Internet Appendix for "Overconfidence and Preferences for Competition"

Ernesto Reuben, Paola Sapienza, and Luigi Zingales[1]

Section I presents a detailed analysis of whether MBAs who consented to the study in 2006 vary systematically from MBAs who did not. Section II presents a similar analysis of whether MBAs who answered the survey in 2015 vary systematically from those who did not. Section III describes in detail the numerous robustness checks reported in the paper but not fully described there due to space constraints. Section IV describes the procedures used to conduct the experiment and survey, including a sample of the instructions used to elicit preferences for competition. Lastly, Section V describes the additional variables used in the robustness checks.

I. Selection into the sample in 2006

In this section, we evaluate whether the 409 participants who consented to the analysis of all their data (including their earnings) differ from the 129 participants who consented to the analysis of only some of their data. In the top part of Table AI, we present the means and standard deviations of variables related to preferences for competition plus the fraction of women. For each variable, the table also displays the *p*-value obtained when we test whether the two groups of participants are significantly different from each other. Specifically, we use simple *t*-tests for the continuous variables and χ^2 tests for categorical variables. In the bottom part of Table AI, we present the same information for the control variables we will use for the robustness checks in Section III. We describe these variables and how we collected them in Section V.

[1]Reuben Ernesto, Paola Sapienza, and Luigi Zingales, Internet Appendix for "Overconfidence and Preferences for Competition," Journal of Finance, [DOI STRING]. Please note: Wiley is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the authors of the article.

Table AI – Summary statistics depending on consenting to all parts of the study in 2006

Note: Means, standard deviations, and number of observations for variables of interest. The rightmost column displays *p*-values from tests of equality of distributions between people who consented to the analysis of all their data and those who did not (*t*-tests for ordinal variables and χ^2 tests for categorical variables).

	CONSENTED			DID NOT CONSENT			
	mean	s.d.	N	mean	s.d.	N	*p*-value
Competitive	0.52	0.50	409	0.51	0.50	123	0.867
Performance (rank in sums tasks)	2.48	0.77	409	2.52	0.77	123	0.649
Expected rank in sums tasks	2.24	0.77	409	2.31	0.82	123	0.430
Expected experimental earnings	73.84	57.63	409	66.74	50.96	123	0.190
Overconfidence	0.24	0.63	409	0.21	0.76	123	0.697
Risk aversion coefficient	4.74	4.41	409	3.87	4.79	123	0.074
Non-competitive tournament	0.41	0.49	409	0.44	0.50	123	0.513
Fraction of women	0.30	0.46	409	0.34	0.48	129	0.388
Additional control variables							
Age	28.22	2.44	409	28.93	2.72	129	0.009
Fraction non-white	0.55	0.50	409	0.64	0.48	129	0.062
Fraction US residents	0.77	0.42	409	0.74	0.44	129	0.584
Fraction married before MBA	0.26	0.44	409	0.22	0.41	129	0.362
Fraction religious	0.47	0.50	409	0.42	0.50	129	0.312
GMAT Quantitative percentile	81.91	12.81	406	80.84	16.06	129	0.489
GMAT Verbal percentile	88.02	11.45	406	85.31	12.75	129	0.033
GMAT Analytic percentile	71.91	21.75	383	68.7	22.63	112	0.184
GPA	3.33	0.34	391	3.18	0.42	99	0.002
CRT score	2.44	1.33	409	2.43	1.35	129	0.979
RMET score	0.75	0.10	409	0.74	0.10	129	0.469
Discount rate	0.05	0.04	376	0.05	0.04	108	0.718
Trust	0.38	0.30	409	0.34	0.29	123	0.183
Reciprocity	0.36	0.20	409	0.33	0.20	123	0.151
Cooperation	0.33	0.47	409	0.29	0.46	123	0.436
Survey overconfidence	0.90	4.56	391	1.92	4.42	99	0.044
Survey risk aversion (general)	6.44	1.89	409	6.57	2.19	129	0.541
Survey risk aversion (monetary)	1.49	1.01	409	1.62	0.95	129	0.182

By and large, we find no significant differences between the participants who fully consented to the study and those who did not. If we use an unadjusted significance threshold of 5%, we find a significant difference in four out of twenty-five variables (age, GMAT verbal percentile, GPA, and the survey measure of overconfidence). However, if we adjust *p*-values with the Benjamini-Hochberg method to account for multiple comparisons (Benjamini and Hochberg 1995), we find a significant difference in only one variable (GPA). Importantly for this paper, neither the fraction of women nor the fraction of participants who chose the tournament are significantly different. Moreover, if we test for each gender whether the fraction of individuals who chose the tournament differs between those who fully consented and those who did not, we do not find a statistically significant difference for men ($p = 0.794$) or women ($p = 0.704$).

Finally, to test whether preferences for competition differ between participants who fully consented to the study and those who did not, we run a probit regression with the participants' tournament choice as the dependent variable. In line with the regressions in Table II, as independent variables, we include the participants' gender, performance, overconfidence, and risk aversion coefficient. In addition, we also add a dummy variable equal to one for participants who did not fully consent. We find that the estimated marginal effect of the dummy variable is minimal (0.003) and is not statistically significant ($p = 0.957$).

II. Attrition in the 2015 follow-up survey

Of the 409 participants who consented to the analysis of their data in 2006, 263 (64.3%) answered the follow-up survey in 2015. To evaluate whether the 263 survey respondents differ from the nonresponding 146 participants, in the top part of Table AII, we present the means and standard deviations of the variables in Table I for which we have data for both samples. For each variable, the table also displays the p-value obtained when we test whether the two samples are significantly different from each other (t-tests for continuous variables and χ^2 tests for categorical variables). In the bottom part of Table AII, we present the same information for the control variables we will use for the robustness checks in Section III. We describe these variables and how we collected them in Section V.

For most variables, there are no statistically significant differences between the participants who answered the survey and those who did not. If we use an unadjusted significance threshold of 5%, then we find a significant difference in three out of the fifteen variables in the top part of Table AII (overconfidence, one-off bonuses, and gender) and in six out of the nineteen variables in the bottom part of Table AII (donations to University of Chicago, discount rate, fraction of US residents, fraction of white individuals,

Table AII – Summary statistics depending on responding to the follow-up survey in 2015

Note: Means, standard deviations, and number of observations for variables of interest. The rightmost column displays *p*-values from tests of equality of distributions between people who responded to the follow-up survey in 2015 and those who did not (*t*-tests for ordinal variables and χ^2 tests for categorical variables).

	RESPONDENT			NON-RESPONDENT			
	mean	s.d.	N	mean	s.d.	N	*p*-value
Competitive	0.52	0.50	263	0.52	0.50	146	0.994
Performance (rank in sums tasks)	2.43	0.78	263	2.58	0.76	146	0.063
Expected rank in sums tasks	2.25	0.78	263	2.23	0.76	146	0.735
Expected experimental earnings	76.91	60.41	263	68.31	51.98	146	0.132
Overconfidence	0.18	0.63	263	0.35	0.64	146	0.008
Risk aversion coefficient	4.52	4.40	263	5.13	4.42	146	0.182
Non-competitive tournament	0.41	0.49	263	0.40	0.49	146	0.957
Total income in 2008	180.77	161.02	263	164.11	144.32	146	0.284
Base salary in 2008	108.15	18.65	263	105.40	16.62	146	0.127
Total bonus in 2008	72.63	153.66	263	58.70	140.26	146	0.354
One-off bonus in 2008	43.65	30.50	263	37.30	24.37	146	0.022
Guaranteed performance bonus in 2008	28.98	144.84	263	21.40	137.19	146	0.600
Number of competing job offers	0.47	0.87	263	0.32	0.71	146	0.068
Fraction working in finance in 2008	0.50	0.50	263	0.53	0.50	146	0.674
Fraction working in consulting in 2008	0.26	0.44	263	0.22	0.42	146	0.674
Fraction of women	0.27	0.44	263	0.36	0.48	146	0.041
Additional control variables							
Age	28.04	2.30	263	28.54	2.65	146	0.057
Fraction non-white	0.50	0.50	263	0.64	0.48	146	0.005
Fraction US residents	0.81	0.39	263	0.68	0.47	146	0.003
Fraction married before MBA	0.25	0.43	263	0.27	0.45	146	0.552
Fraction religious	0.48	0.50	263	0.45	0.50	146	0.600
GMAT Quantitative percentile	81.34	12.61	262	82.96	13.14	144	0.228
GMAT Verbal percentile	89.34	9.80	262	85.61	13.67	144	0.004
GMAT Analytic percentile	73.47	21.34	245	69.13	22.27	138	0.064
GPA	3.35	0.34	255	3.29	0.35	136	0.072
CRT score	2.54	1.33	263	2.26	1.32	146	0.045
RMET score	0.75	0.10	263	0.74	0.10	146	0.120
Discount rate	0.05	0.04	244	0.06	0.05	132	0.002
Trust	0.38	0.30	263	0.40	0.30	146	0.475
Reciprocity	0.36	0.20	263	0.37	0.21	146	0.828
Cooperation	0.35	0.48	263	0.30	0.46	146	0.358
Survey overconfidence	0.93	4.51	255	0.85	4.66	136	0.858
Survey risk aversion (general)	6.54	1.83	263	6.26	1.98	146	0.155
Survey risk aversion (monetary)	1.46	1.02	263	1.53	1.01	146	0.526
Donations to University of Chicago	196.74	207.20	263	154.12	184.49	146	0.033

GMAT verbal percentile, and CRT scores). However, if we account for multiple comparisons by adjusting

p-values with the Benjamini-Hochberg method, none of the variables shows a significant difference between the participants who answered the survey and those who did not.

We also test whether preferences for competition differ between participants who consented to the analysis of all their data and those who did not. To do so, we run a probit regression with the participants' tournament choice as the dependent variable and the participants' gender, performance, overconfidence, and risk aversion coefficient as independent variables. In addition, we also included a dummy variable equal to one for participants who did not respond to the follow-up survey. We find that the estimated marginal effect of the dummy variable is small (0.022) and is not statistically significant ($p = 0.690$).

In conclusion, although there is no clear-cut evidence of strong selection effects in responding to the follow-up survey, there might be reasons for worry. In particular, if we do not correct p-values for multiple testing, we see significant differences in three important variables for the paper: overconfidence, one-off bonuses, and gender. Moreover, although the difference is not statistically significant, the total income in 2008 of respondents to the survey is noticeably higher than that of non-respondents (17K or 10% more). For this reason, in Section III, we perform a series of robustness checks where we account for selection into the follow-up survey.

III. Robustness checks

A. Gender differences in preferences for competition

Measurement error and misspecification

Our first robustness check addresses concerns about using a residual measure for preferences for competition (Gillen, Snowberg, and Yariv 2019; van Veldhuizen 2018). As Niederle and Vesterlund (2007), we measure preferences for competition by looking at whether individuals choose a tournament payment scheme. However, since there are several reasons why an individual might choose tournament pay, we interpret the choice of tournament pay as indicating stronger preferences for competition only after we control for the individual's performance, overconfidence, and risk aversion. This way of measuring preferences for competition has recently come under scrutiny because it is not a direct measure of the trait of interest. More precisely, if there is measurement error in the control variables or they are incorrectly specified in the regression, it is possible for there to be a bias in the estimated effects of preferences for competition.

If the concerns raised by Gillen et al. (2019) are valid in our data, they could steer us toward incorrect inferences. First, in regressions where we evaluate whether there are gender differences in preferences for competition (Table II), a significant coefficient for the gender dummy might be due to (residual) gender differences in risk aversion or overconfidence and might not be due to differences in preferences for

Table AIII – Robustness of the gender gap in preferences for competition to measurement error and misspecification

Note: Regressions of the choice of tournament pay. Marginal effects from probit regressions and standard errors in parenthesis. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10.

	(1)	(2)	(3)	(4)
Woman	–0.148**	–0.183***	–0.147**	–0.170**
	(0.060)	(0.061)	(0.064)	(0.066)
Obs.	409	409	409	409
χ^2 test	97.11***	120.03***	95.53***	117.98***

competition. Second, in regressions where we test the effect of preferences for competition on income, a significant coefficient for the competitive dummy might again be due to (residual) effects of risk aversion or overconfidence. In this subsection, we take a closer look at the identification of gender differences in preferences for competition. In subsection III.B, we do the same for the effect of preferences for competition on income.

To evaluate whether there is bias in the identification of gender differences in preferences for competition, we reran the probit regressions in Table II with additional control variables. We report the resulting marginal effects in Table AIII. In column (1), we simply reproduce the last regression of Table II, where we regress the participants' choice between tournament and piece-rate pay on the participants' gender, performance, overconfidence, and risk preferences. In this regression, there is a significant gender gap in the choice of tournament pay of 14.8% ($p = 0.013$), which we interpret in the paper as a gender difference in preferences for competition. Given the large number of dummy variables in subsequent regressions, we limit the table to the coefficient of the gender dummy.

In column (2), we control for the original control variables non-parametrically. Specifically, we split each control variable into six equally sized bins (sextiles), which we introduce as dummy variables. Introducing these variables allows us to capture non-linear relations between choosing tournament pay and performance, risk aversion, and overconfidence, which can be a source of bias in the identification of gender differences in the preferences for competition.[1] As can be seen in Table AIII, the coefficient of the gender dummy increases by 4.6 percentage points to 18.3% and remains statistically significant ($p = 0.003$).

In column (3), we use answers to the initial survey to include additional measures of the participants' risk aversion and overconfidence. These variables are bound to be noisier than our lab measures since we did not elicit them with incentive-compatible methods. However, as demonstrated by Gillen et al. (2019), they can capture some of the potential measurement error of the incentive-compatible variables. For risk aversion, we use a commonly-used survey measure of general attitudes toward risk (Falk et al. 2018) and

[1] We also tried specifications using the squared value of the control variables. We obtain very similar results.

another self-reported measure of risk attitudes in the monetary domain. For overconfidence, we use the participants' expected GPA decile (estimated in 2006) minus their actual GPA decile (in 2008). We provide descriptive statistics for these variables in Tables AI and AII and a detailed description in Section V. Once again, the coefficient of the gender dummy is robust to the introduction of all these variables: it decreases by 0.1 percentage points to 14.7% ($p = 0.022$).

Finally, in column (4), we include all the additional control variables from the regressions in columns (2) and (3). Compared to the regression in column (1), we find an increase in the gender coefficient of 2.2 percentage points to 17.0% ($p = 0.010$). Hence, overall, we do not find evidence that the gender difference in preferences for competition is due to measurement error or misspecification.

Additional control variables

Here, we test the robustness of the gender difference in preferences for competition to the inclusion of a large set of control variables. Specifically, we include all the fifteen additional control variables seen in Tables AI and AII These variables include demographic characteristics (e.g., age, race, residence, marital status, religiosity), measures of different kinds of abilities (e.g., mathematical, verbal, and analytical skills, capacity for cognitive reflection, and emotional intelligence), and other standard experimental measures (e.g., willingness to trust, reciprocate, and cooperate with others). We describe all these variables in detail in Section V. This exercise allows us to evaluate whether preferences for competition describe variation between individuals that is not captured by typically measured observables. In addition, if risk aversion and overconfidence affect variables such as GMAT scores, trust, and cooperation, then the inclusion of these variables should reduce the effects of measurement error (as reasoned above).

The impact of including all these control variables on the gender dummy is seen in Table AIV. As above, in column (1), we simply reproduce the last regression of Table II. In column (2), we include the fifteen additional control variables, while in column (3), we also add the variables used in column (4) of Table AIII to reduce bias due to measurement errors and misspecification. We can see that the inclusion of all these control variables has a moderate effect on the magnitude of the gender gap in tournament pay. In column (2), it shrinks by 2.9 percentage points (around 0.47 standard errors) to 11.9% ($p = 0.069$) and by 1.0 percentage points (around 0.16 standard errors) to 13.8% ($p = 0.057$).

Table AIV – Robustness of the gender gap in preferences for competition to additional controls

Note: Regressions of the choice of tournament pay. Marginal effects from probit regressions and standard errors in parenthesis. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10.

	(1)	(2)	(3)
Woman	–0.148**	–0.119*	–0.138*
	(0.060)	(0.065)	(0.073)
Additional controls	No	Yes	Yes
Measurement error controls	No	No	Yes
Obs.	409	409	409
χ^2 test	97.11***	112.66***	146.33***

We think that this is compelling evidence that, by and large, competitiveness captures individual variation that would otherwise remain unobserved. Moreover, even though the shrinking of the gender gap (and the increase in its *p*-value) might suggest that there is a bias due to measurement error, we should point out that these regressions are not as appropriate to test measurement error as the regressions in Table AIII (where the coefficient does not shrink). The reason is that the variables in Table AIII are measures of risk preferences and overconfidence unrelated to preferences for competition. In contrast, the additional control variables in these regressions could be related to this trait. For instance, it is conceivable that preferences for competition, which have been shown to affect educational choices (Buser, Niederle, and Oosterbeek 2014; Reuben, Wiswall, and Zafar 2017), have a direct effect on variables measuring ability, such as GMAT scores, which would also explain the attenuation of the coefficient.

Preferences for high rewards

Recall that, like in the third period of the experiment, in the fourth period, participants had to choose whether they wanted to be compensated for their past performance according to the piece rate or tournament payment schemes. Unlike in the third period, however, they did not have to perform the adding task again as their decision applied to their past piece-rate performance. As Niederle and Vesterlund (2007) argue that this decision is akin to a choice between a certain payoff and a lottery with ambiguous probabilities and is not affected by the participants' attitudes towards competition. If this is the case, it is interesting to check whether men and women choose differently in this period and, if they do, whether these differences help us account for the gender difference in preferences for competition.

Table AV – Gender gap in preferences for competition and preferences for high rewards

Note: Regressions of the choice of tournament pay in the fourth period of the experiment (columns (1), (2), and (3)) and the choice of tournament pay in the third period of the experiment (columns (4) and (5)). Marginal effects from probit regressions and standard errors in parenthesis. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10.

	(1)	(2)	(3)	(4)	(5)
Woman	–0.220***	–0.175***	–0.104*	–0.148**	–0.133**
	(0.049)	(0.052)	(0.058)	(0.060)	(0.064)
Performance		0.154***	0.267***	0.272***	0.209***
		(0.026)	(0.032)	(0.034)	(0.036)
Overconfidence			0.217***	0.203***	0.152***
			(0.032)	(0.033)	(0.034)
Risk aversion			–0.056**	–0.083***	–0.073**
			(0.027)	(0.028)	(0.029)
Non-competitive tournament					0.330***
					(0.056)
Obs.	409	409	409	409	409
χ^2 test	17.36***	49.87***	96.97***	97.11***	129.08***

Columns (1) through (3) of Table AV reproduce the regressions of Table II but use as the dependent variable a dummy variable that equals one if the participant chooses tournament pay in the fourth period of the experiment. We label this variable "non-competitive tournament." Without any controls (column (1)), the gender gap in choosing the non-competitive tournament equals 22.0%. Controlling for the participants' performance reduces the gender gap by 4.5 percentage points to 17.5% (column (2)). Further controlling for the participants' overconfidence and risk preferences reduces the gender gap to 10.4% (column (3)). Hence, like with the decision to compete, gender plays a role in choosing non-competitive tournament pay.

For our paper, more important is to test whether the participants' choice of non-competitive tournament pay explains the gender gap in choosing the competitive tournament (reported in Table II), which we interpret as a gender difference in preferences for competition. Columns (4) and (5) of Table AV contain regressions that use the competitive dummy as the dependent variable. In column (4), we reproduce the last regression of Table II, where the decision to compete is explained by the participants' performance, overconfidence, and risk aversion. In column (5), we add the non-competitive dummy as an independent variable. We can see that even though the coefficient of the non-competitive dummy is statistically significant, controlling for this variable does not have a remarkable effect on the gender gap in the decision to compete. The coefficient of the gender dummy shrinks by 1.5 percentage points but is still economically and statistically significant. This result supports our conclusion that there is a gender difference in preferences for competition among our participants.

Table AVI – Robustness of the effect of preferences for competition on income in 2008 to measurement error and misspecification

Note: Regressions of the log of total income in 2008 in column (1) and of the log of base salary in 2008 in column (2). Hurdle model of the likelihood of receiving a bonus in column (3) and its magnitude in column (4). Hurdle model of the likelihood of receiving a guaranteed performance bonus in column (5) and its magnitude in column (6). Linear estimates in columns (1), (2), (3), and (5). Marginal effects in columns (4) and (6). Standard errors in parenthesis. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10.

	Total income	Base salary	Total bonus		Guaranteed performance bonus	
			Received	Amount	Received	Amount
	(1)	(2)	(3)	(4)	(5)	(6)
Woman	–0.088**	–0.004	0.034	–0.331***	–0.057	–0.546**
	(0.038)	(0.018)	(0.026)	(0.083)	(0.061)	(0.261)
Competitive	0.074**	0.018	0.012	0.162**	0.011	0.547***
	(0.036)	(0.017)	(0.024)	(0.079)	(0.057)	(0.211)
Obs.	409	409	409	380	409	153
F test / χ^2 test	2.12***	1.43	24.25	52.91***	24.13	27.15

B. Effect of preferences for competition on income in 2008

Next, we provide a series of robustness checks for the effect of preferences for competition on income in 2008 (reported in Table V). Throughout this subsection, we focus solely on regressions without industry fixed effects to keep the number and size of the tables at a reasonable level. In line with the main body of the paper, our results do not change much if we control for industry. We can provide the regressions with industry fixed effects upon request.

Measurement error and misspecification

We start by examining whether the effect of preferences for competition on income in 2008 is overestimated due to measurement error or misspecification of the control variables (Gillen et al. 2019). Similar to subsection III.A, in Table AVI, we reran the more interesting regressions from Table V, adding the two survey measures of risk aversion and overconfidence (see Tables AI and AII for descriptive statistics of these variables and Section V for a detailed description), as well as substituting performance, risk aversion, and overconfidence with their respective six dummy variables. For space considerations, we limit the table to the coefficients of the gender and competitive dummies.

In column (1), the dependent variable is the log of total income in 2008. The inclusion of the control variables slightly decreases the coefficient of the competitive dummy from 0.079 to 0.074, but it remains both economically and statistically significant. In column (2), the dependent variable is the log of base salaries in 2008. The additional control variables do not change the result of no relationship between base salaries and preferences for competition (the coefficient of the competitive dummy remains unchanged at

0.022 and remains statistically insignificant). In columns (3) and (4), we reran the two-step hurdle model used to estimate the probability of getting a bonus (column (3)) and the magnitude of the bonus received (column (4)). In columns (5) and (6), we repeat the same regression but solely for guaranteed performance bonuses. Once again, we do not find that the inclusion of the control variables negatively impacts the coefficient of the competitive dummy. It changes from 0.158 to 0.162 when considering the magnitude of all bonuses and from 0.571 to 0.547 when considering the magnitude of guaranteed performance bonuses.

In summary, we do not find evidence that the relationship between the various forms of income in 2008 and the competitive dummy is due to measurement error or misspecification, which suggests that it is indeed driven by preferences for competition.

Additional control variables

Here, we test the robustness of the effect of preferences for competition on income in 2008 to the inclusion of a large set of control variables. Like in subsection III.A, in Table AVII, we reran a selection of the regressions from Table V, including the fifteen additional control variables seen in Tables AI and AII (described in Section V). These variables include demographic characteristics (e.g., age, race, residence, marital status, religiosity), measures of different kinds of abilities (e.g., mathematical, verbal, and analytical skills, capacity for cognitive reflection, and emotional intelligence), and other common experimental measures (e.g., willingness to trust, reciprocate, and cooperate with others). This robustness check lets us evaluate whether preferences for competition describe variation in income that is not captured by typically measured observables. Once again, we limit the table to the coefficients of the gender and competitive dummies.

In column (1), the dependent variable is the log of total income in 2008. Including the additional control variables does not affect the coefficient of the competitive dummy (it changes from 0.079 to 0.078, which is less than 0.03 standard errors). In column (2), the dependent variable is the log of base salaries in 2008. The additional control variables do not modify the result of no relationship between base salaries and preferences for competition (the coefficient changes from 0.022 to 0.024 and remains statistically insignificant). In columns (3) and (4), we reran the two-step hurdle model used to estimate the probability of getting a bonus (column (3)) and the magnitude of the bonus received (column (4)). In columns (5) and (6), we repeat the same regression but solely for the guaranteed performance bonuses. Once again, we do not find that the inclusion of the control variables has a large effect on the coefficient of the competitive dummy. It slightly decreases from 0.158 to 0.152 (less than 0.08 standard errors) when considering the magnitude of all bonuses and from 0.571 to 0.528 (from $13K to $12K, around 0.20 standard errors) when considering the magnitude of guaranteed performance bonuses.

Table AVII – Robustness of the effect of preferences for competition on income in 2008 to additional controls

Note: Regressions of the log of total income in 2008 in column (1) and of the log of base salary in 2008 in column (2). Hurdle model of the likelihood of receiving a bonus in column (3) and its magnitude in column (4). Hurdle model of the likelihood of receiving a guaranteed performance bonus in column (5) and its magnitude in column (6). Linear estimates in columns (1), (2), (3), and (5). Marginal effects in columns (4) and (6). Standard errors in parenthesis. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10.

	Total	Base	Total bonus		Guaranteed performance bonus	
	income	salary	Received	Amount	Received	Amount
	(1)	(2)	(3)	(4)	(5)	(6)
Woman	–0.083**	0.007	0.035	–0.356***	–0.051	–0.575**
	(0.038)	(0.018)	(0.026)	(0.085)	(0.062)	(0.248)
Competitive	0.077**	0.024	0.006	0.150*	0.035	0.524**
	(0.036)	(0.017)	(0.024)	(0.080)	(0.058)	(0.213)
Obs.	409	409	409	380	409	153
F test / χ^2 test	2.64***	1.80**	32.12**	48.04***	31.00*	32.37**

In summary, we do not find that the relationship between preferences for competition and income in 2008 is affected by including a large set of control variables. Therefore, it seems likely that preferences for competition explain variance in earnings that would otherwise remain unexplained.

Preferences for high rewards

In this subsection, we analyze the effect of choosing tournament pay without having to perform under competitive conditions. Niederle and Vesterlund (2007) posit that this decision is unaffected by the participants' preferences for competition. If this is the case, it is interesting to analyze whether this variable is also a good predictor of the participants' income in 2008. To test the effect of a 'preference for high rewards,' we reran a selection of the regressions from Table V, including the non-competitive tournament dummy in Table AVIII. In columns (1) through (4), we use the non-competitive tournament dummy instead of the competitive dummy. In columns (5) through (8), we use both the non-competitive tournament and competitive dummies. Once again, we limit the table to the coefficients of the gender and competitive dummies.

In columns (1) and (5), the dependent variable is the log of total income in 2008. When the non-competitive tournament dummy is included alone (in column (1)), its coefficient is positive but is less than half the coefficient of competitive in Table V (0.036 vs. 0.079, which is a difference of around one standard error), and it is not statistically different from zero ($p = 0.322$). When both the non-competitive tournament and competitive dummies are included, the coefficient for competitive is economically and statistically significant (around \$12K, $p = 0.046$). By contrast, the coefficient for the non-competitive tournament is small and far from statistical significance (around \$2K, $p = 0.709$). In columns (2) and (6), the dependent

Table AVIII – Income in 2008 and preferences for competition or preferences for high rewards

Note: Regressions of the log of total income in 2008 in columns (1) and (5) and of the log of base salary in 2008 in columns (2) and (6). Hurdle model of the likelihood of receiving a bonus in columns (3) and (7), and of its magnitude (in logs) in columns (4) and (8). Linear estimates in columns (1), (2), (4), (5), (6), and (8). Marginal effects in columns (3) and (7). Standard errors in parenthesis. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10.

	Total income	Base salary	Total bonus		Total income	Base salary	Total bonus	
			Received	Amount			Received	Amount
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Woman	–0.114***	–0.010	0.034	–0.397***	–0.106***	–0.009	0.034	–0.379***
	(0.036)	(0.016)	(0.032)	(0.079)	(0.036)	(0.017)	(0.030)	(0.079)
Non-competitive	0.036	0.027	–0.016	0.019	0.014	0.023	–0.020	–0.032
tournament	(0.036)	(0.017)	(0.032)	(0.081)	(0.038)	(0.018)	(0.029)	(0.085)
Competitive					0.075**	0.015	0.024	0.167**
					(0.038)	(0.017)	(0.041)	(0.083)
Obs.	409	409	409	380	409	409	409	380
F test / χ^2 test	2.89**	1.16	3.77	27.79***	3.09***	1.09	4.16	32.14***

variable is the log of base salaries in 2008. Consistent with the results reported in the main body of the paper, neither the coefficient of non-competitive tournament nor of competitive display a significant association with base salaries. Finally, in columns (3) and (4), as well as (7) and (8), we reran the two-step hurdle model used to estimate both the probability of getting a bonus (columns (3) and (7)) and the magnitude of the bonus received (columns (4) and (8)). Once again, we do not find that the non-competitive tournament dummy is significantly associated with the magnitude of the bonus, while the competitive dummy is.

In summary, we find compelling evidence that the coefficient of the competitive dummy variable is indeed capturing a relationship between the participants' preferences for competition and income that is not related to the choice of a tournament payment scheme *per se*.

Guaranteed performance bonus

In this subsection, we test the robustness of the results for the guaranteed performance bonus component. Recall that we group the various bonuses into two components: the one-off bonus component, which includes relocation, tuition, sign-on, and retention bonuses, and the guaranteed performance bonus component, which includes stock options, profit sharing, guaranteed performance, and other bonuses. Since it is not entirely clear what bonuses are classified as "other" in Table AIX, we reran the two-step hurdle model used to estimate both the probability of getting some guaranteed performance bonus (column (6) in Table V) and the magnitude of the guaranteed performance bonus (column (7) in Table V) excluding "other" bonuses. The exclusion of other bonuses decreases the coefficient of the competitive dummy from 0.571 to 0.461 when considering the magnitude of the guaranteed performance bonus, but it remains large and

Table AIX – Robustness of the guaranteed performance bonus to the exclusion of bonuses classified as "other"

Note: Hurdle model of the likelihood of receiving a guaranteed performance bonus in column (1) and of its magnitude (in logs) in column (2). Marginal effects in column (1) and linear estimates in column (2). Standard errors in parenthesis. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10.

	Received	Amount
	(1)	(2)
Woman	–0.030	–0.474**
	(0.054)	(0.238)
Competitive	0.043	0.461**
	(0.053)	(0.219)
Obs.	409	128
F test / χ^2 test	1.38	12.19**

statistically significant. The exclusion of other bonuses also leaves unchanged the lack of a significant correlation between receiving some guaranteed performance bonus and being competitive.

Boasting

In this subsection, we analyze the correlation between boasting and income in 2008 (and 2015). As mentioned in the main body of the paper, two years after the initial experiment, 95 of the MBAs in our sample participated in another study. In that experiment, they were asked, in private, to recall the number of additions they had answered correctly. Subsequently, they were asked to communicate that number to a group of colleagues. We use the difference between the recalled performance and the performance communicated to others as a measure of boasting. The details for this experiment are available in Reuben et al. (2012).

In Table AX, we look at the effect of boasting on income in 2008 (and 2015). Column (1) reruns the baseline specification from Table V using only the 95 MBAs who participated in the second experiment. Column (2) runs the same specification, but it substitutes the measure of overconfidence with the measure of boasting. Column (3) includes both the overconfidence and the boasting measures. Columns (4), (5), and (6) do the same exercise but using the log of income in 2015 as the dependent variable. In these regressions, we start with the baseline regression of Table VIII, which includes the interaction between overconfidence and preferences for competition. We use this specification since it is the one that displays a significant effect for preferences for competition.

Table AX – Determinants of income controlling for propensity to boast

Note: Regressions of the log of total income in 2008 in columns (1) to (3) and the log of total income in 2015 in columns (4) to (6). Regressions in columns (2), (3), (5), and (6) restrict the sample to the participants for whom we measured the propensity to boast. Overconfidence, risk aversion, and boasting are standardized to have a mean of zero and a standard deviation of one. All regressions include performance as a control. Linear estimates with standard errors in parenthesis. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10.

	Total Income in 2008			Total Income in 2015		
	(1)	(2)	(3)	(4)	(5)	(6)
Woman	–0.116	–0.114	–0.117	–0.171	–0.231	–0.205
	(0.076)	(0.075)	(0.076)	(0.218)	(0.223)	(0.219)
Competitive	0.043	0.043	0.047	–0.024	–0.046	–0.003
	(0.069)	(0.066)	(0.069)	(0.180)	(0.180)	(0.182)
Risk aversion	–0.041	–0.035	–0.036	–0.049	–0.029	–0.039
	(0.033)	(0.033)	(0.033)	(0.085)	(0.089)	(0.086)
Overconfidence	–0.010		–0.008	0.102		0.113
	(0.035)		(0.035)	(0.103)		(0.103)
Boasting		–0.035	–0.035		–0.126	–0.166
		(0.030)	(0.031)		(0.144)	(0.139)
Overconfidence x Competitive				–0.419**		–0.461***
				(0.167)		(0.169)
Boasting x Competitive					0.179	0.251
					(0.169)	(0.165)
Obs.	95	95	95	73	73	73
F test / χ^2 test	1.15	1.42	1.18	1.60	0.68	1.50
R^2	0.061	0.074	0.075	0.127	0.058	0.158

Comparing the baseline regression of Table V to that in column (2), we see that the gender coefficient is roughly the same in both regressions, but the coefficient for preferences for competition is roughly half in the smaller sample. It is also evident from the large standard errors that we do not have sufficient power to reach clear conclusions. That being said, comparing column (1) with columns (2) and (3) is still illustrative of the impact of using the boasting variable as a control. In columns (2) and (3), we can see that the coefficient of the boasting variable is negative. In addition, introducing boasting does not reduce the size of the coefficient for preferences for competition (compared to column (1)). In other words, there is no support for the idea that boasting is responsible for the positive relationship between preferences for competition and income in 2008. The results for income in 2015 are analogous. Remarkably, the interaction between overconfidence and preferences for competition is still statistically significant in the reduced sample (column (4)). In columns (5) and (6), the coefficient for boasting is negative, and its introduction has no effect on the interaction between overconfidence and preferences for competition (column (6)). Hence, once again, we do not find evidence that the association between preferences for competition and income in 2015 is impacted by a tendency to boast.

C. Effect of preferences for competition on income in 2015

Next, we provide a series of robustness checks for the effect of preferences for competition on income in 2015. Once again, we focus on regressions without industry fixed effects since results do not change much by controlling for industry (regressions with industry fixed effects are available upon request).

Attrition in the 2015 follow-up survey

As reported in Section II, there is some evidence of selection into the 2015 follow-up survey. With p-values uncorrected for multiple comparisons, we find significant differences in three important variables: overconfidence, one-off bonuses, and gender. Hence, in Table AXI, we re-estimated regressions from Tables VII and VIII, correcting for selection into the follow-up survey using Heckman's two-step procedure (Heckman 1979).[2] Panel A shows the coefficients of the second stages, while Panel B shows the coefficients of the respective first stages.

In the first stage, we include the same independent variables as in the second stage (i.e., competitive and gender dummies, overconfidence, risk aversion, and performance). In addition, to limit potential problems caused by collinearity between the correction term and the independent variables, we include two exclusion restrictions (Puhani 2000). The first exclusion restriction is the log of participants' donations to the class gift to the University of Chicago. As we saw in Section II, the more money a student donated to the class gift, the more likely they are to respond to the follow-up survey (a one standard deviation increase in donations predicts a 5.6 percentage point increase in the probability of responding to the follow-up survey). This effect is probably due to how much individuals identify with the university and not their income or

[2] We obtain similar results we maximum likelihood estimates. Also, note that 263 individuals answered the 2015 follow-up survey. However, among the respondents there were 13 who were not employed. Since our dependent variable is employment income, we dropped these individuals from the analysis. However, including them gives very similar results.

Table AXI – Robustness of the effect of preferences for competition and overconfidence on income in 2015 to selection into the follow-up survey

Note: Regressions of the log of total income in 2015 in columns (1) and (2), the log of base salary in 2015 in columns (3) and (4), and the log of the performance bonus in 2014 in columns (5) and (6). Panel A contains linear estimates corrected for selection into the follow-up survey using Heckman's two-step procedure. Panel B reports the marginal effects of the respective selection equations. Standard errors in parenthesis. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10.

	Total income		Base salary		Bonus amount	
	PANEL A: SELECTION-CORRECTED ESTIMATES					
	(1)	(2)	(3)	(4)	(5)	(6)
Woman	–0.426***	–0.427***	–0.237***	–0.238***	–0.915***	0.925***
	(0.129)	(0.127)	(0.088)	(0.087)	(0.262)	(0.259)
Competitive	0.026	0.028	–0.005	–0.005	0.104	0.073
	(0.099)	(0.098)	(0.068)	(0.068)	(0.201)	(0.200)
Overconfidence	–0.097	–0.006	–0.068*	–0.031	0.018	0.198
	(0.059)	(0.072)	(0.040)	(0.050)	(0.130)	(0.149)
Overconfidence × Competitive		–0.214**		–0.087		–0.435**
		(0.088)		(0.061)		(0.179)
Risk aversion	–0.114**	–0.108**	–0.063**	–0.061*	–0.098	-0.091
	(0.047)	(0.046)	(0.032)	(0.032)	(0.097)	(0.096)
	PANEL B: SELECTION INTO THE SURVEY					
	(1)	(2)	(3)	(4)	(5)	(6)
Woman	–0.132**	–0.132**	–0.132**	–0.132**	–0.145***	–0.145***
	(0.052)	(0.052)	(0.052)	(0.052)	(0.055)	(0.055)
Competitive	–0.034	–0.035	–0.034	–0.035	–0.040	–0.041
	(0.054)	(0.054)	(0.054)	(0.054)	(0.057)	(0.057)
Overconfidence	–0.054**	–0.066*	–0.054**	–0.066*	–0.076***	–0.076**
	(0.028)	(0.035)	(0.028)	(0.035)	(0.029)	(0.036)
Overconfidence × Competitive		0.027		0.027		0.002
		(0.048)		(0.048)		(0.050)
Risk aversion	–0.032	–0.031	–0.032	–0.031	–0.040	–0.040
	(0.024)	(0.024)	(0.024)	(0.024)	(0.025)	(0.025)
Discount rate	–0.072***	–0.071***	–0.072***	–0.071***	–0.080***	–0.079***
	(0.023)	(0.024)	(0.023)	(0.024)	(0.025)	(0.025)
Donations to University of Chicago	0.048**	0.049**	0.048**	0.049**	0.041*	0.041*
	(0.023)	(0.023)	(0.023)	(0.023)	(0.024)	(0.024)
Uncensored obs.	396	396	396	396	364	364
Censored obs.	146	146	146	146	146	146
χ^2 test	17.03**	25.73***	11.62**	14.96**	21.46***	27.98***

preferences for competition. The second exclusion restriction is the participants' elicited discount rate. Higher discount rates are strongly associated with a lower likelihood of responding to the follow-up survey (a one standard deviation increase in the discount rate predicts a 7.3 percentage point decrease in the

Table AXII – Robustness of the effect of preferences for competition and overconfidence on income in 2015 to measurement error and misspecification

Note: Regressions of the log of total income in 2015 in column (1) and of the log of base salary in 2015 in column (2). Hurdle model of the likelihood of receiving a bonus in column (3) and of its magnitude (in logs) in column (4). Regression of income growth, the log of total income in 2015 minus the log of total income in 2008, in column (5). Linear estimates in columns (1), (2), (4), and (5). Marginal effects in column (3). Standard errors in parenthesis. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10.

	Total income	Base salary	Bonus Received	Bonus Amount	Income Growth
	(1)	(2)	(3)	(4)	(5)
Woman	–0.379***	–0.220**	–0.002	–0.947***	–0.264**
	(0.119)	(0.092)	(0.038)	(0.255)	(0.121)
Competitive	0.014	–0.023	0.020	0.076	–0.048
	(0.093)	(0.061)	(0.034)	(0.179)	(0.100)
Overconfidence	0.009	–0.005	–0.019	0.108	–0.017
	(0.051)	(0.031)	(0.019)	(0.098)	(0.057)
Overconfidence × Competitive	–0.187**	–0.078	–0.043	–0.382**	–0.182***
	(0.082)	(0.055)	(0.029)	(0.157)	(0.086)
Obs.	250	250	250	218	250
F test / χ^2 test	2.68***	1.28	47.25***	65.82***	1.98**

probability of responding to the follow-up survey). This association is most likely an effect of discount rates, which have been linked theoretically and empirically to procrastination in filling out questionnaires (Ariely and Wertenbroch 2002; O'Donoghue and Rabin 1999; Reuben, Sapienza, and Zingales 2015), that is unrelated to income in 2015 or preferences for competition.[3]

In columns (1) and (2), the dependent variable of the second stage is the log of total income in 2015. Correcting for selection into the follow-up survey slightly decreases the coefficient of the competitive dummy (by around 1 log point in column (1) and 3 log points in column (2)). The coefficient of overconfidence is similar in magnitude but is no longer significant in the specification of column (1). However, the interaction between overconfidence and competitive remains negative, large, and highly statistically significant. Hence, we conclude that selection into the follow-up survey had little effect on the estimated effects of preferences for competition, overconfidence, and their interaction. It is important to note that these results are not the consequence of a weak first stage (Panel B). Both exclusion restrictions are good predictors of answering the survey. In columns (3) and (4), the dependent variable of the second stage is the log of base salaries in 2015. Correcting for selection into the follow-up survey had a negligible effect on the coefficient of the competitive dummy, overconfidence, and their interaction. In columns (5)

[3] The correlation coefficients between donations to the university or discount rates with either the competitive dummy or total income in 2008 are low (less than 0.056 for donations and 0.010 for discount rates) and are not statistically significant.

and (6), the dependent variable of the second stage is the log of the performance bonus. Note that, to run these regressions, we dropped the survey respondents who did not receive a bonus. For this reason, the coefficients in Table AXI are not entirely comparable to those in Tables VII and VIII, which are based on a two-step hurdle model. Nonetheless, it is telling that the coefficients of preferences for competition, overconfidence, and the interaction term are similar in both regressions.

Next, we provide a series of robustness checks for the effect of preferences for competition and its interaction with overconfidence on income in 2015 (reported in Table VIII).

Measurement error and misspecification

We start by looking at whether the effect of preferences for competition on income in 2015 is affected by measurement error or misspecification of the control variables. Similar to subsection III.A and III.B, in Table AXII, we reran the regressions from Table VIII, adding the two survey measures of risk aversion, the survey measure of overconfidence, and substituting performance and risk aversion with six dummy variables. Since we are interested in the interaction between overconfidence and preferences for competition, we did not transform overconfidence into six dummy variables. For space considerations, we limit the table to the coefficients of the gender and competitive dummies, overconfidence, and the interaction of overconfidence and the competitive dummy.

In column (1), the dependent variable is the log of total income in 2015. In column (2), the dependent variable is the log of base salaries in 2015. In columns (3) and (4), we reran the two-step hurdle model used to estimate the probability of getting a bonus (column (3)) and the magnitude of the bonus received (column (4)). Lastly, in column (5), the dependent variable is income growth, defined as the log of total income in 2015 minus the log of total income in 2008. Once again, we do not find that the inclusion of the additional variables has an important qualitative effect on the results reported in the paper. The coefficient of the competitive dummy is similar in all regressions, and it never reaches statistical significance. By contrast, the interaction between the competitive dummy and overconfidence is of similar magnitude and statistical significance to the coefficients seen in Table VIII. In short, we do not find evidence that the results concerning income in 2015 are affected by measurement error or misspecification.

Table AXIII – Robustness of the effect of preferences for competition and overconfidence on income in 2015 to additional controls

Note: Regressions of the log of total income in 2015 in column (1) and of the log of base salary in 2015 in column (2). Hurdle model of the likelihood of receiving a bonus in column (3) and of its magnitude (in logs) in column (4). Regression of income growth, the log of total income in 2015 minus the log of total income in 2008, in column (5). Linear estimates in columns (1), (2), (4), and (5). Marginal effects in column (3). Standard errors in parenthesis. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10.

	Total income	Base salary	Bonus Received	Bonus Amount	Income Growth
	(1)	(2)	(3)	(4)	(5)
Woman	–0.381***	–0.216**	–0.043	–0.795***	–0.289**
	(0.130)	(0.099)	(0.045)	(0.274)	(0.131)
Competitive	0.039	0.013	0.004	0.045	–0.026
	(0.095)	(0.064)	(0.037)	(0.165)	(0.108)
Overconfidence	0.014	–0.007	–0.038	0.138	–0.022
	(0.056)	(0.033)	(0.030)	(0.111)	(0.059)
Overconfidence × Competitive	–0.223***	–0.112**	–0.043	–0.461***	–0.206**
	(0.083)	(0.056)	(0.038)	(0.171)	(0.085)
Obs.	250	250	250	218	250
F test / χ^2 test	3.48***	1.40	38.53**	67.75***	1.88**

Additional control variables

Next, we test the robustness of the effect of preferences for competition, overconfidence, and their interaction on income in 2015 to the inclusion of a large set of control variables. As in subsection III.A, in Table AXIII, we reran the regressions from Table VIII, including the fifteen additional control variables seen in Tables AI and AII (described in Section V). These variables include demographic characteristics (e.g., age, race, residence, marital status, religiosity), measures of different kinds of abilities (e.g., mathematical, verbal, and analytical skills, capacity for cognitive reflection, and emotional intelligence), and other standard experimental measures (e.g., willingness to trust, reciprocate, and cooperate with others). For space considerations, we limit the table to the coefficients of the gender and competitive dummies, overconfidence, and the interaction of overconfidence and the competitive dummy.

In column (1), the dependent variable is the log of total income in 2015. In column (2), the dependent variable is the log of base salaries in 2015. In columns (3) and (4), we reran the two-step hurdle model used to estimate the probability of getting a bonus (column (3)) and the magnitude of the bonus received (column (4)). Lastly, in column (5), the dependent variable is the log of total income in 2015 minus the log of total income in 2008. Once again, the coefficient of the competitive dummy is not statistically significant in any regression, and the interaction between the competitive dummy and overconfidence is of similar magnitude and statistical significance to the coefficients in Table VIII. In summary, the relationship between

Table AXIV – Income in 2015 and preferences for competition or for high rewards

Note: Regressions of the log of total income in 2015 in column (1) and of the log of base salary in 2015 in column (2). Hurdle model of the likelihood of receiving a bonus in column (3) and of its magnitude (in logs) in column (4). Regression of income growth, the log of total income in 2015 minus the log of total income in 2008, in column (5). Linear estimates in columns (1), (2), (4), and (5). Marginal effects in column (3). Standard errors in parenthesis. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10.

	Total income	Base salary	Bonus Received	Bonus Amount	Income Growth
	(1)	(2)	(3)	(4)	(5)
Woman	–0.391***	–0.198**	–0.004	–1.022***	–0.265**
	(0.105)	(0.082)	(0.046)	(0.238)	(0.106)
Competitive	0.011	–0.023	0.020	0.055	–0.041
	(0.097)	(0.067)	(0.051)	(0.179)	(0.114)
Non-competitive tournament	0.137	0.112	–0.001	0.142	0.104
	(0.101)	(0.072)	(0.052)	(0.192)	(0.118)
Overconfidence	–0.004	–0.026	–0.030	0.119	–0.026
	(0.053)	(0.030)	(0.034)	(0.109)	(0.056)
Overconfidence × Competitive	–0.215**	–0.100*	–0.029	–0.513**	–0.212**
	(0.091)	(0.059)	(0.044)	(0.191)	(0.099)
Overconfidence × Non-competitive	–0.007	0.010	–0.056	0.187	–0.007
tournament	(0.100)	(0.068)	(0.048)	(0.207)	(0.105)
Obs.	250	250	250	218	250
F test / χ^2 test	4.79***	2.07**	15.12*	36.11***	3.30***

preferences for competition, overconfidence, and their interaction with income in 2015 is unaffected by the inclusion of a large set of control variables.

Preferences for high rewards

In this subsection, we analyze again the effect of choosing tournament pay without having to perform under competitive conditions, a decision that is arguably unaffected by the participants' preferences for competition. We evaluate the effect of a 'preference for high rewards' on the participants' income in 2015 in Table AXIV. The table presents selected regressions from Table VIII, including the non-competitive tournament dummy, which we also interact with overconfidence.

Table AXV – Interaction of preferences for competition with risk aversion and performance

Note: Regressions of the log of total income in 2015 in columns (1) to (4). Hurdle model of the likelihood of receiving a bonus (not shown) and of its magnitude (in logs) in columns (5) to (8). Linear estimates in all columns. Standard errors in parenthesis. ***, **, and * indicate statistical significance at 0.01, 0.05, and 0.10.

	Total income				Bonus amount			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Woman	–0.391***	–0.366***	–0.386***	–0.372***	–1.008***	–0.954***	–0.988***	–0.975***
	(0.106)	(0.105)	(0.108)	(0.102)	(0.240)	(0.241)	(0.247)	(0.236)
Competitive	0.055	0.042	0.056	0.069	0.089	0.107	0.107	0.117
	(0.089)	(0.091)	(0.091)	(0.090)	(0.170)	(0.172)	(0.172)	(0.171)
Overconfidence	0.011	–0.084*	–0.084*	–0.012	0.154	–0.022	–0.026	0.135
	(0.051)	(0.048)	(0.046)	(0.054)	(0.105)	(0.101)	(0.100)	(0.112)
Risk aversion	–0.099***	–0.164***	–0.105***	–0.163***	–0.109	–0.221**	–0.120	–0.221**
	(0.043)	(0.056)	(0.043)	(0.054)	(0.086)	(0.104)	(0.087)	(0.104)
Overconfidence	–0.219**			–0.176*	–0.425***			–0.398**
x Competitive	(0.077)			(0.093)	(0.162)			(0.196)
Risk aversion		0.130		0.132		0.238		0.244
x Competitive		(0.084)		(0.081)		(0.168)		(0.164)
Performance			0.191**	0.093			0.264	0.063
x Competitive			(0.090)	(0.106)			(0.181)	(0.217)
Obs.	250	250	250	250	250	250	250	250
F test / χ^2 test	5.92***	5.16***	5.41***	4.66***	33.80***	30.76***	33.29***	39.13***

In column (1), the dependent variable is the log of total income in 2015. In column (2), the dependent variable is the log of base salaries in 2015. In columns (3) and (4), we reran the two-step hurdle model used to estimate the probability of getting a bonus and its magnitude. In column (5), the dependent variable is the log of total income in 2015 minus the log of total income in 2008. In all regressions, the coefficients of the non-competitive tournament dummy and its interaction with overconfidence are not statistically significant and small in magnitude. Their inclusion does not change the lack of significance of the competitive dummy or the magnitude and significance of the interaction between competitive and overconfidence. In summary, we do not find that controlling for the non-competitive dummy changes the conclusions reported in the paper.

Interaction of preferences for competition with other variables

In this subsection, we analyze whether the interaction between preferences for competition and other variables predicts the MBAs' income in 2015. Although there is a compelling argument for why preferences for competition interact with overconfidence. One might wonder whether this occurs for all determinants of tournament entry in the experiment. Hence, in Table AXV, we rerun selected regressions from Table VIII, including these additional interactions. We concentrate on the dependent variables that exhibit the strongest interaction between preferences for competition and overconfidence: total income in columns (1) through (4) and the magnitude of the realized bonus in columns (5) through (8). In columns (1) and (5), we reproduce the regressions of Table VIII with the interaction between preferences for competition and overconfidence. In columns (2) and (6), instead of overconfidence, we interact preferences for competition with our measure of risk aversion (the standardized CRRA coefficient). In columns (3) and (7), we interact preferences for competition with standardized performance. Finally, in columns (4) and (8), we include all these interaction terms simultaneously.

We find that the interaction between preferences for competition and risk aversion is positive, suggesting that the negative association between risk aversion and income is weaker for competitive individuals. However, the coefficient of this interaction is never statistically significant. The interaction between preferences for competition and performance is positive and statistically significant for income when it is included alone. However, the coefficient of this interaction is much smaller and no longer statistically significant when the other interactions are included in the regression. Importantly, the interaction between preferences for competition and overconfidence is not affected much by the inclusion of these other interactions. Specifically, the magnitude of the coefficient is similar and remains statistically significant.

IV. Procedures for the initial survey and experiment

This section describes the procedures used to conduct the initial survey and the experiment. We concentrate on the parts of the survey and experiment relevant to the paper. Further details can be found in Reuben, Sapienza, and Zingales (2008), including all survey questions and experimental instructions.

A. The initial survey

Participants completed the online survey in the fall of 2006. The deadline to complete the survey was the day participants took part in the experiment. Completing the survey was a requirement to pass one of the MBA core courses and took approximately one hour. The survey included questions on demographic characteristics and standard questionnaires of personality traits. We do not use the survey variables in the main body of the paper, but we do use them in the robustness checks. In Section V, we describe the variables used in these checks.

B. The experiment

We ran the experiment in October 2006 in four sessions of around 140 participants. It lasted for about 90 minutes. Participation in the experiment was a requirement of one of the MBAs' core courses. The experiment was programmed and run using zTree (Fischbacher 2007).

The experiment consisted of eight parts: three decision problems and five games. Participants played the eight parts in the following order: lottery with losses, asset market game, trust game, preferences for competition game, chocolate auction, social dilemma game, lottery without losses, and discount rate elicitation task. We gave the instructions for each part before the start of the respective part (the only exception being the instructions for the asset market game, which they received before their arrival). Importantly, participants received no information about the outcome of the games or lotteries during the experiment. Instead, they received feedback on their performance in specific games and on the behavior of other participants a few days later through an email.

Participants received a $20 show-up fee, which could be used to cover potential losses during the experiment. Also, we paid participants the amount they earned in one randomly chosen part (we did the randomization only among six parts since we always paid the lottery with losses and discount rate elicitation tasks). We paid participants who earned more than the show-up fee with a check delivered to their mailbox. Including the show-up fee, participants earned $99 on average (the standard deviation was $63).

In the main body of the paper, we describe the parts of the experiment used to measure preferences for competition. Below, we provide the instructions for these parts of the experiment. Moreover, in Section V, we describe the parts of the experiment used to measure the additional control variables utilized in the robustness checks.

Instructions for the sums tasks

This game is divided into 4 periods. At the beginning of the game, you will be divided into groups of four. The participants in your group will be the same throughout the 4 periods.

In each of the first 3 periods, you will be given a series of *addition tasks* (sums of four 2-digit numbers like the one below). You will have 150 seconds to answer as many questions as you want. The computer will record the number of sums that you answer correctly. You may use paper and pencil, but you *cannot* use a calculator. In each period, the rules for the payment are different and will be explained in detail before the start of the respective period.

One of the 4 periods will be randomly selected by the computer to determine your earnings for Game 3. In addition, after period 4, there will be a bonus section consisting of four questions. Any money earned in the bonus section will be added to this experiment's earnings.

Instructions for the piece-rate period

In this period, you will be paid $4 for each correct answer you give.

Example: If you answer 6 questions correctly, your earnings for period 1 equal $24. Remember, you can write down the numbers on a piece of paper, but you *cannot* use a calculator.

Instructions for the tournament period

In this period, you will compete against the other *three participants* in your group. Your payment is contingent on you having the highest number of correct answers. You will be paid *$16* for each correct answer if you have the *highest* number of correct answers in your group. If you do not have the highest number of correct answers, you will earn *$0* in this period. If there are two or more group members tied in first place, one of them will be randomly selected to be paid $16 for each correct answer (the others are paid $0). Note that all group members will face the same difficulty. That is, everyone will face the same sequence of numbers.

Example: Suppose that the other three participants in your group answer 5, 9, and 12 questions correctly. If you answer 11 questions correctly, your earnings in this period would equal $0. If you answer 13 questions correctly, your earnings in this period would equal $208. Remember, you can write down the numbers on a piece of paper, but you *cannot* use a calculator.

Instructions for the choice period

In this period, you will replay the same game, but you choose the rule according to which you will be paid. You can be paid with Rule 4 or with Rule 16:

Rule 4: If you choose this rule, you will be paid $4 for each correct answer regardless of what others do.

Rule 16: If you choose this rule, you will be paid according to your performance relative to the performance of the other three group members. You will earn $16 for each correct answer if you have more correct answers than the other group members had in period 2. If you do not have more correct answers than the other group members had, you will earn $0 in this period. If you tie in first place, a random draw will determine whether you are paid $16 for each correct answer or $0.

Remember, you can write down the numbers on a piece of paper, but you *cannot* use a calculator.

Instructions for the uncompetitive choice period

In this period, you do not have to repeat the addition task, but you have the choice to be paid *again* for your period 1 performance in two ways. You can choose to be paid according to Rule 4 or Rule 16.

Rule 4: If you choose this rule, you will be paid $4 for each question answered correctly in period 1 regardless of what others did.

Rule 16: If you choose this rule, you will be paid $16 for each correct answer in period 1 if you have more correct answers than the other three group members had in period 1. If you did not have more correct answers than the other group members had, you will earn $0 in this period. If you tie in first place, a random draw will determine whether you are paid $16 for each correct answer or $0.

Recall that in period 1, you correctly answered *XX questions*. Note that this choice determines your period 4 earnings; it does not affect your earnings from period 1.

Instructions to elicit the participants' expected rank in each period

In this screen, we would like you to estimate your performance relative to that of the other three players. For each of the first three periods, indicate whether you think you ranked first, second, third, or fourth. You will receive *$2* for every period in which you correctly estimate your rank. In case of a tie, you will receive the $2 if there is a way of resolving the tie that makes your estimate correct.

Example: Suppose that in period 1 you had 8 correct answers and the other three group members had 6, 8, and 11 correct answers. You will receive $2 if you guess that your rank is second or third in period 1.

V. Description of additional control variables

This section describes the additional control variables used in the robustness checks. We divide them depending on the data source: administrative data from the University of Chicago, the initial survey, or the experiment.

A. Administrative data

In addition to gender, the business school supplied us with the following variables:

- Age (in months).
- Race, which we used to construct a dummy variable indicating non-white individuals.
- Visa status, which we used to construct a dummy variable indicating whether an individual is a US resident (citizen or legal resident).
- Marital status, which we used to construct a dummy variable indicating a married individual.
- GMAT percentile scores. Both the aggregate score and the score of each of the three components: quantitative, verbal, and analytic.
- GPA at graduation.
- The number of dollars students donated to their class gift to the University of Chicago.

Note that all these variables, except for GPA and donations to the university, were collected in 2006, before the students started their MBA. We collected the last two variables in 2008 at graduation.

B. Initial survey

We use the following variables from the initial survey in the robustness checks.

- Religiosity, which we measured with the yes/no question: "Are you religious now?"
- A self-reported measure of the participants' general attitude toward risk, which has been shown to correlate with incentivized measures of risk aversion and is commonly used in the literature (Falk et al. 2018). It consists of the question: "Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please select a number between 0 and 10 where 0 means unwilling

to take risks and 10 means fully prepared to take risks." For this variable to measure risk aversion, we reverse-coded it so that higher numbers imply more aversion to risk.

- Another self-reported measure of the participants' attitudes toward risk was elicited in the monetary domain. Specifically, we asked participants to indicate: "What is the maximum price you are willing to pay for a ticket in a lottery that pays you $5K with 50% probability and nothing with 50% probability?" For this variable to measure risk aversion, we use $2.5K minus their answer to the question so that higher values imply more aversion to risk.
- We elicited the participants estimated academic performance by asking them: "In your future exams at the University of Chicago, in which decile of the GPA distribution do you expect yourself to be?" We then used their answer to this question minus their actual GPA decile at graduation to create the non-incentivized survey measure of overconfidence.
- We measured the participants' tendency to suppress intuitive responses using the Cognitive Reflection Test or CRT (Frederick 2005). We simplified the original test to four questions: (i) A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost? (ii) If you flipped a fair coin 3 times, what is the probability that it would land "heads" at least once? (iii) If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? (iv) Two cars are on a collision course, traveling toward each other in the same lane. Car A is traveling 70 miles an hour. Car B is traveling 80 miles an hour. How far apart are the cars one minute before they collide? The CRT score consists of the number of correct answers.
- To measure the participants' ability to detect emotions (a key component of emotional intelligence), we use the "reading the mind in the eyes" test or RMET (Baron-Cohen et al. 2001). It consists of correctly recognizing the emotions of various individuals by looking at pictures of their eyes. The RMET score consists of the fraction of correct answers.

C. Experiment

From the experiment, we use the following measures of important individual characteristics.

Discount rate

To measure time preferences, we gave participants a series of choices of the following form: receive x dollars today or receive $(1 + y)x$ dollars in two weeks, where x equaled their earnings in the experiment. Each subject answered thirteen such questions where y varied from 0 to 0.12 in steps of 0.01. After that, one of the questions was randomly selected and paid. We always paid participants by dropping a check into their mail folder during a day in which they had to attend class.

Trust and reciprocity

We measure the participants' propensity to trust and reciprocate by having them play a trust game (Berg, Dickhaut, and McCabe 1995). In the game, a first-mover is endowed with $50 and decides how much to send to a second-mover (in multiples of $5). Any amount sent is multiplied by three. The second mover then decides how much to return to the first mover.

Each participant played two trust games. First, they played as the first mover and then as the second mover. Participants made their second-mover decision using the strategy method. They indicated how much to return for each possible sent amount without knowing how much the first mover actually sent. They received no feedback between decisions and knew they were not playing with the same participant. We use the fraction of the $50 sent as first movers as the participants' measure of trust and the fraction they returned conditional on receiving $150 as their measure of reciprocity.

Cooperation

To measure their willingness to cooperate, participants played a variation of the design used by Fischbacher, Gächter, and Fehr (2001). Specifically, participants were randomly assigned to groups of eight, given an endowment of $50, and asked to make two contribution decisions to a linear public good game: an "unconditional" and a "conditional" decision. For their unconditional decision, each participant i indicated whether they are willing to contribute $c_i \in \{0, 50\}$ to the public good. For their conditional decision, each participant i indicated whether they are willing to contribute $c_i(x) \in \{0, 50\}$ given that $x \in \{0, 1, 2, 3, 4, 5, 6, 7\}$ other group members contribute. To determine the final contributions to the public good, seven unconditional decisions were selected at random and were used to determine the conditional decision of the remaining group member. Participant i's earnings equaled $50 - c_i + 0.3 \times \sum_j c_j$.
We use the unconditional contribution as a participant's willingness to cooperate.

References

Ariely, Dan and Klaus Wertenbroch. 2002. "Procrastination, Deadlines, and Performance: Self-Control by Precommitment." *Psychological Science* 13: 219–24.

Baron-Cohen, Simon, Sally Wheelwright, Jacqueline Hill, Yogini Raste, and Ian Plumb. 2001. "The 'Reading the Mind in the Eyes' Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-Functioning Autism." *Journal of Child Psychology and Psychiatry* 42(2): 241–51.

Benjamini, Yoav and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society Series B (Methodological)* 57: 289–300.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10(1): 122–42.

Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz. 2010. "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors." *American Economic Journal: Applied Economics* 2(3): 228–55.

Buser, Thomas, Muriel Niederle, and H. Oosterbeek. 2014. "Gender, Competitiveness, and Career Choices." *The Quarterly Journal of Economics* 129(3): 1409–47.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. 2018. "Global Evidence on Economic Preferences." *The Quarterly Journal of Economics* 133(4): 1645–92.

Fischbacher, Urs. 2007. "Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics* 10(2): 171–78.

Fischbacher, Urs, Simon Gächter, and Ernst Fehr. 2001. "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment." *Economics Letters* 71(3): 397–404.

Frederick, Shane. 2005. "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives* 19(4): 25–42.

Gillen, Ben, Erik Snowberg, and Leeat Yariv. 2019. "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study." *Journal of Political Economy* 127(4): 1826–63.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1): 153.

Niederle, Muriel and Lise Vesterlund. 2007. "Do Women Shy Away From Competition? Do Men Compete Too Much?" *The Quarterly Journal of Economics* 122(3): 1067–1101.

O'Donoghue, Ted and Matthew Rabin. 1999. "Doing It Now or Later." *American Economic Review* 89(1): 103–24.

Puhani, Patrick. 2000. "The Heckman Correction for Sample Selection and Its Critique." *Journal of Economic Surveys* 14(1): 53–68.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales. 2008. A Description of the Templeton Chicago MBAs Longitudinal Study. University of Chicago.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales. 2015. "Procrastination and Impatience." *Journal of Behavioral and Experimental Economics* 58: 63–76.

Reuben, Ernesto, Matthew Wiswall, and Basit Zafar. 2017. "Preferences and Biases in Educational Choices and Labour Market Expectations: Shrinking the Black Box of Gender." *The Economic Journal* 127(604): 2153–86.

van Veldhuizen, Roel. 2022. "Gender Differences in Tournament Choices: Risk Preferences, Overconfidence or Competitiveness?" *Journal of the European Economic Association* 20(4): 1595–1618.