Investments (Urban Sample)

	Health	Education	Clothing	Parents home tutor
Early health shocks	1.149***	-0.328***	-0.018	-0.962
	[0.374]	[0.116]	[0.021]	[1.678]
Birth weight(kg): <2	0.548	-0.107	0.008	1.155
	[0.342]	[0.086]	[0.062]	[2.153]
Birth weight(kg): 2-2.5	0.344	0.001	0.060	0.100
	[0.286]	[0.057]	[0.045]	[2.030]
Birth weight (kg): 2.5-3	0.339	-0.025	0.030	-0.473
	[0.271]	[0.057]	[0.034]	[1.838]
Gender (boy=1)	-0.006	-0.046	-0.042	0.392
	[0.145]	[0.049]	[0.039]	[1.434]
# Observations	688	688	688	687
R-squared	0.020	0.035	0.005	0.002

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 49: Within-Twin-Pair FE Estimates of the Determinants of Family Investments (Male Sample)

	Health	Education	Clothing	Parents home tutor
Early health shocks	1.085**	-0.171**	-0.108	-2.393
	[0.426]	[0.074]	[0.091]	[2.565]
Birth weight(kg): <2	0.561	0.059	0.002	0.455
	[0.348]	[0.066]	[0.046]	[2.712]
Birth weight(kg): 2-2.5	0.269	0.011	0.008	1.115
	[0.240]	[0.029]	[0.054]	[2.332]
Birth weight (kg): 2.5-3	0.094	-0.020	-0.044	0.548
	[0.226]	[0.013]	[0.037]	[1.905]
Gender (boy=1)				
# Observations	541	541	541	539
R-squared	0.023	0.025	0.007	0.002

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 50: Within-Twin-Pair FE Estimates of the Determinants of Family Investments (Female Sample)

	Health	Education	Clothing	Parents home tutor
Early health shocks	2.080***	-0.410**	-0.028	0.868
	[0.708]	[0.188]	[0.030]	[0.628]
Birth weight(kg): <2	0.277	0.019	0.019	1.018
	[0.374]	[0.097]	[0.079]	[1.954]
Birth weight(kg): 2-2.5	0.291	0.114*	0.056	2.117
	[0.319]	[0.060]	[0.059]	[1.535]
Birth weight (kg): 2.5-3	0.393	0.053	0.031	2.172*
	[0.293]	[0.057]	[0.045]	[1.161]
Gender (boy=1)				
# Observations	560	560	560	556
R-squared	0.045	0.051	0.004	0.004

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 51: Within-Twin-Pair FE Estimates of the Determinants of Family Investments (Mixed-Gender Sample)

	Health	Education	Clothing	Parents home tutor
Early health shocks	0.824**	-0.006	-0.024	-2.695
	[0.332]	[0.053]	[0.032]	[2.333]
Birth weight(kg): <2	0.664*	-0.076*	0.012	1.120
	[0.391]	[0.046]	[0.033]	[2.455]
Birth weight(kg): 2-2.5	0.834***	-0.066	0.024	-2.232
	[0.311]	[0.042]	[0.025]	[2.081]
Birth weight (kg): 2.5-3	0.768***	-0.064	0.000	-3.675**
	[0.264]	[0.043]	[0.016]	[1.809]
Gender (boy=1)	0.111	-0.032	-0.029	0.184
	[0.091]	[0.022]	[0.018]	[0.784]
# Observations	360	360	360	356
R-squared	0.038	0.009	0.010	0.019

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Parents home tutor is measured in minutes per day.

Table 52: Within-Twin-Pair FE Estimates of the Determinants of Health (Whole Sample)

	Height (*z*-score)	Weight (*z*-score)	BMI (*z*-score)	Health Status	Hospital Visits
Early health shocks	-0.005	-0.270**	-0.204*	-0.449***	0.163***
	[0.102]	[0.115]	[0.121]	[0.111]	[0.051]
Birth weight(kg): <2	-0.389***	-0.524***	-0.367***	-0.078	0.030
	[0.087]	[0.080]	[0.095]	[0.055]	[0.041]
Birth weight(kg): 2-2.5	-0.329***	-0.323***	-0.141*	-0.081*	0.080**
	[0.074]	[0.065]	[0.081]	[0.044]	[0.036]
Birth weight (kg): 2.5-3	-0.228***	-0.207***	-0.054	-0.037	0.046
	[0.065]	[0.054]	[0.068]	[0.037]	[0.032]
Gender (boy=1)	0.027	0.058	0.031	-0.032	0.047**
	[0.052]	[0.041]	[0.052]	[0.026]	[0.020]
# Observations	1423	1435	1411	1455	1451
R-squared	0.021	0.048	0.021	0.047	0.019

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 53: Within-Twin-Pair FE Estimates of the Determinants of Health (Rural Sample)

	Height (*z*-score)	Weight (*z*-score)	BMI (*z*-score)	Health Status	Hospital Visits
Early health shocks	-0.114	-0.494***	-0.418***	-0.522***	0.085
	[0.165]	[0.154]	[0.121]	[0.164]	[0.060]
Birth weight(kg): <2	-0.393***	-0.433***	-0.226*	-0.141*	0.028
	[0.133]	[0.108]	[0.126]	[0.084]	[0.050]
Birth weight(kg): 2-2.5	-0.379***	-0.228***	0.043	-0.105*	0.107***
	[0.108]	[0.080]	[0.105]	[0.060]	[0.037]
Birth weight (kg): 2.5-3	-0.219**	-0.158**	0.048	-0.024	0.060*
	[0.093]	[0.066]	[0.086]	[0.046]	[0.033]
Gender (boy=1)	0.054	0.087	0.095	-0.033	0.039
	[0.071]	[0.054]	[0.064]	[0.038]	[0.025]
# Observations	740	755	745	773	771
R-squared	0.027	0.052	0.031	0.055	0.020

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 54: Within-Twin-Pair FE Estimates of the Determinants of Health (Urban Sample)

	Height (z-score)	Weight (z-score)	BMI (z-score)	Health Status	Hospital Visits
Early health shocks	0.079	-0.091	-0.040	-0.409***	0.221***
	[0.127]	[0.158]	[0.184]	[0.151]	[0.076]
Birth weight(kg): <2	-0.383***	-0.640***	-0.548***	-0.015	0.024
	[0.107]	[0.121]	[0.143]	[0.071]	[0.073]
Birth weight(kg): 2-2.5	-0.282***	-0.445***	-0.368***	-0.059	0.043
	[0.094]	[0.108]	[0.124]	[0.067]	[0.068]
Birth weight (kg): 2.5-3	-0.232***	-0.273***	-0.188*	-0.048	0.026
	[0.084]	[0.092]	[0.106]	[0.061]	[0.062]
Gender (boy=1)	-0.002	0.023	-0.057	-0.032	0.058*
	[0.075]	[0.064]	[0.087]	[0.033]	[0.033]
# Observations	683	680	666	682	680
R-squared	0.019	0.058	0.032	0.046	0.023

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 55: Within-Twin-Pair FE Estimates of the Determinants of Health (Male Sample)

	Height (z-score)	Weight (z-score)	BMI (z-score)	Health Status	Hospital Visits
Early health shocks	0.038	-0.073	0.102	-0.441**	0.205**
	[0.100]	[0.186]	[0.157]	[0.186]	[0.084]
Birth weight(kg): <2	-0.449***	-0.473***	-0.256*	-0.042	0.045
	[0.114]	[0.127]	[0.152]	[0.083]	[0.078]
Birth weight(kg): 2-2.5	-0.332***	-0.360***	-0.168	-0.034	0.056
	[0.103]	[0.110]	[0.141]	[0.068]	[0.063]
Birth weight (kg): 2.5-3	-0.210**	-0.143*	-0.019	0.013	0.034
	[0.087]	[0.080]	[0.111]	[0.050]	[0.058]
Gender (boy=1)					
# Observations	524	527	519	538	535
R-squared	0.039	0.048	0.015	0.057	0.018

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 56: Within-Twin-Pair FE Estimates of the Determinants of Health (Female Sample)

	Height (z-score)	Weight (z-score)	BMI (z-score)	Health Status	Hospital Visits
Early health shocks	0.107	-0.251**	-0.355**	-0.327*	0.115
	[0.154]	[0.122]	[0.159]	[0.191]	[0.081]
Birth weight(kg): <2	-0.071	-0.525***	-0.571***	-0.085	0.049
	[0.098]	[0.110]	[0.127]	[0.078]	[0.065]
Birth weight(kg): 2-2.5	-0.065	-0.297***	-0.304***	-0.072	0.099*
	[0.073]	[0.086]	[0.101]	[0.059]	[0.059]
Birth weight (kg): 2.5-3	-0.066	-0.226***	-0.200**	-0.031	0.054
	[0.056]	[0.077]	[0.087]	[0.047]	[0.051]
Gender (boy=1)					
# Observations	548	553	544	558	558
R-squared	0.002	0.056	0.054	0.027	0.014

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 57: Within-Twin-Pair FE Estimates of the Determinants of Health (Mixed-Gender Sample)

	Height (z-score)	Weight (z-score)	BMI (z-score)	Health Status	Hospital Visits
Early health shocks	-0.303	-0.641**	-0.497	-0.641***	0.152
	[0.315]	[0.262]	[0.312]	[0.172]	[0.104]
Birth weight(kg): <2	-0.699***	-0.613***	-0.310	-0.110	-0.015
	[0.230]	[0.179]	[0.215]	[0.122]	[0.072]
Birth weight(kg): 2-2.5	-0.617***	-0.344**	0.027	-0.161*	0.090
	[0.180]	[0.134]	[0.166]	[0.097]	[0.065]
Birth weight (kg): 2.5-3	-0.376**	-0.254**	0.037	-0.094	0.055
	[0.156]	[0.111]	[0.137]	[0.085]	[0.055]
Gender (boy=1)	0.019	0.061	0.040	-0.031	0.046**
	[0.053]	[0.041]	[0.052]	[0.026]	[0.020]
# Observations	351	355	348	359	358
R-squared	0.047	0.064	0.024	0.068	0.032

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 58: Within-Twin-Pair FE Estimates of the Determinants of Academic Performance (Whole Sample)

	Literature (self-reported)	Mathematics (self-reported)	Literature (exam record)	Mathematics (exam record)
Early health shocks	-0.345**	-0.529***	-5.158***	-5.384**
	[0.150]	[0.147]	[1.659]	[2.659]
Birth weight(kg): <2	-0.184*	-0.182	-2.335*	-2.454
	[0.102]	[0.114]	[1.380]	[1.653]
Birth weight(kg): 2-2.5	-0.081	-0.038	-1.281	-0.395
	[0.075]	[0.090]	[1.142]	[1.329]
Birth weight (kg): 2.5-3	-0.050	0.011	-1.155	-0.151
	[0.065]	[0.075]	[0.961]	[1.136]
Gender (boy=1)	-0.179***	0.014	-2.659***	-0.633
	[0.047]	[0.054]	[0.713]	[0.843]
# Observations	1431	1425	1362	1343
R-squared	0.022	0.017	0.022	0.009

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 59: Within-Twin-Pair FE Estimates of the Determinants of Academic Performance (Rural Sample)

	Literature (self-reported)	Mathematics (self-reported)	Literature (exam record)	Mathematics (exam record)
Early health shocks	-0.186	-0.598***	-5.441**	-2.604
	[0.160]	[0.168]	[2.572]	[2.666]
Birth weight(kg): <2	-0.107	-0.255	-3.175	-2.489
	[0.148]	[0.157]	[2.041]	[2.247]
Birth weight(kg): 2-2.5	-0.012	-0.012	-2.361	-0.543
	[0.090]	[0.112]	[1.626]	[1.865]
Birth weight (kg): 2.5-3	0.023	0.120	-2.244	-1.317
	[0.068]	[0.082]	[1.385]	[1.597]
Gender (boy=1)	-0.178***	-0.096	-3.231***	-2.027*
	[0.062]	[0.066]	[1.039]	[1.161]
# Observations	759	757	711	705
R-squared	0.017	0.034	0.027	0.009

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 60: Within-Twin-Pair FE Estimates of the Determinants of Academic Performance (Urban Sample)

	Literature (self-reported)	Mathematics (self-reported)	Literature (exam record)	Mathematics (exam record)
Early health shocks	-0.511**	-0.510**	-4.632**	-7.330*
	[0.236]	[0.227]	[2.214]	[4.279]
Birth weight(kg): <2	-0.289*	-0.166	-1.123	-2.135
	[0.152]	[0.172]	[1.797]	[2.384]
Birth weight(kg): 2-2.5	-0.181	-0.116	0.221	0.127
	[0.134]	[0.151]	[1.511]	[1.822]
Birth weight (kg): 2.5-3	-0.154	-0.151	0.291	1.250
	[0.125]	[0.140]	[1.235]	[1.540]
Gender (boy=1)	-0.180**	0.173**	-1.893**	1.203
	[0.074]	[0.088]	[0.914]	[1.213]
# Observations	672	668	651	638
R-squared	0.036	0.022	0.018	0.024

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 61: Within-Twin-Pair FE Estimates of the Determinants of Academic Performance (Male Sample)

	Literature (self-reported)	Mathematics (self-reported)	Literature (exam record)	Mathematics (exam record)
Early health shocks	-0.401*	-0.653***	-5.630**	-6.272*
	[0.242]	[0.229]	[2.527]	[3.718]
Birth weight(kg): <2	-0.163	-0.069	0.985	-3.225
	[0.181]	[0.191]	[2.342]	[2.431]
Birth weight(kg): 2-2.5	-0.201*	-0.096	-1.438	-2.844
	[0.118]	[0.145]	[1.736]	[1.916]
Birth weight (kg): 2.5-3	-0.194*	-0.079	-1.846	-3.046**
	[0.111]	[0.128]	[1.292]	[1.336]
Gender (boy=1)				
# Observations	527	523	507	500
R-squared	0.018	0.022	0.019	0.015

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 62: **Within-Twin-Pair FE Estimates of the Determinants of Academic Performance (Female Sample)**

	Literature (self-reported)	Mathematics (self-reported)	Literature (exam record)	Mathematics (exam record)
Early health shocks	-0.461*	-0.457*	-2.020	-3.776
	[0.241]	[0.270]	[1.940]	[5.815]
Birth weight(kg): <2	-0.213	-0.014	-4.843*	-0.483
	[0.147]	[0.169]	[2.531]	[2.988]
Birth weight(kg): 2-2.5	-0.085	0.068	-2.808	0.156
	[0.108]	[0.134]	[2.351]	[2.542]
Birth weight (kg): 2.5-3	-0.097	0.064	-2.142	-0.949
	[0.090]	[0.114]	[2.194]	[2.331]
Gender (boy=1)				
# Observations	551	549	519	511
R-squared	0.017	0.011	0.010	0.004

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 63: Within-Twin-Pair FE Estimates of the Determinants of Academic Performance (Mixed-Gender Sample)

	Literature (self-reported)	Mathematics (self-reported)	Literature (exam record)	Mathematics (exam record)
Early health shocks	-0.133	-0.470*	-8.104**	-6.843
	[0.292]	[0.247]	[3.773]	[4.354]
Birth weight(kg): <2	-0.204	-0.586**	-3.897	-5.630*
	[0.206]	[0.230]	[2.411]	[3.041]
Birth weight(kg): 2-2.5	0.061	-0.071	0.448	1.866
	[0.163]	[0.189]	[2.007]	[2.524]
Birth weight (kg): 2.5-3	0.136	0.067	0.340	3.632*
	[0.126]	[0.144]	[1.476]	[2.150]
Gender (boy=1)	-0.180***	0.001	-2.632***	-0.632
	[0.047]	[0.055]	[0.723]	[0.843]
# Observations	353	353	336	332
R-squared	0.055	0.046	0.067	0.048

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 64: Within-Twin-Pair FE Estimates of the Determinants of Schooling Performance (Whole Sample)

	Good Student Awards	Awards in Contests	Grade repetition	Parents interviewed	Minor actions in class
Early health shocks	-0.185***	-0.087**	0.109***	0.100*	0.279**
	[0.069]	[0.039]	[0.042]	[0.061]	[0.131]
Birth weight(kg): <2	-0.070	0.017	-0.010	0.033	-0.002
	[0.050]	[0.029]	[0.026]	[0.040]	[0.087]
Birth weight(kg): 2-2.5	-0.034	0.018	-0.014	0.030	-0.079
	[0.041]	[0.027]	[0.022]	[0.030]	[0.067]
Birth weight (kg): 2.5-3	-0.016	0.011	-0.008	0.035	0.040
	[0.037]	[0.024]	[0.019]	[0.027]	[0.056]
Gender (boy=1)	-0.086***	-0.031**	0.003	0.083***	0.293***
	[0.027]	[0.014]	[0.014]	[0.021]	[0.045]
# Observations	1461	1461	1461	1459	1445
R-squared	0.019	0.010	0.010	0.019	0.050

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 65: Within-Twin-Pair FE Estimates of the Determinants of Schooling Performance (Rural Sample)

	Good Student Awards	Awards in Contests	Grade repetition	Parents interviewed	Minor actions in class
Early health shocks	-0.096	0.006	0.083	-0.039	0.396*
	[0.086]	[0.006]	[0.055]	[0.042]	[0.233]
Birth weight(kg): <2	-0.131**	0.003	-0.024	-0.003	0.015
	[0.060]	[0.028]	[0.043]	[0.042]	[0.123]
Birth weight(kg): 2-2.5	-0.099**	0.006	-0.014	0.032	-0.054
	[0.046]	[0.022]	[0.034]	[0.032]	[0.086]
Birth weight (kg): 2.5-3	-0.082**	0.006	0.004	0.050**	0.084
	[0.037]	[0.022]	[0.029]	[0.025]	[0.067]
Gender (boy=1)	-0.110***	-0.028**	0.002	0.041*	0.377***
	[0.033]	[0.013]	[0.021]	[0.024]	[0.061]
# Observations	773	773	773	771	766
R-squared	0.028	0.007	0.006	0.011	0.084

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 66: Within-Twin-Pair FE Estimates of the Determinants of Schooling Performance (Urban Sample)

	Good Student Awards	Awards in Contests	Grade repetition	Parents interviewed	Minor actions in class
Early health shocks	-0.236**	-0.163**	0.125**	0.197**	0.169
	[0.105]	[0.068]	[0.062]	[0.100]	[0.143]
Birth weight(kg): <2	0.016	0.037	0.000	0.057	-0.021
	[0.086]	[0.058]	[0.026]	[0.072]	[0.127]
Birth weight(kg): 2-2.5	0.059	0.036	-0.019	0.018	-0.112
	[0.075]	[0.055]	[0.022]	[0.058]	[0.108]
Birth weight (kg): 2.5-3	0.071	0.019	-0.025	0.011	-0.010
	[0.071]	[0.049]	[0.020]	[0.055]	[0.095]
Gender (boy=1)	-0.054	-0.036	0.004	0.144***	0.174***
	[0.045]	[0.028]	[0.018]	[0.037]	[0.066]
# Observations	688	688	688	688	679
R-squared	0.019	0.017	0.023	0.047	0.021

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 67: Within-Twin-Pair FE Estimates of the Determinants of Schooling Performance (Male Sample)

	Good Student Awards	Awards in Contests	Grade repetition	Parents interviewed	Minor actions in class
Early health shocks	-0.252**	-0.131*	0.127*	0.045	0.203
	[0.103]	[0.068]	[0.068]	[0.108]	[0.182]
Birth weight(kg): <2	-0.179**	0.043	-0.040	0.082	0.060
	[0.081]	[0.047]	[0.049]	[0.061]	[0.145]
Birth weight(kg): 2-2.5	-0.118*	0.042	-0.019	0.082	-0.098
	[0.061]	[0.043]	[0.037]	[0.052]	[0.093]
Birth weight (kg): 2.5-3	-0.123**	0.013	-0.014	0.074	0.033
	[0.060]	[0.038]	[0.039]	[0.047]	[0.072]
Gender (boy=1)					
# Observations	541	541	541	541	532
R-squared	0.031	0.014	0.020	0.006	0.014

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 68: Within-Twin-Pair FE Estimates of the Determinants of Schooling Performance (Female Sample)

	Good Student Awards	Awards in Contests	Grade repetition	Parents interviewed	Minor actions in class
Early health shocks	-0.223*	-0.054	0.109	0.111	0.282
	[0.129]	[0.056]	[0.074]	[0.075]	[0.207]
Birth weight(kg): <2	0.022	-0.006	-0.023	0.012	-0.088
	[0.073]	[0.049]	[0.046]	[0.059]	[0.132]
Birth weight(kg): 2-2.5	0.055	-0.007	-0.044	-0.001	-0.103
	[0.064]	[0.046]	[0.044]	[0.043]	[0.107]
Birth weight (kg): 2.5-3	0.035	-0.001	-0.038	0.004	0.024
	[0.055]	[0.038]	[0.039]	[0.039]	[0.082]
Gender (boy=1)					
# Observations	560	560	560	559	554
R-squared	0.012	0.002	0.016	0.006	0.013

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 69: Within-Twin-Pair FE Estimates of the Determinants of Schooling Performance (Mixed-Gender Sample)

	Good Student Awards	Awards in Contests	Grade repetition	Parents interviewed	Minor actions in class
Early health shocks	-0.052	-0.055	0.074	0.181	0.437
	[0.124]	[0.075]	[0.075]	[0.119]	[0.333]
Birth weight(kg): <2	-0.034	0.010	0.022	-0.002	0.063
	[0.114]	[0.060]	[0.055]	[0.091]	[0.182]
Birth weight(kg): 2-2.5	-0.037	0.022	0.018	0.012	-0.005
	[0.086]	[0.050]	[0.035]	[0.061]	[0.143]
Birth weight (kg): 2.5-3	0.045	0.021	0.022	0.019	0.056
	[0.072]	[0.048]	[0.018]	[0.052]	[0.126]
Gender (boy=1)	-0.086***	-0.031**	0.005	0.080***	0.291***
	[0.028]	[0.014]	[0.015]	[0.022]	[0.046]
# Observations	360	360	360	359	359
R-squared	0.035	0.018	0.004	0.049	0.118

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 70: Within-Twin-Pair FE Estimates of the Determinants of Noncognitive Skills (Whole Sample)

	Feel lonely	Anxious	Easily distracted	Careless	Easily Frightened	Emotional instability
Early health shocks	0.165***	0.121***	0.151***	0.142***	0.144***	0.112**
	[0.062]	[0.045]	[0.055]	[0.053]	[0.048]	[0.051]
Birth weight(kg): <2	0.002	-0.035	-0.006	0.029	-0.011	0.020
	[0.030]	[0.033]	[0.038]	[0.021]	[0.035]	[0.017]
Birth weight(kg): 2-2.5	0.013	-0.003	-0.023	0.026	0.007	0.009
	[0.023]	[0.027]	[0.033]	[0.018]	[0.027]	[0.015]
Birth weight (kg): 2.5-3	0.009	0.000	-0.025	0.017	-0.011	0.011
	[0.019]	[0.023]	[0.028]	[0.015]	[0.022]	[0.013]
Gender (boy=1)	0.000	0.069***	0.074***	0.021	-0.006	0.010
	[0.015]	[0.016]	[0.020]	[0.013]	[0.016]	[0.008]
# Observations	1461	1461	1461	1461	1461	1455
R-squared	0.018	0.032	0.023	0.022	0.015	0.026

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 71: Within-Twin-Pair FE Estimates of the Determinants of Noncognitive Skills (Rural Sample)

	Feel lonely	Anxious	Easily distracted	Careless	Easily Frightened	Emotional instability
Early health shocks	0.253***	0.071	0.148**	0.165**	0.092	0.044
	[0.089]	[0.056]	[0.071]	[0.075]	[0.057]	[0.074]
Birth weight(kg): <2	-0.014	0.016	0.028	0.044	-0.084	0.010
	[0.044]	[0.051]	[0.054]	[0.028]	[0.053]	[0.018]
Birth weight(kg): 2-2.5	0.001	0.022	0.000	0.027	-0.011	0.008
	[0.027]	[0.042]	[0.039]	[0.018]	[0.039]	[0.012]
Birth weight (kg): 2.5-3	0.005	0.022	0.004	0.011	-0.030	0.013
	[0.020]	[0.034]	[0.032]	[0.010]	[0.030]	[0.010]
Gender (boy=1)	0.004	0.076***	0.083***	0.013	-0.006	0.010
	[0.019]	[0.022]	[0.026]	[0.015]	[0.024]	[0.007]
# Observations	773	773	773	773	773	769
R-squared	0.037	0.031	0.026	0.032	0.015	0.012

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 72: **Within-Twin-Pair FE Estimates of the Determinants of Noncognitive Skills (Urban Sample)**

	Feel lonely	Anxious	Easily distracted	Careless	Easily Frightened	Emotional instability
Early health shocks	0.099	0.160**	0.154*	0.125*	0.191***	0.163**
	[0.084]	[0.068]	[0.082]	[0.074]	[0.073]	[0.069]
Birth weight(kg): <2	0.022	-0.097***	-0.053	0.015	0.071*	0.029
	[0.046]	[0.037]	[0.058]	[0.037]	[0.043]	[0.034]
Birth weight(kg): 2-2.5	0.030	-0.039	-0.058	0.028	0.034	0.008
	[0.043]	[0.028]	[0.056]	[0.036]	[0.037]	[0.032]
Birth weight (kg): 2.5-3	0.013	-0.031	-0.066	0.024	0.019	0.009
	[0.038]	[0.025]	[0.051]	[0.032]	[0.033]	[0.029]
Gender (boy=1)	-0.007	0.062***	0.064**	0.031	-0.008	0.010
	[0.026]	[0.023]	[0.033]	[0.022]	[0.018]	[0.017]
# Observations	688	688	688	688	688	686
R-squared	0.009	0.041	0.024	0.018	0.038	0.039

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 73: Within-Twin-Pair FE Estimates of the Determinants of Noncognitive Skills (Male Sample)

	Feel lonely	Anxious	Easily distracted	Careless	Easily Frightened	Emotional instability
Early health shocks	0.040	0.093	0.036	0.035	0.117*	0.047
	[0.092]	[0.057]	[0.070]	[0.070]	[0.071]	[0.074]
Birth weight(kg): <2	-0.023	-0.067	0.059	0.050**	-0.047	-0.020
	[0.057]	[0.054]	[0.060]	[0.025]	[0.050]	[0.020]
Birth weight(kg): 2-2.5	0.036	-0.033	0.031	0.057**	-0.022	-0.003
	[0.042]	[0.039]	[0.050]	[0.025]	[0.049]	[0.011]
Birth weight (kg): 2.5-3	0.012	-0.008	0.023	0.028	-0.051	0.005
	[0.037]	[0.030]	[0.041]	[0.024]	[0.043]	[0.003]
Gender (boy=1)						
# Observations	541	541	541	541	541	539
R-squared	0.009	0.013	0.003	0.012	0.023	0.007

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets: * significant at 10%: ** significant at 5%; *** significant at 1%.

Table 74: **Within-Twin-Pair FE Estimates of the Determinants of Noncognitive Skills (Female Sample)**

	Feel lonely	Anxious	Easily distracted	Careless	Easily Frightened	Emotional instability
Early health shocks	0.394***	0.228**	0.279***	0.219**	0.282***	0.238**
	[0.113]	[0.099]	[0.105]	[0.098]	[0.106]	[0.103]
Birth weight(kg): <2	0.049	-0.057	-0.107**	0.049	0.041	0.025
	[0.032]	[0.040]	[0.054]	[0.042]	[0.044]	[0.037]
Birth weight(kg): 2-2.5	0.034	-0.023	-0.080*	0.029	0.037	0.015
	[0.032]	[0.034]	[0.046]	[0.037]	[0.034]	[0.036]
Birth weight (kg): 2.5-3	0.045*	-0.017	-0.068*	0.024	0.043	0.026
	[0.024]	[0.033]	[0.035]	[0.032]	[0.029]	[0.031]
Gender (boy=1)						
# Observations	560	560	560	560	560	557
R-squared	0.144	0.036	0.031	0.041	0.053	0.090

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 75: **Within-Twin-Pair FE Estimates of the Determinants of Noncognitive Skills (Mixed-Gender Sample)**

	Feel lonely	Anxious	Easily distracted	Careless	Easily Frightened	Emotional instability
Early health shocks	0.070	0.049	0.189*	0.218*	0.006	0.073
	[0.076]	[0.076]	[0.104]	[0.118]	[0.017]	[0.070]
Birth weight(kg): <2	-0.030	0.016	0.058	-0.019	-0.043	0.067
	[0.076]	[0.079]	[0.085]	[0.049]	[0.092]	[0.041]
Birth weight(kg): 2-2.5	-0.034	0.053	-0.036	-0.004	0.014	0.015
	[0.044]	[0.063]	[0.068]	[0.033]	[0.055]	[0.025]
Birth weight (kg): 2.5-3	-0.031	0.026	-0.038	-0.003	-0.020	0.004
	[0.033]	[0.051]	[0.062]	[0.019]	[0.039]	[0.024]
Gender (boy=1)	-0.001	0.073***	0.076***	0.017	-0.005	0.012
	[0.015]	[0.016]	[0.021]	[0.013]	[0.016]	[0.008]
# Observations	360	360	360	360	360	359
R-squared	0.004	0.057	0.056	0.036	0.007	0.030

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 76: Within-Twin-Pair FE Estimates of the Determinants of the Parent-Child Relationship (Whole Sample)

	Parents perception			Childs perception				
	Expected education level	Expected college level	Parents-children relationship	Discuss an issue	Tell them school life	Tell them what you think	Play games together	Watch TV together
Early health shocks	-0.491***	-0.147***	-0.141**	0.152	-0.009	0.036	-0.087	-0.104
	[0.146]	[0.055]	[0.070]	[0.119]	[0.158]	[0.144]	[0.112]	[0.087]
Birth weight(kg): <2	-0.073	-0.052	-0.012	0.081	0.078	-0.032	0.058	0.056
	[0.087]	[0.033]	[0.023]	[0.086]	[0.099]	[0.095]	[0.091]	[0.084]
Birth weight(kg): 2-2.5	-0.042	-0.030	0.004	0.003	0.061	-0.067	0.037	0.128*
	[0.068]	[0.026]	[0.020]	[0.069]	[0.081]	[0.076]	[0.071]	[0.069]
Birth weight (kg): 2.5-3	-0.054	-0.037	-0.001	-0.085	0.002	-0.041	0.025	0.048
	[0.061]	[0.023]	[0.018]	[0.061]	[0.069]	[0.064]	[0.060]	[0.061]
Gender (boy=1)	-0.098**	-0.025	-0.031**	0.082*	0.154***	-0.002	-0.022	0.091**
	[0.042]	[0.016]	[0.014]	[0.047]	[0.049]	[0.051]	[0.043]	[0.044]
# Observations	1447	1447	1453	1449	1459	1451	1450	1455
R-squared	0.026	0.018	0.026	0.009	0.008	0.001	0.001	0.007

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 77: Within-Twin-Pair FE Estimates of the Determinants of the Parent-Child Relationship (Rural Sample)

	Parents perception			Childs perception				
	Expected education level	Expected college level	Parents-children relationship	Discuss an issue	Tell them school life	Tell them what you think	Play games together	Watch TV together
Early health shocks	-0.243	0.003	-0.036	0.051	-0.162	-0.006	-0.266	-0.144
	[0.159]	[0.060]	[0.038]	[0.108]	[0.156]	[0.197]	[0.199]	[0.107]
Birth weight(kg): <2	-0.035	-0.059	0.014	0.169	0.084	-0.036	0.035	0.086
	[0.132]	[0.049]	[0.029]	[0.111]	[0.147]	[0.136]	[0.119]	[0.108]
Birth weight(kg): 2-2.5	-0.064	-0.038	0.034	0.087	0.077	-0.140	-0.021	0.171**
	[0.094]	[0.035]	[0.023]	[0.082]	[0.110]	[0.105]	[0.089]	[0.087]
Birth weight (kg): 2.5-3	-0.100	-0.060**	0.020	-0.064	-0.066	-0.044	-0.019	0.086
	[0.083]	[0.029]	[0.023]	[0.066]	[0.095]	[0.085]	[0.071]	[0.074]
Gender (boy=1)	-0.129**	-0.043*	-0.041**	0.121**	0.077	0.008	-0.013	0.083
	[0.057]	[0.023]	[0.018]	[0.061]	[0.067]	[0.067]	[0.047]	[0.054]
# Observations	768	768	771	768	772	769	766	771
R-squared	0.015	0.013	0.020	0.016	0.010	0.004	0.004	0.010

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 78: Within-Twin-Pair FE Estimates of the Determinants of the Parent-Child Relationship (Urban Sample)

	Parents perception			Childs perception				
	Expected education level	Expected college level	Parents-children relationship	Discuss an issue	Tell them school life	Tell them what you think	Play games together	Watch TV together
Early health shocks	-0.700***	-0.268***	-0.236**	0.264	0.138	0.074	0.071	-0.080
	[0.226]	[0.082]	[0.119]	[0.197]	[0.257]	[0.210]	[0.119]	[0.131]
Birth weight(kg): <2	-0.095	-0.037	-0.046	-0.021	0.079	-0.021	0.100	0.006
	[0.110]	[0.043]	[0.037]	[0.141]	[0.131]	[0.131]	[0.143]	[0.134]
Birth weight(kg): 2-2.5	-0.001	-0.011	-0.034	-0.095	0.058	0.011	0.113	0.065
	[0.097]	[0.040]	[0.035]	[0.122]	[0.117]	[0.110]	[0.118]	[0.113]
Birth weight (kg): 2.5-3	0.005	-0.009	-0.033	-0.118	0.084	-0.032	0.091	-0.009
	[0.088]	[0.036]	[0.030]	[0.114]	[0.098]	[0.094]	[0.107]	[0.105]
Gender (boy=1)	-0.059	-0.003	-0.019	0.032	0.261***	-0.009	-0.031	0.102
	[0.061]	[0.022]	[0.023]	[0.075]	[0.070]	[0.081]	[0.080]	[0.073]
# Observations	679	679	682	681	687	682	684	684
R-squared	0.060	0.056	0.053	0.008	0.019	0.001	0.002	0.005

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 79: Within-Twin-Pair FE Estimates of the Determinants of the Parent-Child Relationship (Male Sample)

	Parents perception			Childs perception				
	Expected education level	Expected college level	Parents-children relationship	Discuss an issue	Tell them school life	Tell them what you think	Play games together	Watch TV together
Early health shocks	-0.573***	-0.254***	-0.164	0.023	-0.234	-0.239	-0.254*	-0.137
	[0.204]	[0.089]	[0.125]	[0.185]	[0.253]	[0.156]	[0.148]	[0.125]
Birth weight(kg): <2	-0.082	-0.010	-0.010	0.049	0.109	0.000	0.077	0.054
	[0.140]	[0.055]	[0.031]	[0.158]	[0.152]	[0.145]	[0.147]	[0.147]
Birth weight(kg): 2-2.5	-0.093	-0.017	-0.019	-0.013	0.098	0.003	0.020	0.134
	[0.101]	[0.043]	[0.031]	[0.095]	[0.119]	[0.124]	[0.094]	[0.110]
Birth weight (kg): 2.5-3	-0.049	-0.024	-0.012	-0.047	0.012	-0.099	0.032	0.064
	[0.091]	[0.039]	[0.030]	[0.076]	[0.098]	[0.116]	[0.069]	[0.094]
Gender (boy=1)								
# Observations	537	537	538	537	541	538	538	541
R-squared	0.038	0.045	0.034	0.001	0.004	0.006	0.004	0.004

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 80: **Within-Twin-Pair FE Estimates of the Determinants of the Parent-Child Relationship (Female Sample)**

	Parents perception			Childs perception				
	Expected education level	Expected college level	Parents-children relationship	Discuss an issue	Tell them school life	Tell them what you think	Play games together	Watch TV together
Early health shocks	-0.607**	-0.119	-0.171	0.519**	0.217	0.124	-0.010	-0.039
	[0.254]	[0.079]	[0.120]	[0.207]	[0.302]	[0.311]	[0.252]	[0.191]
Birth weight(kg): <2	0.023	-0.025	0.003	0.103	0.030	0.012	0.072	0.067
	[0.150]	[0.047]	[0.017]	[0.154]	[0.193]	[0.162]	[0.156]	[0.155]
Birth weight(kg): 2-2.5	0.085	0.009	0.013	0.023	0.025	-0.060	0.045	0.201
	[0.113]	[0.033]	[0.010]	[0.141]	[0.161]	[0.130]	[0.137]	[0.139]
Birth weight (kg): 2.5-3	0.001	0.000	-0.003	0.003	-0.019	-0.003	0.011	0.158
	[0.102]	[0.029]	[0.006]	[0.130]	[0.143]	[0.109]	[0.118]	[0.133]
Gender (boy=1)								
# Observations	553	553	557	553	558	555	556	558
R-squared	0.033	0.014	0.036	0.015	0.003	0.002	0.001	0.009

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 81: Within-Twin-Pair FE Estimates of the Determinants of the Parent-Child Relationship (Mixed-Gender Sample)

	Parents perception			Childs perception				
	Expected education level	Expected college level	Parents-children relationship	Discuss an issue	Tell them school life	Tell them what you think	Play games together	Watch TV together
Early health shocks	-0.185	0.011	-0.047	-0.005	0.104	0.367	0.124	-0.091
	[0.326]	[0.103]	[0.072]	[0.214]	[0.204]	[0.307]	[0.148]	[0.124]
Birth weight(kg): <2	-0.126	-0.108*	-0.038	0.116	0.110	-0.107	0.015	0.129
	[0.143]	[0.058]	[0.067]	[0.150]	[0.172]	[0.185]	[0.176]	[0.145]
Birth weight(kg): 2-2.5	-0.094	-0.070	0.033	0.031	0.075	-0.124	0.058	0.071
	[0.137]	[0.055]	[0.051]	[0.127]	[0.151]	[0.152]	[0.142]	[0.116]
Birth weight (kg): 2.5-3	-0.094	-0.078*	0.021	-0.204*	0.015	-0.015	0.044	-0.075
	[0.119]	[0.047]	[0.043]	[0.107]	[0.122]	[0.104]	[0.122]	[0.089]
Gender (boy=1)	-0.106**	-0.030*	-0.033**	0.084*	0.154***	-0.010	-0.026	0.089*
	[0.042]	[0.017]	[0.015]	[0.047]	[0.049]	[0.052]	[0.043]	[0.045]
# Observations	357	357	358	359	360	358	356	356
R-squared	0.020	0.018	0.025	0.036	0.028	0.009	0.003	0.022

Source: CCTS. Notes: Each column comes from a separate regression. Robust standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

9 Summary Statistics of Parent-Child Relationship

Table 82: Summary Statistics of Parental-Child Relationship

	Whole	Rural	Urban	Male	Female
Parent-child relationship					
Reported by both parents					
Expected highest education level of the child (1: middle school: 7: doctor)	3.47	2.84	4.19	3.46	3.59
Expected highest education level higher than college (dummy)	0.54	0.38	0.73	0.53	0.58
Relationship between you and the child (1: bad; 2: average; 3: good; 4: excellent)	3.08	3.02	3.13	3.06	3.12
Reported by the child					
Measured as 1: never; 2: seldom; 3: sometimes; 4: always; 5: every day					
Discuss an issue with your parents	3.37	3.60	3.11	3.39	3.28
Tell your parents your school life	2.67	2.77	2.56	2.77	2.52
Tell your parents what you think	3.25	3.31	3.19	3.32	3.11
Play game with your parents	3.68	3.95	3.38	3.71	3.54
Watch TV with your parents	2.60	2.69	2.49	2.58	2.56

References

BEHRMAN, J., R. POLLAK, AND P. TAUBMAN (1982): "Parental preferences and provision for progeny," *The Journal of Political Economy*, 90(1), 52–73.

FAMILY PLANNING COMMISSION OF YUNNAN PROVINCE (2003): "Regulations on Population and Family Planning in Yunnan Province," *China Population Press*.

HECKMAN, J. J. (2007): "The Economics, Technology and Neuroscience of Human Capability Formation," *Proceedings of the National Academy of Sciences*, 104(3), 13250–13255.

Appendix for Essay Two

Education and Preferences:
Experimental Evidences from Chinese Adult Twins

Appendix I: Experiment Instruction

Game One (Moderate Prospect)

We will randomly draw one card from a deck of 20 cards: 10 red and 10 black. You have two options.

- Guess the color of the card drawn. You will receive RMB40 if your guess is correct; and nothing if your guess is wrong.
- Receive a sure amount of money if you do not wish to guess.

Tick "√" your choice. Tick only one. You will be paid based on your decision.

1.) Guess: Red ___ Black ___
You will receive RMB40 if your guess is correct; and nothing if your guess is wrong.

2.) Receive RMB20 ___

Note: This is on risk attitude toward moderate prospects. If subjects choose 1), they are classified as risk tolerant, otherwise risk averse.

Game Two (Moderate Hazard)

We will randomly draw one card from a deck of 20 cards: 10 red and 10 black. You have two options.

- Guess the color of the card drawn. You will lose RMB10 if your guess is wrong; and nothing if your guess is correct.
- Lose a sure amount of money if you do not wish to guess.

Tick "√" your choice. Tick only one. You will be paid based on your decision.

1.) Guess: Red ___ Black ___

You will lose RMB10 if your guess is wrong; and nothing if your guess is correct.

2.) Lose RMB5 ___

Note: This is on risk attitude toward moderate hazards. If subjects choose 1), they are classified as risk tolerant, otherwise risk averse.

Game Three (Longshot Prospect)

You have the following three options:	
"Pick 7 out of 36"	The market price for this lottery ticket is RMB 2. The maximal prize is 5 million
"Permutation 5"	The market price for this lottery ticket is RMB2. The maximal prize is 0.1 million
"RMB 2 for sure"	

Tick "√" your choice. You will be paid based on your decision.

1.) ____ "One in 100"
2.) ____ "One in 10"
3.) ____ "RMB2 for sure"

Other than your first choice, tick "√" your choice from the two remaining. You will not be paid in this decision.

1.) ____ "One in 100"
2.) ____ "One in 10"
3.) ____ "RMB2 for sure"

Note: This is on risk attitude toward longshot prospects. We used real lottery tickets in the experiment. Subjects are classified as risk tolerant in the sense of exhibiting longshot preference, when 1 is preferred to 2, which is in turn preferred to 3.

Game Four (Longshot Hazard)

You have the folloing two options:

- Lose RMB2 for sure.
- Draw one card randomly from a deck of 10 cards numbered from 1 to 10 for three times. If you get the card with number 1 each time, you lose RMB2000 and nothing if you do not get the card with number 1 each time.

Tick "√" your choice. Tick only one. This is a hypothetical choice.

1.) Lose RMB2 for sure

2.) Draw the cards

Note: This is on risk attitude toward longshot hazards. If subjects choose 1), they are classified as risk averse, otherwise risk tolerant.

Game Five (Allais-Type Behavior)

For either alternative, a card will be drawn at random from a deck of 10 cards numbered from 1 to 10. These are all hypothetical choice. Please choose the one you like.

1. Which would you prefer? Tick "√" your choice

A. Receiving RMB100 if #1 to #8 is drawn. Receiving 0 if #9 or #10 is drawn.

B. Receiving RMB100 if #1 to #9 is drawn. Paying RMB80 if #10 is drawn.

2. Which would you prefer? Tick "√" your choice

A. Receiving 0 if #1 or #2 is drawn. Paying RMB80 if #3 to #10 is drawn.

B. Receiving RMB100 if #1 is drawn. Paying RMB80 if #2 to #10 is drawn

Note: This is on Allais-type behavior. If subjects choose AA, or BB, we classify these subjects as expected utility type behavior. If subjects choose AB, we classify these subjects as Allais-type behavior.

Game Six (Ambiguity Aversion)

You have the following two options:

T1: Guess the color of a card we draw randomly from a deck of 20 cards – 10 red and 10 black. You will receive RMB10 if your guess is correct; and nothing if your guess is wrong.

T2: Guess the color of a card we draw randomly from a deck of 20 cards with unknown proportions red and black cards. You will receive RMB12 if your guess is correct; and nothing if your guess is wrong.

Tick "√" your choice. Tick only one. You will be paid based on your decision.

Participate in T1	Participate in T2
Red ___ Black ___	Red ___ Black ___

Note: This is on ambiguity aversion. If subjects bet on T1, they are ambiguity averse.

Game Seven (Familiarity Bias)

You have the following two options:

T1: Guess whether the high temperature recorded in Beijing on ____ 2008 was odd or even. You will receive RMB11 if your guess is correct; and nothing if your guess is wrong.

T2: Guess whether the high temperature recorded in Tokyo on ____ 2008 was odd or even. You will receive RMB13 if your guess is correct; and nothing if your guess is wrong.

Tick "√" your choice. Tick only one. You will be paid based on your decision.

Participate in T1	Participate in T2
Odd ___ Even ___	Odd ___ Even ___

Note: This is on familiarity bias. If subjects bet on T1, they are familiarity biased. (Beijing is the capital of China, and Tokyo is the capital of Japan. We assume that our Beijing subjects are more familiar with Beijing than Tokyo)

Game Eight (Impatience and Hyperbolic Discounting)

The questions here are all hypothetical. Your decision will not have any real financial consequence. Please answer the following questions supposed you need to make decisions facing such situations.

1. Suppose that you can get RMB100 tomorrow, or you can get RMB 120 eight days later. Which one do you prefer:
   (1) Get RMB100 tomorrow;
   (2) Get RMB120 eight days later.

2. Suppose that you can get RMB100 91 days later, or you can get RMB120 98 days later. Which one do you prefer:
   (1) Get RMB100 91 days later;
   (2) Get RMB120 98 days later.

Note. This is on time discounting. If subjects choose RMB100 today in the first case, they are impatient. If they prefer RMB100 today over getting RMB120 seven days, and prefer getting RMB120 98 days later over getting RMB100 91 days later, they exhibit hyperbolic discounting behavior.

Game Nine (Anticipation)

Suppose you will have dinner with your favorite star. You could choose having it today, or three days later. You would choose:

(1) Today;

(2) 3 days later.

Note. This is on anticipation. If subjects choose 3 days later, they have preference of anticipation.

Game Ten (Dread)

Suppose you will take a non-lethal 110 volt shock. You could choose taking it today, or three months later. You would choose:

(1) Today;

(2) 3 months later.

Note. This is on dread. If subjects choose today, they have preference of dread.

Game Eleven (Hopefulness)

Suppose your relative has pregnant for three months. In a regular body check, the sex of the baby could be detected. You would prefer to:

(1) know it immediately;

(2) delay until it is born.

Note. This is on hopefulness. If subjects choose to delay, they have preference of hopefulness.

Game Twelve (Anxiousness)

Suppose you have 90% chance of winning RMB1000. The uncertainty is supposed to resolve today. Something happens, so it is delayed until 2 weeks later. If you pay RMB2, you could resolve the uncertainty immediately. Would you pay RMB2?

(1) Yes;

(2) No.

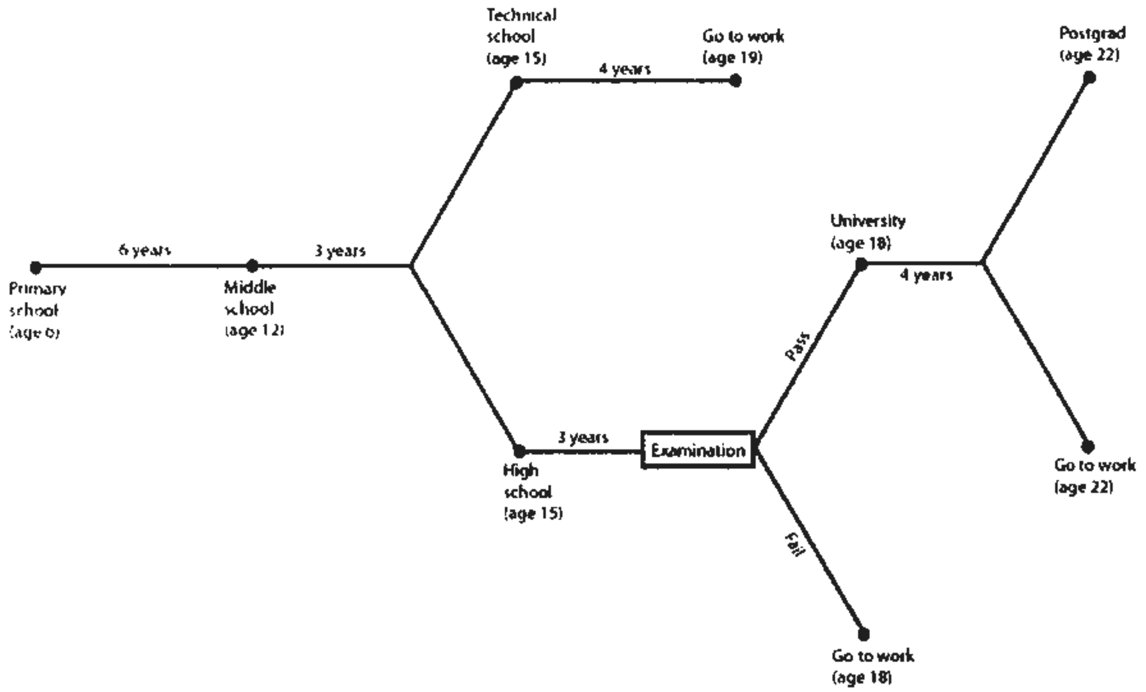
Note. This is on anxiousness. If subjects choose to pay, they have preference of anxiousness.

Appendix II: Experimental Variable Construction

Preference	Instruction	meaning	binary coding 0 and 1
Decision Making under Risk and Uncertainty			
Moderate Prospect	GAME ONE	1 risk tolerant; 2 risk averse	0 risk averse; 1 risk tolerant
Moderate Hazard	GAME TWO	1 risk tolerant; 2 risk averse	0 risk averse; 1 risk tolerant
Longshot Prospect	GAME THREE	First choice 1 and second choice 2, 1; otherwise 0	0 risk averse; 1 risk tolerant
Longshot Hazard	GAME FOUR	1 buy insurance; 2 risk tolerant	0 risk averse; 1 risk tolerant
Allais-type Behavior	GAME FIVE	0: AA or BB, expected utility; 1 otherwise.	0 expected utility behavior; 1 otherwise
Ambiguity Aversion	GAME SIX	1 ambiguity averse; 2 not ambiguity averse	0 not ambiguity averse; 1 ambiguity averse
Familiarity Bias	GAME SEVEN	1familiarty bias; 2 not familiarity bias	0 not familiarity bias; 1familiarty bias
Decision Making Involving Time			
Impatience	GAME EIGHT	0 if choosing (1) in Question 1, patient; =1 otherwise, impatient	0 patient; 1 impatient
Hyperbolic Discounting	GAME EIGHT	0 others, 1 if choosing (1) in Question 1, and 2 in Question 2	0 others; 1 hyperbolic
Anticipation	GAME NINE	1 no anticipation; 2 anticipation	0 no anticipation; 1 anticipation
Dread	GAME TEN	1 dread; 2 no dread	0 no dread; 1 dread
Hopefulness	GAME ELEVEN	1 no hope; 2 hope	0 no hope; 1 hopefulness
Anxiousness	GAME TWELVE	1 anxiety; 2 not anxiety	0 no anxiousness; 1 anxiousness

Note: Appendix I gives the experiment instruction.

Appendix III: The Education System in China



Note: (a) the entrance age for primary school is not fixed at 6. Most of children enter at primary school at age 5-7. (b) Before 1986, the primary education was 5 years in most areas; after 1986, the country began to promote a 6-years primary education system. In the early 1990s, almost all areas adopted the 6-year primary education system. (c)There are some high school graduates enter into technical school. In our sample, 36% of the technical school graduates have studied high school before entering technical school. The remaining 64% technical school graduates only studied middle school before entering technical school.

Appendix IV: OLS and Within-Twin-Pair Fixed-Effects Estimates of Education and Decision Making under Risk and Uncertainty

Table A1: OLS and Within-Twin-Pair Fixed-Effects Estimates of Education and Decision Making under Risk and Uncertainty

	Dependent variables					
	Moderate prospect			Moderate hazard		
	OLS		FE	OLS		FE
	(1)	(2)	(3)	(4)	(5)	(6)
High school	-0.087	-0.099	0.081	-0.130	-0.073	-0.057
	(0.80)	(0.89)	(0.42)	(1.28)	(0.70)	(0.29)
Technical school	0.002	0.074	0.199	-0.090	0.042	0.183
	(0.02)	(0.53)	(0.87)	(0.74)	(0.32)	(0.79)
College and above	0.004	0.060	0.438**	-0.060	0.016	0.386*
	(0.034)	(0.43)	(2.03)	(0.54)	(0.12)	(1.77)
Age	0.033	0.055		0.064**	0.072*	
	(1.00)	(1.41)		(2.05)	(1.97)	
Age-squared	-0.041	-0.066*		-0.081**	-0.091**	
(1/100)	(1.18)	(1.65)		(2.48)	(2.41)	
Male	0.0330	-0.029		0.004	-0.017	
	(0.39)	(0.33)		(0.06)	(0.21)	
Father middle school		0.012			-0.193*	
		(0.10)			(1.68)	
Father high school		0.159			0.098	
		(0.79)			(0.51)	
Father technical school		0.143			-0.282	
		(0.60)			(1.26)	
Father college or above		0.302*			-0.029	
		(1.83)			(0.19)	
Mother middle school		-0.194			0.007	
		(1.41)			(0.06)	
Mother high school		-0.377**			-0.294*	
		(2.12)			(1.75)	
Mother technical school		-0.520***			-0.089	
		(2.65)			(0.48)	
Mother college or above		-0.201			0.046	
		(0.83)			(0.20)	
Observations	140	140	140	140	140	140
Twin pairs			70			70
R-squared	0.11	0.19	0.08	0.11	0.17	0.09

Note: Absolute values of t-statistics computed based on the permutation procedure (Freedman and Lane, 1983) are in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; the omitted educational group for the individual is middle school and below, and the omitted educational group for parental education is primary school; a city dummy is included in each OLS estimation.

Table A1: OLS and Within-Twin-Pair Fixed-Effects Estimates of Education and Decision Making under Risk and Uncertainty (Cont.)

	Dependent variables					
	Longshot prospect			Longshot hazard		
	OLS		FE	OLS		FE
	(7)	(8)	(9)	(10)	(11)	(12)
High school	0.094	0.129	0.202	-0.187	-0.167	-0.331
	(0.85)	(1.13)	(0.96)	(1.60)	(1.43)	(1.55)
Technical school	0.180	0.234	0.338	0.0186	0.081	-0.165
	(1.36)	(1.62)	(1.36)	(0.14)	(0.57)	(0.67)
College and above	0.061	0.158	0.482**	-0.085	-0.064	-0.061
	(0.50)	(1.10)	(2.06)	(0.69)	(0.44)	(0.26)
Age	0.031	0.013		0.041	0.044	
	(0.90)	(0.33)		(1.17)	(1.07)	
Age-squared	-0.048	-0.034		-0.057	-0.064	
(1/100)	(1.35)	(0.82)		(1.55)	(1.52)	
Male	-0.029	-0.008		-0.042	-0.076	
	(0.34)	(0.095)		(0.46)	(0.81)	
Father middle school		-0.193			-0.150	
		(1.56)			(1.16)	
Father high school		-0.073			-0.229	
		(0.35)			(0.97)	
Father technical school		-0.446*			0.244	
		(1.83)			(1.02)	
Father college or above		-0.340**			0.321*	
		(2.01)			(1.86)	
Mother middle school		-0.059			-0.137	
		(0.42)			(0.92)	
Mother high school		0.051			0.004	
		(0.28)			(0.024)	
Mother technical school		0.049			-0.225	
		(0.24)			(1.13)	
Mother college or above		0.296			-0.406*	
		(1.19)			(1.67)	
Observations	140	140	140	128	128	128
Twin pairs			70			64
R-squared	0.14	0.21	0.03	0.14	0.25	0.05

Note: Absolute values of t-statistics computed based on the permutation procedure (Freedman and Lane, 1983) are in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; the omitted educational group for the individual is middle school and below, and the omitted educational group for parental education is primary school; a city dummy is included in each OLS estimation.

Table A1: OLS and Within-Twin-Pair Fixed-Effects Estimates of Education and Decision Making under Decision Making under Risk and Uncertainty (Cont.)

	Dependent variables					
	Allais-type behavior			Ambiguity aversion		
	OLS		FE	OLS		FE
	(13)	(14)	(15)	(16)	(17)	(18)
High school	0.092	0.094	0.295	0.176	0.077	0.372*
	(0.97)	(0.96)	(1.64)	(1.56)	(0.67)	(1.72)
Technical school	0.064	0.056	0.274	0.369***	0.240*	0.242
	(0.59)	(0.47)	(1.33)	(2.73)	(1.66)	(0.95)
College and above	0.177*	0.206*	0.417**	0.213*	0.230	0.175
	(1.77)	(1.73)	(2.15)	(1.73)	(1.60)	(0.73)
Age	-0.060**	-0.098***		0.020	0.020	
	(2.11)	(2.84)		(0.59)	(0.50)	
Age-squared	0.070**	0.107***		-0.017	-0.018	
(1/100)	(2.32)	(3.00)		(0.47)	(0.44)	
Male	-0.128*	-0.134*		-0.075	-0.076	
	(1.73)	(1.71)		(0.85)	(0.85)	
Father middle school		0.107			0.303**	
		(1.03)			(2.44)	
Father high school		-0.246			-0.033	
		(1.28)			(0.16)	
Father technical school		-0.034			0.112	
		(0.18)			(0.46)	
Father college or above		0.037			0.055	
		(0.26)			(0.32)	
Mother middle school		-0.125			-0.191	
		(1.05)			(1.36)	
Mother high school		-0.108			0.184	
		(0.75)			(1.01)	
Mother technical school		0.145			-0.353*	
		(0.90)			(1.75)	
Mother college or above		-0.098			-0.083	
		(0.50)			(0.33)	
Observations	122	122	122	140	140	140
Twin pairs			61			70
R-squared	0.15	0.21	0.08	0.07	0.17	0.05

Note: Absolute values of t-statistics computed based on the permutation procedure (Freedman and Lane, 1983) are in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; the omitted educational group for the individual is middle school and below, and the omitted educational group for parental education is primary school; a city dummy is included in each OLS estimation.

Table A1: OLS and Within-Twin-Pair Fixed-Effects Estimates of Education and Decision Making under Risk and Uncertainty (Cont.)

	Dependent variables Familiarity bias		
	OLS		FE
	(19)	(20)	(21)
High school	0.080	0.092	0.319*
	(0.85)	(0.98)	(1.65)
Technical school	-0.160	-0.171	0.065
	(1.43)	(1.43)	(0.28)
College and above	-0.125	-0.093	0.007
	(1.22)	(0.78)	(0.032)
Age	0.001	-0.004	
	(0.032)	(0.12)	
Age-squared	0.009	0.0161	
(1/100)	(0.31)	(0.48)	
Male	0.0176	0.0845	
	(0.24)	(1.14)	
Father middle school		-0.081	
		(0.79)	
Father high school		0.047	
		(0.28)	
Father technical school		-0.139	
		(0.69)	
Father college or above		-0.380***	
		(2.73)	
Mother middle school		0.050	
		(0.43)	
Mother high school		0.264*	
		(1.76)	
Mother technical school		0.287*	
		(1.73)	
Mother college or above		0.115	
		(0.56)	
Observations	140	140	140
Twin pairs			70
R-squared	0.10	0.21	0.06

Note: Absolute values of t-statistics computed based on the permutation procedure (Freedman and Lane, 1983) are in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; the omitted educational group for the individual is middle school and below, and the omitted educational group for parental education is primary school; a city dummy is included in each OLS estimation.

Appendix V: OLS and Within-Twin-Pair Fixed-Effects Estimates of Education and Decision Making Involving Time

Table A2: OLS and Within-Twin-Pair Fixed-Effects Estimates of Education and Preferences Involving Time

	Dependent variables					
	Impatience			Hyperbolic discounting		
	OLS		FE	OLS		FE
	(1)	(2)	(3)	(4)	(5)	(6)
High school	0.047	0.017	-0.004	-0.064	-0.055	-0.133
	(0.42)	(0.15)	(0.020)	(0.66)	(0.53)	(0.66)
Technical school	-0.133	-0.177	-0.413	0.010	-0.004	-0.057
	(1.00)	(1.19)	(1.56)	(0.088)	(0.032)	(0.24)
College and above	-0.211*	-0.214	-0.520**	-0.116	-0.188	-0.382*
	(1.73)	(1.44)	(2.09)	(1.09)	(1.45)	(1.70)
Age	-0.011	-0.011		-0.002	0.001	
	(0.33)	(0.27)		(0.08)	(0.03)	
Age-squared	0.015	0.015		-0.007	-0.009	
(1/100)	(0.42)	(0.35)		(0.23)	(0.24)	
Male	-0.256***	-0.248***		-0.178**	-0.175**	
	(2.93)	(2.69)		(2.33)	(2.17)	
Father middle school		0.059			0.024	
		(0.46)			(0.22)	
Father high school		-0.148			0.045	
		(0.69)			(0.24)	
Father technical school		0.075			0.072	
		(0.30)			(0.33)	
Father college or above		0.044			0.099	
		(0.25)			(0.65)	
Mother middle school		0.003			0.050	
		(0.020)			(0.39)	
Mother high school		0.210			0.073	
		(1.11)			(0.44)	
Mother technical school		-0.028			0.187	
		(0.14)			(1.03)	
Mother college or above		-0.221			-0.160	
		(0.86)			(0.72)	
Observations	140	140	140	140	140	140
Twin pairs			70			70
R-squared	0.13	0.16	0.10	0.07	0.10	0.07

Note: Absolute values of *t*-statistics computed based on the permutation procedure (Freedman and Lane, 1983) are in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; the omitted educational group for the individual is middle school and below, and the omitted educational group for parental education is primary school; a city dummy is included in each OLS estimation.

Table A2: OLS and Within-Twin-Pair Fixed-Effects Estimates of Education and Preferences Involving Time (Cont.)

	Dependent variables					
	Anticipation			Dread		
	OLS		FE	OLS		FE
	(7)	(8)	(9)	(10)	(11)	(12)
High school	0.068	0.058	0.048	-0.301***	-0.284**	-0.220
	(0.61)	(0.51)	(0.23)	(2.84)	(2.55)	(1.14)
Technical school	0.127	0.058	0.136	-0.215*	-0.233	-0.515**
	(0.98)	(0.41)	(0.56)	(1.70)	(1.65)	(2.22)
College and above	0.290**	0.152	0.208	-0.132	-0.214	-0.389*
	(2.49)	(1.09)	(0.92)	(1.17)	(1.56)	(1.81)
Age	-0.046	-0.028		-0.008	-0.008	
	(1.34)	(0.70)		(0.25)	(0.21)	
Age-squared	0.042	0.021		0.002	-0.000	
(1/100)	(1.17)	(0.50)		(0.053)	(0.0024)	
Male	0.155*	0.111		0.017	-0.015	
	(1.73)	(1.22)		(0.20)	(0.18)	
Father middle school		-0.053			0.022	
		(0.42)			(0.19)	
Father high school		0.010			0.029	
		(0.051)			(0.14)	
Father technical school		-0.169			-0.345	
		(0.73)			(1.49)	
Father college or above		0.252			0.135	
		(1.43)			(0.82)	
Mother middle school		0.014			0.018	
		(0.11)			(0.14)	
Mother high school		0.350**			0.054	
		(2.03)			(0.31)	
Mother technical school		-0.139			0.057	
		(0.72)			(0.30)	
Mother college or above		-0.009			0.068	
		(0.037)			(0.29)	
Observations	130	130	130	136	136	136
Twin pairs			65			68
R-squared	0.23	0.31	0.02	0.08	0.11	0.07

Note: Absolute values of *t*-statistics computed based on the permutation procedure (Freedman and Lane, 1983) are in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; the omitted educational group for the individual is middle school and below, and the omitted educational group for parental education is primary school; a city dummy is included in each OLS estimation.

Table A2: OLS and Within-Twin-Pair Fixed-Effects Estimates of Education and Preferences Involving Time (Cont.)

	Dependent variables					
	Hopefulness			Anxiousness		
	OLS		FE	OLS		FE
	(13)	(14)	(15)	(16)	(17)	(18)
High school	-0.201*	-0.186	-0.225	-0.062	-0.069	-0.130
	(1.71)	(1.51)	(1.08)	(0.64)	(0.67)	(0.60)
Technical school	-0.367***	-0.336**	-0.440*	0.011	-0.013	0.036
	(2.64)	(2.15)	(1.80)	(0.09)	(0.10)	(0.14)
College and above	-0.069	-0.145	-0.478**	0.079	0.044	-0.085
	(0.56)	(0.99)	(2.14)	(0.74)	(0.34)	(0.35)
Age	-0.003	0.014		-0.010	-0.017	
	(0.08)	(0.32)		(0.34)	(0.48)	
Age-squared	0.000	-0.018		0.021	0.028	
(1/100)	(0.00)	(0.39)		(0.65)	(0.75)	
Male	0.252***	0.191*		-0.021	-0.039	
	(2.72)	(1.93)		(0.28)	(0.48)	
Father middle school		0.012			0.141	
		(0.089)			(1.26)	
Father high school		0.086			0.112	
		(0.40)			(0.60)	
Father technical school		0.113			-0.144	
		(0.46)			(0.66)	
Father college or above		0.337*			0.054	
		(1.89)			(0.35)	
Mother middle school		0.047			-0.030	
		(0.33)			(0.24)	
Mother high school		-0.179			-0.169	
		(0.97)			(1.03)	
Mother technical school		-0.052			0.012	
		(0.25)			(0.064)	
Mother college or above		-0.153			0.172	
		(0.61)			(0.77)	
Observations	128	128	128	140	140	140
Twin pairs			64			70
R-squared	0.14	0.19	0.07	0.11	0.14	0.01

Note: Absolute values of t-statistics computed based on the permutation procedure (Freedman and Lane, 1983) are in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; the omitted educational group for the individual is middle school and below, and the omitted educational group for parental education is primary school; a city dummy is included in each OLS estimation.

Appendix VI: Between-Families and Within-Twin-Pair Correlations of Education and Other Variables

Table A3: Between-Families and Within-Twin-Pair Correlations of Education and Other Variables

Between-family correlations		Within-twin-pair correlations	
	Education		ΔEducation
Married	-0.1445*** (<0.01)	ΔMarried	-0.0173 (0.70)
Spousal education	0.6172*** (<0.01)	ΔSpousal education	0.1518** (0.02)
Party member	0.2571*** (<0.01)	ΔParty member	0.1166** (0.02)
Working in foreign firm dummy	0.0904* (0.06)	ΔWorking in foreign firm dummy	0.0214 (0.66)
Tenure	-0.2614*** (<0.01)	ΔTenure	-0.1253*** (0.01)

Notes: The significance levels are in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. The between-family correlations are the correlations between average family education (average of the twins) and average family characteristics, and the within-twin-pair correlations are the correlations between the within-twin-pair differences in education and the within-twin-pair differences in other characteristics.

Source: Li, H.; Liu, P. and Zhang, J. "Estimating Returns to Education Using Twins in Urban China." 2010, *Journal of Development Economics*, forthcoming.