J Risk Uncertain (2016) 53:163–200
DOI 10.1007/s11166-016-9246-7

Education and anomalies in decision making: Experimental evidence from Chinese adult twins

Soo Hong Chew[1] · Junjian Yi[1] · Junsen Zhang[2] · Songfa Zhong[1]

Published online: 8 March 2017

Abstract We estimate the effects of education on two dimensions of decision making behavior—risk and time—beyond those considered to be normal-ranged to encompass behavioral anomalies with respect to expected utility as well as time consistency. We conduct a number of incentivized choice experiments on Chinese adult twins to measure decision making behavior, and use a within-twin-pair fixed-effects estimator to deal with unobservable family-specific effects. The estimation results show that a higher education level tends to reduce the degree of risk aversion towards moderate prospects, moderate hazards, and longshot prospects. For anomalies under risk and uncertainty, college graduates exhibit significantly more Allais-type behavior compared to high school dropouts, while high school graduates exhibit more ambiguity aversion as well as a familiarity preference relative to high school dropouts. For decision making involving time, a higher education level tends to reduce the degree of impatience, and to reduce behavioral anomalies including hyperbolic discounting, dread, and hopefulness. The experimental observations suggest that people with a higher education level tend to exhibit more behavioral anomalies in risk attitudes but fewer behavioral anomalies involving time, hence implying that education has multi-functions in preference formation and human capability building. This study contributes to the understanding of the nature of these behavioral anomalies and the roles of education in human decision making.

Keywords Risk preference · Time preference · Risk and time · Behavioral anomalies · Education · Field experiment · Twins

Electronic supplementary material The online version of this article (doi:10.1007/s11166-016-9246-7) contains supplementary material, which is available to authorized users.

✉ Songfa Zhong
zhongsongfa@gmail.com

[1] Department of Economics, National University of Singapore, Singapore, Singapore

[2] Department of Economics, Chinese University of Hong Kong, Sha Tin, Hong Kong

JEL Classifications C91 · D81 · D91

"*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.*" [Oskar Morgenstern 1979]

1 Introduction

At the heart of economic analysis is our understanding of decision making behavior ranging from decision making under risk and uncertainty to decision making involving time. The classical models of expected utility and exponential discounted utility have been challenged by a number of behavioral anomalies.[1] In decision making under risk and uncertainty, the Allais paradox (1953) and the Ellsberg paradox (1961) have been particularly well studied and have inspired 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 behavioral anomalies involving time, e.g., hyperbolic discounting (Laibson 1997), incidence of consumption (Loewenstein 1987), and the timing of uncertainty resolution (Kreps and Porteus 1978; Chew and Ho 1994).

In the literature on decision making, there are debates on whether behavioral anomalies represent the true preferences of the decision makers or reflect biases due to cognitive errors on the part of the decision makers.[2] The former view is reflected in the revealed preference approach that aims to deliver a more general preference framework to accommodate the supposed behavioral anomalies. The latter view is reflected in the opening quote by Oskar Morgenstern, which suggests that behavioral anomalies are more akin to errors that can be rectified via positive intervention such as education.

To arrive at a deeper understanding of the nature of behavioral anomalies, we investigate the impact of education on a range of decision making behavior under risk and uncertainty and involving time. Among the factors influencing decision making behavior, education appears especially important since we spend many years in formal schooling, and we learn and develop different ways of thinking and acting. Becker and Mulligan (1997) argue that schooling focuses students' attention on the future as schooling can communicate images of the situations and difficulties of adult

[1] In the preamble to a series of "Anomalies" in the *Journal of Economic Perspectives*, Richard Thaler wrote, "An empirical result qualifies as an anomaly if it is difficult to "rationalize", or if implausible assumptions are necessary to explain it within the paradigm." In this paper, "behavioral anomalies" refer to choice behaviors that are not compatible with the implications of classical preference theory.

[2] Savage (1954) argues 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"). Moreover, Charness et al. (2010) show that interaction with other subjects decreases the frequency of the conjunction fallacy, and suggest that deliberating the alternative actions with other subjects helps to improve the understanding of the decision problem and to reduce decision errors.

life. In addition, through repeated practice at problem solving, schooling helps children learn to simulate scenarios. Thus educated people should be more productive at reducing the remoteness of future pleasures. One of the key challenges in investigating how education may influence behavior is determining causality. Unobservable family background and individual heterogeneity may simultaneously determine educational outcomes and preference formation. In other words, education may be correlated with unobservable family background and the effects of an individual's endowment, which would render any correlation between education and behavior spurious.[3]

Our primary goal in this paper is to empirically identify the causal effect of education on behavioral anomalies in addition to those considered to be normal-ranged including risk attitudes across domains and time discounting factor. We first conduct a number of incentivized choice experiments on adult twins to observe their behavior. Then we use a within-twin-pair fixed-effects estimator to carry out the identification (Ashenfelter and Krueger 1994; Behrman et al. 1996; Behrman and Rosenzweig 1999; Oreopoulos and Salvanes 2011). Given that twins have similar family backgrounds, and monozygotic twins are genetically identical,[4] the effects of unobserved family background and genetic endowment should be similar for both twins. Taking the within-twin-pair difference will reduce the unobservable family background and individual endowment effects, which are the main causes of the bias in ordinary least squares estimation. Intuitively, by comparing experimentally observed choice behavior of twins with different educational attainment, we may be more confident that an observed correlation between education and behavior is not due to a spurious 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 two important dimensions of decision making behavior—choice under risk and uncertainty and involving time. We find that a higher education level tends to reduce the degree of risk aversion towards moderate prospects, moderate hazards, and longshot prospects. In terms of decision making anomalies, college graduates exhibit significantly more Allais–type behavior compared to high school dropouts, while high school graduates exhibit more ambiguity aversion as well as familiarity bias relative to high school dropouts. For decision making involving time, a higher education level tends to reduce the degree of impatience, and to reduce a number of behavioral anomalies including hyperbolic discounting, dread, and hopefulness. However, anxiousness does not appear to be sensitive to educational attainment.

The estimation results are robust with respect to sensitivity analyses (i) using the instrumental variables (IV) method to account for potential measurement errors; and (ii) controlling for birth weight and restricting the estimation sample to monozygotic (MZ) twins to account for possible biases arising from omitted variables. In summary, our experimental evidence from Chinese adult twins suggests that people with a higher

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

[4] Gorseline (1932) seems to be the first attempt to look at sibling data in economics. 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 estimators. They conclude that within-twin-pair fixed-effects estimators require much weaker identification assumptions than within-sibling fixed-effects estimators.

education level tend to exhibit more anomalous behavior in decision making under risk and less anomalous behavior in decision making involving time.

Our findings contribute to behavioral economics and labor economics in a number of aspects. First, our results shed light on whether the observed behavioral anomalies represent actual preferences or reflect cognitive biases. Intuitively, education would increase our cognitive skills and reduce cognition errors. Hence behavioral anomalies are more likely to represent actual preferences if education increases the tendency to exhibit such behavior. Otherwise, behavioral anomalies may reflect cognitive biases if education decreases the tendency to exhibit the observed choice behavior. Specifically, our results suggest that Allais and Ellsberg behavior are more likely to represent preferences rather than cognitive errors of the decision makers, while anomalies in temporal decision making are more likely to reflect biases rather than preferences of the decision makers. Furthermore, these findings are consistent with intuitions underpinning prescriptive tools, such as the so-called save more tomorrow savings program (Thaler and Benartzi 2004), which provide behavioral mechanisms to help individuals with time inconsistency problems make better decisions.

Second, this paper contributes to recent studies on the relationship between cognitive skills and experimentally observed behavior (e.g., Frederick 2005; Dohmen et al. 2010; Benjamin et al. 2013; Burks et al. 2009; Taylor 2013). Frederick (2005) finds that student subjects who scored higher cognitive skills measured by the cognitive reflection tasks tend to be less risk averse and more patient. Dohmen et al. (2010) conduct experiments measuring risk aversion, impatience, and cognitive ability with 1000 German adults, and find that high cognitive ability is associated with less risk aversion and more patience. Burks et al. (2009) use a sample of 1000 trainee truckers, and show that individuals with better cognitive skills are more patient in both the short run and long run, and are more willing to take calculated risks. However, these studies establish a correlation—rather than a causal relationship—between cognitive skills and behavior. Benjamin et al. (2013) similarly find that high school students with higher standardized test scores are less risk averse and more patient. In order to demonstrate the causal impact, they further provide laboratory interventions on cognitive resources and show that cognitive ability affects risk attitudes and temporal discounting. Here, we investigate the effect of education on decision making behavior on a wider range of behavior. Moreover, we adopt a twin design to address issues of causality.

Third, studying the relationship between education and decision making behavior sheds light on the literature on the formation of preference. Stable preferences, maximizing behavior, and market equilibrium have been regarded as fundamental assumptions underpinning the analytic framework of economic analysis. While we may reasonably assume that basic preferences do not change rapidly over a short period of time, we need to allow preferences to change gradually over an extended duration, e.g., it makes sense to treat preferences as being endogenous from a life-cycle perspective.[5] Previous studies have analyzed theoretically the processes of preference formation and preference change (e.g., Becker 1992, 1996; Becker and Mulligan 1997). Theories have been proposed about endogenous determination of preferences through

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

wealth (Becker and Mulligan 1997), market institutions (Bowles 1998), and culture (Bisin and Verdier 2000).

Fourth, this paper also contributes to the broad literature on human capital. Human capital theory, integrating new developments in neuroscience, psychology, and behavioral science, has expanded substantially in recent years (see, e.g., Rutter 2006; Heckman 2007; ter Weel 2008). Non-cognitive skills or personality traits are now accepted as an important dimension of human capital (Heckman and Rubinstein 2001; Borghans et al. 2008). If preferences are shaped by education, then there is merit to the argument that preferences should be treated as endogenous in human capital theory. Our results thus suggest a potentially fruitful direction in future research to explore the mechanisms underlying the interaction between human capital formation and economic behavior.

Finally, the estimation of the returns to education has been one of the major subjects in economics for decades (Card 1999). Yet, there have been limited studies to explore the mechanisms underlying the relationship between education and socioeconomic success. Oreopoulos and Salvanes (2011) investigate non-pecuniary returns of education including job satisfaction, happiness, and preferences. In particular, using within-twin estimates and conditioning on income, they find that siblings with more schooling are less likely to be divorced, less likely to give birth as teenagers, and more likely to have educated spouses. Our identified relationships between educational attainment and decision making under risk and uncertainty and involving time will help enhance our understanding of the non-pecuniary returns to education. Moreover, risk attitudes underpin a wide range of economic behavior such as portfolio choice and insurance purchase, while time preference has been linked to economic behavior such as saving and consumption as well as physical exercise and smoking. Should education affect behavior in these domains, education is likely to have long-term economic consequences for individuals. Our results support the notion that education has multiple functions in preference formation and human capability building.

2 Experimental design involving adult twins

We use two sources of data in this study. The first is derived from the Chinese Adult Twin Survey (CATS) that was conducted in 2002. The second is derived from the experiments that were conducted in 2008 on a subsample of the twins in CATS. This section describes CATS, the experiments, and the summary statistics.

2.1 The Chinese adult twins survey (CATS)

The socioeconomic variables in our analysis are derived from CATS,[6] which 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—Chengdu, Chongqin, Harbin, Hefei, and Wuhan. CATS is based on existing twin questionnaires from the United States and elsewhere, and covers a wide range of demographic, social, and economic information.

[6] Li et al. (2007), Huang et al. (2009), Li et al. (2010), Li et al. (2012), and Rosenzweig and Zhang (2013) provide a detailed description of the CATS data.

The questionnaire was designed by one of the authors of this paper in close consultation with Mark Rosenzweig and Chinese experts at the NBS. Adult twins between 18 and 65 years of age were identified by the local statistical bureaus. The questionnaires were completed through face-to-face personal interviews. Several site checks of the survey work were made, and the data input process was closely monitored.

CATS is the first socioeconomic twin survey in China and perhaps the first in Asia. We consider a pair of twins to be identical (monozygotic, MZ) or non-identical (dizygotic, DZ) based on their hair color, physical characteristics, and gender. There are a total of 1501 pairs (3002 individuals) in the dataset. 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.

2.2 Decision making experiment

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 CATS.[7] Due to budget constraints, we conducted the experiments in two cities: Hefei and Wuhan, the capitals of Anhui and Hubei provinces, respectively. We were not able to reach a substantial proportion of twins who have changed their addresses and contact information between 2002 and 2008. Furthermore, because participation in the experiments was voluntary, some individuals in CATS did not take part in the experiments. Eventually, we recruited 70 pairs (140 individuals) of twins for our experiments. As the identification of the causal effect is based on within-twin variations, the selection of the twins in the experiment is not a major problem.

In the experiment, each individual who took part in the experiment was paid a participation fee of RMB60.[8] In addition, there were various payoffs in each experiment. We incentivized the choices of the subjects on risk attitudes towards moderate prospect, moderate hazard, and longshot prospect, as well as ambiguity aversion and familiarity bias. We did not provide incentives for the rest including attitude towards longshot hazard, Allais paradox, impatience, hyperbolic discounting, dread, anticipation, anxiety, and hopefulness, due to the difficulty in implementation. Most individuals completed the experiment within one hour. The maximum time spent is close to one and a half hours. The money was paid in cash after participants had finished the experiment. On average, individuals earned RMB42 in addition to the participation fee. These earnings are higher than the average hourly wage of RMB9.5 in these two cities (National Bureau of Statistics (NBS) 2009). The experimental design and instructions are presented in Web Appendix I. The review of behavioral decision models and related evidence is detailed in Web Appendix II.

2.2.1 Attitudes towards fourfold risks

Tversky and Kahneman (1992) examine attitudes towards fourfold risks and observe that decision makers are risk averse towards moderate prospects and longshot hazards, and risk seeking towards moderate hazards and longshot prospects. In assessing risk

[7] We also hired several experiment assistants in conducting the experiments.
[8] The exchange rate is US \$1 ≈ RMB ¥6.8 in 2008.

attitudes towards moderate prospects (Game 1 in Appendix I), subjects choose between an even-chance lottery of receiving RMB40 and receiving zero, versus receiving the expected outcome of RMB20 with certainty. 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 attitudes towards moderate hazards (Game 2 in Appendix I), subjects begin by choosing between an even-chance lottery of losing RMB10 and losing zero, versus losing the expected outcome of RMB5 with certainty. Subjects were incentivized, i.e., losses were deducted from subjects' participation 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 3 in Appendix I), subjects order the value of three items: (A) an RMB2 lottery ticket with 1 in 58,433,760 chance of winning 5 million; (B) an RMB2 lottery ticket with 1 in 100,000 chance of winning 0.1 million; and (C) RMB2 for sure. We paid subjects their preferred 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 4 in Appendix I), subjects are classified as being disposed to insure if they prefer losing RMB2 with certainty over losing RMB2000 with 0.1% probability. For the Fourfold Risks, subjects were incentivized with their preferred choices for Games 1–3, but not for Game 4.

Table 1 defines the experimental measures of preferences. All variables are 0–1 dummy variables. Specifically, the four variables of moderate prospect, moderate hazard, longshot prospect, and longshot hazard are coded as 0 if the subject exhibits risk aversion in Games 1–4; otherwise, the variables are coded as 1.

2.2.2 Allais-type behavior

Allais (1953) suggests a choice situation in which subjects' choices tend to violate the independence axiom. Chew and Waller (1986) study binary choices designed to test the independence axiom's parallelism implication on the behavior of indifference curves in a probability triangle (Game 5 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. Following Chew and Waller (1986), we adopt two of the four pairs of binary choices—H (high) and L (low). In our design, the H pair involves subjects choosing between receiving RMB100 with 80% probability and receiving nothing with 20% probability (Option A) versus receiving RMB100 with 90% probability and losing RMB80 with 10% probability (Option B). The L pair involves subjects choosing between losing RMB80 with 80% probability and receiving nothing with 20% probability (Option A) versus losing RMB80 with 90% probability and receiving RMB100 with 10% probability (Option B). We classify subjects as the expected utility type if they either always choose A or always choose B in both the H and L pairs as implied by the independence axiom. Otherwise, we classify subjects as being the Allais type if they choose A in the H pair and choose B in the L pair. This choice pattern implies that their indifference curves fan out in the probability triangle, i.e., their preferences satisfy Machina's (1982) Hypothesis II. We classify subjects as being the reverse Allais type if they choose B in the H pair and choose A in the L pair.

Table 1 The construction and definition of experimentally measured preference variables.

Behavior	Instruction	Meaning	Variable Definition (binary coding 0 and 1)
Decision Making under Uncertainty			
Moderate Prospects	GAME ONE	1 risk tolerant; 2 risk averse	0 risk averse; 1 risk tolerant
Moderate Hazards	GAME TWO	1 risk tolerant; 2 risk averse	0 risk averse; 1 risk tolerant
Longshot Prospects	GAME THREE	First choice 1 and second choice 2, 1; 0 otherwise	0 risk averse; 1 risk tolerant
Longshot Hazards	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	1 familiarity bias; 2 not familiarity bias	0 not familiarity bias; 1 familiarity bias
Decision Making over 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 hope; 2 no hope	0 no hope; 1 hopefulness
Anxiousness	GAME TWELVE	1 anxiety; 2 not anxiety	0 no anxiousness; 1 anxiousness

Web Appendix I provides the experiment instructions

We code the variable Allais-type behavior as 0 if the subject exhibits expected utility behavior; otherwise, we code the variable as 1 (Table 1). The Allais-type behavior was included in the questionnaire without actual incentives.

2.2.3 Ambiguity aversion and familiarity bias

Most experimental studies on the original Ellsberg paradox (1961) involve choosing between betting on an urn with a known probability distribution and betting on an urn with an unknown probability distribution. Betting correctly in either case would pay the same. It is found that people tend to bet on the urn with the known probability distribution (see, e.g., Camerer and Weber 1992). In order to generate a more even split of individual differences 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, which we determine by running pretests. In the ambiguity aversion task (Game 6 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.

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 the temperature in Beijing on a specific historical day would be odd or even, and betting on the temperature in Tokyo in a similar manner (Game 7 in Appendix I). Our design induces the same objective probability (Machina 2004) of 0.5 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 the temperature in Beijing pays RMB11, which is RMB2 less than betting on the temperature in Tokyo. The variables of ambiguity aversion and familiarity bias are coded as 1 if the subject exhibits ambiguity aversion and familiarity bias; otherwise, the variables are coded as 0 (Table 1). Subjects' choices for the tasks of ambiguity aversion and familiarity bias were incentivized.

2.2.4 Impatience and hyperbolic discounting

Contrary to the prediction of exponential discounting, it is commonly observed that people tend to exhibit hyperbolic discounting, i.e., they are being more impatient in the immediate future, and more patient in the distant future (see, e.g., Frederick et al. 2002 for a review). We make use of a simple hypothetical choice task (Game 8 in Appendix I). In Situation 1, subjects choose between receiving RMB100 today (A) and receiving RMB120 seven days later (B). In Situation 2, subjects choose between receiving RMB100 91 days later (A) and receiving RMB120 98 days later (B). If subjects choose A in Situation 1, the variable of impatience is coded as 1; otherwise, the variable equals 0. If subjects choose B in Situation 2, the variable of hyperbolic discounting is coded as 1; otherwise, the variable equals 0 (Table 1). The time discounting task was included in the questionnaire without actual incentives.

2.2.5 Anticipation and dread

Another dimension of decision making over time concerns the timing of consumption, specifically, the idea of anticipation and dread. Loewenstein (1987) runs an experiment in which subjects indicated how much they would pay to obtain (avoid) outcomes that would occur either immediately or after one of several delays. We adopt a similar design to Loewenstein (1987) for timing-of-consumption preference with both desirable and aversive outcomes: having dinner with a movie star of one's choice, and receiving a non-lethal electric shock (Games 9 and 10 in Appendix I). Subjects are asked whether they prefer to have the dinner today or three days later. If they choose three days later, we classify them as experiencing anticipation, and the variable of anticipation is coded as 1. Subjects are asked whether they prefer to receive a non-lethal electric shock today or six months later. If they choose today, we classify them as experiencing dread, and the variable of dread is coded as 1 (Table 1). Subject's choices for these two tasks were not incentivized.

2.2.6 Hopefulness and anxiousness

The timing of uncertainty resolution is another dimension of decision making over time. Timing may matter for two reasons: planning advantage of early resolution (Kreps and Porteus 1978) and anticipatory feelings such as hopefulness in the case of late resolution (Chew and Ho 1994). We adopt a similar design to Chew and Ho (1994) and Lovallo and Kahneman (2000) for the timing of uncertainty resolution (Games 11 and 12 in Appendix I). In one task, subjects are presented with the hypothetical scenario that one of his/her relatives is pregnant. Subjects decide whether to pay RMB2 to delay (by 6 months) the resolution of uncertainty about the gender of the baby or to pay nothing and resolve the uncertainty immediately. We classify the subjects as experiencing hopefulness if they prefer to pay to delay the resolution of uncertainty. In this case, the variable of hopefulness is coded as 1; otherwise, the variable equals 0 (Table 1).

In another task, subjects are presented with the hypothetical scenario of facing a lottery of receiving RMB1000 with 90% probability and zero with 10% probability. Subjects decide whether to pay RMB2 to resolve the uncertainty immediately or to pay nothing and wait two weeks to resolve the uncertainty. If they choose to resolve the uncertainty now, we classify them as experiencing anxiousness, although there may be some value for planning. In this case, the variable of anxiousness is coded as 1; otherwise, the variable equals 0 (Table 1). Naturally, we did not incentivize subjects' choices for these two tasks.

2.3 Summary statistics of main variables

Summary statistics of the variables are reported in Tables 2a and 2b. Table 2a first reports the individual's education levels and other socioeconomic variables. Educational attainment is categorized into five levels. The first three concern general education while the next two involve professional education. We use education levels rather than years of education because years of education in high school and in technical school are not comparable. In China, the student faces two choices after graduating from middle school: technical school or high school. If the student enters technical school, she will undergo four years of vocational education and then enter the labor market. If the student enters high school, she will undergo three years of general education before taking the college entrance examination. If she passes the examination, she will go to college. Otherwise, she will seek employment.[9] Few technical school graduates take the college entrance examination. The qualitative difference between technical school and high school is confirmed by Li et al. (2010). Using the CATS data, they find that the returns to high school is between 4.0 and 4.5%. In contrast, the returns to technical school is between 20.6% and 22.5%. Finally, given that there were only a few subjects with only primary school education, we group them with subjects with middle school education, and treat those with primary school or middle school education as a single baseline group.

The adult twins in the sample are relatively old with a mean age of 46 and a minimum age of 28. Thus, it makes sense to use educational attainment in the CATS of 2002 because those twins who attend colleges would have graduated by 2002. In the OLS regression, we include parental education levels to check how family background affects individuals' preferences. All twin siblings are of the same gender, and 47% are males. By comparing the OLS estimates with and without parental education levels and within-twin-pair fixed-effects estimates, we are able to infer how family background and individual heterogeneity affect adult preferences.

We include birth weight to control for pre-birth differences between twins.[10] To detect the channels through which education affects preference, we include family annual income and health in the regression.[11] It has been suggested that educational attainment increases income, and that those with higher education levels have better health (Grossman 1975). We study whether education affects preferences through income or health.

Table 2b reports the summary statistics for the experimentally observed measures of preferences. Column (6) 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 variation, ranging from 34.50% to 54.39%, for all preference variables. The last column indicates whether the games conducted to elicit the preferences are monetarily incentivized. The existence of incentives neither affects the frequency nor the distribution of the longshot prospect and longshot hazard. The game used to measure the longshot prospect is monetarily incentivized, whereas the game used to measure the longshot hazard is not. Column (1) shows that 57% of the subjects show risk tolerance in these two games. We further compare the frequency of the tasks with incentives (including ambiguity aversion and familiarity bias) with the frequency of the tasks without incentives (including Allais behavior and hyperbolic discounting). We find that 38% of subjects are neutral to ambiguity, 21% are neutral to familiarity, 82% are consistent with expected utility, and 77% are time-consistent. A *t*-test shows that the differences in means between

[9] Appendix III describes the Chinese education system.

[10] Behrman and Rosenzweig (2004) find birth weight affects a series of short- and long-run individual outcomes, including health, academic performance, educational attainment, and earnings.

[11] In CATS, the self-reported health status is categorized into five levels: poor, fair, good, very good, and excellent. We categorize health into a dummy variable that equals 1 if the individual reported the health status as good, very good, or excellent, and 0 otherwise.

Table 2a Summary statistics of educational attainment and other variables.

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

The sample size is 140

incentivized and not incentivized experimental outcomes in decision anomalies are statistically significant. However, because these experiments are used to measure difference types of preferences, we are not sure whether the differences are due to monetary incentives or differences in preferences and their elicitation methods.

3 Empirical strategy

Our empirical analysis focuses on the estimation of the following equations:

$$y_{1i} = \alpha + E_{1i}\beta + X_i\gamma + Z_{1i}\delta + \mu_i + \varepsilon_{1i} \quad (1)$$

$$y_{2i} = \alpha + E_{2i}\beta + X_i\gamma + Z_{2i}\delta + \mu_i + \varepsilon_{2i} \quad (2)$$

where $y_{ji}(j = 1, 2)$ is the experimentally observed choice behavior of twin j in family i. $E_{ji}(j = 1, 2)$ is a vector containing dummies for education levels of twin j in family i; X_i is the set of family background variables that are observable and varying across families

Table 2b Summary statistics of experimental measures of decision making behavior.

Variables	# Obs. (1)	Mean (2)	S.D. (3)	Min (4)	Max (5)	Within-twin variation (6)	Incentivized (7)
Decision making under risk and uncertainty							
Moderate prospects	140	0.64	0.48	0	1	37.33%	Yes
Moderate hazards	140	0.73	0.45	0	1	44.42%	Yes
Longshot prospects	140	0.57	0.50	0	1	40.83%	Yes
Longshot hazards	128	0.57	0.50	0	1	36.66%	No
Allais-type behavior	122	0.18	0.38	0	1	44.37%	No
Ambiguity aversion	140	0.62	0.49	0	1	44.02%	Yes
Familiarity bias	140	0.79	0.41	0	1	50.91%	Yes
Decision making over time							
Impatience	140	0.54	0.50	0	1	47.39%	No
Hyperbolic discounting	140	0.23	0.42	0	1	52.65%	No
Anticipation	130	0.43	0.50	0	1	34.50%	No
Dread	136	0.73	0.45	0	1	42.69%	No
Hopefulness	128	0.41	0.49	0	1	35.62%	No
Anxiousness	140	0.76	0.43	0	1	54.39%	No

Web Appendices I and II provide experimental instructions and variable construction. Column (6) gives the percentage of the within-twin-pair variation to the total variation for each behavioral variable

but not across twins; Z_{ji} (j = 1, 2) is a set of observed variables that vary across twins; and μ_i represents a set of unobservable variables at the family level that may also affect preferences.

All experimentally measured preference variables y_{ji} are 0–1 dummy variables. However, we use a linear probability model and ordinary least squares (OLS) method to estimate Equation (1) (or Equation (2)) because the OLS estimator facilitates our within-twin-pair fixed-effects estimation discussed below and the interpretation of the estimated coefficients (Angrist and Pischke 2008). The OLS estimate β of the educational effect on preferences in the equations above is generally biased. The bias arises because we do not usually have a perfect measure of μ_i, which is likely to correlated with E_{ji} and y_{ji} simultaneously. Thus, we apply a within-twin-pair fixed-effects estimator for twins based on the first difference between (1) and (2):

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

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 having a bias caused by the omitted variable μ_i.

The within-twin-pair fixed effects estimation has long been used to identify causal relationships in labor economics, following the seminal study of economic returns to education by Behrman and Taubman (1976). Griliches (1979) first provides a rigorous treatment of the sibling models and data in econometric research. However, this method has its limitations (see Behrman and Rosenzweig 1999 and Bound and Solon 1999 for a critical assessment). In response to the criticisms, Ashenfelter and Rouse (1998) have extensively examined and validated the assumption of within-twin-pair random deviation from optimal schooling in the within-twin-pair fixed-effects estimator using U.S. data. As is standard, our identification hinges on the assumption that the within-twin-pair differences in schooling levels are uncorrelated with any omitted variables, which may themselves affect the formation of preferences in the future. We systematically examine the robustness of the within-twin-pair fixed-estimates in Section 6 below.

The small size of our sample of 140 subjects may cast doubt on the validity of the results from classical tests such as the *t*-statistics and *F*-statistics. Micceri (1989) offers an extensive survey and concludes that classical test 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 known to be 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 3000 draws under the random permutation procedure.[12]

4 Results on education and preferences

We first present the estimated effects of education on decision making under risk and uncertainty, followed by the estimated effects of education on decision making involving time.

4.1 Education and decision making under risk and uncertainty

Tables 3a, 3b, 3c, and 3d report both the OLS and within-twin-pair fixed-effects estimates of education and decision making under risk and uncertainty. As discussed above, we categorize educational attainment into four groups. The estimated coefficients on the three educational groups in the regression equation are relative to the omitted baseline group of middle-school-and-below, respectively. Columns (1)–(3) report the estimation results for risk attitudes towards 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's and mother's educational attainment, Column (2) shows an increase in the magnitude of the estimated coefficients on technical school and college level experience. Column (2) also shows that father's education level increases risk tolerance while mother's education level decreases risk tolerance, although the coefficients on paternal education levels are statistically insignificant. This finding may help rationalize a bargaining or collective household model with multiple decision makers rather than a unitary household model involving a dictator or a dominant preference in the family.[13] From the magnitudes of the estimated coefficients, the negative effect of mother's educational attainment seems to dominate the positive effect of father's

[12] We also compute the standard errors using the standard procedure, and find that our regression results remain robust.

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

Table 3a OLS and within-twin-pair fixed-effects estimates of education and decision making under risk and uncertainty – moderate prospects and moderate hazards.

	Dependent variables					
	Moderate prospects			Moderate hazards		
	OLS		Fixed Effects	OLS		Fixed Effects
	(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

Absolute values of *t*-statistics computed based on the permutation procedure (Freeman and Lane 1983) are in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; the omitted educational group 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. The two dependent variables are dummy variables. They equal 1 if the subject is risk tolerant towards moderate prospects and moderate hazards in Games 1–2; they equal 0 if the subject exhibits risk aversion. See Web Appendix I for the experiment instructions and Table 1 for the definition of experimentally measured preference variables educational attainment on the child's risk attitudes towards moderate prospects, but the differences are statistically insignificant.[14]

Column (3) reports the within-twin-pair fixed-effects estimation results, in which both observed and unobserved family characteristics are cancelled out. In this case, we find that the education level at college-and-above significantly increases risk tolerance towards moderate prospects. The coefficient on college-and-above is statistically significant at the 5% level. The effect of education on risk attitudes towards moderate prospects is economically substantial. Specifically, relative to the baseline group with middle-school-and-below education, the probability of subjects with college-and-above education choosing the even-chance lottery of receiving RMB40 and receiving zero (versus receiving the expected outcome of RMB20 with certainty) in Game 1 increases by 43.8 percentage points.

Columns (4)–(6) report the estimation results for risk attitudes towards moderate hazards. Similar to the results for risk attitudes towards moderate prospects, 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. The coefficient on college-and-above is statistically significant at the 10% level. The effect of education on risk attitudes towards moderate hazards is substantial. Specifically, relative to the baseline group with middle-school-and-below education, the probability of subjects with college-and-above education choosing the even-chance lottery of losing RMB10 and losing zero (versus losing the expected outcome of RMB5 with certainty) in Game 2 increases by 38.6 percentage points.

In Columns (4)–(5), the age effect on risk attitudes towards moderate hazards seems nonlinear. Age increases risk tolerance with respect to moderate hazards in the beginning, and then declines with old age. Both estimates on age and age-squared (1/100) are statistically significant at the 5% level. Based on the estimates, a subject's degree of risk tolerance with respect to moderate hazards peaks at age 39. This nonlinear pattern of the age effect exists also in the estimation of risk attitudes towards moderate prospects (Columns (1)–(2)). Although the estimated coefficients are only marginally significant, the magnitude of the estimated coefficient shows that a subject's degree of risk tolerance with respect to moderate prospects peaks at age 42. The estimated nonlinear relationship between age and attitude towards moderate risks may help reconcile a controversy in the literature.

Columns (7)–(9) report the estimates for attitudes towards longshot prospects. Similar to the results for risk attitudes towards moderate prospects and moderate hazards, the fixed-effects estimation results in Column (9) show that the education level of college-and-above increases risk tolerance for longshot prospects. The coefficient on college-and-above is statistically significant at the 5% level. To measure longshot prospects, the subject was asked to order the value of three choices in Game 3: (A) an RMB2 lottery ticket with a very small probability of winning five million, (B) an RMB2 lottery ticket with a small probability of winning 0.1 million, and (C) RMB2 with certainty. The magnitude of the coefficient on college-and-above in Column (9) implies that relative to the baseline group with middle-school-and-below

[14] The Wald test result fails to reject the null hypothesis that the sum of the coefficients on paternal and maternal education levels equals zero in Column (2).

Table 3b OLS and within-twin-pair fixed-effects estimates of education and decision making under risk and uncertainty – longshot prospects and longshot hazards.

	Dependent variables					
	Longshot prospects			Longshot hazards		
	OLS		Fixed Effects	OLS		Fixed Effects
	(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

Absolute values of *t*-statistics computed based on the permutation procedure (Freeman and Lane 1983) are in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; the omitted educational group 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. The two dependent variables are dummy variables. They equal 1 if the subject is risk tolerant towards longshot prospects and longshot hazards in Games 3–4; they equal 0 if the subject exhibits risk aversion. See Web Appendix I for the experiment instructions and Table 1 for the definition of experimentally measured preference variables education, the probability of subjects with college-and-above education preferring A to B to C increases by 48.2 percentage points.

On the contrary, Columns (10)–(12) do not show a significant effect of education on risk attitudes towards longshot hazards. We note that Game 4—used to measure risk attitudes towards longshot hazards—is not monetarily incentivized. However, the statistically insignificant effect of education on risk attitudes towards longshot hazards is not likely to be driven by the non-monetary incentivization. Table 1 shows no difference in the distribution of risk attitudes between longshot prospects and hazards, although Game 3—used to measure risk attitudes towards longshot prospects—is incentivized. Therefore, the difference between the effect of education on longshot prospects and on longshot hazards requires further exploration.

Columns (13)–(21) report 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 (13)–(15). It is interesting to find that those who are more educated are more likely to exhibit Allais-type behavior. Both the OLS and within-twin-pair fixed-effects estimates are statistically significant at least at the 10% level for the group with college-and-above education. Column (15) shows that relative to the group with middle-school-and-below education, the probability of the Game 6 choices of subjects with college-and-above education violating the independence axiom's parallelism implication increases by 41.7 percentage points. Column (14) shows that males are less likely to exhibit Allais-type behavior. The age effect on Allais-type behavior is nonlinear; the presence of Allais-type behavior declines with age at an increasing rate.

We present the estimated results for ambiguity aversion and familiarity bias in Columns (16)–(21). Relative to the baseline group with middle-school-and-below education, the fixed-effects estimates in Columns (18) and (21) show that more educated people are consistently more ambiguity averse as well as more biased towards 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. The coefficients on high school in both Columns (18) and (21) are statistically significant at the 10% level, whereas the coefficients on technical school and college-and-above are statistically insignificant. Relative to the baseline group with middle-school-and-below education, the probability of high school subjects making the ambiguity aversion choice in Game 7 increases by 37.2 percentage points, and the probability of making the familiarity bias choice in Game 8 increases by 31.9 percentage points.

In summary, Tables 3a, 3b, 3c, and 3d suggest that education increases subjects' risk tolerance towards moderate prospects, moderate hazards, and longshot prospects. In addition, more educated people are more likely to deviate from the prediction of expected utility theory, and are more likely to display Allais-type behavior. High school educated subjects also exhibit more ambiguity aversion as well as familiarity bias relative to pre-high school subjects. Finally, the big differences between the OLS estimates and within-twin-pair fixed-effects estimates for each preference measure in Tables 2a, 2b and 3a, 3b, 3c, 3d confirm that it is important to control for cross-family heterogeneity.

Table 3c OLS and within-twin-pair fixed-effects estimates of education and decision making under decision making under risk and uncertainty – Allais behavior and ambiguity aversion.

	Dependent variables					
	Allais behavior			Ambiguity aversion		
	OLS		Fixed Effects	OLS		Fixed Effects
	(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

Absolute values of *t*-statistics computed based on the permutation procedure (Freeman and Lane 1983) are in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; the omitted educational group 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. The two dependent variables are dummy variables. The dependent variable of Allais equals 1 if the subject exhibits Allais-type behavior in Game 5; it equals 0 if the subject exhibits expected utility behavior. The dependent variable of ambiguity aversion equals 1 if the subject exhibits ambiguity aversion in Game 6; otherwise, it equals 0. See Web Appendix I for the experiment instructions and Table 1 for the definition of experimentally measured preference variables

4.2 Education and decision making involving time

Tables 4a, 4b, and 4c report the OLS and within-twin-pair fixed-effects estimates of education on decision making involving time. Columns (1)–(3) suggest that subjects with college-and-above education are significantly more patient than other groups. The within-twin-pair fixed-effects estimation results in Column (3) show that the coefficient on college-and-above is statistically significant at the 5% level. Relative to the group with middle-school-and-below education, subjects with college-and-above education are 52.0 percentage points more likely to choose receiving RMB120 seven days later over receiving RMB100 today in Game 8. In addition, the within-twin-pair fixed-effects estimate in 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. The coefficient on college-and-above is statistically significant at the 10% level. Moreover, the effect of education on hyperbolic discounting behavior is economically substantial. Relative to the group with middle-school-and-below education, the probability of subjects with college-and-above education making time inconsistent choices in Game 8 decreases by 38.2 percentage points.

Columns (1)–(2) and (4)–(5) suggest that there exist substantial gender differences in time preferences. Males are more patient and less disposed to exhibiting hyperbolic discounting than females. The coefficients on gender (male indicator) in Columns (1)–(2) are statistically significant at the 1% level. Relative to females, males are about 25 percentage points less likely to choose receiving RMB100 today over receiving RMB120 seven days later in Game 8. The coefficients on gender in Columns (4)–(5) are statistically significant at the 5% level. The probability of making time inconsistent choice in Game 8 is 18 percentage points lower for males than females.

Columns (7)–(9) report the estimated effects of education on anticipation. We find that education has a positive effect on anticipation. However, the within-twin-pair fixed-effect estimated coefficient on college-and-above is statistically insignificant (Column (9)), although the OLS estimate is statistically significant at the 5% level. By contrast, Columns (12) and (15) indicate that more education significantly decreases both dread and hopefulness. Finally, subjects with college-and-above education are less likely to be anxious, although the estimates are not statistically significant (Column (18)).

In summary, the within-twin-pair fixed-effects estimates in Tables 4a, 4b, and 4c suggest that education decreases impatience, hyperbolic discounting, dread, hopefulness, and anxiousness. While education seems to increase anticipation, the fixed-effects estimate is statistically insignificant. In contrast to choice under risk and uncertainty, people with higher levels of education tend to exhibit less “bias” in preference regarding time. Tables 4a, 4b, and 4c also indicate 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 of addressing the causality between education and preferences.

5 Robustness

This section reports the results from a robustness analysis. First, to deal with potential measurement errors with educational attainment, we conduct an instrumental variable

Table 3d OLS and within-twin-pair fixed-effects estimates of education and decision making under risk and uncertainty – familiarity bias.

	Dependent variables		
	Familiarity bias		
	OLS		Fixed Effects
	(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

Absolute values of *t*-statistics computed based on the permutation procedure (Freeman and Lane 1983) are in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; the omitted educational group 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. The dependent variable is a dummy. It equals 1 if the subject exhibits familiarity bias in Game 7; otherwise, it equals 0. See Web Appendix I for the experiment instructions and Table 1 for the definition of experimentally measured preference variables 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 twins. Third, we estimate the effects of educational attainment on preferences controlling for income and health. Fourth, we check for the possibility of reversal causality. Fifth, we repeat the estimations with a larger sample. Finally, we examine any remaining potential biases in the within-twin-pair estimates.

5.1 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 a measurement error bias. This paper follows Ashenfelter and Krueger (1994) in obtaining good instrumental variables to deal with possible measurement error problems. 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 as the cross-reported education should be correlated with the true education of a twin but should not directly affect the latter's left-hand-side variables.

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 for $E_1^1 - E_2^2$ 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) demonstrates, the measurement error term in $E_1^2 - E_2^1$ and the measurement error term 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 canceled out.

Measurement error should not be a major concern in this study as there are only six individuals in the whole sample whose self-reported education levels are different from those reported by co-twins. The potential rate of misreports of education is only 4.29%, which is lower than the misreport rate of X in Ashenfelter and Krueger (1994). Considering the low misreport rate, we expect that the IV estimates will not be very different from the within-twin-pair fixed-effects estimates reported in the previous section. This prediction is confirmed by Table 5, which shows the instrumental variables within-twin-pair fixed-effects estimates of education and preferences. In summary, our within-twin-pair fixed-effects estimates of the education effects on preferences are robust to the measurement error problem.

5.2 Omitted variables

While twins share a similar family environment, there may be unobservable heterogeneity between twins. For example, twins may differ in nutrition intake

Table 4a OLS and within-twin-pair fixed-effects estimates of education and anomalies in decision making involving time – impatience and hyperbolic discounting.

	Dependent variables					
	Impatience			Hyperbolic discounting		
	OLS		Fixed Effects	OLS		Fixed Effects
	(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

Absolute values of *t*-statistics computed based on the permutation procedure (Freeman and Lane 1983) are in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; the omitted educational group 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. The two dependent variables are dummy variables. If the subject exhibits impatience and hyperbolic discounting behavior in Game 8, the two variables equal 1; otherwise, they equal 0. See Web Appendix I for the experiment instructions and Table 1 for the definition of experimentally measured preference variables before birth and have different birth weights. Behrman and Rosenzweig (2004) find that birth weight affects a series of short- and long-run outcomes such as health, education, and income. Therefore, we include birth weight to control for pre-birth endowment as a robustness test. Table 6 reports the within-twin-pair fixed-effects estimates controlling for birth weight. The estimated effects of educational attainment on preferences are very similar to those in Tables 3a, 3b, 3c, 3d and 4a, 4b, 4c, indicating that our results are robust to the inclusion of birth weight as a control variable. In Table 6, we further find that birth weight does not have a statistically significant effect on preferences.

Given the small sample size, we have included both the MZ and DZ twins in our estimation. Although DZ twins share an identical family environment, they only share half the genetic endowment. Thus, it may be argued that the within-twin-pair fixed-effects estimation is not as 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 36 pairs of MZ twins, the effects of educational attainment on preferences remains similar to those reported in the previous section. Despite the small sample size, the education level at college-and-above significantly increases risk tolerance towards moderate prospects and decreases impatience.

5.3 Estimates controlling for income and health status

It has been argued that socioeconomic variables such as income and health affect preferences, while education affects income and health. Thus, we have included income and health in the regression to check whether the effects of education on preferences are acting through income or through health. Tables 8 and 9 report the within-twin-pair fixed-effects estimates of educational attainment on preferences by controlling for income and health, respectively. We do not find any significant change in the estimates.

5.4 Reverse causality

Another potential problem with our within-twin-pair fixed-effects estimates is reverse causality bias. Although unobservable family factors and individual heterogeneity may have been taken care of by using the within-twin-pair fixed-effects estimator, there may be a reversal causality problem running from preference to educational attainment. However, for our subjects, preferences are experimentally measured at, on average, 45 years of age while education is completed before age 22.[15] Moreover, our within-twin-pair fixed-effects estimates are less likely to be biased by the reversal causality problem. Since twin siblings, particular MZ twin siblings, share common family background and genetic endowment, it is unlikely that twins have differences in preferences at the early stages of their lives that influence their subsequent educational attainment.

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

Table 4b OLS and within-twin-pair fixed-effects estimates of education and anomalies in decision making involving time – anticipation and dread.

	Dependent variables					
	Anticipation			Dread		
	OLS		Fixed Effects	OLS		Fixed Effects
	(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

Absolute values of *t*-statistics computed based on the permutation procedure (Freeman and Lane 1983) are in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; the omitted educational group 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. The two dependent variables are dummy variables. The dependent variable of anticipation equals 1 if the subject exhibits anticipation in Game 9; otherwise, it equals 0. The dependent variable of dread equals 1 if the subject exhibits dread in Game 10; otherwise, it equals 0. See Web Appendix I for the experiment instructions and Table 1 for the definition of experimentally measured preference variables

5.5 Additional sample

We test the robustness of our results with a larger sample. Prior to the experiments reported in the main results, we conducted some of the experimental measurements for 63 pairs of twins using other channels such as twin festivals (see Zhong et al. 2009 for details). The experimental measurements include moderate prospects, moderate hazards, longshot prospects, longshot hazards, impatience, and time inconsistency with the same experimental instructions. We did not include the other measurements on time preferences for this sample. Including this additional sample, we have 133 pairs of the twins, and the results remain qualitatively similar to those reported in the previous section. The estimation results are reported in Table A1 of Web Appendix IV.

5.6 Potential biases of within-twin-pair fixed-effects estimates

Ashenfelter and Rouse (1998) emphasize that there are no genetic differences between identical twins except measurement errors. They argue that different schooling levels of identical twins are mainly due to random deviations that are not related to the determinants of schooling choices. However, the within-twin-pair estimation may not completely eliminate the bias of the conventional cross-sectional estimation, although our within-twin-pair fixed-effects estimates are consistently robust in a series of sensitivity analyses discussed above. 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 unobserved within-twin-pair heterogeneity contributes to endogenous variation in education, then the within-twin-pair estimation is still subject to an endogeneity bias. Thus, the major issue of concern with the within-twin-pair estimate is whether it is less biased than the cross-sectional OLS estimate, and correspondingly a better estimate.

Note that the bias in the cross-sectional OLS 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, $\frac{\mathrm{cov}(E_i,\mu_i+\varepsilon_i)}{\mathrm{var}(E_i)}$. 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, $\frac{\mathrm{cov}(\Delta E_i,\Delta\mu_i+\Delta\varepsilon_i)}{\mathrm{var}(\Delta E_i)}$. 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 OLS estimator. In that case, we can credit the within-twin-pair estimates as being preferred to OLS estimates.

Ashenfelter and Rouse (1998) suggest 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. (2012) conduct 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 OLS regression is likely to be larger than the bias in the within-twin-pair regression. Li et al.

Table 4c OLS and within-twin-pair fixed-effects estimates of education and anomalies in decision making involving time – hopefulness and anxiousness.

	Dependent variables					
	Hopefulness			Anxiousness		
	OLS		Fixed Effects	OLS		Fixed Effects
	(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

Absolute values of *t*-statistics computed based on the permutation procedure (Freeman and Lane 1983) are in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%; the omitted educational group 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. The two dependent variables are dummy variables. The dependent variable of hopefulness equals 1 if the subject exhibits hopefulness in Game 11; otherwise, it equals 0. The dependent variable of anxiousness equals 1 if the subject exhibits anxiousness in Game 12; otherwise, it equals 0. See Web Appendix I for the experiment instructions and Table 1 for the definition of experimentally measured preference variables

(2012) find that the between-family correlations all exceed the within-twin-pair correlations, suggesting that the within-twin-pair estimation of the returns to education may indeed be less affected by omitted individual heterogeneity than the cross-sectional OLS estimation.[16] Given that we also use the CATS data, Li et al. (2012) provide suggestive evidence that our within-twin-pair fixed-effects estimates are less biased than cross-sectional OLS estimates.

6 Discussion and concluding remarks

Using survey and experimental data, we present a systematic empirical study of the effect of educational attainment on two dimensions of decision making behavior—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 identify the effect of education on these two dimensions of preference behavior. Our fixed-effects estimates indicate that people with higher levels of education are less risk averse towards moderate prospects, moderate hazards, and longshot prospects. These findings are in line with previous findings on risk attitudes and cognitive ability (Frederick 2005; Dohmen et al. 2010; Benjamin et al. 2013; Burks et al. 2009; Taylor 2013). Furthermore, we extend the findings to moderate hazards and longshot prospects. In addition, we find that more educated people tend to be more disposed to exhibit Allais-type behavior and ambiguity aversion. In terms of preferences involving time, our findings suggest that those who are more educated tend to be more patient and exhibit less hyperbolic discounting.

Since the earlier work of Allais (1953), there have been debates on whether behavioral anomalies reflect the true preferences or the cognitive biases of the decision makers. Should anomalies reflect the limitations of cognitive ability, increased understanding of these anomalies ought to increase the frequency of "truly" normative responses (Savage 1954), and decision makers may then benefit from prescriptive measures to correct their own behavior. Contrary to Morgenstern's (1979) view, we find that education increases anomalous decision making behavior under risk and uncertainty, which suggests that these violations of expected utility reflect the decision maker's actual preference. By contrast, we find that education decreases anomalous decision making behavior across time, which suggests that time inconsistency could in part reflect cognitive biases of the decision makers. Furthermore, this finding supports the observation that decision makers voluntarily demand precommitment devices to reduce their self-control problem in decision making over time (i.e., Thaler and Benartzi 2004). However, similar mechanisms are much less discussed when it comes to Allais– type behavior or ambiguity aversion in decision making under risk. There is scope to investigate the mechanisms driving the differential effects of education on anomalous decision making behavior in the two dimensions of risk and time.

Although the relationships between demographic and socioeconomic variables, cognitive and non-cognitive skills, and risk attitudes have been extensively investigated

[16] Table A2 in Web Appendix V cites the between-families and within-twin-pair correlations of education and other variables in Li et al. (2012).

Table 5 Instrumental variables within-twin-pair fixed-effects estimates of education and decision making behavior.

	Dependent variables						
	Decision making under uncertainty						
	Moderate prospects	Moderate hazards	Longshot prospects	Longshot hazards	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	

Absolute values of *t*-statistics computed based on the permutation procedure (Freeman and Lane 1983) are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; the omitted educational group is middle-school-and-below. The notes under Tables 2a, 2b and 3a, 3b, 3c, 3d define each dependent variable

Table 6 Within-twin-pair fixed-effects estimates of education and decision making behavior controlling for birth weight.

	Dependent variables						
	Decision making under uncertainty						
	Moderate prospects	Moderate hazards	Longshot prospects	Longshot hazards	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)	

Table 6 (continued)

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

Absolute values of *t*-statistics computed based on the permutation procedure (Freeman and Lane 1983) are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; the omitted educational group is middle-school-and-below. The notes under Tables 2a, 2b and 3a, 3b, 3c, 3d define each dependent variable

Table 7 Within-twin-pair fixed-effects estimates of education and decision making behavior using Mz twins pairs.

	Dependent variables						
	Decision making under uncertainty						
	Moderate prospects	Moderate hazards	Longshot prospects	Longshot hazards	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	

Absolute values of *t*-statistics computed based on the permutation procedure (Freeman and Lane 1983) are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; the omitted educational group is middle-school-and-below. The notes under Tables 2a, 2b and 3a, 3b, 3c, 3d define each dependent variable

Table 8 Within-twin-pair fixed-effects estimates of education and decision making behavior controlling for income.

	Dependent variables						
	Decision making under uncertainty						
	Moderate prospects	Moderate hazards	Longshot prospects	Longshot hazards	Allais-type behavior	Ambiguity aversion	Familiarity bias
	(1)	(2)	(3)	(4)	(5)	(6)	(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	Hyperbolic discounting	Anticipation	Dread	Hopefulness	Anxiousness	
	(8)	(9)	(10)	(11)	(12)	(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	

Absolute values of *t*-statistics computed based on the permutation procedure (Freeman and Lane 1983) are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; the omitted educational group is middle-school-and-below. The notes under Tables 2a, 2b and 3a, 3b, 3c, 3d define each dependent variable

Table 9 Within-twin-pair fixed-effects estimates of education and decision making behavior controlling for health.

	Dependent variables						
	Decision making under uncertainty						
	Moderate prospects	Moderate hazards	Longshot prospects	Longshot hazards	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)	

Table 9 (continued)

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

Absolute values of *t*-statistics computed based on the permutation procedure (Freeman and Lane 1983) are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; the omitted educational group is middle-school-and-below. The notes under Tables 2a, 2b and 3a, 3b, 3c, 3d define each dependent variable

in the literature, the issue of causality has rarely been addressed. Our overall results suggest that the OLS estimated correlations between education and preferences are far from being causal. The substantial differences between OLS estimates and within-twin-pair fixed-effects estimates for each category of preference imply that some omitted variables relating to family background may confound those observed correlations between educational attainments and preferences.

To the best of our knowledge, this paper seems to be the first study exploring the causal effect of education on decision making behavior. However, this paper has limitations. One concerns the small sample size of 140 subjects (70 pairs of twins), despite our proper treatment of the problem using a permutation-based inference procedure. As a field experiment on adult twins, 140 subjects seem moderate in size, relative to other behavioral and experimental studies in the literature. As an empirical study, we do not attempt to model the mechanism through which education affects individuals' decision making under risk and uncertainty and involving time. We anticipate that our new empirical findings would give rise to further research to explore the theoretical mechanisms underpinning the interplay between education and preference formation in general, and between education and decision making anomalies in particular.

Acknowledgment We gratefully acknowledge financial support from the Chinese University of Hong Kong (Focused Innovations Scheme), the Hong Kong University of Science and Technology (Project No. AOE-MG/H-03/06), and National University of Singapore (FASS Start-up Grant No. R-122-000-211-133 and HSS Grant No. R-122-000-232-646).

References

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'ecole Americaine [Rational Man Behavior in the Presence of Risk: Critique of the Postulates and Axioms of the American School]. *Econometrica, 21*, 503–546.

Angrist, J. D., & Pischke, J. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Ashenfelter, O., & Krueger, A. (1994). Estimates of the economic return to schooling for a new sample of twins. *American Economic Review, 84*(5), 1157–1173.

Ashenfelter, O., & Rouse, C. (1998). Income, schooling, and ability: Evidence from a new sample of identical twins. *Quarterly Journal of Economics, 113*(1), 253–284.

Becker, G. S. (1992). Habits, addictions, and traditions. *Kyklos, 45*(3), 327–345.

Becker, G. S. (1996). *Accounting for Tastes*. Cambridge: Harvard University Press.

Becker, G. S., & Mulligan, C. B. (1997). The endogenous determination of time preference. *Quarterly Journal of Economics, 112*(3), 729–758.

Behrman, J. R. (1997). Intrahousehold distribution and the family. In Rosenzweig, M. R. & Stark, O. (Eds.), *Handbook of Population and Family Economics*, Volume 1A. Amsterdam: Elsevier Science, North-Holland, 125–87.

Behrman, J. R., & Rosenzweig, M. R. (1999). 'Ability' biases in schooling returns and twins: A test and new estimates. *Economics of Education Review, 18*(2), 159–167.

Behrman, J. R., & Rosenzweig, M. R. (2004). Returns to birthweight. *The Review of Economics and Statistics, 86*(2), 586–601.

Behrman, J. R., & Taubman, P. (1976). Intergenerational transmission of income and wealth. *American Economic Review, 66*(2), 436–440.

Behrman, J. R., Taubman, P., Wales, T., & Hrubec, Z. (1977). *Inter- and intragenerational determination of socioeconomic success with special reference to genetic endowment and family and other environment.* Mimeo: University of Pennsylvania.

Behrman, J. R., Rosenzweig, M. R., & Taubman, P. (1996). College choice and wages: Estimates using data on female twins. *The Review of Economics and Statistics, 73*(4), 672–685.

Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (2013). Who is 'behavioral'? Cognitive ability and anomalous preferences. *Journal of the European Economic Association, 11*(6), 1231–1255.

Bisin, A., & Verdier, T. (2000). A model of cultural transmission, voting and political ideology. *European Journal of Political Economy, 16*, 5–29.

Borghans, L., Duckworth, A. L., Heckman, J. J., & ter Weel, B. (2008). The economics and psychology of personality traits. *NBER Working Paper No., 13*, 810.

Bound, J., & Solon, G. (1999). Double trouble: On the value of twins-based estimation of the return to schooling. *Economics of Education Review, 18*(2), 169–182.

Bowles, S. (1998). Endogenous preferences: The cultural consequences of markets and other economic institutions. *Journal of Economic Literature, 36*(1), 75–111.

Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences, 106*(19), 7745–7750.

Camerer, C., & Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty, 5*(4), 325–370.

Card, D. (1999.) The causal effect of education on earnings. In Ashenfelter, O. & Card, D. (Eds.), *Handbook of Labor Economics*, Volume 3A. Amsterdam: Elsevier Science, North-Holland, 1801–63.

Charness, G., Karni, E., & Levin, D. (2010). On the conjunction fallacy in probability judgment: New experimental evidence regarding Linda. *Games and Economic Behavior, 68*(2), 551–556.

Chew, S. H., & Ho, J. L. (1994). An empirical study of attitude toward the timing of uncertainty resolution. *Journal of Risk and Uncertainty, 8*(3), 267–288.

Chew, S. H., & Waller, W. S. (1986). Empirical tests of weighted utility theory. *Journal of Mathematical Psychology, 30*(1), 55–72.

Dohmen, T. J., Falk, A., Huffman, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review, 100*(3), 1238–1260.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics, 75*(4), 643–669.

Fox, C. R., & Tversky, A. (1995). Ambiguity aversion and comparative ignorance. *Quarterly Journal of Economics, 110*(3), 585–603.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25–42.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature, 40*(2), 351–401.

Freeman, D., & Lane, D. (1983). A nonstochastic interpretation of reported significance level. *Journal of Business and Economic Statistics, 1*, 292–298.

Gorseline, D. W. (1932). *The effect of schooling upon income*. Bloomington: Indiana University Press.

Griliches, Z. (1979). Sibling models and data in economics: Beginnings of survey. *Journal of Political Economy, 5*, 37–64.

Grossman, M. (1975). The correlation between health and schooling. In Terleckyj, N. E. (Ed.), *Household Production and Consumption: 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.

Heckman, J. J. (2007). The technology and neuroscience of capacity formation. *Proceedings of the National Academy of Sciences, 104*(33), 13,250–13,255.

Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review, 91*(2), 145–149.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P., & Yavitz, A. (2010). Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quantitative Economics, 1*, 1–46.

Huang, C., Li, H., Liu, P., & Zhang, J. (2009). Why does spousal education matter for incomes? Assortative mating and cross-productivity. *Journal of Labor Economics, 27*(4), 633–652.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–292.

Kreps, D. M., & Porteus, E. L. (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica, 46*(1), 185–200.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics, 112*(2), 443–477.

Li, H., Liu, P., Zhang, J., & Ma, N. (2007). Economic returns to Communist Party membership: Evidence from urban Chinese twins. *Economic Journal, 523*, 1504–1520.

Li, H., Rosenzweig, M., & Zhang, J. (2010). Altruism, favoritism, and guilt in the allocation of family resources: Sophie's Choice in Mao's mass send-down movement. *Journal of Political Economy, 118*(1), 1–38.

Li, H., Liu, P., & Zhang, J. (2012). Estimating returns to education using twins in urban China. *Journal of Development Economics, 97*(2), 494–504.

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *Economic Journal, 3*, 666–684.

Lovallo, D., & Kahneman, D. (2000). Living with uncertainty: Attractiveness and resolution timing. *Journal of Behavioral Decision Making, 13*(2), 179–190.

Machina, M. J. (1982). Expected utility analysis without the independence axiom. *Econometrica*, 277–323.

Machina, M. J. (2004). Almost-objective uncertainty. *Economic Theory, 24*(1), 1–54.

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin, 105*, 156–166.

Morgenstern, O. (1979). Some reflections on utility. In M. Allais & O. Hagen (Eds.), *Expected Utility and the Allais Paradox*. Dodrecht : D. Reidel.

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

Oreopoulos, P., & Salvanes, K. G. (2011). Priceless: The nonpecuniary benefits of schooling. *The Journal of Economic Perspectives, 25*(1), 159–184.

Rosenzweig, M. R., & Zhang, J. (2013). Economic growth, comparative advantage, and gender differences in schooling outcomes: Evidence from the birthweight differences of Chinese twins. *Journal of Development Economics, 104*, 245–260.

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

Savage, L. J. (1954). *The Foundations of Statistics*. New York: Wiley.

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature, 38*(2), 332–382.

Stigler, G. J., & Becker, G. S. (1977). De gustibus non est. disputandum. *American Economic Review, 67*(2), 76–90.

Taubman, P. (1976a). The determinates of earnings: Genetics, family and other environments; a study of white male twins. *American Economic Review, 66*(5), 858–870.

Taubman, P. (1976b). Earnings, education, genetics, and environment. *Journal of Human Resources, 11*(4), 447–461.

Taylor, M. P. (2013). Bias and brains: Risk aversion and cognitive ability across real and hypothetical settings. *Journal of Risk and Uncertainty, 46*(3), 299–320.

ter Weel, B. (2008). The noncognitive determinants of labor market and behavioral outcomes: Introduction to the symposium. *Journal of Human Resources, 43*(4), 729–737.

Thaler, R. H., & Benartzi, S. (2004). Save More Tomorrow(TM): Using behavioral economics to increase employee saving. *Journal of Political Economy, 112*(1), 164–187.

Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *Economic Journal, 113*(485), 3–33.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297–323.

Zhong, S., Chew, S. H., Set, E., Zhang, J., Xue, H., Sham, P. C., Ebstein, R. P., & Israel, S. (2009). The heritability of attitude toward economic risk. *Twin Research and Human Genetics, 12*(01), 103–107.