Can Experimental Economics Explain Competitive Behavior Outside the Lab?*

Y. JANE ZHANG†

August, 2012

Abstract

In light of the growing literature on competitive inclination measured in economic experiments, this study examines whether competitive inclination derived using the now standard Niederle and Vesterlund (2007) experiment is predictive of a subject's subsequent propensity to take a competitive and highly consequential high school entrance exam in rural China. Individual parameters of competitive inclination are estimated from lab data using a discrete choice mixed logit model. Including a matrilineal and a patrilineal ethnic minority group in the subject pool allowed for the testing of ethnicity-gender effects. Results indicate that a middle school student with an inclination toward competition one standard deviation above the mean is 7.2 percentage points more likely to take the exam, controlling for prior test scores. Contrary to previous experimental results using adult populations, no gender differences were found in competitive inclination in any of the ethnic groups. This is consistent with the finding of no gender differences in exam taking behavior, controlling for prior test scores.

*I am deeply grateful to Shachar Kariv and Edward Miguel for their guidance and support throughout this research project. I thank David Card, Vinci Chow, Fred Finan, Bryan Graham, Jeremy Magruder, Ulrike Malmendier, Betty Sadoulet, Changcheng Song, Ken Train; seminar participants at the UC Berkeley Development Seminar, Development Lunch, Pysch and Econ Non-Lunch, and IGERT seminar; seminar participants at the University of Western Ontario, Williams College, National University of Singapore, Hong Kong University of Science and Technology, and Santa Clara University. I also thank Zhao Xinghan, Yang Haitao and Mao Falin for their invaluable support and assistance in the field.

†Hong Kong University of Science and Technology

1 Introduction

A growing experimental literature on competition has found that men are much more willing than women to choose to compete in a tournament rather than take a piece-rate payment, even after controlling for performance in the experimental task (e.g., Niederle and Vesterlund (2007); Booth and Nolen (2012); Niederle et al. (2010); Sutter and Rützler (2010); Healy and Pate (2011); Dargnies (2011)). A related strand of the literature finds that males also perform better than females under competitive situations relative to non-competitive situations (e.g., Gneezy et al. (2003); Gneezy and Rustichini (2004); Paserman (2010)). These results resonate with the conventional wisdom that men enjoy competitive environments more than women do (see, for example, Tierney (2005) and Varian (2006)), and have the potential to explain the underrepresentation of women in areas where competition is intense, such as in upper-level corporate management (Bertrand and Hallock, 2001; Wolfers, 2006) and parliamentary-level politics (United Nations, 2010).[1]

The labor market implications of the experimental results rest upon the assumption that lab measures of competitive inclination are meaningful indicators of decisions involving competition in the world outside of the lab. This study aims to explicitly test the assumption. Using the now standard Niederle and Vesterlund (2007) experimental design, I test whether experimentally derived measures of competitive inclination can predict a subsequent real world decision to take the competitive and highly consequential high school entrance exam in rural China. This exam is relevant both because stakes are high (admissions to secondary school, where returns to education have been estimated at 11.5% per year (De Brauw and Rozelle, 2008), are based on performance in this exam, and further schooling is practically impossible without a secondary school education) and because competition is especially fierce in rural China, where the continuation rate from middle school to high school is 33% (Ministry of Education, 2007; 2008).

Experiments were conducted with 248 middle school students in rural Yunnan Province in September 2008 and January 2010. A post-experiment survey was administered to capture demographic and socioeconomic background characteristics. Individual risk aversion data were collected separately using the Binswanger (1980) instrument in January and March 2010. Aca-

[1]Bertrand and Hallack find that between 1992 and 1997, of the top 5 highest paid executives in a large set of US public corporations, 2.5% were women; Wolfers finds that between 1992 and 2004, women occupied the position of CEO of the S&P 1500 companies 1.3% of the time. The UN reports that in 2010, 18% of parliamentarians worldwide are women, including seats reserved specifically for female politicians. In developing regions, this figure is 15%.

demic performance data were collected from the school administration. Since regular term middle school tests are standardized across the county and designed to mimic high school entrance exams, performance on these tests serves as a good proxy for academic ability as it applied to high school entry.[2] The subjects' real world competitive behavior was observed in a set of county-level administrative records that I collected and assembled following each entrance exam cycle.[3]

The lab data on task performance and risk aversion were used to estimate a structural discrete choice mixed logit model of the decision to compete in the tournament, from which I obtained individual-specific parameters of competitive inclination using a method similar to the one proposed in Revelt and Train (2000). While previous studies in the literature have largely focused on the existence (or the lack) of average gender differences in competitive inclination, generating individual estimates of competitive inclination is critical for examining the link between competitive inclination in and outside the lab at the individual level.

Results show that even after accounting for academic performance, the lab measure of competitive inclination is a statistically significant predictor of the propensity to take the exam, with an economically meaningful magnitude. Subjects with a taste for competition one standard deviation above the mean are 7.2 percentage points more likely to take the entrance exam, controlling for prior academic performance. Taken from the baseline, this represents a reduction of 36% in the exam truancy rate. To my knowledge, this finding is the first direct evidence linking lab measures of competitive inclination with competitive behavior outside the lab at an individual level.

The location of this study was chosen for its ethnic diversity, which allows me to explore how gender differences in competitive inclination and real world competitive behavior may vary with ethnic gender norms, in a naturally controlled environment. A previous study on competitive inclination in unilineal societies in India and Tanzania found matrilineal women to be more competitively inclined than matrilineal men, whereas patrilineal women were less competitively inclined than patrilineal men (Gneezy et al., 2009). In this setting, the cross-cultural comparisons can be made within the same county. The Han Chinese, along with a matrilineal minority group (Mosuo) and a patrilineal minority group (Yi), live together in an area one-third the geographical size of the San Francisco Bay Area and, moreover, attend schools under the same national public school

[2]In contrast, in the United States, preparation for the SATs, which are a set of standardized exams used in college admissions, is generally not a part of the standard high school curriculum.

[3]Using administrative data side-steps the problem of potentially unreliable self-reports.

system.[4]

Using school rosters, I randomly selected experimental subjects, who are on average 15.7 years old, from each of the three ethnic groups in roughly equal numbers, balanced across gender. I find no statistically significant gender difference in competitive inclination or in exam-taking behavior, controlling for prior test scores, among any of the three ethnic groups. This result adds to the mixed findings on the willingness to compete among adolescents and children around the world (Booth and Nolen, 2012; Sutter and Rützler, 2010; Dreber et al., 2011; Cárdenas et al., 2012).

The lack of gender difference in competitive inclination contrasts with the findings from Zhang (2012), which finds that among high school students in this region, both matrilineal and patrilineal students exhibit a significant gender gap in competitive inclination, although the Han Chinese did not. The differences in the findings between high school and middle school subjects would be consistent with a developmental explanation. However, due to sample selection factors, the two sets of results are not directly comparable.

The remainder of the paper proceeds in Section 2 to discuss the background literature. Section 3 describes the experimental setting. Section 4 describes the data collection procedures. Section 5 develops a structural discrete choice mixed logit model. Section 6 presents results from the structural estimation, findings on external validity, and findings on gender differences in competitive inclination across ethnic groups. Section 7 concludes with a discussion of the results and their potential implications.

2 Related Literature

This study contributes to a literature exploring the generalizability of findings from competition experiments to real world competitive behavior. One approach in this literature analyzes data from real competitions that take place outside of the lab, with the outcome of interest often being performance responses to competitive environments or high stakes, i.e., similar to the Gneezy et al. (2003) research question. For example, Paserman (2010) finds that male professional tennis players perform better than female ones do during critical points of the match, relative to performance during less critical points. Ors et al. (2008) finds that while women outperformed men in

[4]For another study using Mosuo and Yi subjects, see Gong and Yang (2012), which examines gender differences in risk attitudes among adults.

high school completion exams in France, the same men outperformed the same women in a highly competitive entrance exam two years later. Attali et al. (2011) finds that the performance differential on real GRE questions versus experimental ones that did not affect GRE scores was higher for men relative to women. Lavy (2008) documents potentially contrary evidence: no gender difference was found in the performance of high school teachers who were rewarded according to the relative performance of their students on subject exams. It is unclear, however, how the individual teachers performed before the reward system was put in place.

Another approach in this literature uses a natural field experiment where incentives are similar to laboratory incentives but participants do not realize that they are in an experiment. In Flory et al. (2012), researchers posted job openings online and randomized the payment schemes that were disclosed to respondents to the ads. They found that of those who responded, men were more likely than women to officially apply for the jobs in the treatment where the payment scheme is based on a tournament with high stakes, relative to the treatment where payment is a fixed hourly rate.

Given that the aim of the current study is to assess the external validity of the influential and widely used Niederle and Vesterlund (2007) competition experiment, the approach is to directly link individual-level competitive inclination elicited in the lab to individual decisions involving a real world competition of consequence. As such, this study contributes to a nascent literature on validating experimental economics results in the broader world.[5] Most of this literature involves social preference experiments. Karlan (2005) finds that those Peruvian microfinance borrowers who were more trustworthy in a trust game (i.e. returned more money to their partners) were more likely to repay their microloans. Benz and Meier (2008) finds that Swiss students who donated more money to charities in a classroom experiment also donate more money to charities in a natural setting. Finan and Schechter (2010) finds that Paraguayans who exhibited more reciprocity in an experimental setting were more likely to vote for the party from whom they accepted a gift. To my knowledge, the current study is the first to test the relationship between lab and non-lab behavior involving competition experiments.[6]

[5]See Falk and Fehr (2003) for a discussion on the generalizability of lab experiments and Leavitt and List (2007) for a discussion on external validity as it applies to social preference experiments.

[6]For studies that have correlated survey responses to questions about competitive attitudes to labor market outcomes see Fortin (2005) and Manning and Swaffield (2008).

3 Experimental Setting

All subjects are recruited from two middle schools in Ninglang county. Ninglang is a mountainous county located in the border province of Yunnan, with GDP per capita at $630 of 2008 (China County Statistics, 2008).[7]

3.1 Education in Ninglang County

In 1986, China passed the Law on Compulsory Education, mandating six years of primary and three years of middle school education for all children (Ministry of Education, 1986). Schools in Ninglang County, as elsewhere in China, follow a uniform standard for textbooks, curricula, and exams. Middle schools and high schools are boarding schools, although students whose homes are nearby may choose to commute. Admission to high school is competitive and depends almost exclusively on an entrance exam.[8] At the very least, one must take the entrance exam in order to gain admission – I did not encounter a single case of someone not taking the exam and continuing on to high school. In the current study, about 80% of middle school students take the entrance exam, and less than 30% go on to high school without repeating.[9]

As is common in rural China, the two high schools in Ninglang are both located in the county seat. The few who score sufficiently high for a prefectural level high school may apply to it once exam scores are made public, but by default everyone who passes the exam is admitted to one of the two schools in Ninglang county. The process for assigning the admitted pool of students into one of the two county high schools is essentially random.

The fees for the entrance exam is around $70 and mainly covers three days of food and lodgings in the county seat, where the exam is administered. Although this is not a small sum for this population, school administrators insist that the fee does not prevent students from taking the exam. Among the administrators and students I spoke with, the consensus was that parents almost invariably want the student to take the entrance exam – if the student does not take the

[7] At the exchange rate of $1 = RMB 6.8.

[8] Admissions may be extended to those who just missed the cutoff and could pay an extra fee, but I did not find written guidelines for this process.

[9] Although some students repeat grade 9, it is rare for someone who does not continue directly to high school to pass the exam in a later year. In the sample of grade 9 students in 2008, 5 out of 75 who did not go on to high school directly passed the exam on the subsequent try.

exam, it was his or her own choice against their parents' wishes. For a more systematic understanding of the issue, I consulted a study on the barriers to education in rural Gansu province that surveyed over 2,000 students and families (Hannum and Adams, 2009). For children aged 13-16 who drop out of school, the survey found that the top two reasons given were consistent across responses given by the child, their mother, and their village leader: poor academic performance and simply not wanting to go to school. The comments from in-depth interviews conducted in Gansu echo the sentiments I heard in Ninglang. In that study, one student said: "[My parents tell me:] If you pass entrance exams, even if we have to sell our house and vehicle, we will, in order to support your schooling." A mother said: "In this village, if you do not study, you are in for a hard life...but if your child refuses to learn, we, as parents, really cannot do anything (Hannum and Adams, 2009)."

3.2 Ethnic minorities in Ninglang County

With a population of 230,000, Ninglang's three main ethnic groups are the Yi, the Han Chinese, and the Mosuo, comprising 62%, 20%, and 9% of the population, respectively (China Census, 2000).[10] Historical records show that the Han Chinese, the Yi, and the Mosuo have coexisted as separate ethnic groups in Ninglang for at least one hundred fifty years.[11] Each ethnic group has a distinct language, although the language of instruction in schools is Mandarin, a dialect of Han Chinese.

The Yi (and the Han Chinese traditionally) are patrilineal (see, for example, articles in Harrell (2001)), while the Mosuo are matrilineal (see, for example, Cai (2001); Shih (1993); Walsh (2001)). Patriliny and matriliny refers to "the gender direction of the transmission of associations, rights, and duties from one generation to the next (Harrell, 2002)."[12] While Yi households tend to be headed by a man, Mosuo households generally have two heads, one male and one female (Cai, 2001: 122-123). A unique feature of the Mosuo matrilineal society is their sexual visitation system called the "walking marriage," where the man travels to visit the woman in the evenings. It

[10]In modern China, each citizen is categorized into one of 56 official ethnic groups. Ethnic classification is displayed on the national identification card, along with other basic information such as gender and date of birth. Membership in an ethnic group can only pass from parent to child (or sometimes from grandparent to grandchild).

[11]For a more detailed discussion of the origins of the ethnic groups in Ninglang see Zhang (2012).

[12]In contrast, patriarchy and matriarchy refers to the gender which typically holds political power. Anthropologists have yet to find a single matriarchal society (Stone, 1997, p. 110-111).

normally does not involve cohabitation and is "nonexclusive, noncontractual, and nonobligatory (e.g. Shih, 1993)."

In the last 60 years of Communist Party rule in China, radical policies have promoted women's participation in the paid labor force for both ideological and economic reasons (Croll, 1983: 2; Wolf, 1985: 81; Yang, 1965).[13] Today, 67.4% of Chinese women over the age of 15 are in the labor force, a higher percentage than that in all the OECD countries, with the exception of Iceland (ILO, 2009).[14] Accompanying and supporting the labor policies is a national Marriage Law which weakened the traditional Han Chinese patrilineal family (Hershatter, 2004: 999). Minority ethnic groups, however, were granted exemptions from the legislation (Dreyer, 1976: 119), which allowed traditional cultural gender norms to be preserved, to some degree.

A short written survey was administered to the subjects after the experimental session, with the purpose of understanding what key socioeconomic or demographic factors may vary across ethnic groups, and also as a check on the anthropological evidence. The questions were written to correspond to the Chinese Census wherever possible to maximize clarity.[15] Table 1 shows selected background characteristics. As expected, the ethnic correlates correspond to the anthropological evidence, with the majority of the Mosuo having either a female head of household or one that is related to the subject maternally. The Mosuo are also the most likely to have parents participating in a walking marriage. The subjects are slightly older compared with students in the same grades in the United States, as is common in rural China, and for the most part are past the age of puberty. The Yi have the most children, although both the Mosuo and the Han Chinese have not restricted themselves to only one child, indicating that fertility policies are less restrictive in this area.[16] All three ethnic groups are predominantly agricultural and educational levels are low, with the subjects on average having more education than their heads of households. The Yi and Mosuo heads of households have lower educational attainment than their Han Chinese counterparts, possibly in part due to the language barrier. The subsequent analyses will include both demographic and socioeconomic controls.

[13]Other formerly planned economies shared in the same objectives (Therborn, 2004), which have resulted in high female labor force participation in these countries as well (ILO, 2009).

[14]In comparison, US female labor force participation is 58.4% (ILO, 2009).

[15]Nonetheless, verbal clarifications were required for some of the questions, such as the definition of brother and sister in a household, since often cousins lived together under the same roof.

[16]See Zhang (2012) for a detailed discussion of the fertility policy in this region.

4 Data collection

4.1 Administrative and educational outcome data

Scores from the most recent comprehensive exams taken by each subject were collected from each school and matched to the subjects by name. High school entrance exam records came from the county bureau of education. For all students who did not have an exam record, I confirmed with their middle school teachers that they had not taken the exam. In some cases there had been a typographical error or name change, and the "matchable" name was obtained from the teachers. In less than 5% of the cases, the student transferred out of the province and teachers did not know whether they took the entrance exam in their new province.

4.2 Experimental procedures

Experiments were conducted in two rounds – fall semester of 2008 and fall semester of 2009. In 2008, middle school subjects in grades 8 and 9 were recruited from two middle schools. In 2009, additional middle school subjects in grade 9 were recruited from one of the middle schools that participated in 2008. The middle schools are located in townships approximately 50 kilometers apart.

Subjects were randomly recruited from the schools' rosters within ethnicity and gender such that each session consisted of one ethnic group and was evenly divided across gender. Session size ranged from 8 to 24 subjects. There were a total of 16 sessions. All sessions took place during the school day, either during normal breaks, or during times that administrators deemed appropriate. Laboratories were set up in vacant classrooms, which were generally rooms designated for taking exams. Absentees were replaced by the first students on the randomized roster that matched on ethnicity and gender. All experimental instructions were read out loud by the author in Mandarin, which is the national language as well as the official language of instruction. Hard copies of the instructions were also distributed to everyone. Sessions lasted around an hour and a half. Experimental responses were either captured electronically using the Z-tree program, or subjects recorded their responses on paper and graders assessed these responses during each session. Scratch paper was provided in all sessions.

The experiment closely follows Niederle and Vesterlund (2007). The task used throughout the experiment was to add sets of five two-digit numbers and to do as many as possible in five minutes.[17] The number of problems correctly solved is the subject's "score" in the subsequent discussion. See below for a sample problem.

12	34	41	87	64	

The experiment consists of four stages, throughout which subjects were randomly seated in groups of four (two males and two females) and were not allowed communication although they could see one another.

Stage 1: Piece-rate - subjects are compensated RMB 0.5 for each problem solved.[18]

Stage 2: Compulsory tournament - subjects who solve the most problems in their group of 4 receive RMB 2 for each correctly solved problem, while the others receive no payment.[19]

Stage 3: Discretionary tournament - subjects first choose which of the two types of compensation scheme (piece-rate or tournament) they would like to apply to their performance in this stage. If they choose piece-rate, they are paid RMB 0.5 per problem solved. If they choose tournament, they receive RMB 2 per problem if they score highest in their group of four, and nothing if there is someone in their group who scores higher than they do.

Stage 4: Submit piece-rate performance - subjects choose which compensation scheme (piece-rate or tournament) they would like to apply their Stage 1 piece-rate performance. If they choose the piece-rate compensation scheme, they receive the same payment as they did in Stage 1. If they choose to submit their performance to a tournament, they are paid RMB 2 per problem if they scored the highest in Stage 1 in their group of four, and nothing if someone in their group scored higher than they did.

Following Niederle and Vesterlund (2007), if the subject chooses tournament in stage 3, their score is compared to the scores of the other three group members in stage 2 (the compulsory tournament stage), rather than their score in stage 3. This ensures that participants choosing the tournament option are competing against the scores of others also performing under the tournament payout conditions, and rules out reasons for choosing the piece-rate scheme such as not wanting

[17]For the experimental instructions, see the web appendix at ihome.ust.hk/~janezhang/.
[18]A meal at the school cafeteria costs RMB 3.5.
[19]In the case of a tie all those tied for highest score are paid 2 RMB per problem.

to impose negative externality on others or strategic response to beliefs about other participants' choices.

Subjects receive their scores from the previous stage before they begin the next stage. However, they do not know their relative ranking within their group. After the third stage, subjects are asked to guess their rank in the compulsory tournament. This information will be used in the analysis to assess the accuracy of their beliefs toward their relative performance.

Following standard experimental practice, one unpaid practice stage was administered before stage 1 to familiarize subjects with the task. At the end of the experiment, one of the stages was randomly chosen for payout, to minimize wealth effects across the stages. A written survey was distributed as students waited for their payment. The show up fee is RMB 2.

4.3 Tournament entry results

Raw rates of tournament entry from this study are shown in context with results from select experimental studies in Table 2. Middle school students exhibit very little gender difference, with point estimates between 2 - 6 percent. In contrast, the smallest gender difference from the other studies is 15 percent. These raw rates of tournament entry will be used in the subsequent analyses to derive individual competitive inclination.

4.4 Risk aversion data

Because payoffs are uncertain under the tournament payment scheme given performance, the decision to enter the tournament could depend on one's risk preferences. To elicit risk preferences, the Binswanger (1980) instrument is used. This instrument has been shown to generate less noise and more time-consistency, especially for subjects with low cognitive ability, as compared with finer but more complex risk instruments Dave et al. (2010). See Zhang (2012) for an extensive discussion. The risk instrument was presented at least one month after the competition experiment. The time gap serves to minimize wealth effects from earnings in the experiment, although such wealth effects have not been found to be important in prior research.[20]

The instrument consists of six lottery choices, from which subjects choose one. Each lottery is

[20] Holt and Laury (2002) found no such wealth effects between the 3rd and 4th rounds within the same session, even though the 3rd round had 50 times or 90 times the stakes of the 4th round.

a pair of payouts, each with 50% probability of occurring. Assuming a constant relative risk aversion utility function over income, each choice implies a range of values for the CRRA coefficient. See figure 1. Figure 2 presents a histogram of the CRRA coefficients, along with the associated lotteries.[21] Although there is a mass at risk neutrality, there is a non-trivial distribution in the risk averse range. The top row of Figure 3 shows the distribution of risk aversion across ethnicity, by gender. A Mann-Whitney test of the overall sample shows no statistically significant difference in the distributions of risk aversion by gender (p-values=0.19), in contrast to the literature on adult populations, where men are typically found to be less risk averse than women (see, for example, Croson and Gneezy (2009)).

Risk aversion data were collected in the second round of data collection, during the fall semester of 2009. Unfortunately, data could not be obtained for those who were no longer in school, mainly those in grade 9 in 2008 who did not continue on to high school. 40% of risk aversion data is missing for this reason. I employ a standard imputation method described in Cameron and Trivedi (2005) to impute the missing data.[22] The validity of this method relies on the assumption that the data are missing at random, meaning that the value of the missing variable is not correlated with the fact that it is missing. In this case, because we know that the missingness is due to being out of school, the missing at random assumption can be tested. Limiting the sample to middle school subjects recruited in 2009, the difference in the distribution of risk aversion between those who eventually enroll in high school and those who do not are insignificant (a two-sided Fisher's exact test yields a p-value of 0.934). The subsequent analyses will utilize the imputed measures of risk aversion. The substantive results are robust to assuming risk neutrality for everyone and also to using an alternative measure of risk aversion, discussed below. The bottom row in Figure 3 displays the imputed risk aversion distributions.

4.5 Overconfidence data

Using the guesses on the compulsory tournament rank, I construct a proxy for overconfidence by subtracting the guessed rank from the actual rank, with the best rank being 4 and the worst

[21] The distribution corresponds remarkably well to the distribution of comparable choices from the small stakes game in Binswanger (1980).

[22] I impute once using an ordered logit regression of choice in the instrument on gender-ethnicity main and interactions effects, age, and academic performance.

1. This proxy takes on integer values between -3 and 3, with positive values signifying overconfidence, negative values signifying underconfidence, and zero representing a correct guess. The subjects in my sample are on average correct about their rank, despite the fact the guesses were not incentivized monetarily (see Figure 4). In contrast, subjects in Niederle and Vesterlund (2007) were on average overconfident by more than one full rank, with men more overconfident compared with women, although women were also overconfident. Figure 5 breaks the sample up by ethnicity and gender. The average value of the proxy for overconfidence for each ethnic group is statistically indistinguishable from zero. This is contrary to studies conducted in the US and the UK. The fact that the current subjects are in the same schools and grades are public knowledge in China likely contributes to the accuracy of their guessed ranks, which naturally reduces overconfidence as a factor in the tournament entry decision. The difference in the distribution of overconfidence across gender, however, is statistically significant (Mann-Whitney p-value = 0.002), with boys more overconfident than girls.

5 Structural Model

This section presents a binary mixed logit discrete choice model of the tournament entry choice in the lab. Introducing individual heterogeneity in taste for competition allows for the estimation of individual parameters of competitive inclination, which will then be related to the subsequent real world decision to take the high school entrance exam.

The optimal tournament entry decision is modeled in a random utility framework with the observed portion of utility characterized by expected utility and the unobserved portion of utility characterized by $b(\cdot)$:

$$U_i = \sum_{j=1}^{J} p_{ij} V_i(I_{ij}) \cdot b(\alpha_j + \epsilon_{ij}) \tag{1}$$

where discrete alternatives are indexed by j, and preferences may vary across individuals, indexed by i. α_j is an alternative specific constant summarizing the effect on utility of all unobserved factors and ϵ_{ij} is an error term reflecting individual unobserved heterogeneity. In this formulation, $b(\cdot)$ can be interpreted as the behavioral portion of utility.[23]

[23] The behavioral portion of utility enters multiplicatively instead of additively for reasons of modeling convenience, as will become evident in the following discussion.

As in Holt and Laury (2002), preferences are characterized by constant relative risk aversion: $V = I^{1-\gamma}/(1-\gamma)$, where I is payout[24] and γ is the coefficient of constant relative risk aversion. I make the additional assumption that $V(0) = 0$. The CRRA functional form for utility was chosen for its tractability within the structural model. Let U_{it} and U_{ipr} be individual i's utility from choosing the tournament and piece rate options, respectively:[25]

$$U_{it} = \frac{(p_i q_i)(2\hat{s}_i)^{(1-\gamma_i)}}{1-\gamma_i} \cdot \exp(\bar{\alpha} + \epsilon_{it}) \tag{2}$$

$$U_{ipr} = \frac{(0.5\hat{s}_i)^{(1-\gamma_i)}}{1-\gamma_i} \cdot \exp(\epsilon_{ipr}) \tag{3}$$

where the tournament option is chosen iff $U_{it} > U_{ipr}$.

$p_i q_i \in (0,1)$ represents the subjective probability of winning the discretionary tournament, with p_i denoting the true probability of winning and q_i denoting the ratio of the subjective probability of winning to the objective probability of winning. $\hat{s}_i$, the scale of the stakes, is assumed to be anchored on the known compulsory tournament score.[26] γ_i is the individual coefficient of constant relative risk aversion collected using the Binswanger instrument.[27] $\bar{\alpha}$ represents the population average inclination or disinclination to compete over and above what can be accounted for by risk preferences and the subjective probability of winning. When q_i is assumed to be 1, $\bar{\alpha}$ represents the degree to which individuals deviate on average from expected utility maximizing behavior given the objective probability of winning, and captures all reasons for that deviation, including, for example, overconfidence.

The unobserved portion of utility, ϵ_t and ϵ_{pr}, are assumed to be distributed i.i.d. extreme value I, and represent the heterogeneity in individual taste for the tournament and piece-rate options, respectively. All other parameters of the model are also allowed to vary across individuals such

[24] The choice to use payout instead of wealth is supported by Binswanger (1980), which finds that risk aversion coefficients estimated using an experimental game in rural India did not vary with the wealth of the subjects, despite large variations in wealth. Additionally, one can appeal to the fact that since the subjects are rural middle school students, their wealth is approximately zero, so that the expected utility over payout model approximates the expected utility over wealth model.

[25] When $\gamma = 1$, the utility function is defined to be $\ln I$. However, actual values of γ in the present study sample obviates this consideration.

[26] What the value of this number is will not matter, as we will see shortly; the assumption that will have bite is that the scale of the stakes is known at the time of decision making.

[27] Following Binswanger (1980) and Binswanger (1981), I use the geometric mean of the endpoints of each range for the point estimate of γ.

that $\beta \sim f(\beta|\mu_\beta, \sigma_\beta)$, where $f(\cdot)$ is a mixing distribution with mean μ_β and variance σ_β^2.

In the specifications that are reported below, I assume that $q_i = 1$ and that $f(\cdot)$ is a normal mixing distribution. I also work with log utility which preserves utility ordering and has the advantage of being linear in parameters. Given the assumptions above, the probability that individual i chooses to enter the tournament is given by:

$$\Pr(\ln U_{it} > \ln U_{ipr}) = \int \frac{\exp(\bar{\alpha} + \beta_1 \ln p_i + \beta_2(1-\gamma_i))}{1+\exp(\bar{\alpha} + \beta_1 \ln p_i + \beta_2(1-\gamma_i))} f(\beta_1, \beta_2|\mu_\beta, \sigma_\beta) d\beta \tag{4}$$

where $\bar{\alpha}$, the average competitive inclination, is the alternative specific constant for choosing tournament. The individual heterogeneity in competitive inclination is implicitly captured by $\epsilon = \epsilon_t - \epsilon_{pr}$, which is distributed i.i.d. standard logistic. In the subsequent discussion I will refer to individual competitive inclination as $\alpha = \bar{\alpha} + \epsilon$.

Realized scores in the experiment were used to obtain the objective probability of winning, p_i. I assume that the potential score s_{pi} is known up to some noise around the realized score s_{ri}:

$$s_{pi} = s_{ri} + k\eta_i \tag{5}$$

where η_i is an i.i.d. extreme value type I noise term, and k is inversely proportional to the standard deviation of the noise. Then we arrive at the following closed-form expression for p_i, the true probability of winning in a group g:

$$p_i = \Pr(s_{pi} > s_{pj}, j \neq i) = \frac{\exp(ks_{ri})}{\sum_{j \in g} \exp(ks_{rj})} \tag{6}$$

Since winning in the discretionary tournament is defined by scoring higher than the other three group members did in the compulsory tournament stage, s_{ri} denotes the score in the discretionary tournament stage and s_{rj}, $j \neq i$, denotes the scores in the compulsory tournament stage. k, the non-linear scale parameter was estimated separately using a standard logit counterpart of Equation 4.[28] Note that the subjects are students in the same middle schools, where grades (as is common practice in China) are public knowledge and often posted along with student names in the

[28]Ten starting values of k corresponding to standard deviations of η_i ranging from 1 to 5.5 in increments of 0.5 were utilized, and all converged to the same maximum likelihood estimate of k, accurate to the thousandth place. Even so, the possibility that a local rather than a global maximum has been found cannot be completely ruled out.

classroom. If, instead, subjects were drawn from a large university and are virtually anonymous to each other, one could make the alternative assumption that subjects know the overall distribution and knowing their realized score would tell them their relative standing. p_i can then be fully proxied by one's own realized score s_{ri}.

Figure 6 shows the empirical cummulative distribution functions of performance in the compulsory tournament stage by gender for each ethnic group. Visual inspection shows some gender differences in performance but Mann-Whitney tests reveal no significant gender differences, at the 10% level or greater, across any of the ethnic groups. The lack of significant gender differences in the performance in the compulsory tournament is consistent with other studies using the math task (e.g. Niederle and Vesterlund, 2007). Figure 7 shows distributions of the probabilities of winning by gender for each ethnic group. The overlap by gender is substantial, and, again, Mann-Whitney tests reveal no significant gender differences, at the 10% level or greater, across any of the ethnic groups.

6 Results

6.1 Results from structural estimation

Estimation of Equation 4 was by the method of maximum simulated likelihood, using 500 draws of β from $f(\cdot)$ for each individual. The mixed logit parameter estimates are reported in Table 3. All estimates take the expected sign. Estimates of $\bar{\alpha}$ are positive and significant, indicating that middle school students on average have an inclination to compete above and beyond what can be accounted for through their objective probability of winning and their risk aversion.

6.2 Individual parameter of competitive inclination

An important benefit of the structural estimation approach in this study is that it allows for estimates of the individual competitive inclination, α. Taking an approach conceptually similar to Revelt and Train (2000), the conditional distribution of preference parameters for an individual who made the tournament entry choice y_i while facing independent variables x_i, $h(\alpha, \beta | y_i, x_i, \hat{\mu}_\beta, \hat{\sigma}_\beta)$

, can be simulated using an accept-reject procedure.[29] Specifically, 10,000 draws were simultaneously drawn from the distribution of ϵ and from $f(\beta|\hat{\mu}_\beta, \hat{\sigma}_\beta)$ for each individual. The draws that are consistent with the actual tournament entry choice as well as the values of the independent variables were retained ("accepted") while the rest where discarded. The first 1,000 retained draws for each individual comprise the simulated dataset of conditional distributions $h(\alpha, \beta|y_i, x_i, \hat{\mu}_\beta, \hat{\sigma}_\beta)$. Various statistics can be derived from these simulated distributions. In particular, I follow Revelt and Train (2000) to base the measure of individual competitive inclination on the conditional mean, which is simply the averages from the individual simulated conditional distributions. Although this measure would be more correctly notated by $\bar{\alpha}_i$, in the subsequent discussion I will drop the bar and simply refer to this measure as α_i in the interest of notational simplicity.[30]

Figure 8 plots a histogram of the individual α_is. The distribution is clearly bimodal. The left cluster represents those who chose piece-rate and the right cluster represents those who chose tournament. The shape of the distribution implies that competitive inclination is largely driven by the choice of tournament versus piece-rate.[31]

6.3 Linking lab and real world behavior

If we believe that the competition experiment widely used in the literature and replicated in this study reveals a stable inclination toward competition, and that the choice model has been correctly specified, then α_i should be able to predict future competitive behavior outside the lab. I test this hypothesis by linking the individual α_is to the subject's subsequent propensity to take a competitive high school entrance exam.

[29] The Revelt and Train (2000) method uses Bayes' rule to derive the conditional distribution of the slope coefficients, but it does not describe how to derive the conditional distribution of the alternative specific constant, which requires knowledge of the conditional distribution of the utility errors. The accept-reject procedure described here is more general: it is asymptotically equivalent to Revelt and Train (2000) for the random slope coefficients and yet also simulates the utility error terms. See Train (2003) for a discussion on accept-reject procedures.

[30] The entire conditional distribution of α_i for each individual can also be utilized, that is, the variation in α_i can be integrated out in the regression linking high school entrance exam taking and competitive inclination. Substantive results remain unchanged when using this approach. Regressions available upon request.

[31] Although the population distribution of competitive inclination was assumed to be logistically distributed, there is no theoretical reason to expect the distribution of the conditional means to take any particular shape.

6.3.1 Results

Table 4 reports results from probit regressions of the decision to take the entrance exam on α_i. To facilitate interpretation, the α_i have been normalized by its population standard deviation of $\sqrt{\frac{\pi^2}{3}}$. The first column is a simple probit regression of taking the entrance exam on α_i. Column 2 controls for prior academic performance and its quadratic. As expected, prior academic performance is significantly predictive of the propensity to take the entrance exam, although its quadratic is not. Even so, competitive inclination remains significant at the 5% level. A one standard deviation increase in the inclination to compete is associated with a 7.2 percentage point increase in the propensity to take the entrance exam. Given that around 20% of the population do not take the exam, a 7.2 percentage point increase in the exam participation rate implies a reduction of the non-participating population by 36%.

Columns 3-5 add controls for the observable background characteristics that are conventionally thought to be associated with educational continuation, namely ethnicity-gender main effects and interactions, demographics, and socioeconomic status (Hannum and Adams, 2009). By adding the covariates, I am able to examine whether competitive inclination measured experimentally remains predictive of taking the entrance exam after controlling for the more traditional explanatory variables. Results from columns 3-5 confirm that it is.

Column 6 enters α_i as an indicator variable that takes the value of 1 if the individual has an inclination to compete ($\alpha_i > 0$) and 0 if the individual has a disinclination to compete ($\alpha_i < 0$). The results indicate that, controlling for the full set of covariates, on average people who have an inclination to compete are 8 percentage points more likely to take the entrance exam than people who have a disinclination to compete.

6.3.2 Alternative explanation - ability

Given that the magnitude of the coefficient decreased from the first column to the second column in Table 4, we can infer that α_i is positively correlated with ability. A concern is that there remains unobserved ability as it relates to performance on the entrance exam that has not been fully controlled for, and that it may be picked up by α_i. Table 5 is an attempt to address that concern by regressing the entrance exam score on α_i, using the same set of controls as that in Table 4. Clearly,

α_i is not a significant predictor of performance on the entrance exam, at least for the 80% of subjects who took the exam. Unfortunately I cannot observe how those who did not take the exam would have performed, but this analysis may go some way to allay the concern that competitive inclination is purely a proxy for unobserved ability.

6.3.3 Robustness checks - non-structural approach

The structural approach assumes that tournament entry decisions are based on the objective probability of winning. However, if people are overconfident about their abilities, they may appear to be more competitively inclined in the experiment and also be more likely to take the entrance exam, even conditional on their prior academic performance. This would lead to spurious correlation between α_i and taking the entrance exam. In this section I test the robustness of the results linking lab and real world behavior by taking a generalized residual approach to estimating individual measures of competitive inclination. Robustness is tested both in the sense that this is a different econometric approach from the structural mixed logit model developed above, and also in the sense that I will be able to weigh the importance of alternative explanations for the association between the tournament entry decision and taking the high school entrance exam.

The generalized residual approach proceeds in two stages.[32] In the first stage, I estimate a standard probit model of the tournament entry decision y_i^1: $\Pr(y_i^1 = 1|x_i) = E(y_i^1|x_i) = \Phi(x_i'\beta)$ where Φ is the standard normal cdf and x_i is a full set of regressors including both predictors of the choice to enter the tournament and predictors of the propensity to take the entrance exam. From this regression I obtain the generalized residual, corrected for heteroskedasticity:

$$\hat{r}_i = \frac{y_i^1 - \hat{E}[y_i^1|x_i]}{\sqrt{(1-\hat{p}(y_i^1=1|x_i))\cdot(\hat{p}(y_i^1=1|x_i)}} = \frac{y_i^1 - \Phi(x_i'\hat{\beta})}{\sqrt{(1-\Phi(x_i'\hat{\beta}))\cdot\Phi(x_i'\hat{\beta})}} \quad (7)$$

$\hat{r}_i$ measures the residual taste for competition, after accounting for the full set of explanatory variables in x_i.[33]

In the second stage, I estimate a probit model of the decision to take the entrance exam y_i^2,

[32] See Cameron and Trivedi (2005) for a discussion of generalized residuals.

[33] Both α_i and $\hat{r}_i$ capture variation across individuals in the propensity to enter the tournament. But because $\hat{r}_i$, unlike α_i, is not a structural parameter, its sign does not indicate whether the subject has a preference or dispreference for competition.

adding $\hat{r}_i$ to the set of regressors in x_i: $\Pr(y_i^2 = 1|x_i, \hat{r}_i) = \Phi(x_i'\beta + \hat{r}_i\pi)$. Thus, π is a measure of the correlation between the unobserved factors contributing to entry into the tournament, which may be interpreted as laboratory competitive inclination, with the unobserved factors contributing to taking the high school entrance exam, which may be interpreted as real world competitive inclination.

Table 6 reports the results from the generalized residual approach. Panel A reports the first stage results and Panel B reports the second stage results. All specifications include the traditional explanatory variables for taking the entrance exam discussed above. In all three specifications of Panel B, the estimates of π are statistically significant at the 5% level, which is consistent with the interpretation of the structural estimation results. In column 1 the only experimental control is the probability of winning the tournament. The specifications in columns 2 and 3 progressively add risk aversion and overconfidence but the change in the point estimate and the standard error of π from column 1 to column 3 is minimal. This further indicates that the correlation between competitive behavior in the lab and competitive behavior in the real world is not due to risk preferences or overconfidence as measured in the lab.

To address the concern that risk preferences may be inaccurately measured or imputed, I test the robustness of the results to using an alternative measure of risk aversion in Table 7. Following Niederle and Vesterlund (2007), the Stage 4 decision of submitting the piece-rate performance to a tournament is used. This decision incorporates risk aversion along with perceived relative performance, which is factored out when controlling for the probability of winning and overconfidence.[34] Table 7 replicates the specifications of Table 6 using the Stage 4 decision in place of the risk aversion measure obtained using the Binswanger instrument. The results are very similar to those obtained in Table 6.

[34]Niederle and Vesterlund (2007) argues that this measure also incorporates feedback aversion. The drawback of using this measure in the structural analysis is that it cannot be readily translated into a coefficient of constant relative risk aversion.

6.4 Group differences in competitive inclination

6.4.1 Structural parameters

Figure 9 plots histograms of individual α_is by gender, for each ethnic group. Casual observation reveals substantial overlap by gender across male and female distributions in competitive inclination. Mann-Whitney tests were performed on each of the six sub-samples. There were no significant gender differences in any of the three ethnic groups at the 10% level or higher.

Another advantage of modeling the tournament entry decision structurally is that it defines a threshold for inclination and disinclination toward competition. This allows me to test whether people on average are competing too much or if they are avoiding competition. This is an open question in the literature on competitive inclination, which thus far has relied on reduced form methods. I find that of the 6 gender-ethnicity groups, all have positive average α_is. There is no evidence that any of the groups are avoiding competition on average.

6.4.2 Regression analysis

The reduced form estimation equation for testing gender differences across ethnic groups adapted from Equation 4 is as follows:

$$y_i = \beta_0 + \beta_1 male_i + \sum_{j=2}^{3} \beta_j ethn_{ji} \times male_i + \sum_{j=1}^{3} \beta_{j+3} ethn_{ji} + \delta \ln p_i + \tau(1-\gamma_i) + \lambda q_i + X_i' B + v_i \quad (8)$$

where $y_i = 1$ if the subject chooses tournament and 0 if the subject chooses piece-rate. p_i and γ_i are measured as above. q_i is the proxy for overconfidence. $ethn_{1i}$, $ethn_{2i}$, and $ethn_{3i}$ are indicator variables for Han Chinese, Mosuo and Yi, respectively. $male_i$ is an indicator variable taking the value of 1 for males and 0 for females. X_i is a vector of other controls. The coefficients of interest are $\beta_2(\beta_3)$ which indicates the gender difference in the Mosuo (Yi), as compared with the Han Chinese, the omitted group.

Table 8 reports the results from this estimation. The specifications in Columns 2 and 3 add demographic and socioeconomic status controls, respectively. There are no significant gender differences in tournament entry across the three specifications, consistent with the Mann-Whitney test results of the structural parameters.

The lack of a gender difference in tournament entry after controlling for performance in the math task, risk aversion, and overconfidence is consistent with findings from Sweden for adolescents grades 10-12 (Dreber et al., 2011), and in both Sweden and Columbia, for children aged 9-12 (Cárdenas et al., 2012). They contrast, however, with findings from Austria for children aged 9-18 (Sutter and Rützler, 2010), and from the UK for students in co-ed schools, just under the age of 15 (Booth and Nolen, 2012).

The results in the current study also contrast with findings among high school students in this region, where both the Mosuo and the Yi displayed significant gender differences in competitive inclination, although the high school Han Chinese did not (Zhang, 2012). There are at least two reasons the two sets of results are not directly comparable, however. The first is that middle school is compulsory but high school admission is based on an exam. Therefore, the high school students represent a selected sample of the population. Second, since high school students may come from any of the sixteen middle schools in the county, whereas the current study only draws students from two of the middle schools, to the extent that average characteristics of students differ from middle school to middle school, these differences may contribute to the differences in the two sets of results. Therefore, although the two sets of results are consistent with the interpretation that gender differences in competitive inclination emerges in the teenage years (at least for the minority ethnic groups), further research will be needed to isolate the effect of maturation, ideally using panel data or perhaps repeated cross sections from the same population.

7 Conclusion

Leavitt and List (2007) assert that the generalizability of findings from the lab to the world outside is perhaps the most fundamental question in experimental economics. This study, to my knowledge, provides the first direct evidence showing that experimental measures of competitive inclination are predictive of competitive behavior outside the lab. While some recent studies have shown gender differences in the willingness to enter a tournament to be sensitive to experimental design (see, for example, Shurchkov (2011); Healy and Pate (2011); Dargnies (2011)), the findings in this study indicate that competitive inclination measured from the pioneering experiment developed in Niederle and Vesterlund (2007) is meaningful for understanding real world competitive behavior, at least in the context of a competitive exam taking decision facing tens of millions of middle school students each year in China. These findings, therefore, support the continued use of the now standard Niederle and Vesterlund (2007) experimental design in studying competitive inclination, and perhaps encourage testing of the external validity of the variations on the experimental design in this literature.

No gender differences in competitive inclination were found in the three ethnic groups studied, consistent with the fact that no gender differences exist in exam taking, controlling for prior test scores. Where large gender differences in competitive inclination do exist, however, as they do in the United States, further research will be needed to assess whether the gender gap in laboratory competitive inclination can predict subsequent gender differences in educational or labor market outcomes. As a policy tool, the findings from this study suggest that the competition experiment may help educators to identify students that may need extra encouragement in taking advantage of opportunities that involve competition. Some of these opportunities, as we have seen, may have lifelong consequences.

References

2010. Millenium development goals database. United Nations Statistics Division.

Attali, Yigal, Zvika Neeman, and Analia Schlosser. 2011. Rise to the challenge or not give a damn: differential performance in high vs. low stakes tests. *Unpublished manuscript*.

Benz, Matthias, and Stephan Meier. 2008. Do People Behave in Experiments as in the Field? Evidence from Donations. *Experimental Economics* 11:268–281.

Bertrand, Marianne, and Kevin F. Hallock. 2001. The Gender Gap in Top Corporate Jobs. *Industrial and Labor Relations Review* 55:3–21.

Binswanger, Hans P. 1980. Attitudes toward Risk: Experimental Measurement in Rural India. *American Journal of Agricultural Economics* 62(3):396–407.

Booth, Alison, and Patrick Nolen. 2012. Choosing to Compete: How Different are Girls and Boys? *Journal of Economic Behavior and Organization* 81:542–555.

Cai, Hua. 2001. *A Society without Fathers or Husbands*. Zone Books.

Cameron, Colin, and Pravin Trivedi. 2005. *Microeconometrics*. Cambridge University Press.

Cárdenas, Juan-Camilo, Anna Dreber, Emma von Essen, and Eva Ranehill. 2012. Gender differences in competitiveness and risk taking: Comparing children in Colombia and Sweden. *Journal of Economic Behavior and Organization* 83:11–23.

Croll, Elisabeth. 1983. *Chinese Women Under Mao*. M.E. Sharpe.

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

Dargnies, Marie-Pierre. 2011. Men Too Sometimes Shy Away from Competition: The Case of Team Competition. *Management Science, forthcoming*.

Dave, C., C. C Eckel, C. A Johnson, and C. Rojas. 2010. Eliciting risk preferences: When is simple better? *Journal of Risk and Uncertainty* 41(3):219–243.

De Brauw, A., and S. Rozelle. 2008. Reconciling the Returns to Education in Off-Farm Wage Employment in Rural China. *Review of Development Economics* 12(1):57–71.

Dreber, Anna, Emma von Essen, and Eva Ranehill. 2011. In Bloom: Gender Differences in Preferences among Adolescents. *Unpublished manuscript*.

Dreyer, June T. 1976. *China's Forty Millions*. Harvard University Press.

Falk, A., and E. Fehr. 2003. Why Labour Market Experiments? *Labour Economics* 10(4):399–406.

Finan, Frederico, and Laura Schechter. 2010. Vote-Buying and Reciprocity. *SSRN Working Paper Series*.

Flory, Jeffrey, Andreas Leibbrandt, and John List. 2012. Do Competitive Work Places Deter Female Workers? A Large-Scale Natural Field Experiment on Gender Differences in Job-Entry Decisions. *NBER Working Paper No.16546*.

Fortin, Nicole. 2005. Gender Role Attitudes and Women's Labour Market Outcomes Across OECD Countries. *Oxford Review of Economic Policy* 21:416–438.

Gneezy, U., K.L. Leonard, and J.A. List. 2009. Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica* 77(5):1637–1664.

Gneezy, U., and A. Rustichini. 2004. Gender and competition at a young age. *American Economic Review* 94(2):377–381.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. Performance in Competitive Environments: Gender Differences. *Quarterly Journal of Economics* 118(3):1049–1074.

Gong, Binglin, and Chun-Lei Yang. 2012. Gender differences in risk attitudes: Field experiments on the matrilineal Mosuo and the patriarchal Yi. *Journal of Economic Behavior and Organization* 83:59–65.

Hannum, Emily, and Jennifer Adams. 2009. Beyond Cost: Rural Perspectives on Barriers to Education. In *Creating Wealth and Poverty in China*, ed. Deborah Davis and Wang Feng, –. Stanford University Press.

Harrell, Stevan, ed. 2001. *Perspectives on the Yi of Southwest China*. Berkeley: University of California Press.

Harrell, Stevan. 2002. Patriliny, Patriarchy, Patrimony: Surface Features and Deep Structures in the Chinese Family System. Draft.

Healy, Andrew, and Jennifer Pate. 2011. Can Teams Help to Close the Gender Competition Gap? *The Economic Journal* 121:1192–1204.

Hershatter, Gail. 2004. State of the fileld: women in China's long twentieth century. *Journal of Asian Studies* 63(4):991–1065.

Holt, Charles A., and Susan K. Laury. 2002. Risk Aversion and Incentive Effects. *The American Economic Review* 92(5):1644–1655.

International Labour Organization (ILO). 1980-2009. Key Indicators of the Labour Market.

Karlan, Dean. 2005. Using Experimental Economics to Measure Social Capital and Predict Financial Decisions. *The American Economic Review* 95:1688–1699.

Lavy, Victor. 2008. Gender Differences in Market Competitiveness in a Real Workplace: Evidence from Performance-based Pay Tournaments among Teachers. *NBER Working Paper No. 14338*.

Leavitt, Steven, and John A. List. 2007. What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World? *The Journal of Economic Perspectives* 21:153–174.

Manning, Alan, and Joanna Swaffield. 2008. The gender gap in early-career wage growth. *Economic Journal* 118:983–1024.

Ministry of Education. 1986. People's Republic of China Law on Compulsory Education. Beijing: Ministry of Education.

Ministry of Education. 2007. Normal Junior Secondary School Enrollment. Beijing: Ministry of Education.

Ministry of Education. 2008. Normal Senior Secondary School Enrollment. Beijing: Ministry of Education.

National Bureau of Statistics. 2000. China Census. Beijing.

———. 2008. China County Statistics. Beijing.

Niederle, Muriel, Carmit Segal, and Lise Vesterlund. 2010. How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Working Paper, Stanford University*.

Niederle, Muriel, and Lise Vesterlund. 2007. Do Women Shy Away Fom Competition? Do Men Compete Too Much? *Quarterly Journal of Economics* 122:1067–1101.

Ors, Evren, Frederic Palomino, and Eolic Peyrache. 2008. Performance Gender-Gap: Does Competition Matter? *Working Paper, HEC, Paris*.

Paserman, D. 2010. Gender differences in performance in competitive environments: evidence from professional tennis players. *Working Paper, Boston University*.

Revelt, D., and K. Train. 2000. *Customer-specific Taste Parameters and Mixed Logit: Households' Choice of Electricity Suppliers*. Citeseer.

Shih, C.K. 1993. The Yongning Moso : Sexual Union, Household Organization, Gender and Ethnicity in a Matrilineal Duolocal Society in Southwest China. Ph.D. thesis, Stanford University, Stanford, California.

Shurchkov, Olga. 2011. Under Pressure: Gender Differences in Output Quality and Quantity under Competition and Time Constraints. *Journal of the European Economic Association*.

Sutter, Matthias, and Daniela Rützler. 2010. Gender differences in competition emerge early in life. *IZA Discussion Paper 5015*.

Therborn, G. 2004. *Between Sex and Power: Family in the World, 1900-2000*. Routledge.

Tierney, John. 2005. What Women Want. *The New York Times*.

Train, Kenneth. 2003. *Discrete Choice Methods with Simulation*. Cambridge Univ Press.

Varian, Hal R. 2006. The Difference Between Men and Women, Revisited: It's About Competition. *The New York Times*.

Walsh, Eileen. 2001. The Mosuo : Beyond the Myths of Matriarchy : Gender Transformation and Economic Development. Ph.D. thesis, Temple University.

Wolf, Margery. 1985. *Revolution postponed: Women in contemporary china*. Stanford University Press.

Wolfers, Justin. 2006. Diagnosing discrimination: Stock returns and CEO gender. *Journal of the European Economic Association* 4:531–541.

Yang, CK. 1965. *Chinese Communist Society: The Family and The Village*. The MIT Press.

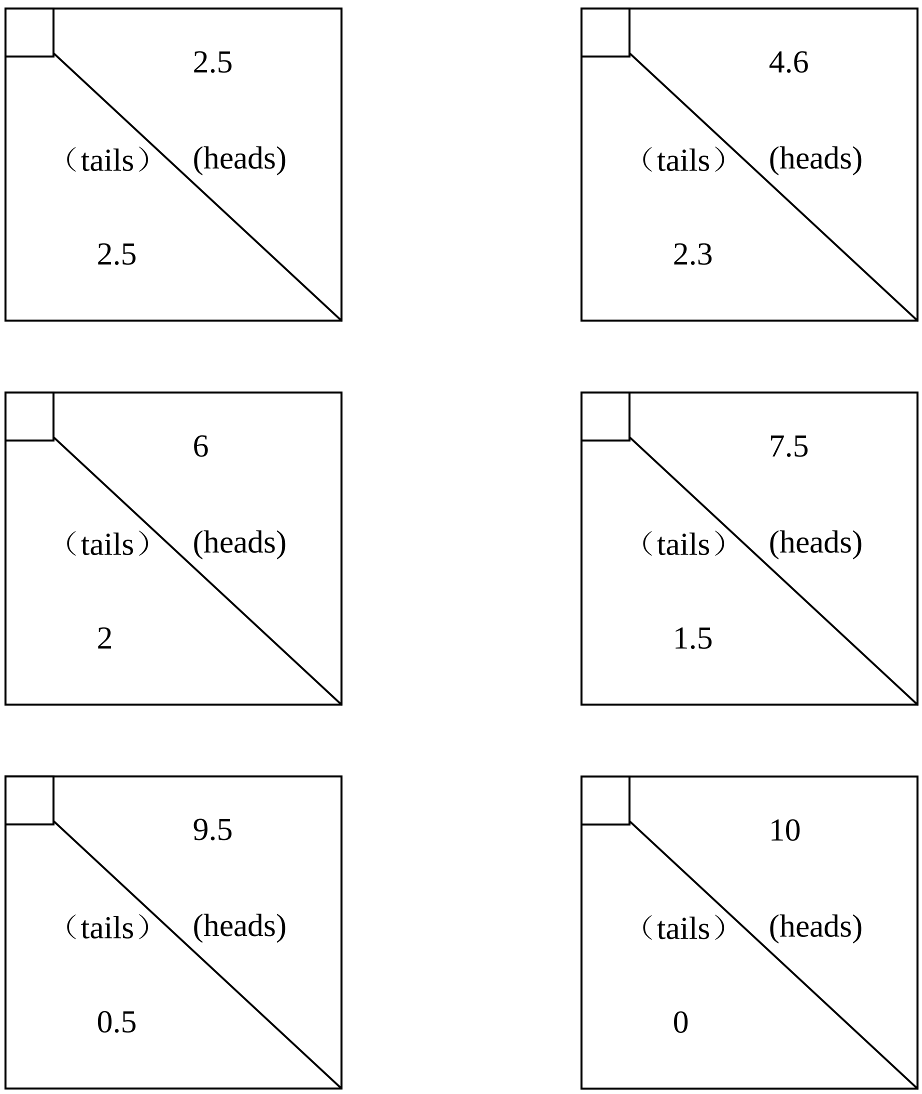
Zhang, Y. Jane. 2012. Narrowing the Gender Gap in Competitive Inclination: Evidence from The Communist Experiment in China. *Unpublished manuscript*.

Figure 1: Binswanger Risk Instrument

Student ID __________________

In this game there are six ways to win money, represented by the 6 pictures below. In each picture, there are two amounts. You may choose one of the six pictures. When you've made your choice, we will determine your payout by a coin toss. If the coin lands on heads, you will receive the amount in the upper right half of your chosen picture; if the coin lands on tails, you will receive the amount in the lower left. As we know, the probability of a coin landing on heads and the probability of it landing on tails is each 50%.

Please make your choice by marking the upper left corner of the picture with a "✓". Please let us know if you have any questions.



(All amounts are in Chinese RMB)

Figure 2: Distribution of CRRA Coefficients with Associated Lottery

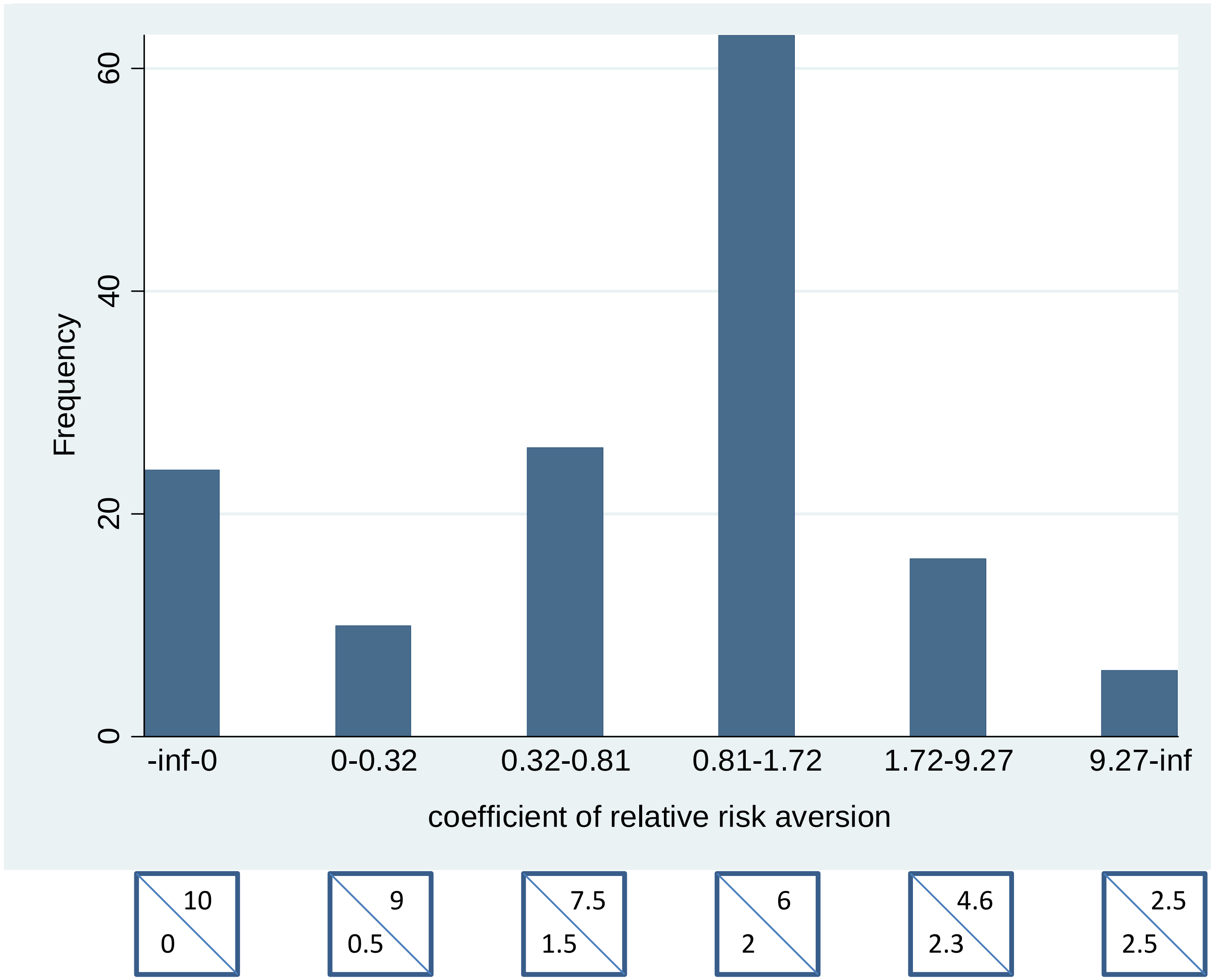


Figure 3: Distribution of Actual and Imputed CRRA Coefficients across Ethnicity, by Gender

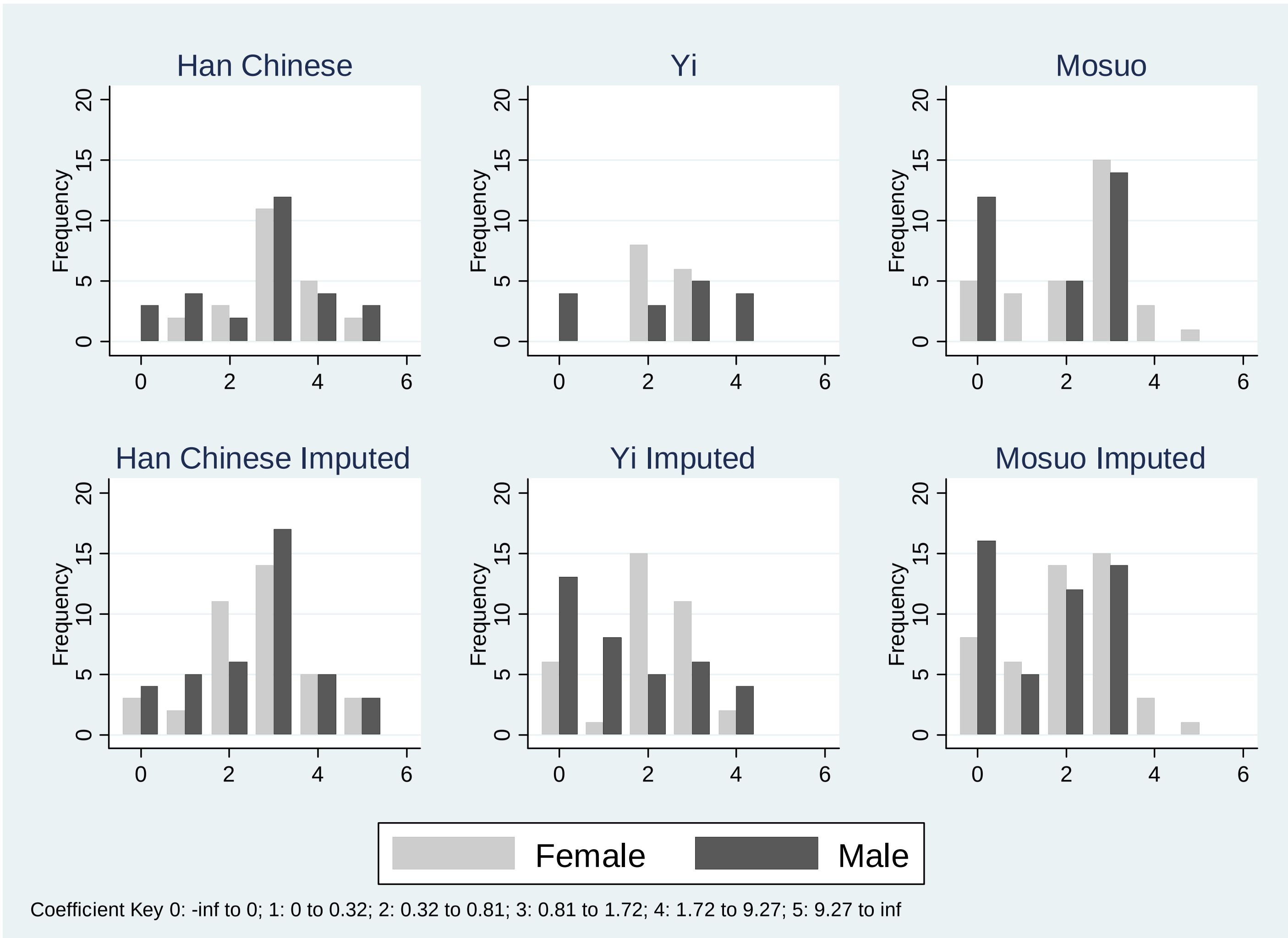


Figure 4: Distribution of the Proxy for Overconfidence

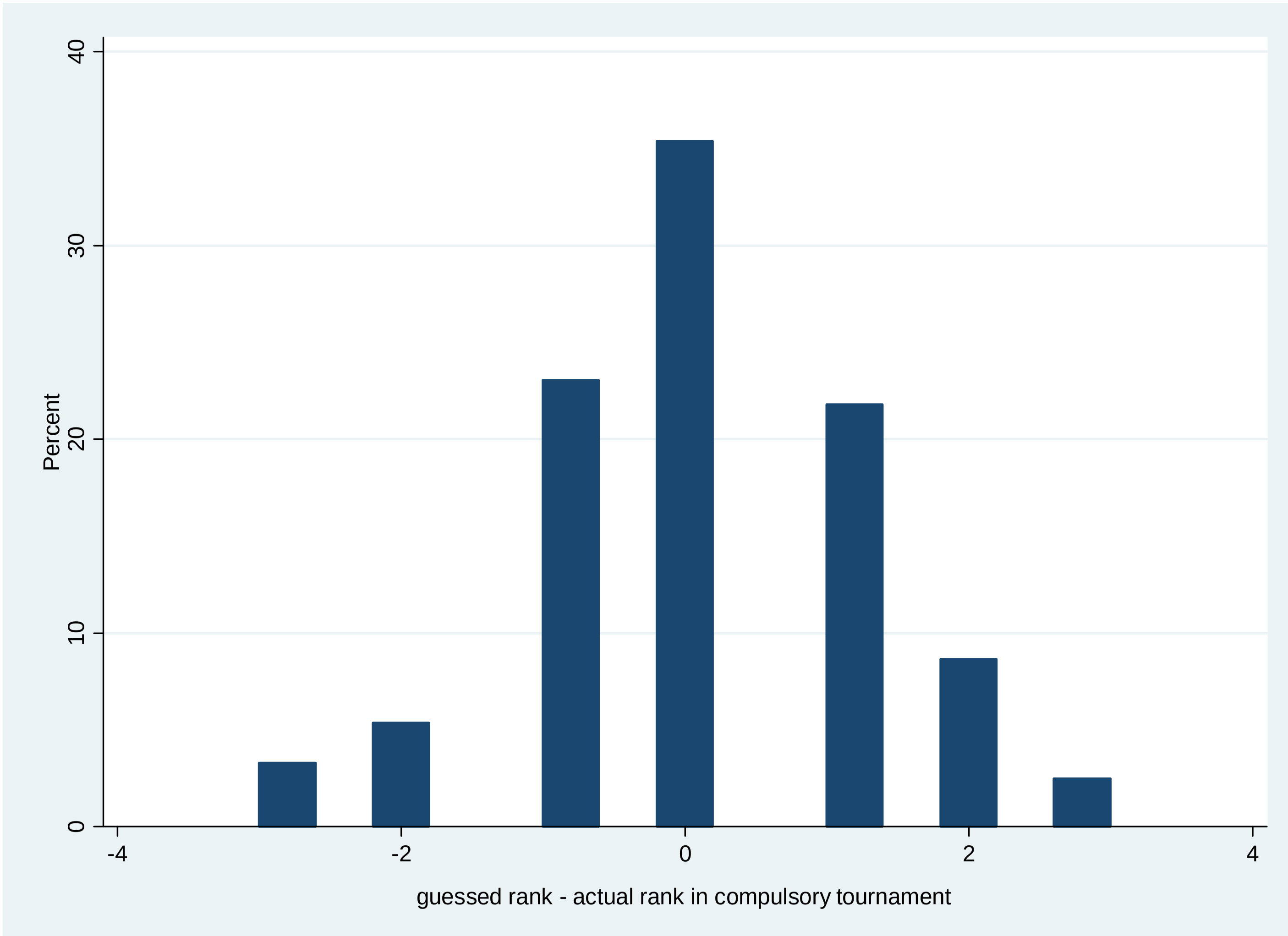


Figure 5: Distribution of the Proxy for Overconfidence across Ethnicity, by Gender

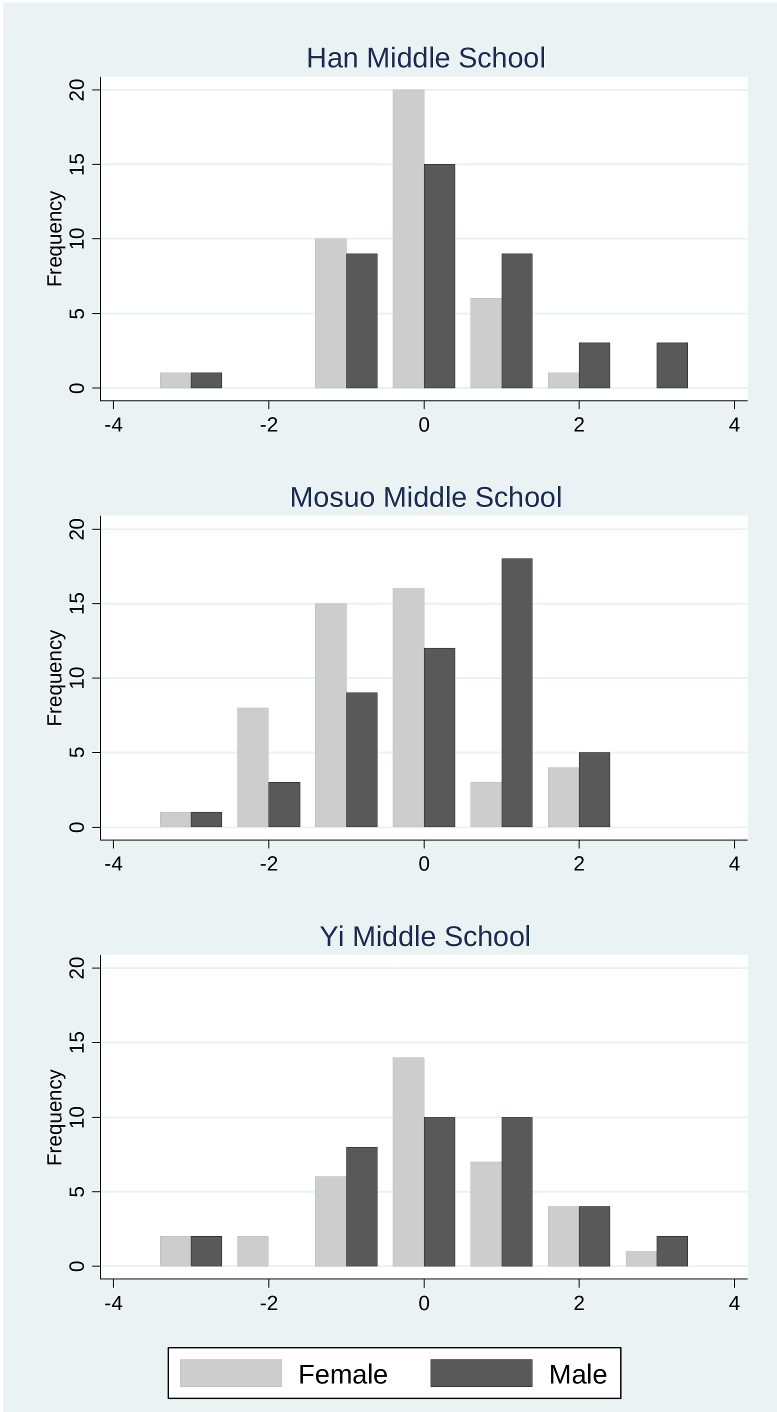


Figure 6: Distribution of the Probability of Winning the Discretionary Tournament across Ethnicity in Middle school and High school, by Gender

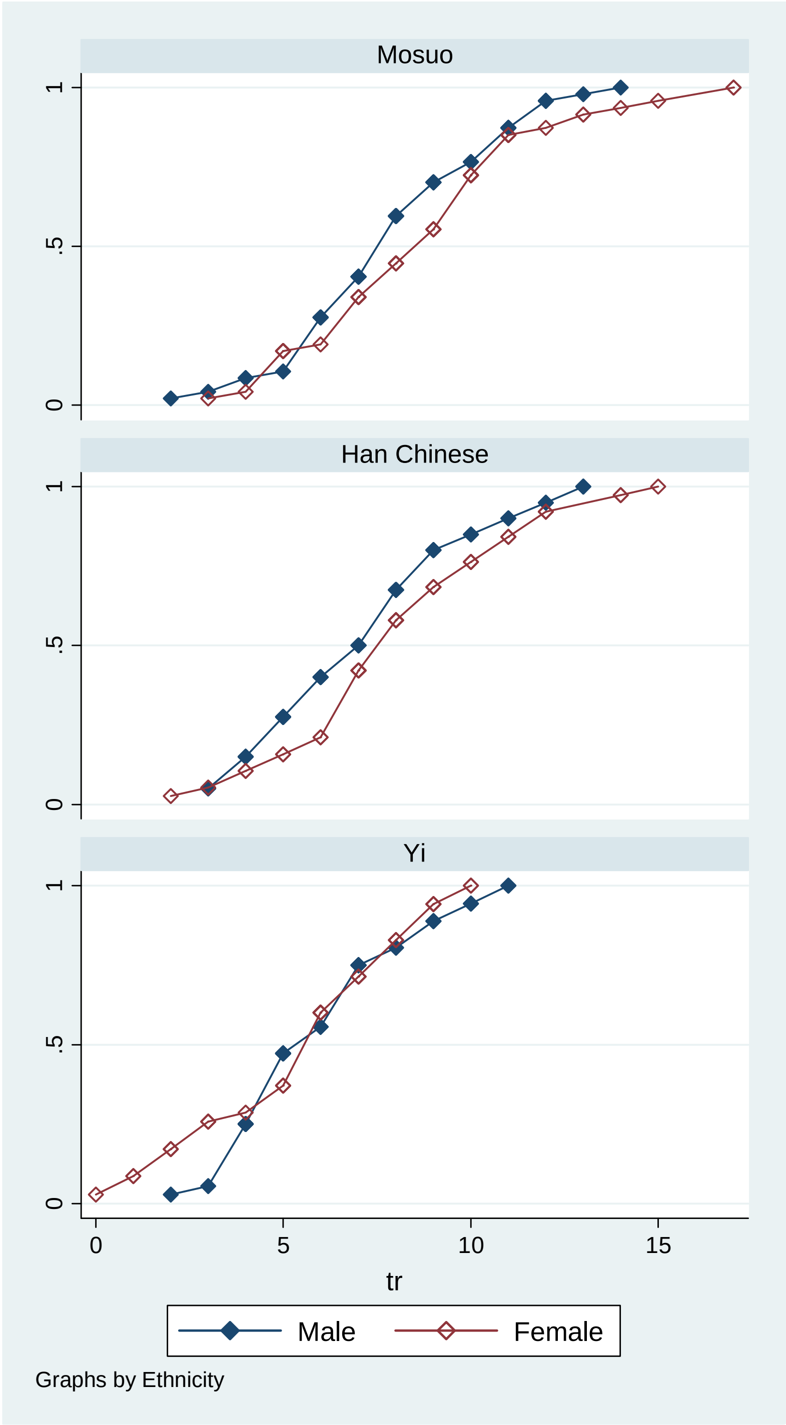


Figure 7: Distribution of the Probability of Winning the Discretionary Tournament across Ethnicity, by Gender

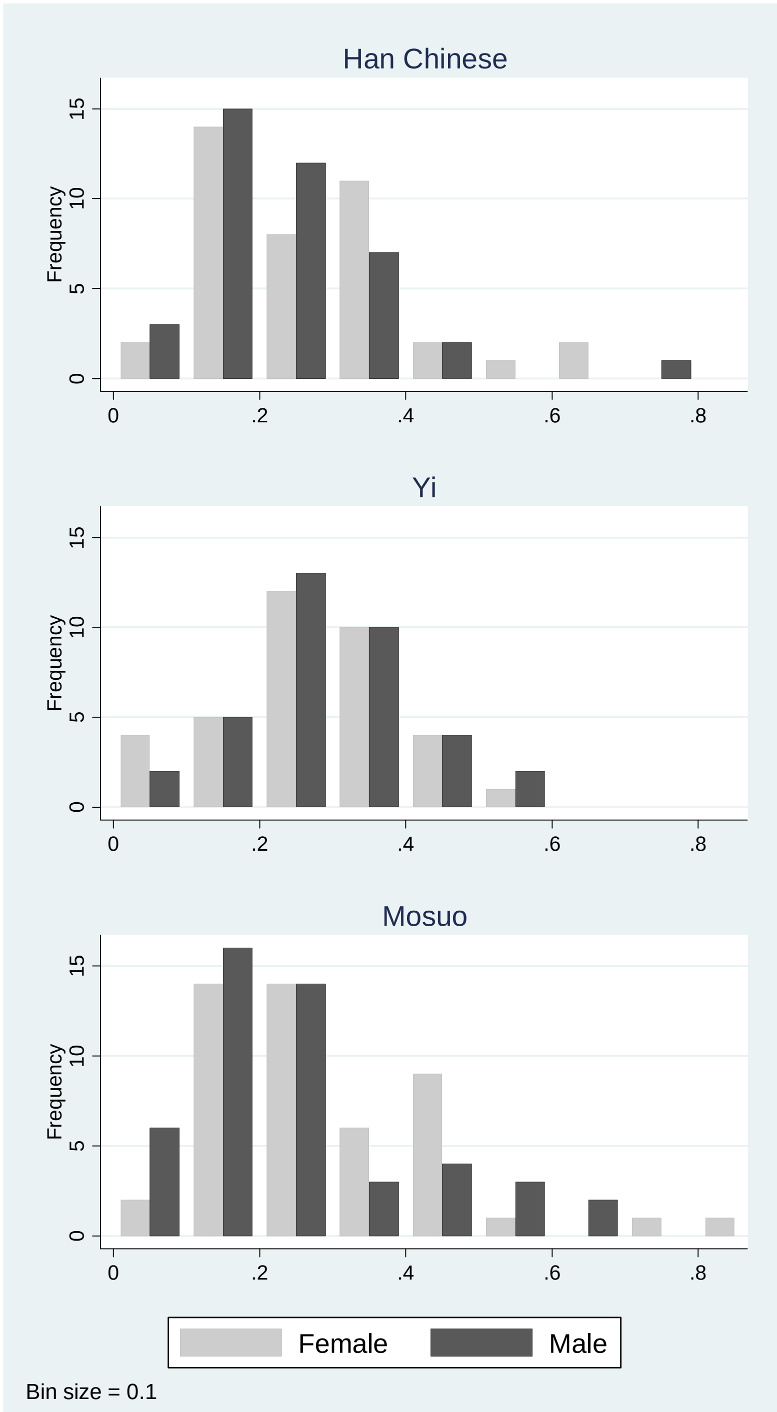


Figure 8: Distribution of Alpha

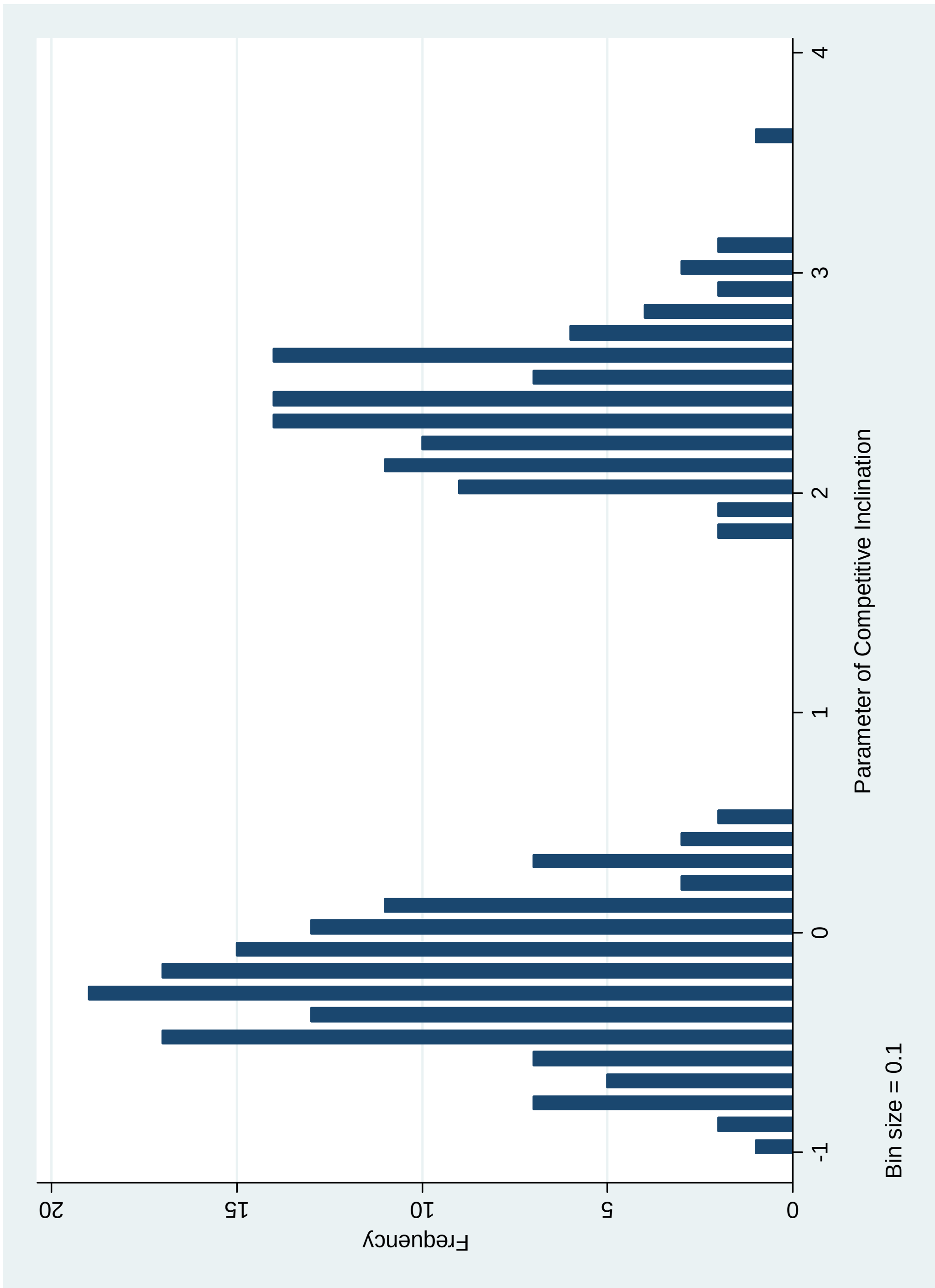


Figure 9: Distribution of Alpha across Ethnicity, by Gender

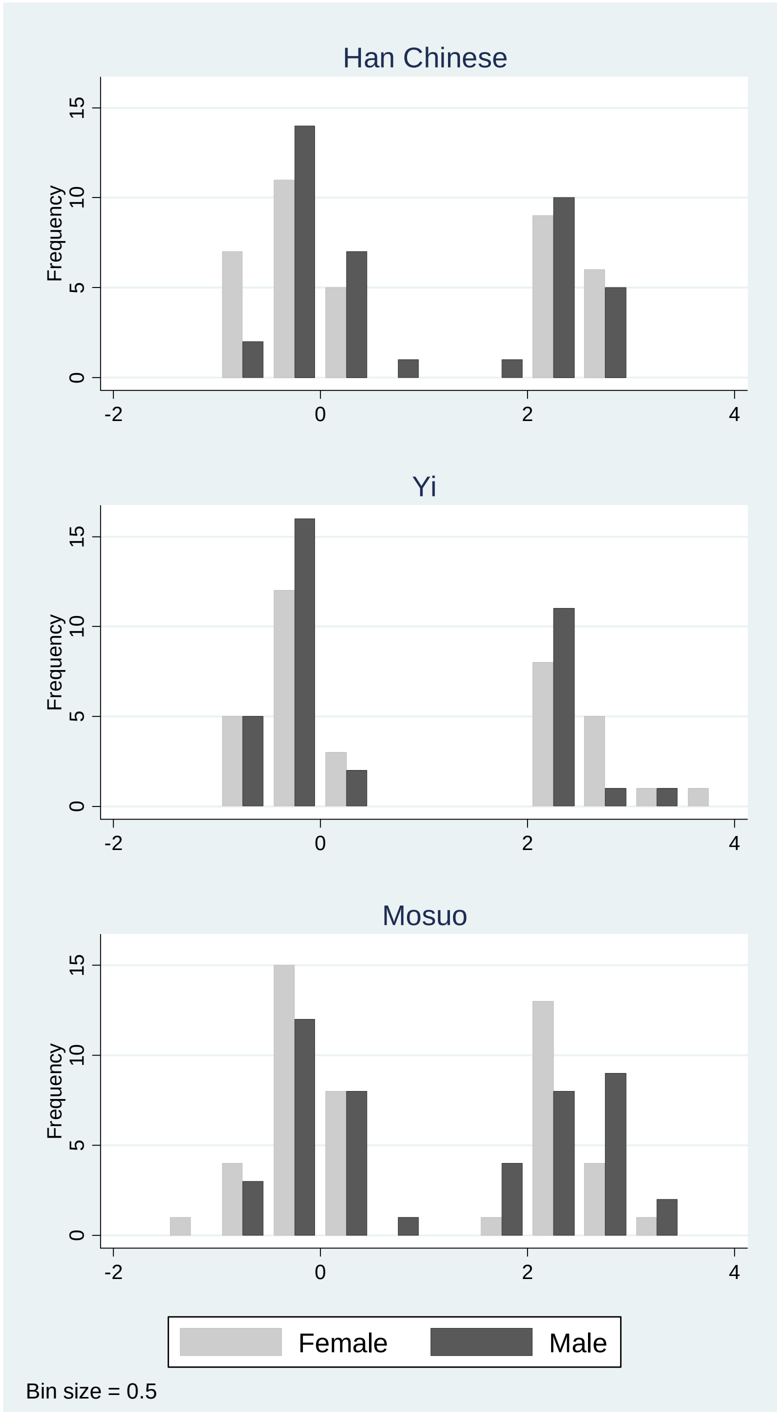


Table 1: Selected Descriptive Statistics (Means; Standard Deviation in Parentheses)

	Han Chinese	Yi Patrilineal	Mosuo Matrilineal	F-stat	p-value
Ethnicity correlates					
sex of h.h. (1 = male)	0.94	0.91	0.53	31.829	0.000
	(0.25)	(0.29)	(0.50)		
head of household is male but related maternally	0.00	0.03	0.09	23.198	0.000
	(0.00)	(0.17)	(0.29)		
parents participating in walking marriage	0.06	0.08	0.46	31.620	0.000
	(0.24)	(0.28)	(0.50)		
Demographics					
age	15.67	15.50	15.84	1.986	0.139
	(0.98)	(1.17)	(1.08)		
# brothers	0.91	1.54	0.94	8.656	0.000
	(0.88)	(0.98)	(1.19)		
# sisters	1.01	1.43	0.98	3.615	0.028
	(0.97)	(1.10)	(1.26)		
Socioeconomic Status					
h.h. engaged in agriculture	0.99	0.97	0.87	5.950	0.003
	(0.11)	(0.17)	(0.34)		
h.h. educational attainment (years)	6.97	4.46	5.10	9.647	0.000
	(3.04)	(3.83)	(4.00)		
Observations	80	72	96		

Table 2: Raw Tournament Entry Statistics

	Female	Male	Sex Difference	p-value	Obs (Female)	Obs (Male)
Tournament Entry						
Middle school (avg age 15.7)						
Han (Chinese)	38%	40%	2%	0.890	39	40
Mosuo (Chinese Matrilineal)	42%	48%	6%	0.543	48	48
Yi (Chinese Patrilineal)	42%	36%	-6%	0.635	36	36
Total					123	124
High school (avg age 18.4)						
Zhang (2012)						
Han (Chinese)	48%	63%	15%	0.154	48	48
Mosuo (Chinese Matrilineal)	50%	73%	23%	0.015	52	52
Yi (Chinese Patrilineal)	38%	60%	22%	0.025	48	48
Total					148	148
University undergraduates						
Niederle & Vesterlund (2007)						
United States	35%	73%	38%	0.002	40	40
Adults (avg age 33.9)						
Gneezy, Leonard, and List (2009)						
Maasai (Tanzanian Patrilineal)	26%	50%	24%	0.040	34	40
Khasi (Indian Matrilineal)	54%	39%	-15%	0.201	52	28
Total					86	68

Table 3: Mixed Logit Estimates

Variable	
α: paramenter of competitive inclination	
mean	0.885**
	(0.418)
log prob of winning tournament	
mean	0.859***
	(0.249)
std dev	0.013
	(0.017)
1-risk aversion	
mean	0.097
	(0.197)
std dev	0.241
	(0.373)
Choice situations	243
Log likelihood	-158.4

Robust standard errors in parentheses, clustered by session

*** $p<0.01$, ** $p<0.05$, * $p<0.10$

Table 4: Structural Approach to Linking Lab and Real World Behavior

Dependent Variable: Take Entrance Exam

	(1)	(2)	(3)	(4)	(5)	(6)
α_i^1	0.104***	0.072**	0.074**	0.063**	0.066**	
	(0.034)	(0.029)	(0.029)	(0.027)	(0.028)	
$\alpha_i > 0^1$						0.079**
						(0.036)
regular grades (percentile)		1.154***	1.092***	1.100***	1.073***	1.071***
		(0.436)	(0.399)	(0.374)	(0.386)	(0.399)
regular grades ^{2} (percentile)		-0.726	-0.625	-0.655	-0.611	-0.596
		(0.532)	(0.473)	(0.452)	(0.468)	(0.483)
ethnicity*gender controls	No	No	Yes	Yes	Yes	Yes
Demographic controls	No	No	No	Yes	Yes	Yes
SES controls	No	No	No	No	Yes	Yes
Observations	231	231	231	226	221	221
Log likelihood	-114.1	-98.97	-96.66	-91.55	-88.70	-89.15
Mean dep var	0.792	0.792	0.792	0.792	0.787	0.787

*** $p<0.01$, ** $p<0.05$, * $p<0.10$

Marginal effects; robust standard errors in parentheses, clustered by session.

Probit regression. Dependent variable = 1 if subject has record of taking the high school entrance exam, 0 if subject was known to have not taken the high school entrance exam.

Demographic controls include age, age squared, # brothers, # sisters

SES controls include household engaged in agriculture, education of household head.

[1] α_i is normalized by the estimated population standard deviation.

Table 5: Regression of Entrance Exam Score on Alpha

Dependent Variable: Score on high school entrance exam

	(1)	(2)	(3)	(4)	(5)	(6)
α_i[1]	16.766	6.255	4.626	4.770	3.496	
	(9.839)	(5.276)	(4.659)	(5.378)	(5.256)	
$\alpha_i > 0$[1]						-2.541
						(7.550)
regular grades (percentile)		304.473***	296.170***	284.407***	291.107***	292.821***
		(61.978)	(54.954)	(53.807)	(57.087)	(57.707)
regular grades ^{2} (percentile)		-61.048	-43.727	-25.891	-39.565	-39.691
		(62.154)	(52.688)	(51.585)	(55.515)	(55.269)
ethnicity*gender controls	No	No	Yes	Yes	Yes	Yes
Demographic controls	No	No	No	Yes	Yes	Yes
SES controls	No	No	No	No	Yes	Yes
Observations	181	181	181	177	172	172
Log likelihood	-1035	-948.9	-943.7	-921.1	-894.2	-894.4

*** $p<0.01$, ** $p<0.05$, * $p<0.10$

Marginal effects; robust standard errors in parentheses, clustered by session.

Linear regression.

Demographic controls include age, age squared, # brothers, # sisters

SES controls include household engaged in agriculture, education of household head.

[1]α_i is normalized by the estimated population standard deviation.

Table 6: Generalized Residual Approach to Linking Lab and Real World Behavior

	(1)	(2)	(3)
Panel A			
Dependent Variable: Choose the Tournament			
log prob of winning tourn.	0.193***	0.193***	0.218***
	(0.054)	(0.054)	(0.060)
1-risk aversion		0.008	0.008
		(0.018)	(0.018)
overconfidence			0.032
			(0.028)
regular grades (percentile)	0.189	0.203	0.167
	(0.457)	(0.458)	(0.466)
regular grades ^{2} (percentile)	0.244	0.233	0.263
	(0.530)	(0.527)	(0.527)
Observations	224	224	224
Panel B			
Dependent Variable: Take Entrance Exam			
Generalized residual (from Panel A)	0.045**	0.045**	0.044**
	(0.022)	(0.022)	(0.022)
log prob of winning tourn.	-0.150**	-0.137**	-0.122*
	(0.059)	(0.055)	(0.063)
1-risk aversion		-0.031**	-0.031**
		(0.015)	(0.015)
overconfidence			0.022
			(0.022)
regular grades (percentile)	1.338***	1.275***	1.240***
	(0.426)	(0.411)	(0.424)
regular grades ^{2} (percentile)	-0.790*	-0.750*	-0.718
	(0.472)	(0.447)	(0.459)
Observations	221	221	221

*** p<0.01, ** p<0.05, * p<0.10

Marginal effects; robust standard errors in parentheses, clustered by session.

Probit regressions. All specifications include controls for ethnicity-gender main effects and interaction effects; demographic controls for age, age squared, # brothers, # sisters; SES controls for whether the household is engaged in agriculture, education of household head.

Table 7: Generalized Residual Approach to Linking Lab and Real World Behavior - Alternative Risk Preference Measure

	(1)	(2)	(3)
Panel A			
Dependent Variable: Choose the Tournament			
log prob of winning tourn.	0.193***	0.159***	0.177***
	(0.054)	(0.052)	(0.060)
submit piece-rate		0.216***	0.210***
		(0.070)	(0.073)
overconfidence			0.021
			(0.031)
regular grades (percentile)	0.189	0.119	0.095
	(0.457)	(0.447)	(0.450)
regular grades ^{2} (percentile)	0.244	0.335	0.354
	(0.530)	(0.522)	(0.518)
Observations	224	224	224
Panel B			
Dependent Variable: Take Entrance Exam			
Generalized residual (from Panel A)	0.045**	0.046**	0.045**
	(0.022)	(0.022)	(0.022)
log prob of winning tourn.	-0.150**	- 0.139***	-0.122**
	(0.059)	(0.052)	(0.058)
submit piece-rate		-0.034	-0.040
		(0.062)	(0.061)
overconfidence			0.023
			(0.023)
regular grades (percentile)	1.338***	1.275***	1.240***
	(0.426)	(0.411)	(0.424)
regular grades ^{2} (percentile)	-0.790*	-0.750*	-0.718
	(0.472)	(0.447)	(0.459)
Observations	221	221	221

*** $p<0.01$, ** $p<0.05$, * $p<0.10$

Marginal effects; robust standard errors in parentheses, clustered by session.

Probit regressions. All specifications include controls for ethnicity-gender main effects and interaction effects; demographic controls for age, age squared, # brothers, # sisters; SES controls for whether the household is engaged in agriculture, education of household head.
Submit piece-rate = 1 if subject chooses to enter their piece-rate performance into a tournament in Stage 4, and = 0 otherwise.

Table 8: Regressions of Tournament Entry, Pooled across Ethnicity

Dependent Variable: Choose the Tournament

	Middle School		
	(1)	(2)	(3)
male	0.025	0.019	-0.001
	(0.154)	(0.150)	(0.143)
male*Mosuo	0.090	0.046	0.059
	(0.183)	(0.176)	(0.171)
Mosuo	-0.038	-0.016	0.039
	(0.134)	(0.132)	(0.142)
male*Yi	-0.137	-0.190	-0.152
	(0.214)	(0.191)	(0.203)
Yi	0.100	0.079	0.086
	(0.118)	(0.123)	(0.118)
log prob of winning tourn.	0.239***	0.264***	0.242***
	(0.062)	(0.060)	(0.055)
1-risk aversion	0.009	0.007	0.005
	(0.017)	(0.017)	(0.017)
overconfidence	0.029	0.035	0.030
	(0.024)	(0.024)	(0.023)
age		0.001	0.006
		(0.484)	(0.486)
age^2		0.002	0.002
		(0.015)	(0.015)
# sisters		-0.028	-0.027
		(0.019)	(0.018)
# brothers		0.043	0.040
		(0.034)	(0.036)
Household head agricultural			0.212**
			(0.092)
Education of household head			0.006
			(0.009)
School fixed effect	-0.102	-0.074	-0.070
	(0.125)	(0.125)	(0.130)
Observations	243	238	232
Log likelihood	-155.9	-148.0	-144.1
Mean dep var	0.416	0.412	0.418

*** p<0.01, ** p<0.05, * p<0.10

Marginal effects; robust standard errors in parentheses, clustered by session.

Probit regression: dependent variable = 1 if subject chooses to enter competition, 0 otherwise.