Incentive schemes, framing, and market behaviour: evidence from an asset-market experiment

Xuegang Cui
Department of Accounting
Business School
Beijing Normal University
No.19 Xinjiekouwai Street
Beijing, China, 100875
cxg@bnu.edu.cn

Nick Feltovich
Department of Economics
Monash Business School
Monash University
Clayton, VIC 3800, Australia
nicholas.feltovich@monash.edu

Kun Zhang*
Department of Economics
Business School
HeBei University of Economics and Business
No.47 Xuefu Road
Shijiazhuang, China, 050061
Kunzhang_uk@163.com

Abstract: We investigate how asset prices and trading behaviour are impacted by the structure and framing of incentives, using a lab experiment. Subjects buy and sell a high-risk asset, a low-risk asset, and riskless cash over 10 rounds. We vary, between-subjects, the incentive scheme (relative versus absolute performance), and how the variable component of incentives is framed (bonus versus penalty), while holding constant the convexity of incentives. Both relative-performance (tournament) incentives and penalty framing are associated with significant increases in the price of the high-risk asset, relative to either its fundamental value or to the price of the low-risk asset. Additional analysis shows significant gender differences in trading behaviour and performance, and evidence that the two may be connected.

JEL classifications: C92, D81, D91, G11, G12, G41

Keywords: asset markets; behavioural finance; bubbles; experiment; tournament; framing; incentives; gender

Declarations of interest: none

* Corresponding author. Xuegang Cui acknowledges funding from the National Natural Science Foundation of China (No. 71772015) and Beijing Social Science Foundation (No. 17GLB016). Kun Zhang acknowledges funding from the National Natural Science Foundation of China (No. 71803037) and Hebei Research Grant for Researchers Returning from Abroad (No. C201869). We thank Lucy Ackert, Zhengyang Bao, Hans Hvide, an editor, and an anonymous referee for helpful comments. Declarations of interest: none.

1. Introduction

Proper functioning of financial markets is essential not only for participants in those markets, but for the economy as a whole. Mispricing of assets distorts resource allocation, reducing efficiency. When financial bubbles burst, spill-overs to the real economy can be severe, as the 2006-10 meltdown of the subprime mortgage market in the US showed. Well-functioning financial markets are also associated with lower borrowing costs and better returns on saving and investments.

One factor with the potential to hinder financial-market performance is the prevalence of tournament incentives in these markets. Traders and other market participants are often vying for prizes such as bonuses and promotions, or to avoid penalties such as demotion, reduction of investment funds at one's disposal, or even job loss, based on performance relative to others. Commentators have expressed concern that these incentives may encourage excessive risk-taking, as well as short-termism more generally.

Understanding the effects of tournament incentives on financial-market behaviour and outcomes is crucial, both for the financial firms themselves and for policy makers. Indeed, this question has received plenty of attention in the academic economics and finance literature (see Section 2 for some examples). However, one complication is that a given incentive scheme can be described in multiple ways. For example, a scheme comprising a low fixed payment and a possible bonus for good (absolute or relative) performance can equivalently be represented as a high fixed payment and a possible penalty for poor performance. Rational traders will ignore this distinction, but many researchers have observed differences in the way gains and losses are treated in a variety of settings, even when the underlying monetary incentives are the same (e.g, Kahneman and Tversky, 1979). A second complication is that introducing a tournament can change the shape of incentives, as well as the source of compensation. A "winner-takes-all" tournament with few winners and many losers makes incentives more convex, and an "elimination contest" tournament with many winners and few losers makes incentives more concave, compared to compensation based on absolute performance. This impedes identification of the effect of relative- versus absolute-performance incentives on its own.

In this paper, we investigate these two manipulations – compensation based on relative versus absolute performance, and bonus versus penalty framing – on asset market behaviour and outcomes. Besides our academic interest, these questions have clear implications for financial-market policy. Recent legislation (such as the 2010 Dodd-Frank Act in the US) has attempted to truncate the right tail of executive-compensation distributions, reducing the convexity of incentives but without affecting the tournament nature of these incentives. If tournament incentives themselves have an effect on risk taking and thence financial stability – beyond the effect of the shape of incentives – then reducing convexity on its own may not solve the problem. Another piece of proposed legislation in the US aims to require banks to set aside a portion of executive compensation, to be clawed back if the bank performs poorly (*Wall Street Journal*, 2019). It is plausible that such lost compensation will be viewed as a penalty, rather than as an unpaid bonus. If so, it is important to understand how changing the frame from reward to penalty will affect behaviour.

We use a lab experiment for our investigation. Evidence from the lab serves to complement other kinds of analysis, due to the lab's particular advantages. Lab experiments give us a high degree of control over aspects of the setting such as market parameters, monetary incentives, information and timing. Our ability to set treatment variables means that we can choose exactly the compensation structures we wish to study, rather than being limited to those already in existence. We can consider the treatments to be exogenous in our data analysis, minimising issues such as selection and endogeneity that would be present in observational data from the field, and allowing us to draw conclusions about causality. Additionally, we can collect a fuller set of data than would be available in the field (e.g., we can observe orders whether or not they are fulfilled, and we can collect additional data from traders such as demographics and attitudes toward risk).

Subjects in our experiment, in the role of traders, buy and sell in anonymous continuous double-auction markets over multiple rounds. There are two risky assets, as well as a safe asset (cash). Each unit (share) of either risky asset pays a single dividend at the end of the final round, with known distributions. (See Section 3 for a discussion of this design aspect.) The risky assets have the same fundamental value (i.e., the expected dividend is the same for both), which is constant over all rounds, but one asset is riskier than the other. Traders can buy and sell via bid or ask offers posted to the market (limit orders) or by accepting an offer posted by another trader (market orders).

Our first treatment variable is how traders are incentivised: either by a tournament in which the top half of performers in a session receives a lump-sum additional payment on top of a fixed payment (relative performance), or by approximately linear incentives (absolute performance).[1] Having half winners and half losers means that the tournament is neither convex nor concave; we refer to this as a "linear" tournament in a slight abuse of terminology. This is in contrast to the convex "winner-takes-all" and concave "elimination contest" formats often studied in the literature. Our use of a linear tournament isolates the effect of moving from compensation based on absolute performance to compensation based on relative performance, separate from the effect of making incentives more or less convex. (See Section 3.1 for details about how our compensation schemes are implemented.)

Our second treatment variable is how the variable component of payment is framed to traders: as a bonus (on top of a low fixed payment) or a penalty (subtracted from a high fixed payment). Notably, the bonus and penalty treatments differ only in framing; the actual shape of incentives (whether relative or absolute) is identical between the two. Thus, both of our manipulations are designed to isolate the effect of one component of financial-market compensation schemes, controlling as much as possible for other factors such as the convexity of the compensation scheme.

Our main result is that both manipulations have significant effects. Either moving from a bonus to a penalty frame, or moving from an absolute incentive scheme to a tournament, leads to higher prices for the riskier asset – both in cash and as a premium over the price of the less risky asset – indicating higher demand for risk in these treatments. Thus, ***even after controlling for the convexity of incentives***, tournaments entail increased risky behaviour (versus payment based on absolute performance), as does loss framing (versus gain framing). These treatment effects are significant not only statistically, but also economically: they imply that tournaments increase the high-risk asset's price by 29 percent of its fundamental value, while loss framing increases it by nearly 13 percent of fundamental value.

In addition to this main result, we observe several secondary results involving gender effects. At the aggregate level, we find that markets with a higher fraction of female traders have higher bid-ask spreads for both of the risky assets. At a disaggregated level, we find that females post fewer limit orders and more market orders, and earn less in the experiment, than males. Finally, we find evidence that these last two results are related, since even after controlling for gender, subjects with a larger fraction of market orders earn less in the experiment. We view these results, while noteworthy, as merely suggestive, since they are not connected to our main research questions.

2. Literature review

The relationship between incentives and behaviour in financial markets and related settings has received substantial attention in the economics and finance literatures. Convex, concave, and approximately linear incentive structures coexist in these settings (Brown et al., 2001; Massa and Patgiri, 2008; Ma et

[1] We use "tournament" and "relative performance" interchangeably throughout the paper.

al., 2019). Convex incentives can arise in many ways: from tournaments with few winners (e.g., competing within a firm for a promotion), from limited liability, from option-type compensation, and so on. Concave incentives can arise from fee structures where fees increase more slowly than assets under management, from tournaments with few losers (e.g., competing within a firm to avoid being terminated), and so on.

Early theoretical examinations of links between convex incentives and increased risk include Carpenter (2000) for option-based compensation and Hvide (2002) for rank-order tournaments (the latter growing out of an earlier literature on incentives and effort; see, for example, Lazear and Rosen, 1981). Hvide and Kristiansen (2003) show that increases in contestant ability can sometimes reduce winners' ability, due to risk taking by weaker contestants. Taylor (2003) shows that when fund managers compete against benchmarks, they track an index after initial success but gamble after initial failure (an empirical analysis supports his finding). Along the same lines, Chen et al. (2009) find that fund managers increase their "tracking error" volatility as their relative performance declines. Similar results have been obtained from theoretical studies involving both convex and concave incentives. Gaba et al. (2004) and Fang et al. (2017) find more aggressive trading in winner-takes-all contests (where only a minority – perhaps of one – wins), but less aggressive trading in elimination contests (where a minority loses), and Sato (2016) finds a similar connection between asset-price bubbles and how much of a winner-takes-all situation fund managers are in. On the other hand, some theorists have argued that the effects of the shape of incentives on risk taking will be moderated by the existence of longer time horizons (Hodder and Jackwerth, 2007; Panageas and Westerfield, 2009).

Empirical support for a relationship between tournament incentives and increased risky trading, using data from the field, has been mixed. Several researchers have reported a positive relationship between convex incentives and risk taking in managers of mutual funds, hedge funds, or other portfolios (Brown et al., 1996; Chevalier and Ellison, 1997; Coval and Shumway, 2005; Elton et al., 2003; Kempf and Ruenzi, 2013; Lee et al., 2019) and in CEOs (Coles et al., 2006). There is also evidence of the corresponding opposite-signed effect from concave incentives (Chevalier and Ellison, 1997; Massa and Patgiri, 2008), along with a combination of the two (Kempf et al., 2009).[2] However, other researchers have concluded that there is no systematic effect (Aragon and Nanda, 2012; Brown et al., 2001; Busse, 2001). We know of no previous studies examining linear tournaments like those used in our study, though of course one can attempt to make inferences about such tournaments based on results from their convex and concave counterparts.

Some economists have attempted to uncover a relationship between tournaments and risky behaviour using lab experiments. A small-scale experiment by James and Isaac (2000) found suggestive evidence that a tournament with convex incentives yields more pronounced asset-price bubbles than linear absolute-performance incentives. Fang et al. (2017) find mixed support for their theoretical result (noted above) in an accompanying lab experiment: risky-asset prices are indeed higher under winner-takes-all than under absolute-payoff incentives, but they are also higher (though less so) in elimination contests than under absolute incentives, and none of these differences are significant. Gaba and Kalra (1999) similarly found greater risk taking as the proportion of winners in a contest decreases, though unlike Fang et al., they did not have a treatment with linear absolute incentives. Eriksen and Kvaløy (2014) report that under convex tournament incentives, subjects take more risks the more frequently their investments are evaluated (the opposite of the implication of myopic loss aversion, noted by Benartzi and Thaler, 1995). Eriksen and Kvaløy (2017) find that subjects in convex tournaments take risks even when the only rationalisable strategy is to take no risk, and they take more risks the more competitive the tournament is. Kuziemko et al. (2014) observe that individuals are last-place averse and take more risk when at the very bottom of a ranking, Feltovich and Ejebu (2014) find that individuals take more risks when provided with payoff-irrelevant ranking information, and Kirchler et al. (2018) find that both

[2] Some researchers have also examined direct links between convex incentives and bank instability (Bai and Elyasiani, 2013) and misreporting (Armstrong et al., 2013).

payoff-irrelevant ranking information and payoff-relevant convex tournament incentives can increase risk taking among underperformers.

There have also been experimental studies on the impacts of bonus and penalty contracts on risky behaviour, with mixed results.[3] Framing payoffs as penalties has been observed to result in higher bidding in contests (Chowdhury et al., 2018) and greater cognitive effort in risky settings (Oblak et al., 2018). Most of the other relevant studies, including those with financial-market settings, varied the shape of incentives along with bonus/penalty framing. Gilpatric (2009) found that adding an explicit penalty for finishing last (a 'stick') to the customary prizes for the winner (a 'carrot') can curb risk taking in a winner-takes-all contest. Bonuses and penalties have also been compared in asset-market settings. Lefebvre and Vieider (2014) find more risk taking by executives when remuneration includes share options (inducing a convex payoff function) than when it only includes shares (so payoffs are linear). Kleinlercher et al. (2014) similarly report that risky-asset price deviations above fundamental value are highest when traders can receive bonuses for high performance and lowest when they receive penalties for low performance, with intermediate prices under linear incentives. (See also Holmén et al., 2014, who observe higher asset prices under bonuses, but did not include a penalty treatment.) On the other hand, Paul et al. (2019) find that deviations above fundamental value are highest under linear incentives, and that offering either a bonus or a penalty reduces them to roughly fundamental value, while offering both bonuses and penalties brings them back up to at least as high as they were under linear incentives.

Our study contributes to this literature in several ways. First, we vary the incentive scheme (absolute payoffs versus tournaments) while holding constant the convexity of the scheme; this contrasts with previous studies that focus on the effects of changes in convexity. Second, we vary the framing of the incentive scheme (gains versus losses) while holding the scheme itself constant; this is in contrast to studies that vary both framing and convexity. Both of these manipulations would likely be impossible to achieve outside of a lab environment, justifying the use of a lab experiment. Moreover, by manipulating both of these variables in a 2x2 design, we can examine whether the effect of one of these variables depends on the other variable. To our knowledge, this has not been done before, though as our results section shows, evidence of an interaction between these variables is not especially strong.

3. Experimental design and research questions

Our experimental design builds on Kleinlercher et al.'s (2014); we deliberately follow most of their design and procedures, to allow comparison of our results with theirs and others that have used similar protocols.[4] Most notably, we follow them by using assets with a constant fundamental value (FV), induced by a single dividend paid at the end of the final round, with no payments before then. This differs from the many previous asset market experiments (e.g., Smith et al., 1988; King et al., 1993; Caginalp et al., 2000; Eckel and Füllbrunn, 2015; Baghestanian et al., 2017) in which dividends are paid in all rounds, so that their assets' FVs are declining rather than constant. Declining FVs give rise to the possibility of "bubbles" in which the asset's price also declines over time, but more slowly than does the asset's FV.[5] Oechssler (2010) and Holt et al. (2017) argue persuasively in favour of using a constant FV in controlled experiments, to avoid the difficulty of differentiating between true bubbles due to

[3] In addition to risk taking, there is a literature on the impacts of bonus and penalty contracts on effort levels, with the typical finding that penalty contracts induce greater effort than bonus contracts. This effect has been observed in lab experiments (Luft, 1994; Hannan et al., 2005), field experiments with factory workers (Hong et al., 2011; Hossain and List, 2012), and field experiments with teachers and students (Fryer et al., 2012; Levitt et al., 2013). Explanations include loss aversion (Kahneman and Tversky, 1979; Cacioppo and Berntson, 1994) and a fear of losing or failing (Morgan and Sisak, 2016).

[4] Thus the results of our experiment will be most directly comparable with Kleinlercher et al.'s (2014), and with other studies using similar designs (e.g., Holmen et al., 2014), and less comparable with other studies that used markedly different designs and/or procedures.

[5] Smith et al. (2000) conduct a controlled experiment where they vary the timing of asset dividends, and report more bubbles when FVs are declining (dividends paid every round) than when they are constant (dividend paid once, after the final round).

speculation versus mispricing due to inertia or traders' failure to recognise the changes in FV, as well as for the sake of increased realism.

Subjects act as fund managers with a portfolio of two risky assets and lab money (called "talers" in the experiment). In each market, ten subjects trade for ten rounds.[6] The assets differ in their dividend distributions. The lower-risk "asset L" pays a final dividend of either 20 or 30 lab-talers with equal probabilities, while the higher-risk "asset H" pays a final dividend of 65 lab-talers with a probability of 20 percent and 15 otherwise. Hence, both assets have equal expected dividends of 25 lab-talers, but they differ in variance (higher for asset H) and skewness (positive for H, zero for L).[7]

At the beginning of the first market round, each subject was endowed with a cash balance of 2000 lab-talers and 20 units of each asset. Subject could buy and sell assets in each round via a continuous double auction with an open order book. Subjects could be active in both markets simultaneously, and both markets were displayed on subjects' screens at the same time. Each trading round lasted 120 seconds. Subjects could not purchase more units than they could afford nor sell more units than they had in their inventories (i.e., no negative cash balances and no short selling). Balances of cash and assets were carried over from round to round. No interest was paid on cash holdings and there were no trading fees or commissions. In the final round, each subject's wealth was equal to her dividends plus her cash balances.

3.1 Experimental treatments

The experiment uses a 2x2 factorial design, where we vary the incentive scheme and its framing.

Cell AB (Absolute performance, Bonus frame): Each subject receives a fixed payment of 20 Chinese yuan (CNY), plus a bonus of 0.01 CNY for each 0.5 lab-talers of final wealth exceeding 1000 lab-talers, up to a maximum of 100 CNY (for wealth of 5000 lab-talers or more).

Cell AP (Absolute performance, Penalty frame): Each subject receives a fixed payment of 100 CNY, with a deduction of 0.01 CNY for each 0.5 lab-talers of final wealth below 5000, down to a minimum of 20 CNY (for wealth of 1000 lab-talers or less).

Cell TB (Tournament, Bonus frame): Each subject receives a fixed payment of 20 CNY, with a bonus of 80 CNY for those in the top 50 percent of final wealth in the session.

Cell TP (Tournament, Penalty frame): Each subject receives a fixed payment of 100 CNY, with a penalty of 80 CNY for those in the bottom 50 percent of final wealth in the session.

Hence cells AB and AP are identical except for framing, as are cells TB and TP. Our AB and AP cells correspond roughly to Kleinlercher et al.'s (2014) cap and penalty treatments respectively. While we know of no examples from the outside world of pairs of incentive schemes differing only in gain/loss framing, examples abound of both bonuses and penalties in remuneration.[8] If such framing matters, this

[6] Following Kleinlercher et al. (2014), as well as Holmen et al. (2014), we told subjects in the instructions that the markets would end at a randomly chosen round between 8 and 14, and pre-determined the realised final round would be round 10. This was done in order to facilitate comparisons across our treatments, as well as with previous work.

[7] The assets' return distributions are the same as was used by Kleinlercher et al. (2014). While positive skewness in our higher-risk asset means that subjects' tastes for skewness may come into play (see, e.g., Eckel and Grossman 2015), the advantage of positive skewness is that it allows us to increase the variance of returns substantially without the risk of negative returns to subjects, which are problematic due to human-ethics considerations and lack of enforceability.

[8] As one example of penalties in finance, Kleinlercher et al. (2014) note that in the 1950s and 1960s, financial professionals' incentives included "hardly any upside potential in the sense of bonus payments, but downside risk" (p. 140), suggesting a penalty frame in those incentives.

will have clear implications for organisations offering them, since changing the framing of incentives is likely to be less costly than changing the incentives themselves.

Less obviously, there is a close correspondence between our AB and TB cells (and similarly between AP and TP). Subject earnings in our absolute-performance incentive scheme are linear over a wide range of wealth levels: between 1000 and 5000 lab-talers, with a ceiling on earnings at 5000 and a floor at 1000. Subjects' initial endowments (2000 lab-talers and 20 units of each asset, with fundamental values of 25 lab-talers each) have an expected value of 3000 lab-talers – the centre of this interval – so subjects would need to take on substantial additional risk in order to raise the possibility that wealth increases beyond 5000 lab-talers or decreases below 1000 lab-talers.[9] The lower and upper earnings bounds under absolute-performance incentives correspond to the earnings for winners and losers in the tournaments.

Our tournament structure, which rewards or punishes half of the traders, is intermediate between the winner-takes-all (only one winner) and elimination-contest (only one non-winner) treatments studied by Fang et al. (2017) and others. Each subject's expected money earnings at the beginning of the experimental session are 60 CNY: exactly the same as expected earnings under absolute-performance incentives. Hence, if no subject were to trade, each subject's expected money earnings would be identical under either absolute-performance or tournament incentives. The treatments differ in how earnings change as a trader moves away from this initial point: linearly under absolute-performance incentives, versus discontinuously in the tournament (a final wealth even one lab-taler above the other traders will earn 100 CNY, while even one lab-taler less than the others will earn only 20 CNY). The similarities between our treatments allow us to identify the effect of moving between compensation based on relative performance versus absolute performance, as distinct from the impact of making the relationship between pay and performance more or less convex.

While the literature has concentrated on tournaments similar to winner-takes-all and elimination-contest, there are many real examples of tournaments in which half or nearly half are successful. This is especially true in sports: about half of teams qualify for post-season play in the US's National Basketball Association (16 of 30 teams), National Hockey League (16 of 31), and top-level college football (80 of 130 teams are invited to bowl games), as well as the Australian Football League (8 of 18) and Japan's Nippon Professional Baseball (6 of 12). In the Scottish Premiership (soccer), there is no post-season, but with five matches remaining in the season, the league "splits", with the top 6 of 12 teams playing the remaining matches amongst themselves, and with only those teams eligible for the premiership.

Such tournaments also exist in the finance industry and in other business settings, though systematic study there is complicated by the confidentiality of many employment contracts. Since 2008, China's finance ministry has maintained a policy of share-based incentive compensation for state-owned companies, in which shares can be awarded to managers if the company's performance is at or above the fiftieth percentile of comparable companies. As a result, even many private companies use this threshold for awarding bonuses. Churchill (1990) cites survey evidence that almost 40 percent of sales contests in US firms had the odds of being a winner between 31 and 70 percent (see Kalra and Shi, 2001, for a discussion). Churchill also notes a common perception among executives that motivation is highest

[9] Trading at fundamental values, a subject could reach 5000 lab-talers in the most favourable outcome (dividends of 30 and 65 for assets L and H) by trading all of her low-risk assets for high-risk assets, and buying an additional 10 units of the high-risk asset with cash. (An even riskier portfolio would be required to have a chance of falling below 1000 lab-talers – even in the worst possible outcome.) Indeed, most decisions in the experiment did take place in the linear region of subjects' incentive schemes. In all subject-rounds of the AB and AP cells, no-one ever held assets worth less than 1000 lab-talers, even in the worst state of the world, and subjects held assets worth more than 5000 lab-talers in the best state in only 1.7 percent of subject-rounds. Even so, it is still possible that the minimum and maximum earnings could have affected the behaviour of subjects who were currently in the linear region of the incentive scheme. For this reason, we call these incentives "approximately linear".

when all salespeople believe they have a reasonable chance to win (and presumably, to lose). Berger et al. (2013) examine a repeated tournament setting at a major hotel chain's reservation centre where the fraction of winners is 40 percent. Finally, these kinds of tournaments are fairly common in experimental studies, again due to the perception that they induce more effort from subjects; see, for example, Tong and Leung (2002), Harbring and Irlenbusch (2008), Knauer (2016) and Kelly et al. (2017).[10]

3.2 Research questions

Our main research questions are driven by our treatment variables. Our first treatment variable is the compensation scheme: absolute performance or relative performance (tournament). Several previous studies (see our literature review above) have found higher levels of risky behaviour in tournaments, though typically these have involved convex incentives.

Research question 1: What is the effect of relative-performance incentives (compared to absolute-performance incentives) on (a) the price of the high-risk asset, (b) the price of the low-risk asset, (c) the price premium for the high-risk asset?

Following the previous literature, we conjecture that risky behaviour will be more prevalent in tournaments, even though our experiment (unlike much of this previous literature) controls for the convexity of incentives. That is, while previous results have been partly attributable to differences in convexity, we expect that they are also partly attributable to tournament incentives per se. For our experiment, this means that the price of asset H, and its premium over asset L (i.e., the difference in the two assets' prices) should be ***higher*** under relative-performance incentives. However, it is less clear what this implies for the price of asset L, which is riskier than cash but less risky than asset H. If an increase in risk simply means a desire to take on more of the highest-risk asset, demand for asset L may be no higher under tournament incentives, so that its price may be similar to that under absolute-performance incentives. By contrast, if an increase in risk means substituting away from cash towards all of the risky assets, demand for asset L (and hence its price) could be higher under tournament incentives.

Our second treatment variable is the framing of incentives: as a bonus or a penalty. As noted in our literature review, previous work has found mixed results. This is illustrated well by the two papers closest to ours in the literature: Kleinlercher et al. (2014) and Paul et al. (2019), who find respectively that risky asset prices are substantially higher (Kleinlercher et al.), or nearly equal (Paul et al.), under bonuses compared to penalties. We surmise that any systematic effect of penalty versus bonus framing may have been swamped by changes to monetary incentives. In both of these previous studies, the treatments varied not only in framing, but in the actual shape of the payoff function, and in different ways in the two studies (see Figure 1, p. 141 in Kleinlercher et al. and Figure 2, p. 451 in Paul et al.). Our AB and AP treatments have identically-shaped payoff functions, and therefore differ only in how these payoffs are framed. The same is true for our TB and TP treatments. Our experiment thus allows a cleaner test of positive versus negative framing.

Research question 2: What is the effect of penalty framing (compared to bonus framing) on (a) the price of the high-risk asset, (b) the price of the low-risk asset, (c) the price premium for the high-risk asset?

Standard theory provides limited insight into these research questions. If traders are risk-neutral expected-utility maximisers, both risky assets should be priced at the fundamental value of 25. Under absolute-performance incentives, identical risk-averse traders will bid the asset prices below this fundamental value (and to a greater degree for the riskier asset H), while identical risk-seeking traders

[10] Additionally, many simple theoretic contest models (e.g., "toy" models used for illustration by theorists, and environments used in lab experiments) involve two people competing for one prize, so that 50 percent are winners. (See, e.g., Grund and Sliwka, 2005.)

will do the opposite. So, if our framing manipulation is successful in changing reference points, and if traders fall prey to the "reflection effect" (Kahneman and Tversky, 1979) where they are risk-averse over gains and risk-seeking over losses, then (a)-(c) will all increase under penalty framing compared to bonus framing – when absolute-performance incentives are being used (our AP versus AB cells). Thus, we conjecture that under absolute-performance incentives, prices of both risky assets and the price premium for the high-risk asset should be ***higher*** under penalty framing.

Under tournament incentives, there are only two possible real-money payments: 100 CNY for tournament winners and 20 CNY for tournament losers. So as long as preferences are over only money amounts, all that matters to traders will be the probability of winning. (In particular, attitudes toward risk and losses are irrelevant, as are outcome-based social preferences like those of Fehr and Schmidt, 1997.) If traders are homogeneous, in equilibrium all will hold the same portfolio, so each will win with probability one-half. Since the low-risk asset earns the higher dividend with probability one-half, buying or selling at the fundamental value will not change the win probability, so its equilibrium price will be the fundamental value of 25. However, the high-risk asset only earns the higher dividend with probability one-fifth, meaning that at a price of 25, a trader can increase her likelihood of winning by selling this asset; hence, its equilibrium price will be ***less*** than 25. In fact, this argument holds at any price higher than the high-risk asset's low dividend of 15, so the equilibrium price will be 15. No differences are predicted between the TB and TP cells.

We are reluctant to take this last set of implications too seriously, as there is plenty of evidence from lab experiments that traders systematically violate theoretical predictions in tournaments. A notable example is Nieken and Sliwka's (2010) observation of excessive risk taking in a tournament setting similar to ours in some ways (choice of risk rather than effort, realisations of the risky outcome perfectly correlated across individuals).[11] Nonetheless, we have no ex-ante conjecture regarding differences in asset prices between bonus and penalty framing when tournament incentives are present. It is possible to construct preferences that would imply higher equilibrium prices for the risky assets, but these would likely have to involve tastes for ranks in addition to money. (E.g., if utility is sufficiently convex in higher placement, traders will want to buy the high-risk asset at the fundamental value, leading to a higher equilibrium price.) Evidence of tastes for rankings has been observed (see, e.g., Kirchler et al., 2018, discussed in the literature review above), and other results provide indirect evidence of such tastes (e.g., Berger et al., 2013, find that social comparisons are amplified when relative-performance information is included in feedback), though a theoretical model incorporating these is beyond the scope of the current paper.

3.3 Experimental procedures

The experiment was conducted at an experimental economics lab in Hebei University of Economics and Business (China). It comprised 36 sessions (markets) – 9 of each cell – with 10 subjects in each market, for a total of 360 subjects. Subjects were recruited from undergraduate and postgraduate courses. Many of the subjects had taken part in previous experiments, but none had participated in asset-market experiments previously, and none participated more than once in our study. The experiment was computerised using z-Tree (Fischbacher, 2007).

In a trading round, each subject could participate as buyer, seller, or both. Offers (bid or ask) for either asset (H or L) comprised price-quantity combinations, where the price could range from 0 to 999 lab-talers (inclusive, with up to two decimal places), and the quantity was a whole number of units. The maximum quantity was determined by the constraint that neither cash nor asset holdings could drop below zero. Once an offer was made, it appeared on everyone's screens, and could be accepted by any other subject with sufficient cash or shares. Thus a subject could trade either by posting a (price,

[11] There is also substantial evidence of other kinds of deviations from equilibrium, such as over-expenditure of effort in tournaments where effort is endogenous. See Sheremeta (2013) and Dechenaux et al. (2015) for surveys.

quantity) offer or by accepting a price posted by another subject (limit orders and market orders, respectively). In the latter case, the subject was able to trade the quantity offered by the poster, or a lower quantity, but not a higher quantity. Accepted offers were executed immediately, with the buyer's cash balance reduced by the expenditure (price x quantity) and his holding of the asset increased by the quantity, and the opposite for the seller.

After 120 seconds, the trading round ended and subjects received feedback that included their current cash and asset holdings. Subjects were also informed of how these holdings would correspond to actual money earnings (hypothetically, if this had been the last round of the session), under each of the four scenarios corresponding to realisations of the asset prices (the 10-percent chance both assets earn the higher dividend, the 40-percent chance asset H earns the lower dividend but asset L earns the higher dividend, and so on). In cells AB and AP, these included calculations of the resulting bonus or penalty, along with the corresponding fixed payment. In cells TB and TP, the bonus or penalty was calculated based on the subject's current ranking within the group.[12] Subjects were given 15 seconds to view their feedback, after which the next round would start.

Each session lasted approximately 80 minutes (including instructions, all tasks, and payment). Before the first market round, we collected information on demographics, risk attitudes (Holt and Laury, 2002) and loss aversion (Abeler et al., 2011).[13] Subjects were paid their earnings (to the nearest 0.01 CNY) immediately after the last market round, privately and in cash. Average payments were just under 60 CNY for the market rounds alone, and about 67 CNY overall, as compared to the local minimum hourly wage of 17 CNY.[14]

4. Results

Of the 360 subjects, roughly two-thirds (68.6 percent) were female, reflecting the gender proportions at the university overall. Nearly all (98.3 percent) were undergraduates, and nearly all were either business students (13.9 percent) or science/engineering students (83.0 percent). We did not observe any significant differences in demographics and attitudes to risk or loss across treatments (Kruskal-Wallis test, session-level data, $p \approx 0.143$ for gender, $p > 0.20$ for all other variables).[15] Also, in contrast to other studies, we do not find large gender differences in attitudes to risk or loss. Females are slightly more risk averse, and slightly less loss averse, than males, but these differences are not significant. Nonetheless, we typically control for demographics and attitudes in the regressions below.

Table 1 summarises some of the aggregate results from the experiment.[16] For each cell, this table shows the average price of each asset (in lab-talers), the volume of transactions of each asset (in units per session-round), and average subject earnings from the main part of the experiment (in local currency). These variables are analysed in more detail below and in the appendix.

[12] That is, subjects were shown, for each of the four scenarios, what their ranking within the group would be.
[13] Both the Holt-Laury and the Abeler et al. task involve ordered lists of binary choices, raising the possibility that a subject will submit non-monotonic responses (e.g., violating stochastic dominance in Holt-Laury). This happened for only 4.4 percent of subjects in the Holt-Laury task and 3.1 percent in the Abeler et al. task. Both tasks were paid, but their realisations were not told to subjects until the end of the session (after the market rounds). Higher values of our risk- and loss-aversion measures correspond to ***less*** tolerance for risk or loss. In Appendix B, we report session-level demographic and attitudinal variables, as well as Kruskal-Wallis tests that confirm these variables are balanced across treatments.
[14] Average payments for the market rounds were exactly 60 CNY in each session of the tournament cells, and averaged 59.10 CNY in the absolute-performance cells. Since average subject payments are similar across treatments, a potential confound (differences in the level of expected earnings) is avoided in our experiment.
[15] See Siegel and Castellan (1988) for descriptions of the non-parametric statistical tests we used, and Feltovich (2005) for critical values of the robust rank-order test statistic used later in this section.
[16] See Appendix B for additional session-level data, including other measures of market mispricing (see Stöckl et al., 2010, for a discussion of these measures).

Table 1: Aggregate results by cell (prices, volumes, and profits from main part of experiment)

Cell	Asset H (high risk)		Asset L (low risk)		Subject earnings (main part of experiment)	
	Price	Volume	Price	Volume	Mean	Std. deviation
AB	26.46	19.93	24.78	20.39	58.0	8.7
AP	30.38	18.74	22.94	16.44	60.2	11.5
TB	30.74	34.48	24.60	25.12	60.0	40.2
TP	37.05	32.46	25.95	25.77	60.0	40.2

Notes: price measured in lab-talers, volume in units per session-round, earnings in CNY.

4.1 Asset prices

Table 2 shows the average price of each asset in the four cells of the experiment, expressed as relative deviations from the fundamental value of 25 (that is, the observed price minus 25, then divided by 25), along with asset H's "price premium" over asset L (the difference in their relative deviations). Also shown are results of non-parametric statistical tests of treatment effects. The price of asset H is closest to its fundamental value (i.e., its relative deviation is closest to zero) in the AB cell. From there, introducing either tournament incentives or penalty framing (TB or AP) significantly increases average prices, and adding the other from there (TP cell) further increases prices, though that last increase is either insignificant or weakly significant. Treatment effects for the asset-H premium are similar: significantly larger in the TB and AP cells than in the AB cell, and larger (but insignificantly so) in the TP cell than in those.

The price of asset L is close to the fundamental value in all four cells. There are some statistically significant differences across the cells, with prices lower in the AP cell compared to the AB cell or the TP cell, though it is questionable whether these differences are economically relevant, given their small magnitudes.[17]

Table 2: Treatment-level average prices (relative deviations from fundamental value, all rounds)

Asset H	Bonus	Penalty	p-value	Asset L	Bonus	Penalty	p-value
Absolute	+5.85	+21.54	***0.011***	Absolute	−0.89	−8.26	0.065
Tournament	+23.00	+48.20	0.119	Tournament	−1.60	+3.80	>0.20
p-value	***0.016***	0.079		p-value	>0.20	***0.002***	

Asset-H premium (asset H – asset L)	Bonus	Penalty	p-value
Absolute	+6.74	+29.79	***<0.001***
Tournament	+24.60	+44.39	0.169
p-value	***0.006***	>0.20	

Note: p-values based on two-tailed robust rank-order tests, session-level data

Figure 1 shows the time series for both risky assets' prices, again expressed as percent deviations from fundamental value (see the Appendix for individual sessions' time series). The only apparent time trend is in the high-risk asset in the AP treatment, though there is some evidence of a declining price for the low-risk asset as well in this treatment. The other prices are fairly stable across rounds.

[17] The significance tests do not change in substantial ways when the final-round data are used instead of all-round averages. When round-10 data are used, the difference in high-risk asset price between the AB and AP cells becomes insignificant, but the differences in high- and low-risk asset prices between the TB and TP cells become significant. The significance of treatment effects on the asset-H premium is unchanged if we use round 10 instead of all rounds.

Figure 1: Time series of asset prices (relative deviations from fundamental value), by cell

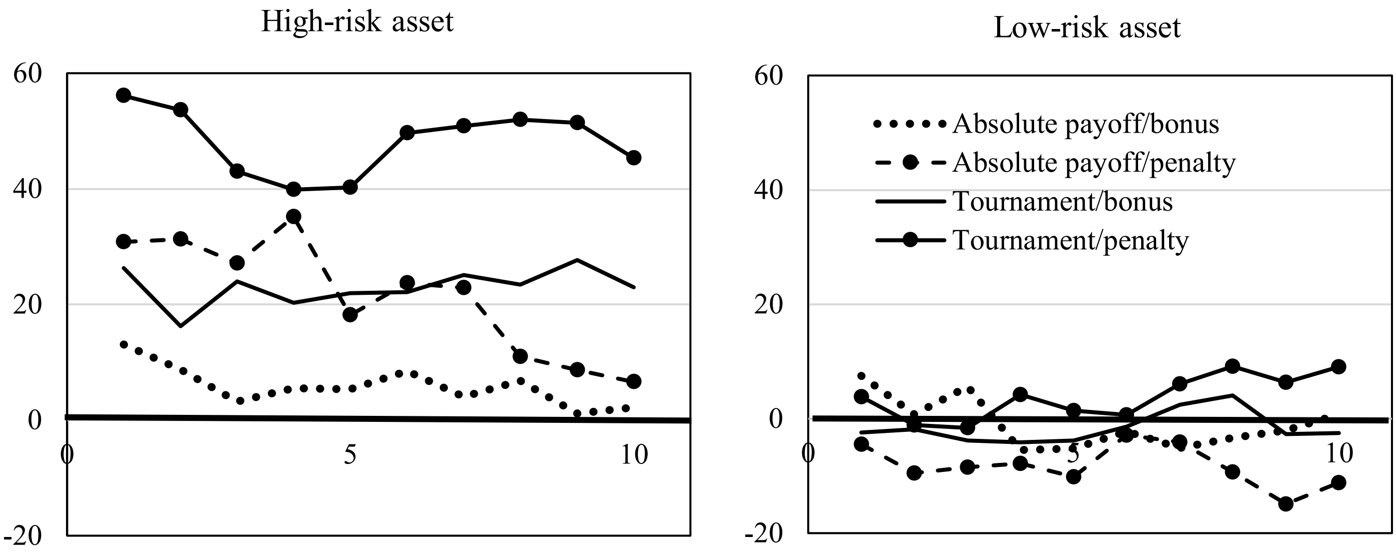


Table 3: Marginal effects (MEs) from panel linear regressions
(session-round data, N=360, std. errors clustered by session)

Dependent variable:	(1) Asset-H price	(2) Asset-L price	(3) Asset-H premium
Penalty treatment indicator…			
…avg. ME	+12.758*** (4.691)	−2.911 (2.777)	+15.669*** (4.265)
…ME in abs. performance treatment	+10.504 (6.447)	−8.943** (4.353)	+19.446*** (6.171)
…ME in tournament treatment	+18.047*** (6.075)	+4.065 (2.505)	+13.982** (5.816)
Signif. of differences	*p > 0.20*	*p ≈ 0.002*	*p > 0.20*
Tournament treatment indicator…			
…avg. ME	+29.249*** (5.417)	+9.851*** (2.583)	+19.398*** (4.747)
…ME in bonus treatment	+19.651*** (4.407)	+0.789 (2.760)	+18.864*** (5.185)
…ME in penalty treatment	+27.195*** (8.590)	+13.797*** (3.369)	+13.398* (8.038)
Signif. of differences	*p > 0.20*	*p ≈ 0.002*	*p > 0.20*
Round	−0.887 (0.607)	+0.010 (0.314)	−0.897 (0.565)
Fraction females in session	−2.464 (13.550)	−18.819** (8.713)	+16.355 (11.700)
Demographics, attitudes on right-hand side?	Yes	Yes	Yes
r-squared	0.497	0.247	0.426

** (**, ***): Significant at 10% (5%, 1%) level.*

Table 3 presents results from three panel linear regressions (generalised least squares with individual subject random effects): one each with the high-risk asset price, the low-risk asset price, and the asset-H premium as the dependent variable. As before, these variables are expressed in terms of relative deviations from fundamental value. The main right-hand-side variables are indicators for the penalty treatment and tournament treatment, the round number (as a single continuous variable), and all two- and three-way interactions of these. We also include session-level demographics and a constant term.[18] The models were estimated by Stata (v. 15), with standard errors clustered by session.

Rather than showing all coefficient estimates, Table 3 displays the most important estimated marginal effects (MEs): average MEs for the penalty-frame and tournament indicators, and MEs for each of these conditional on a particular value for the other (e.g., the ME of the tournament indicator under the bonus frame represents a comparison of the TB and AB cells).

The results largely reinforce what we have seen already. Moving from bonus to penalty framing results in a higher price for the high-risk asset and a larger premium over the low-risk asset: both overall and in the absolute-performance and tournament treatments individually (though one of these differences misses significance, with a p-value of 0.103). Moving from absolute-performance to tournament incentives also results in a higher price and larger premium for the high-risk asset, both overall and in the bonus and penalty treatments individually (though again one of the differences is only borderline significant, with a p-value of 0.096). These effects imply substantial economic significance: the average marginal effect of tournament incentives can be interpreted as an increase of more than 29 percent of the asset's fundamental value, while the corresponding effect for penalty framing is an increase of almost 13 percent of fundamental value.

The effect on the price of asset L is more equivocal. Under penalty framing, it is higher in tournaments, but the corresponding effect under bonus framing is close to zero and insignificant. The effect of penalty framing itself on asset-L prices is negative and significant under absolute performance, and positive but insignificant in tournaments. The economic significance of these point estimates is more limited than their counterparts for the high-risk asset: on the order of 2.9 percent of fundamental value for penalty framing, and 9.9 percent of fundamental value for tournament incentives.

Result 1: The price of the riskier asset – either absolute or relative to that of the less risky asset – is higher under tournament incentives.

Result 2: The price of the riskier asset – either absolute or relative to that of the less risky asset – is higher when the variable component of incentives is framed as a penalty.

There is also some evidence of interactions between our treatment variables. The significant differences in conditional marginal effects for the low-risk asset suggest a positive interaction between penalty framing and tournament incentives (that is, the effect of one variable is larger when the other is present), while there is no corresponding significant difference for the high-risk asset. However, these together would imply that the asset-H premium should show a negative interaction between penalty framing and tournament incentives, and while the sign of the difference between these conditional marginal effects is consistent with this implication (the effect of either penalty framing or tournament incentives is lower when the other is present), the difference is not significant.[19]

4.2 Individual-level results

[18] The demographics were the session-level fractions of females, undergraduates, business students (including economics), average age, average risk tolerance, and average loss tolerance (see section 3.3 for details), along with the products between these last two variables and our treatment indicators. For robustness, we ran additional regressions that included the square of the round number (along with its interactions with the treatment indicators), regressions without the round-number interactions, regressions without the demographic variables, and regressions using session medians instead of means as the dependent variable. These had no qualitative effect on the results (details available from the authors upon request).

[19] To the extent that a difference in effects exists between the low- and high-risk assets, it is similar in spirit to differences observed by Kleinlercher et al. (2014); see, for example, their Table 3 and surrounding discussion.

Table 4: Average marginal effects (MEs) from linear regressions
(subject data, N=360, std. errors clustered by session)

Dependent variable:	Round-10 portfolio balance	Round-10 expected earnings	Round-10 portfolio risk	Total market orders	Total limit orders	Total transaction volume
Penalty treatment	+18.594 (13.114)	−0.384 (0.962)	+0.638 (0.417)	−1.254 (2.120)	+1.009 (3.000)	−6.385 (11.288)
Tournament treatment	+14.800 (12.618)	+0.840 (1.042)	+0.707 (0.489)	+8.455*** (2.330)	+9.751*** (3.050)	+44.004*** (12.131)
Female	−179.188*** (48.035)	−9.181*** (3.345)	−1.034 (0.862)	+6.495*** (1.824)	−8.603*** (2.961)	−9.158 (6.939)
Risk aversion	−18.641* (10.792)	+0.284 (0.650)	−0.025 (0.359)	+0.100 (0.531)	−0.392 (0.760)	−0.652 (2.362)
Loss aversion	+19.175 (14.696)	+3.532*** (1.154)	−0.694 (0.519)	−0.982 (0.912)	+0.473 (0.836)	−1.166 (2.917)
Demographics, attitudes?	Yes	Yes	Yes	Yes	Yes	Yes
r-squared	0.100	0.181	0.049	0.142	0.140	0.148

** (**, ***): Significant at 10% (5%, 1%) level.*

Table 4 shows results from some individual-level regressions. The round-10 portfolio balance is the expected value of a subject's portfolio in lab-talers after all trading has taken place, but before the realised asset dividends become known (i.e., the number of shares of both assets held multiplied by their fundamental value of 25, added to the cash balance). The round-10 expected earnings are based on these asset holdings, but incorporate the incentives of the treatment (bonus or penalty based on rank for the tournament cells, capped cash bonus or penalty for the absolute-performance cells), and are denominated in real money. The round-10 portfolio risk is a unitless measure of the risk of the subject's portfolio, and is given by the formula

$$Risk = 100 \cdot \left[\frac{20x_H + 5x_L}{25x_H + 25x_L + x_C} \right]$$

where x_H, x_L and x_C are the holdings of asset H, asset L and cash respectively. As asset H and asset L have standard deviations of 20 and 5 respectively, and both have expected values of 25, the numerator is a proxy for the portfolio's standard deviation, while the denominator is its expected value. Total market orders are the number of times, over all rounds, that the subject traded by accepting another subject's posted (buy or sell) offer, and total limit orders are the number of times that the subject posted an offer to buy or sell units of an asset. (In both, orders for multiple units are treated as a single order.) Total transaction volume is the number of units of both assets bought or sold (e.g., if a subject bought 2 units of asset H in a round, then sold one of these units in the same round, this would count as 3 units of transaction volume), and is thus a measure of trading activity.

Right-hand-side variables in all of these regressions are indicators for the penalty and tournament treatments, the demographic and attitudinal variables, and the two-way interactions between treatment indicators and demographics/attitudes, along with a constant term. As before, standard errors are clustered by session.

The main result visible in these regression results is the difference between males and females. Females fare worse than males in both measures of performance: expected portfolio balance (lab money) and expected earnings (real money).[20] This seems not to be due to risk attitudes, since (i) we do not observe significant differences between males and females in the risky-choice task, (ii) we control for behaviour in the risky-choice task in our regressions (and it is typically not significant), and (iii) we do not observe gender differences in the measure of portfolio risk.[21] The performance difference may instead be connected to how males and females trade. Though trading volumes are comparable (females trade less than males, but not significantly so), males are more likely to trade by posting offers (limit orders), while females are more likely to trade by accepting an existing posted offer (market orders).

Result 3: Females earn less than males in the experiment, in both lab money and subject payments (real money). Females post fewer limit orders (where they specify the price) and more market orders (where they accept the price posted by another trader).

Table 5: Average marginal effects from linear regressions
(subject data, N=360, std. errors clustered by session)

Dependent variable:	Round-10 portfolio balance	Round-10 expected earnings
Penalty treatment	+4.604 (23.754)	−0.611 (1.216)
Tournament treatment	+72.969** (31.755)	+2.981** (1.457)
Female	−67.471** (32.018)	−2.382 (2.314)
Market orders	−13.251*** (1.627)	−0.757*** (0.108)
Limit orders	+1.262 (0.762)	+0.075** (0.029)
Transaction volume	+1.001** (0.401)	+0.088*** (0.023)
Risk aversion	−14.616* (8.375)	+0.526 (0.626)
Loss aversion	+0.623 (10.904)	+2.571** (1.036)
Other demographics?	No	No
r-squared	0.430	0.360

** (**, ***): Significant at 10% (5%, 1%) level*

Table 5 sheds additional light on the gender differences seen in Table 4, with two additional regressions. The dependent variables are the two performance measures from Table 4: expected final portfolio

[20] This gender difference is arguably economically significant, translating to lower earnings for females of about USD 1.50 per subject (or alternatively, females earning about 87 percent of males). Additional regressions, not reported here, show that females' worse performance is found in all four cells of the experiment, though this gender difference is not significant in the TB cell.

[21] Note from Table 4 that we do not observe many significant average effects of either our loss- or risk-aversion measures on these variables, after controlling for gender and other demographics. Even examination of the interaction effects (not shown, but available from the authors upon request) yields few significant results. One notable exception is the interaction between loss aversion and the tournament treatment, which has a significantly positive effect on expected earnings ($p \approx 0.002$), a weakly positive effect on portfolio balance ($p \approx 0.073$), and a significantly negative effect on portfolio risk ($p \approx 0.042$). (The interaction between risk aversion and the tournament treatment has the same sign of effects on these three variables, but none is significant.)

balance in lab money, and final expected earnings in real money. Our explanatory variables are the indicator for female and the trading-activity variables – total market orders, total limit orders, total transactions – as well as the treatment indicators and a constant term.

The connection between trading style and performance, implied by Table 4, is even more evident in Table 5. Higher numbers of market orders are significantly associated with lower earnings. Higher numbers of limit orders are associated with higher earnings, though not always significantly.[22]

Result 4: Subjects who place more market orders and fewer limit orders earn less (in both lab money and real money), though for limit orders the effect is not always significant.

The table also shows that much of the gender effect on earnings observed in Table 4 is attributable to these behavioural differences. Controlling for market and limit orders reduces the gender effect by about 60 percent (for final portfolio balances) or 75 percent (for money earnings), though in the case of final portfolio balances, the gender effect is still significant even after these are controlled for.

4.3 Additional results

We continue the data analysis by examining some of the other market-level variables. These are not closely connected to our research questions, but appear often enough elsewhere in the literature (e.g., Kleinlercher et al., 2014; Paul et al., 2019) that we include them here for completeness. Table 6 shows, for each treatment, the per-round volume of transactions of assets H and L, and the average bid-ask spread.[23]

The main difference across treatments is in the higher trading of the higher-risk asset in the tournament treatment compared to the absolute-performance treatment, which is consistent with the subject-level results in Table 4. We also observe higher trading of the lower-risk asset, but the difference is not significant.[24] We do not observe systematic differences between treatments in bid-ask spreads in either asset.

Table 6: Additional treatment-level averages (all rounds)

Asset H transactions	Bonus	Penalty	p-value	Asset L transactions	Bonus	Penalty	p-value
Absolute	19.93	18.74	>0.20	Absolute	20.39	16.44	0.17
Tournament	34.48	32.46	>0.20	Tournament	25.12	25.77	>0.20
p-value	***0.038***	***0.009***		p-value	>0.20	0.073	

Asset-H bid-ask spread	Bonus	Penalty	p-value	Asset-L bid-ask spread	Bonus	Penalty	p-value
Absolute	7.85	13.84	>0.20	Absolute	7.41	13.03	0.081
Tournament	9.37	15.18	>0.20	Tournament	9.11	14.58	>0.20
p-value	>0.20	>0.20		p-value	0.18	>0.20	

Note: p-values based on two-tailed robust rank-order tests, session-level data

Table 7 presents results from four panel linear regressions: transaction volumes of the two assets and bid-ask spreads of the two assets. The methodology is similar to that in the regressions of Table 3. As

[22] Additional regressions, not reported here, confirm that market orders for both asset L and asset H are associated with worse performance on both measures.
[23] We follow Kleinlercher et al. (2014) in the definitions of these variables; see there for details.
[24] See the appendix for a discussion of differences in trading of the high-risk asset relative to the low-risk asset, and how this varies by treatment.

was true there, explanatory variables are indicators for the penalty treatment and tournament treatment, the round number, all two- and three-way interactions of these, the session-level demographics and a constant term, and standard errors were clustered by session.

Table 7: Marginal effects (MEs) from panel linear regressions
(session-round data, std. errors clustered by session)

Dependent variable:	Asset H transactions	Asset L transactions	Asset H bid-ask spread	Asset L bid-ask spread
Penalty treatment indicator…				
…avg. ME	−0.919	+0.091	+2.117	+2.148
	(3.533)	(2.680)	(2.315)	(2.139)
…ME in abs. performance treatment	−0.032	−2.028	+2.078	+1.995
	(3.313)	(2.667)	(3.429)	(3.204)
…ME in tournament treatment	−1.805	+2.209	+2.154	+2.286
	(6.191)	(4.455)	(3.806)	(3.505)
Signif. of differences	$p > 0.20$	$p > 0.20$	$p > 0.20$	$p > 0.20$
Tournament treatment indicator…				
…avg. ME	+14.165***	+7.678***	+1.605	+1.576
	(3.848)	(2.832)	(2.934)	(2.783)
…ME in bonus treatment	+15.052***	+5.560*	+1.568	+1.428
	(4.391)	(3.005)	(2.396)	(2.277)
…ME in penalty treatment	+13.279**	+9.797**	+1.644	+1.719
	(5.888)	(4.428)	(5.203)	(4.847)
Signif. of differences	$p > 0.20$	$p > 0.20$	$p > 0.20$	$p > 0.20$
Round	−1.875***	−1.479***	−0.490***	−0.428***
	(0.238)	(0.278)	(0.143)	(0.149)
Fraction females in session	+4.598	+0.064	+13.326*	+13.227**
	(10.328)	(6.930)	(6.928)	(6.625)
Demographics, attitudes on right-hand side?	Yes	Yes	Yes	Yes
N	360	360	357	351
r-squared	0.329	0.246	0.272	0.247

The increased trading of both risky assets under tournament incentives, seen in Table 6, is present in these regression results as well, and the differences are significant for both assets (though a regression using the pooled low- and high-risk asset data shows additionally that the effect of tournament incentives is significantly larger for the high-risk asset than the low-risk asset) and in both bonus and penalty treatments.[25] By contrast, there are no significant differences in either transaction volumes or bid-ask spreads between bonus and penalty framing. The fraction of female traders in a market has a significant positive effect on bid-ask spreads, and an insignificant (though also positive) effect on transaction volumes.

[25] Since both trading volume and asset prices are higher in tournaments, we might conjecture a more systematic positive relationship between the level of trading and asset prices in these markets. However, that turns out not to be the case. While there are positive relationships between asset prices and asset volumes in the aggregate data (pooling all experimental cells yields a correlation coefficient of +0.422 for the high-risk asset and +0.347 for the low-risk asset, $p \approx 0.011$ and 0.361 respectively), these relationships become insignificant after controlling for the cell, with no significant relationships at the session- or session-round-level between volume and prices. Hence, the results in Table 7 are distinct from those observed previously (e.g., in Table 3), and we consider it to be unlikely that there is a causal relationship (in either direction) between price and volume in our data. As further evidence of the lack of a causal relationship, regressions similar to those in Table 3 but including the two assets' volumes and bid-ask spreads yield similar results regarding treatment effects, whereas the volume and spread variables are typically jointly and individually insignificant (details available from the authors).

Result 5: Transaction volumes are higher under tournament incentives.

Result 6: Markets with more female traders have higher bid-ask spreads for both assets, and a higher price premium for the riskier asset.

(The last part of result 6 is supported by results from Table 3.)

4.4 Discussion

While our main results – higher risky-asset prices under penalty framing and under tournament incentives – are compelling, it is less clear what exactly is driving them. One might expect a connection between treatment effects on asset prices and corresponding treatment effects on other variables such as volume or bid-ask spread (e.g., if high asset prices are due to more active trading), but we observe no relationships there that could be considered causal (see especially footnote 25 for a discussion of the price-volume relationship for the high-risk asset).

While this may seem counter-intuitive, Table 8 presents a simple explanation: both buyers' bids and sellers' asks are higher in tournaments and under penalty framing. The table presents regression results with individual bids and asks (i.e., limit orders, irrespective of whether accepted by a counter-party) for the high-risk asset as the dependent variable. (Also included as explanatory variables are the round number and the demographic and attitudinal variables used elsewhere – e.g., Table 4 – though these variables are jointly insignificant.)

Table 8: Marginal effects from linear regressions
(subject data, std. errors clustered by session)

Dependent variable:	Asset H bid (in lab money)	Asset H ask (in lab money)
Penalty treatment	+2.922** (1.327)	+8.177** (3.086)
Tournament treatment	+4.063*** (1.187)	+7.192** (2.844)
Demographics, attitudes, round number on right-hand side?	Yes	Yes
N	3153	4496
r-squared	0.100	0.082

Both bids and asks are significantly higher under penalty framing (compared to bonus framing) and under tournament incentives (compared to absolute-performance incentives), suggesting that our main treatment effects are due to both buyers and sellers setting prices above fundamental value, as opposed to, for example, sellers setting high asks which are accepted by buyers via market orders. The fact that both bids and asks rise in tournaments and under penalty framing is consistent with the main treatment effect (higher price for the high-risk asset in these treatments), and with a lack of any direct connection between prices and either volumes or bid-ask spreads or volume.

Loss aversion, by contrast, may be part of the explanation. Other researchers have conjectured that over-pricing of assets is due to herd behaviour by traders fearing under-performance, which is exacerbated by penalty framing (Paul et al., 2018), suggesting traders exhibit loss aversion. In our setting, selling the high-risk asset raises the possibility of missing out on high earnings in case the asset subsequently turns out to be successful. Under bonus framing, this missing out may lead to a foregone bonus (a reduced gain), but under penalty framing, it instead could lead to an increased penalty (a larger loss). In this latter case, loss-averse traders would be expected to set higher asks for selling the asset (to compensate for the potential loss), while they should also bid more in order to buy the asset (to avoid the potential loss).

Some suggestive evidence for this conjecture is shown in Table 9, which presents additional results from the regressions comprising Table 3 above. These regressions include a session-level measure of loss aversion, and its interactions with the main treatment variables, but those results were not presented in Table 3 for space reasons. Table 9 shows the average marginal effect of loss aversion, and its marginal effects conditional on each treatment variable.

Table 9: Marginal effects (MEs) of loss-aversion measure from Table 3 regressions

Dependent variable:	(1) Asset-H price	(2) Asset-L price	(3) Asset-H premium
Average loss aversion in group…			
…avg. ME	+1.102 (5.514)	−3.773 (2.761)	+4.785 (6.229)
…ME in absolute-performance treatment	−5.880 (9.049)	−4.131 (4.243)	−1.749 (10.466)
…ME in tournament treatment	+4.904 (6.288)	−3.416 (3.083)	+11.320* (6.181)
Signif. of differences	$p > 0.20$	$p > 0.20$	$p > 0.20$
…ME in bonus treatment	−19.051*** (4.914)	−7.801** (3.247)	−11.320** (5.690)
…ME in penalty treatment	+21.075** (9.843)	+0.254 (4.439)	+20.821* (10.699)
Signif. of differences	$p < 0.001$	$p \approx 0.14$	$p \approx 0.006$

** (**, ***): Significant at 10% (5%, 1%) level.*

The main result apparent from the table is that while loss aversion does not have a significant effect on the asset-H price overall, it has substantial effects in individual treatments. Notably, it has a significant positive effect under penalty framing, while under bonus framing it has a significant negative effect. That is, under penalty framing, more loss-averse groups of traders lead to higher prices above fundamental value for the high-risk asset. Similar differences exist between tournament and absolute-performance incentives (positive and negative marginal effects of loss aversion, respectively), though these differences are not significant.

In tournaments, the effect of penalty framing should be at least as strong due to the larger potential gains and losses at stake. However, we have observed that tournament incentives have an effect even under bonus framing, and as noted previously, our design means that this effect is not due simply to changes in the convexity of incentives. While we cannot rule out loss aversion as a driver of these results, Table 9 suggests that its effect may be relatively small, compared to its role in bonus-versus-penalty framing. An alternative conjecture is that subjects have tastes for ranks, or for competition, beyond any monetary gains or losses. Such behaviour has been observed in other settings (e.g., Charness et al., 2014, find that some subjects choose to incur monetary costs to sabotage the ranks of other subjects, even though these rankings carry no monetary benefit).

5. Conclusion

We investigate how financial-market behaviour and outcomes are affected by (i) compensation based on relative performance (tournament) versus absolute performance (approximately linear incentives), and (ii) framing of the variable component of compensation as a bonus versus a penalty. Subjects in our lab experiment trade in markets with a high-risk asset, a low-risk asset, and riskless cash. Both of our treatment variables have substantial effects: using either tournament compensation or penalty framing increases demand for the high-risk asset, as evidenced both by higher prices for that asset and by a higher price premium over the low-risk asset. These effects are not only statistically significant, but also economically significant, corresponding to increases of 29 percent of the high-risk asset's fundamental value due to tournament compensation, and of 13 percent of its fundamental value due to penalty framing.

The treatment effects we observe for the high-risk asset are largely consistent with, and provide some clarification of, results from the existing literature in this area (see Section 2 for details). Previous work has often found increased risky behaviour under convex tournament incentives compared to linear absolute incentives, and more risky behaviour under tournament incentives as convexity increases. We do not know of any previous experiment with linear tournament incentives. However, our results are consistent with what could be inferred from Fang et al.'s (2017) comparison of convex "winner-takes-all" tournaments, concave "elimination contest" tournaments, and linear absolute incentives. They found higher risky-asset prices under either convex or concave tournaments compared to absolute incentives; while these differences were insignificant, they suggest that linear tournament incentives – which presumably will have effects in between those of convex and concave incentives – should also lead to higher asset prices than absolute incentives.[26] Our results also relate to previous work comparing other kinds of behaviour under piece rates and tournaments; for example, Rigdon and D'Esterre (2015) and Feltovich (2019) find more lying in tournaments, even when monetary stakes are comparable to those under piece rates.[27]

The literature has also shown mixed results regarding penalty versus bonus framing. However, inference from these previous results is complicated by the confound between framing (gains versus losses) and changes to the shape of incentives. Our study isolates the pure effect of framing, and we see that the way monetary incentives are framed matters, even when the incentives themselves are held constant.

By contrast, both manipulations have much smaller and often insignificant effects on the low-risk asset's price. Overall, penalty framing has no significant effect on the low-risk asset's price. Tournament compensation does significantly increase the low-risk asset's price, but the size of the effect is much smaller than the corresponding effect for the high-risk asset (about 9.9 percent of the fundamental value). We did not have strong prior beliefs about how these treatments would affect the low-risk asset market, since it can be viewed as either a safe outlet (compared to the high-risk asset) or a risky one (compared to cash). Hence manipulations that alter risk taking were always likely to have an ambiguous effect here.

The policy implications of our main results are straightforward. Our finding of increased risk taking in tournaments has relevance for policies, like some provisions in the US's Dodd-Frank Act, which make tournament incentives less convex. Previous work suggests that such policies should indeed reduce financial risk taking, as intended. Our results reinforce this conclusion, but they also imply that reduced-convexity tournaments will still yield more risk taking than incentives based on absolute performance, so lowering convexity will not completely solve the problem. Our other main finding actually argues *against* a more recent policy proposal (*Wall Street Journal*, 2019) mandating a portion of bank executives' compensation to be deferred, with the potential to be clawed back. It seems likely that the loss of such compensation would be framed as a penalty rather than a foregone bonus, and our results suggest that this would increase risk taking, rather than decreasing it.

[26] The lack of substantial treatment effects on the price of the low-risk asset is also in keeping with previous results in the literature; see, for example, Kleinlercher et al.'s (2014) Figure 3.
[27] These results fit into a larger literature comparing lying in tournaments versus piece rates, but without controlling for monetary incentives.

Some of our secondary results – while not the main focus of this paper – are also worth noting, as they involve potential gender differences. At the aggregate level, markets with more female traders have significantly larger bid-ask spreads for both risky assets. At the individual level, females post fewer limit orders and more market orders, and earn less in the experiment, than males. This last set of results is not coincidental, as we additionally find significant relationships between how individuals trade and their earnings, with a larger fraction of market orders corresponding to lower earnings – even for males and females separately.

These gender effects are mainly in keeping with previous literature where such effects have been examined. It is true that our aggregate-level finding of higher risky-asset prices in markets with more females is the opposite of the negative effect on prices reported by Eckel and Füllbrunn (2015), who compare female-only and male-only markets. However, Cueva and Rustichini (2015) and Holt et al. (2017) actually observe *higher* prices in female-only markets when there are constant fundamental values (as in our setting), though these differences are not significant. Holt et al. (2017) are able to replicate Eckel and Füllbrunn's result, but only in markets with declining fundamental values (like Eckel and Füllbrunn's setting). Regarding our individual-level effects, both Kleinlercher et al. (2014) and Paul et al. (2019) also found that females earned less than males (Kleinlercher et al. noted this result explicitly, while Paul et al. observed that by the end of the session, females held less of both assets and cash than males). Kleinlercher et al. find that females post fewer limit orders (relative to market orders) for the high-risk asset, though there was little difference in the low-risk asset, and Fellner and Maciejovsky (2007) report a similar result in a one-risky-asset setting. Holt et al. (2017) also observe fewer limit orders in female-only markets than in male-only markets.

Some researchers have reported systematic differences in risk attitudes between males and females (however see Nelson, 2014 and 2015). We believe risk attitudes are not driving our gender-difference results, as we do not find significant differences between males and females in a lottery-choice task, and in any case we control for risk attitudes (and loss aversion) in the relevant regressions.[28] A possible alternative explanation is attitudes to competition, which also have been found to vary by gender (e.g., Niederle and Vesterlund, 2007). Placing limit orders (which requires posting a price-and-quantity offer) may be thought of as a more aggressive trading strategy than placing market orders (which only requires accepting someone else's posted offer), and individuals who shy away from competition might prefer the latter. Since posting prices tends to be associated with larger shares of the available surplus – for example in posted-price markets compared to double auctions (Ketcham et al., 1984; Mestelman and Welland, 1988; and Mestelman, 2008, though see also Smith, 1964 and Williams, 1973) and ultimatum bargaining compared to Nash bargaining (Roth, 1995; Fischer et al, 2007) – an aversion to competition may be costing females (and less competitive individuals within either gender).

We did not elicit attitudes to competition – or to rankings generally – in our experiment, but an investigation of this potential connection would be a promising topic for follow-up work. Other follow-up work could look at the robustness of our results. We followed the experimental design of one strand of previous research (especially Kleinlercher et al, 2014) closely, which invites the question of whether our results would hold in other double-auction settings, such as those with a declining rather than constant asset fundamental value, or with a fixed versus unknown number of rounds. Also, our high-risk asset has a positively-skewed distribution of returns, which has the advantage of allowing a large variance without the danger of negative subject payments, but other researchers may wish to examine settings where all assets have symmetric distributions. We strongly encourage future work in these directions.

[28] Additionally, the evidence for gender differences in risk tolerance in Chinese subjects is mixed, even compared to that for western subjects (Eckel and Grossman, 2008). For example, Gong and Yang (2012) report gender differences, while Lam (2015) reports no systematic differences.

References

Aggarwal, R. and N.M. Boyson (2016), "The performance of female hedge fund managers", *Review of Financial Economics* 29, pp. 23-36.

Armstrong, C.S., D.F. Larcker, G. Ormazabal and D.J. Taylor (2013), "The relation between equity incentives and misreporting: the role of risk-taking incentives", *Journal of Financial Economics* 109, pp. 327-350.

Baghestanian, S., P. Gortner and B. Massenot (2017), "Compensation schemes, liquidity provision, and asset prices: an experimental analysis", *Experimental Economics* 20, pp. 481-505.

Bai, G. and E. Elyasiani (2013), "Bank stability and managerial compensation", *Journal of Banking and Finance* 37, pp. 799-813.

Benartzi, S. and R.H. Thaler (1995), "Myopic loss aversion and the equity premium puzzle", *Quarterly Journal of Economics* 110, pp. 73-92.

Berger, L., K.J. Klassen, T. Libby and A. Webb (2013), "Complacency and giving up across repeated tournaments: evidence from the field", *Journal of Management Accounting Research* 25, pp. 143-167.

Brown, S.J., W.N. Goetzmann and J. Park (2001), "Careers and survival: competition and risk in the hedge fund and CTA industry", *Journal of Finance* 41, pp. 1869-1886.

Brown, K.C., W.V. Harlow and L.T. Starks (1996), "Of tournaments and temptations: an analysis of managerial incentives in the mutual fund industry", *Journal of Finance* 51, pp. 85-110.

Busse, J.A. (2001), "Another look at mutual fund tournaments", *Journal of Financial and Quantitative Analysis* 36, pp.53-73.

Cacioppo, J.T. and G.G. Berntson (1994), "Relationship between attitudes and evaluative space: a critical review, with emphasis on the separability of positive and negative substrates", *Psychological Bulletin* 115, pp. 401-423.

Caginalp, G., D. Porter and V.L. Smith (2000), "Overreactions, momentum, liquidity, and price bubbles in laboratory and field asset markets", *Journal of Psychology and Financial Markets* 1, pp. 24-48.

Carpenter, J.N. (2000), "Does option compensation increase managerial risk appetite?" *Journal of Finance* 55, pp. 2311-2331.

Charness, G., D. Masclet and M.C. Villeval (2014), "The dark side of competition for status", *Management Science* 60, pp. 38-55.

Chen, H.-L. and G.G. Pennacchi (2009), "Does prior performance affect a mutual fund's choice of risk? Theory and further empirical evidence", *Journal of Financial and Quantitative Analysis* 44, pp. 745-775.

Chevalier, J. and G. Ellison (1997), "Risk taking by mutual funds as a response to incentives", *Journal of Political Economy* 105, pp. 1167-1200.

Chevalier, J. and G. Ellison (1999), "Career concerns of mutual fund managers", *Quarterly Journal of Economics* 114, pp. 389-432.

Chowdhury, S.M., J.Y. Jeon and A. Ramalingam (2018), "Property rights and loss aversion in contests", *Economic Inquiry* 56, pp. 1492-1511.

Churchill, G. A., N. M. Ford and O. Walker (1990). *Sales Force Management: Planning, Implementation and Control*, 3rd ed. Richard D. Irwin, Homewood, IL.

Coles, J.L., N.D. Daniel and L. Naveen (2006), "Managerial incentives and risk-taking", *Journal of Financial Economics* 79, pp. 431-468.

Coval, J.D. and T. Shumway (2005), "Do behavioral biases affect prices?" *Journal of Finance* 60, pp. 1-34.

Cueva, C. and A. Rustichini (2015), "Is financial instability male-driven? Gender and cognitive skills in experimental asset markets", *Journal of Economic Behavior and Organization* 119, pp. 330-344.

Deaves, R., E. Lüders and G.Y. Luo (2009), "An experimental test of the impact of overconfidence and gender on trading activity". *Review of Finance* 13, pp. 555-575.

Dechenaux, E., D. Kovenock and R.M. Sheremeta (2015), "A survey of experimental research on contests, all-pay auctions and tournaments", *Experimental Economics* 18, pp. 609-669.

Eckel, C.C. and S.C. Füllbrunn (2015), "Thar SHE blows? Gender, competition, and bubbles in experimental asset markets", *American Economic Review* 105, pp. 906-920.

Eckel, C.C. and P.J. Grossman (2008), "Men, women and risk aversion: experimental evidence". In Plott, C.R. and Smith, V.L. (Eds.), *Handbook of Experimental Economics Results, Volume 1* (pp. 1061-1073). Amsterdam: North-Holland.

Elton, E.J., M.J. Gruber and C.R. Blake (2003), "Incentive fees and mutual funds", *Journal of Finance* 43, pp. 779-804.

Eriksen, K.W. and O. Kvaløy (2014), "Myopic risk-taking in tournaments", *Journal of Economic Behavior and Organization* 97, pp. 37-46.

Fang, D., M. Holmén, D. Kleinlercher and M. Kirchler (2017), "How tournament incentives affect asset markets: a comparison between winner-take-all tournaments and elimination contests", *Journal of Economic Dynamics and Control* 75, pp. 1-27.

Fehr, E., and K.M. Schmidt (1999), "A theory of fairness, competition and cooperation", *Quarterly Journal of Economics* 114, pp. 817-868.

Fellner, G. and B. Maciejovsky (2007), "Risk attitude and market behavior: evidence from experimental asset markets", *Journal of Economic Psychology* 28, pp. 338-350.

Feltovich, N. (2005), "Critical values for the robust rank-order test", *Communications in Statistics - Simulation and Computation* 34, pp. 525-547.

Fischbacher, U. (2007), "z-Tree: Zurich toolbox for ready-made economic experiments", *Experimental Economics* 10, pp. 171-178.

Fischer, S., W. Güth, and K. Pull (2007), “Is there as-if bargaining?” *Journal of Socio-Economics* 36, pp. 546-560.

Fryer, R.G., S.D. Levitt, J.A. List and S. Sadoff (2012), “Enhancing the efficacy of teacher incentives through loss aversion: a field experiment”, NBER working paper no. 18237.

Gaba, A. and A. Kalra (1999), “Risk behavior in response to quotas and contests”, *Marketing Science* 18, pp. 417-434.

Gaba, A., I. Tsetlin and R.L. Winkler (2004), “Modifying variability and correlations in winner-take-all contests”, *Operations Research* 52, pp. 384-395.

Gilpatric, S.M. (2009), “Risk taking in contests and the role of carrots and sticks”, *Economic Inquiry* 47, pp. 266-277.

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 & Organization* 83, pp. 59-65.

Grossman, P.J. and C.C. Eckel (2015), “Loving the long shot: risk taking with skewed lotteries”, *Journal of Risk and Uncertainty* 51, pp. 195-217.

Grund, C. and D. Sliwka (2005), “Envy and compassion in tournaments”, *Journal of Economics & Management Strategy* 14, pp. 187-207.

Hannan, R.L., V.B. Hoffman, D.V. and Moser (2005), “Bonus versus penalty: does contract frame affect employee effort?” In Rapoport, A. and R. Zwick (Eds.), *Experimental Business Research* (pp. 151-169). New York: Springer.

Harbring, C. and B. Irlenbusch (2008), “How many winners are good to have? On tournaments with sabotage”, *Journal of Economic Behavior & Organization* 65, pp. 682-702.

Hodder, J.E. and C.J. Jackwerth (2007), “Incentive contracts and hedge fund management”, *Journal of Financial and Quantitative Analysis* 42, pp. 811-826.

Holmén, M., M. Kirchler and D. Kleinlercher (2014), “Do option-like incentives induce overvaluation? Evidence from experimental asset markets”, *Journal of Economic Dynamics and Control* 40, pp. 179-194.

Holt, C.A., M. Porzio and M.Y. Song (2017), “Price bubbles, gender, and expectations in experimental asset markets”, *European Economic Review* 100, pp. 72-94.

Hossain, T. and J.A. List (2012), “The behavioralist visits the factory: increasing productivity using simple framing manipulations”, *Management Science* 58, pp. 2151-2167.

Hvide, H.K. (2002), “Tournament rewards and risk taking”, *Journal of Labor Economics* 20, pp. 877-898.

Hvide, H.K. and E.G. Kristiansen (2003), “Risk taking in selection contests”, *Games and Economic Behavior* 42, pp. 172-179.

James, D. and R.M. Isaac (2000), “Asset markets: how they are affected by tournament incentives for individuals”, *American Economic Review* 90, pp. 995-1004.

Kahneman, D. and A. Tversky (1979), “Prospect theory: an analysis of decision under risk”, *Econometrica* 47, pp. 263-292.

Kalra, A. and M. Shi (2001), “Designing optimal sales contests: a theoretical perspective”, *Marketing Science* 20, pp. 170-193.

Kelly, K., A. Presslee and A. Webb (2017), “The effects of tangible rewards versus cash rewards in consecutive sales tournaments: a field experiment”, *Accounting Review* 92, pp. 165-185.

Kempf, A. and S. Ruenzi (2013), “Tournaments in mutual-fund families”, *Review of Financial Studies* 21, pp. 1013-1036.

Kempf, A., S. Ruenzi and T. Thiele (2009), “Employment risk, compensation incentives, and managerial risk taking: evidence from the mutual fund industry”, *Journal of Financial Economics* 92, pp. 92-108.

Ketcham, J., V.L. Smith and A.W. Williams (1984), “A comparison of posted-offer and double-auction pricing institutions”, *Review of Economic Studies* 51, pp. 595-614.

King, R., V.L. Smith, A.W. Williams and M. van Boening (1993), “The robustness of bubbles and crashes in experimental stock markets”, in Day, R.H. and P. Chen (Eds.), *Nonlinear Dynamics and Evolutionary Economics* (pp. 183-200). Oxford: Oxford University Press.

Kirchler, M., F. Lindner and U. Weitzel (2018), “Rankings and risk-taking in the finance industry”, *Journal of Finance* 73, pp. 2271-2302.

Kleinlercher, D., J. Huber and M. Kirchler (2014), “The impact of different incentive schemes on asset prices”, *European Economic Review* 68, pp. 137-150.

Knauer, T., F. Sommer and A. Wohrman (2016), “Tournament winner proportion and its effect on effort: an investigation of the underlying psychological mechanisms”, *European Accounting Review* 26, pp. 681-702.

Kuziemko, I., R.W. Buell, T. Reich and M.I. Norton (2014), “ ‘Last-place aversion’: evidence and redistributive implications”, *Quarterly Journal of Economics* 129, pp. 105-149.

Lam, D. (2015), "Gender differences in risk aversion among Chinese university students", *Journal of Gambling Studies* 31, pp. 1405–1415.

Lazear, E.P., and S. Rosen (1981). “Rank-order tournaments as optimum labor contracts”, *Journal of Political Economy* 89, pp. 841-864.

Lefebvrea, M. and F.M. Vieider (2014), “Risk taking of executives under different incentive contracts: experimental evidence”, *Journal of Economic Behavior and Organization* 97, pp. 27-36.

Hong, F., T. Hossain and J.A. List (2015), “Framing manipulations in contests: a natural field experiment”, *Journal of Economic Behavior and Organization* 118, pp. 372-382.

Lee, J.H, C. Trzcinka and S. Venkatesan (2019), "Do portfolio manager contracts contract portfolio management?" Forthcoming, *Journal of Finance*.

Levitt, S., J.A. List, S Neckermann and S. Sadoff (2013), "The behavioralist goes to school: leveraging behavioral economics to improve educational performance'', NBER working paper 18165.

Luft, J. (1994), "Bonus and penalty incentives: contract choice by employees", *Journal of Accounting and Economics* 18, pp. 181-206.

Ma, L., Y. Tang and J. Gómez (2019), "Portfolio manager compensation in the US mutual fund industry", *Journal of Finance* 74, pp. 587-638.

Mestelman, S. (2008), "The performance of double-auction and posted-offer markets with advance production", in Plott, C.R. and V.L. Smith (Eds.), *Handbook of Experimental Economics Results, v. 1* (pp. 78-82). Amsterdam: North-Holland.

Mestelman, S. and D. Welland (1988), "Advance production in experimental markets", *Review of Economic Studies* 55, pp. 641-654.

Morgan, J., and D. Sisak (2016), "Aspiring to succeed: a model of entrepreneurship and fear of failure", *Journal of Business Venturing* 31, pp. 1-21.

Nelson, J.A. (2014), "The power of stereotyping and confirmation bias to overwhelm accurate assessment: the case of economics, gender, and risk aversion", *Journal of Economic Methodology* 21, pp. 211-231.

Nelson, J.A. (2015), "Are women really more risk-averse than men? A re-analysis of the literature using expanded methods", *Journal of Economic Surveys* 29, pp. 566-585.

Nieken, P. and D. Sliwka (2010), "Risk-taking tournaments – theory and experimental evidence", *Journal of Economic Psychology* 31, pp. 254-268.

Niederle, M. and L. Vesterlund (2007), "Do women shy away from competition? Do men compete too much?" *Quarterly Journal of Economics* 122, pp. 1067-1101.

Oblak, K., M. Ličen, and S. Slapničar (2018), "The role of cognitive frames in combined decisions about risk and effort", *Management Accounting Research* 39, pp. 35-46.

Oechssler, J. (2010), "Searching beyond the lamppost: let's focus on economically relevant questions", *Journal of Economic Behavior and Organization* 73, pp. 65-67.

Panageas, S. and M. Westerfield (2009), "High-water marks: high-risk appetites? Convex compensation, long horizons, and portfolio choice", *Journal of Finance* 64, pp. 1-36.

Paul, D.J., J. Henker and S. Owen (2019), "The aggregate impacts of tournament incentives in experimental asset markets", *Experimental Economics* 22, pp 441-476.

Rigdon, M.L. and A.P. D'Esterre (2015), "The effects of competition on the nature of cheating behaviour", *Southern Economic Journal* 81, pp. 1012-1024.

Roth, A.E. (1995), "Bargaining experiments", in Kagel, J.H. and A.E. Roth (Eds.), *Handbook of Experimental Economics* (pp. 253-348). Princeton, NJ: Princeton University Press.

Sato, Y. (2016), “Fund tournaments and asset bubbles”, *Review of Finance* 20, pp. 1383-1426.

Sheremeta, R.M. (2013), “Overbidding and heterogeneous behavior in contest experiments”, *Journal of Economic Surveys* 27, pp. 491-514.

Siegel, S.and N.J. Castellan, Jr. (1988), *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill, New York.

Smith, V.L. (1964), “Effect of market organization on competitive equilibrium”, *Quarterly Journal of Economics* 78, pp. 181-201.

Smith, V.L., G.L. Suchanek and A.W. Williams (1988), “Bubbles, crashes, and endogenous expectations in experimental spot asset markets” *Econometrica* 56, pp. 1119-1151.

Smith, V.L., M. van Boening and C.P. Wellford (1988), “Dividend timing and behaviour in laboratory asset markets” *Economic Theory* 16, pp. 567--583.

Stöckl, T., J. Huber and M. Kirchler (2010), “Bubble measures in experimental asset markets”, *Experimental Economics* 13, pp. 284-298.

Taylor, J. (2003), “Risk-taking behavior in mutual fund tournaments”, *Journal of Economic Behavior and Organization* 50, pp. 373-383.

Tong, K.K. and K. Leung (2002), “Tournament as a motivational strategy: extension to dynamic situations with uncertain duration”, *Journal of Economic Psychology* 23, pp. 399-420.

Wall Street Journal (2019), “Limits on Wall Street pay are back on regulators’ agenda”, https://www.wsj.com/articles/limits-on-wall-street-pay-are-back-on-regulators-agenda-11551839333, accessed 25 August 2019.

Williams, F.E. (1973), “The effect of market organization on competitive equilibrium: the multi-unit case”, *Review of Economic Studies* 40, pp. 97-113.

Appendix A: Experimental instructions (English translation)

[*Please note that instructions for the 4 treatments were identical, except the explanation of subsection "Calculation of your payout in RMB". We included the 4 different subsections separately.*]

Instructions for Treatment Absolute performance_BONUS (short for AB), Treatment Absolute performance_PENALTY (short for AP), Treatment TOURNAMENT BONUS (short for TB), and Treatment TOURNAMENT PENALTY (short for TP)

General Information

This experiment is concerned with replicating an asset market where 10 traders can trade two assets (Asset HIGH RISK and Asset LOW RISK) simultaneously. Thereby you act as a portfolio manager who actively manages a portfolio consisting of the two assets. In the end of the experiment, you are paid out according to your performance.

Market Description

Each trader gets an initial endowment of 20 units of asset HIGH RISK and 20 units of asset LOW RISK. Additionally, each trader gets a working capital of 2000 Taler (experimental currency)

- Each of the two assets can be traded separately and it is up to you how many units of each asset you want to trade.
- The trading screen is split into two areas – on for trades of asset HIGH RISK, one for trades of asset LOW RISK (see trading screen below).
- The prices of the two assets are independent and are only determined through the trading activities of the 10 traders.

Information about asset HIGH RISK

The markets will end randomly between periods 8 and 14 with equal probability. At the end of the experiment (after 8-14 periods) the units of asset HIGH RISK you own are bought back by the experimenter. The buyback price is determined randomly and can take on the following values (all traders in the market get the information in the table mentioned below).

Buyback price for each asset HIGH RISK in Taler	Probability of the buyback price
15	80%
65	20%

Information about asset LOW RISK

The markets will end randomly between periods 8 and 14 with equal probability. At the end of the experiment (after 8-14 periods) the units of asset LOW RISK you own are bought back by the experimenter. The buyback price is determined randomly and can take on the following values (all traders in the market get the information in the table mentioned below).

Buyback price for each asset LOW RISK in Taler	Probability of the buyback price
20	50%
30	50%

Trading

Trading is accomplished in form of a double auction, i.e. each trader can appear as buyer and seller at the same time. Asset HIGH RISK and asset LOW RISK are traded on two separate market places (see below). You can submit any quote of asset HIGH RISK and LOW RISK with prices ranging from 0 to a maximum of 999 Taler (with at most two decimal places). For every quote you make, you have to enter the number of assets you intend to trade as well. Note that your Taler and asset holdings cannot drop below zero. If you buy assets, your Taler holdings will be diminished by the respective expenditures (price x quantity) and the number of assets will be increased by the quantity of newly bought assets. Inversely, if you sell assets, your Taler holdings will be increased by the respective revenues (price x quantity) and the number of assets will be decreased by the quantity of newly sold assets.

[Begin Treatment AB]

Calculation of your payout in RMB

Your payout at the end of the experiment (after 8-14 periods) is calculated as follows:

Payout in RMB = Fixed Payment + Bonus Payment

If your wealth in Taler is above 1000 Taler you will get a bonus payment. This bonus payment is added to your fixed payment of 20 RMB and is calculated as follows: You will get 2 RMB for every 100 Taler exceeding 1000 Taler. If your wealth is below 1000 Taler, you will only get the fixed payment of 20 RMB. Please note, your bonus payment cannot be more than 80 RMB, i.e. your maximal payout is capped at 100 RMB (80 RMB bonus payment + 20 RMB fixed payment).

Your wealth in Taler is calculated as follows: the units of asset HIGH RISK you hold are multiplied with the randomly drawn buyback price (15 with 80% probability or 65 with 20% probability) and the units of asset LOW RISK you hold are multiplied with the randomly drawn buyback price (20 with 50% probability or 30 with 50% probability). This sum is then added to your holdings in Taler. If your final wealth exceeds 1000 Taler you get a bonus payment, if not, you only get your fixed payment.

WEALTH IN TALER = ASSETS HIGH RISK* BUYBACK PRICE HIGH RISK + ASSETS LOW RISK* BUYBACK PRICE LOW RISK + TALER

Example 1: At the end of the experiment you hold 25 assets HIGH RISK and 15 assets LOW RISK. Your Taler

holdings are 1800. The randomly drawn buyback price for asset HIGH RISK is 65 and the randomly drawn buyback price for asset LOW RISK is 30.
Wealth in Taler: ((25*65)+(15*30)+1800)=3875
Bonus payment in RMB: (3875-1000)*0.02**=**2875*0.02 =57.5
Fixed payment in RMB: 20
Payout in RMB: 57.5 + 20 = 77.5

Example 2: At the end of the experiment you hold 10 assets HIGH RISK and 15 assets LOW RISK. Your Taler holdings are 2400. The randomly drawn buyback price for asset HIGH RISK is 15 and the randomly drawn buyback price for asset LOW RISK is 20.
Wealth in Taler: ((10*15)+(15*20)+2400)=2850
Bonus payment in RMB: (2850-1000)*0.02**=**1850*0.02 = 37
Fixed payment in RMB: 20
Payout in EUR: 37 + 20= 57

[End Treatment AB]

[Begin Treatment AP]

Calculation of your payout in RMB

Your payout at the end of the experiment (after 8-14 periods) is calculated as follows:

Payout in RMB = Fixed Payment - Penalty for bad performance

If your wealth in Taler reach 5000 Taler or more, you will get your fixed payment of 100 RMB. If your wealth is below 5000 Taler, you will get a deduction of 2 RMB for every 100 Taler below 5000 Taler. The maximum deduction is 80 RMB, i.e. your minimal payout is capped at 20 RMB.

Your wealth in Taler is calculated as follows: the units of asset HIGH RISK you hold are multiplied with the randomly drawn buyback price (15 with 80% probability or 65 with 20% probability) and the units of asset LOW RISK you hold are multiplied with the randomly drawn buyback price (20 with 50% probability or 30 with 50% probability). This sum is then added to your holdings in Taler.

WEALTH IN TALER = ASSETS HIGH RISK* BUYBACK PRICE HIGH RISK + ASSETS LOW RISK* BUYBACK PRICE LOW RISK + TALER

Payout scheme:

Example 1: At the end of the experiment you hold 25 assets HIGH RISK and 15 assets LOW RISK. Your Taler holdings are 1800. The randomly drawn buyback price for asset HIGH RISK is 65 and the randomly drawn buyback pricefor asset LOW RISK is 30.
Wealth in Taler: (25*65)+(15*30)+1800 = 3875
Penalty: (5000-3875)*0.02=1125*0.02=22.5
Fixed payment: 100
Payout in RMB: 100-22.5=77.5

Example 2: At the end of the experiment you hold 10 assets HIGH RISK and 15 assets LOW RISK. Your Taler holdings are 2400. The randomly drawn buyback price for asset HIGH RISK is 15 and the randomly drawn

buyback price for asset LOW RISK is 20.
Wealth in Taler: ((10*15)+(15*20)+2400)= 2850
Penalty: (5000 – 2850)*0.02 =2150*0.02 = 43
Fixed payment: 100
Payout in RMB: 100 - 43= 57

[End Treatment AP]

[Begin Treatment TB]

Calculation of your payout in RMB

Your payout at the end of the experiment (after 8-14 periods) is calculated as follows:

Payout in RMB = Fixed Payment + Bonus Payment

Everyone gets a fixed payment of 20 RMB. However, If your wealth in Taler is among the top 50% of all subjects in the market, you will get a bonus payment of 80 RMB.

Your wealth in Taler is calculated as follows: the units of asset HIGH RISK you hold are multiplied with the randomly drawn buyback price (15 with 80% probability or 65 with 20% probability) and the units of asset LOW RISK you hold are multiplied with the randomly drawn buyback price (20 with 50% probability or 30 with 50% probability). This sum is then added to your holdings in Taler.

WEALTH IN TALER = ASSETS HIGH RISK* BUYBACK PRICE HIGH RISK + ASSETS LOW RISK* BUYBACK PRICE LOW RISK + TALER

Example 1: At the end of the experiment you hold 25 assets HIGH RISK and 15 assets LOW RISK. Your Taler holdings are 1800. The randomly drawn buyback price for asset HIGH RISK is 65 and the randomly drawn buyback price for asset LOW RISK is 30.
Wealth in Taler: ((25*65)+(15*30)+1800)=3875
Fixed payment in RMB: 20
If your wealth in Taler is among the top 50%, you will get 80 RMB for Bonus, thus you final payout = 20+80=100 RMB. Otherwise, your final payout= 20 RMB

Example 2: At the end of the experiment you hold 10 assets HIGH RISK and 15 assets LOW RISK. Your Taler holdings are 2400. The randomly drawn buyback price for asset HIGH RISK is 15 and the randomly drawn buyback price for asset LOW RISK is 20.
Wealth in Taler: ((10*15)+(15*20)+2400)=2850
Fixed payment in RMB: 20
If your wealth in Taler is among the top 50%, you will get 80 RMB for Bonus, thus you final payout = 20+80=100 RMB. Otherwise, your final payout= 20 RMB

[End Treatment TB]

[Begin Treatment TP]

Calculation of your payout in RMB

Your payout at the end of the experiment (after 8-14 periods) is calculated as follows:

Payout in RMB = Fixed Payment - Penalty for bad performance

Everyone gets a fixed payment of 100 RMB. However, If your wealth in Taler is among the bottom 50% of all subjects in the market, you will get a penalty of 80 RMB.

Your wealth in Taler is calculated as follows: the units of asset HIGH RISK you hold are multiplied with the randomly drawn buyback price (15 with 80% probability or 65 with 20% probability) and the units of asset LOW RISK you hold are multiplied with the randomly drawn buyback price (20 with 50% probability or 30 with 50% probability). This sum is then added to your holdings in Taler.

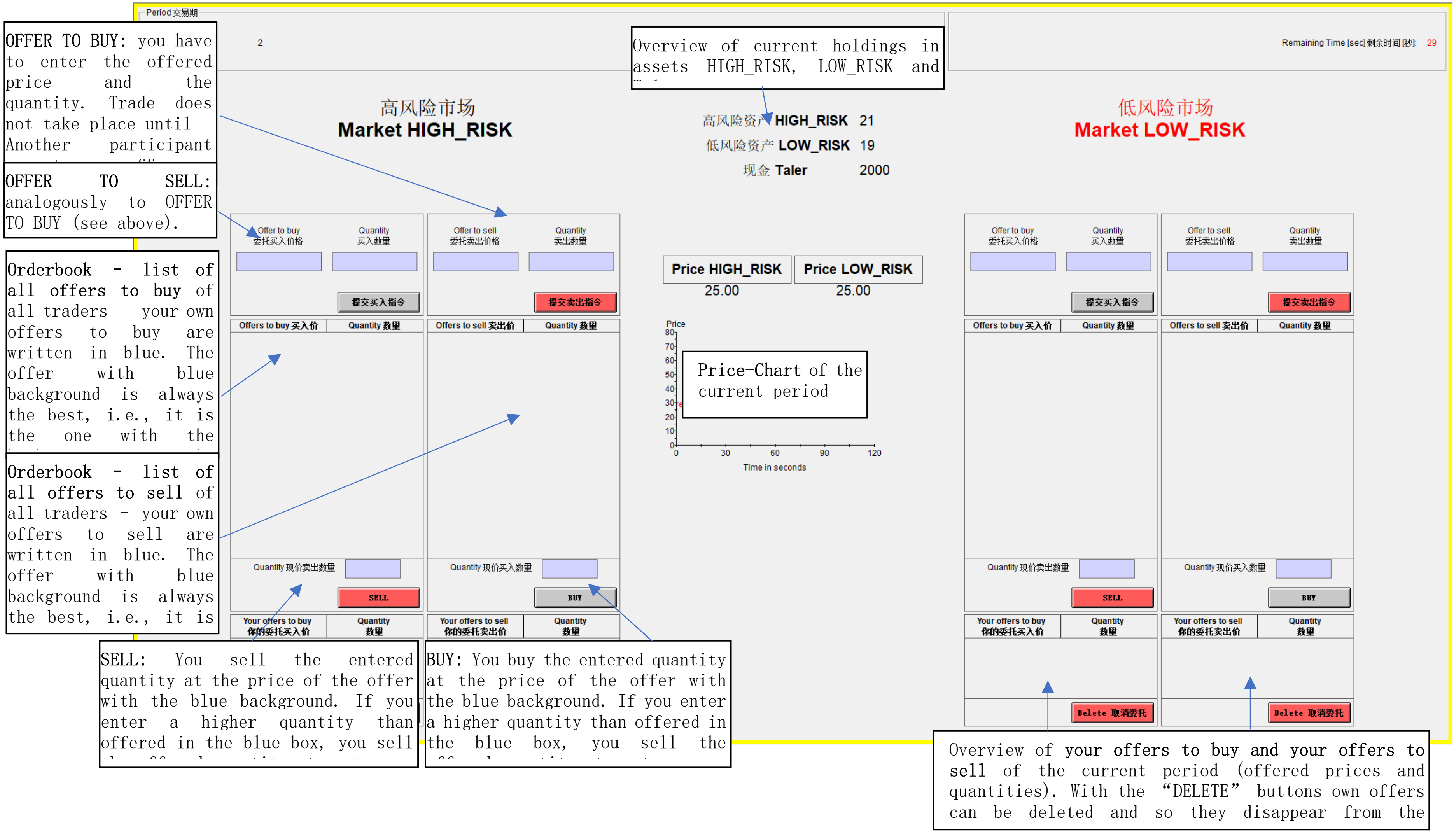
WEALTH IN TALER = ASSETS HIGH RISK* BUYBACK PRICE HIGH RISK + ASSETS LOW RISK* BUYBACK PRICE LOW RISK + TALER

Example 1: At the end of the experiment you hold 25 assets HIGH RISK and 15 assets LOW RISK. Your Taler holdings are 1800. The randomly drawn buyback price for asset HIGH RISK is 65 and the randomly drawn buyback price for asset LOW RISK is 30.
Wealth in Taler: ((25*65)+(15*30)+1800)=3875
Fixed payment in RMB: 100
If your wealth in Taler is among the bottom 50%, you will get a penalty of 80 RMB, thus you final payout = 100-80=20 RMB. Otherwise, your final payout= 100 RMB.

Example 2: At the end of the experiment you hold 10 assets HIGH RISK and 15 assets LOW RISK. Your Taler holdings are 2400. The randomly drawn buyback price for asset HIGH RISK is 15 and the randomly drawn buyback price for asset LOW RISK is 20.
Wealth in Taler: ((10*15)+(15*20)+2400)=2850
Fixed payment in RMB: 100
If your wealth in Taler is among the bottom 50%, you will get a penalty of 80 RMB, thus you final payout = 100-80=20 RMB. Otherwise, your final payout= 100 RMB.

[End Treatment TP]

Trading Screen: In the following graph trading will be explained in detail



Appendix B: Additional data and analysis

Tables B1 and B2 show some session- and treatment-level averages from the experiment. Note that higher values of the risk- and loss-aversion measures correspond to less appetite for risk or loss.

Table B1: session- and cell-level averages, demographic/attitudinal variables and asset prices

Cell/session	Age	Business	Female	Loss aversion	Risk aversion	Undergrad	H price	L price
AB1	20.5	1	0.6	2.9	5.4	1	27.85	25.01
AB2	20.3	0.2	0.7	2.7	6.1	1	22.29	22.59
AB3	19.6	0.6	0.6	2.6	5	1	25.25	25.08
AB4	20	0.9	0.7	2.4	6.2	1	29.2	22.83
AB5	20.8	0.6	0.5	1.8	5.4	1	26.07	24.48
AB6	20.1	1	0.8	2.7	5.9	1	23.99	24.55
AB7	19.9	0.9	0.7	2.6	5.6	1	28.52	24.31
AB8	18.8	0.7	0.5	2.2	6.5	1	29.97	31.02
AB9	20.2	1	0.6	2.5	5.9	1	25.02	23.14
AB-avg	20.02	0.767	0.633	2.489	5.778	1	26.46	24.77
TB1	19.4	0.7	0.5	3.4	5.5	1	29.5	23.9
TB2	19.4	0.9	1	2.3	4.8	0.9	32.02	24.2
TB3	20.1	0.5	0.7	2.8	6.8	1	24.89	23.17
TB4	21.2	1	0.5	2.1	4.3	0.9	33.11	25.83
TB5	21.1	1	0.6	2.7	4.9	1	28.32	26.54
TB6	20.5	0.8	0.8	2.2	6.1	1	33.75	25.58
TB7	20.9	0.7	0.8	2.5	6.1	1	28.17	21.38
TB8	20.8	1	0.8	2.1	5.4	1	31.61	25.83
TB9	20.7	1	0.7	2	6.5	1	35.37	24.96
TB-avg	20.46	0.844	0.711	2.456	5.600	0.978	30.75	24.60
AP1	19.9	0.4	0.7	2.6	6.3	1	30.57	24.14
AP2	19	0.7	0.8	2.8	5.4	1	33.58	20.86
AP3	19.8	0.8	0.3	3	6.7	1	27.64	21.95
AP4	19.2	1	0.5	3	5.4	1	27.06	24.75
AP5	19.6	0.8	0.7	3.2	6.6	1	33.86	22.84
AP6	19.5	1	1	2.6	6.1	1	29.55	21.47
AP7	19.9	0.9	0.8	2.5	6.1	1	26.91	21.81
AP8	18.3	1	0.4	2.3	5.3	1	33.04	23.74
AP9	20.6	0.9	0.5	2.7	6.1	0.9	31.23	24.86
AP-avg	19.53	0.833	0.633	2.744	6.000	0.989	30.38	22.94
TP1	19.4	0.7	0.8	2.5	5.7	1	31.78	25.28
TP2	19.7	0.9	0.9	3.1	6.2	1	46.86	29.34
TP3	20.9	1	0.6	2.6	5.4	1	29.44	24.17
TP4	21.1	0.9	0.8	2.1	6.8	1	41.14	30.8
TP5	21	0.8	0.8	2.5	6.5	1	29.95	22.72
TP6	19.3	0.9	0.8	2.5	5.9	1	44.61	25.67
TP7	20.5	0.9	0.8	2.6	6.8	0.7	32.37	24.82
TP8	20	1	0.8	3.4	5.8	1	46.29	25.66
TP9	19.6	0.8	0.6	2.3	5.7	1	31.01	25.1
TP-avg	20.17	0.878	0.767	2.622	6.089	0.967	37.05	25.95
KW p-value	0.07	0.93	0.13	0.27	0.48	0.56	0.002	0.02

The bottom row of Table B1 shows p-values from Kruskall-Wallis tests of equality across the four treatments. One of the six variables (age) is significantly different across treatments at the 10-percent level, but none are at the 5-percent level. Since one of six is approximately what would be expected under the null hypothesis of no differences across treatments, we conclude that the treatments are balanced demographically and attitudinally.

Table B2 reports additional measures of asset mispricing for individual sessions and for each cell in aggregate. The measures are as follows:

Relative absolute deviation (RAD) is the average magnitude of the deviation of price from fundamental value, normalised by the fundamental value:

$$RAD = \frac{1}{T \cdot \overline{FV}} \sum_{t=1}^{T} \left| \overline{P}_t - FV_t \right|$$

Here, T stands for the total number of rounds, $\overline{P}_t$ is the average trading price in round t, and $\overline{FV}$ is the mean of fundamental value over all rounds.

Relative deviation (RD) is similar to RAD, but with signed deviations rather than magnitudes only:

$$RD = \frac{1}{T \cdot \overline{FV}} \sum_{t=1}^{T} \left(\overline{P}_t - FV_t \right)$$

A positive value for RD is an indicator that the market on average overvalues the asset, whereas negative values are an indicator for undervaluation. (Note that RD is an affine transformation of the price measure we use in this paper; hence results using it are identical to those reported in section 4.1 of the main text.)

Price amplitude (PA) is defined as the difference between peak and trough of mean rounds prices relative to the fundamental value, normalised by the initial fundamental value FV_1:

$$PA = \frac{1}{FV_1} \left[\max_t \left(\overline{P}_t - FV_t \right) - \min_t \left(\overline{P}_t - FV_t \right) \right]$$

A high PA indicates large price swings relative to the fundamental value.

Normalised turnover (TR) is the total number of transactions over the life of the asset $\sum_t n_t$ (where n_t denotes the number of transactions in round t), divided by the total number of units outstanding (TSU), normalised by the total number of rounds T:

$$TR = \frac{1}{T \cdot TSU} \sum_{t=1}^{T} n_t$$

For example, a value of 0.10 indicates that on average 10 percent of all units are traded in each round. Thus, a high TR value indicates a high volume of trade. (Note that TR is an affine transformation of the number of transactions; hence results using it are identical to those reported in section 4.3 of the main text.)

Table B2: Cell and session-level mispricing measures (all rounds)

Cell	Session	Asset H RAD	Asset H RD	Asset H PA	Asset H TR	Asset L RAD	Asset L RD	Asset L PA	Asset L TR
AB	1	0.122	0.114	0.394	0.850	0.022	0.000	0.131	1.015
	2	0.109	-0.109	0.197	0.930	0.096	-0.096	0.050	0.835
	3	0.029	0.010	0.182	1.495	0.012	0.003	0.057	1.185
	4	0.196	0.168	0.412	1.105	0.159	-0.087	0.613	1.085
	5	0.061	0.043	0.186	1.375	0.022	-0.021	0.058	1.720
	6	0.061	-0.040	0.235	0.905	0.018	-0.018	0.068	0.960
	7	0.182	0.141	0.530	0.620	0.033	-0.028	0.094	0.455
	8	0.199	0.199	0.227	0.750	0.241	0.241	1.028	1.270
	9	0.084	0.001	0.566	0.940	0.079	-0.075	0.160	0.650
	Mean	0.116	0.058	0.325	0.997	0.076	-0.009	0.251	1.019
AP	1	0.223	0.223	0.242	1.115	0.052	-0.034	0.192	1.160
	2	0.419	0.343	1.338	0.895	0.198	-0.165	0.730	0.905
	3	0.119	0.106	0.470	0.975	0.122	-0.122	0.209	0.790
	4	0.083	0.083	0.190	0.955	0.045	-0.010	0.250	0.710
	5	0.355	0.355	0.838	1.095	0.098	-0.086	0.249	0.730
	6	0.210	0.182	0.847	0.510	0.143	-0.141	0.407	0.425
	7	0.111	0.076	0.365	0.980	0.127	-0.127	0.150	0.885
	8	0.322	0.322	0.474	0.905	0.050	-0.050	0.116	0.850
	9	0.260	0.249	0.632	1.005	0.073	-0.006	0.210	0.945
	Mean	0.233	0.215	0.600	0.937	0.101	-0.083	0.279	0.822
TB	1	0.180	0.180	0.160	0.870	0.044	-0.044	0.082	0.810
	2	0.281	0.281	0.412	2.265	0.074	-0.032	0.269	1.630
	3	0.034	-0.004	0.134	1.730	0.077	-0.073	0.164	1.235
	4	0.325	0.325	0.504	1.370	0.034	0.033	0.056	1.145
	5	0.175	0.133	0.492	0.820	0.062	0.062	0.130	0.765
	6	0.350	0.350	0.352	1.680	0.036	0.023	0.123	1.310
	7	0.127	0.127	0.301	1.110	0.145	-0.145	0.218	0.845
	8	0.264	0.264	0.276	2.820	0.065	0.033	0.238	1.885
	9	0.415	0.415	1.242	2.850	0.134	-0.002	0.581	1.680
	Mean	0.239	0.230	0.430	1.724	0.075	-0.016	0.207	1.256
TP	1	0.277	0.271	0.548	0.715	0.027	0.011	0.140	0.495
	2	0.874	0.874	0.128	1.355	0.173	0.173	0.477	1.390
	3	0.178	0.178	0.388	1.085	0.033	-0.033	0.078	1.015
	4	0.646	0.646	0.970	1.400	0.232	0.232	0.300	0.910
	5	0.198	0.198	0.397	1.795	0.091	-0.091	0.096	1.805
	6	0.784	0.784	0.767	2.165	0.059	0.027	0.212	1.695
	7	0.295	0.295	0.250	1.170	0.035	-0.007	0.174	0.630
	8	0.852	0.852	1.122	2.425	0.082	0.026	0.341	1.925
	9	0.240	0.240	0.306	2.495	0.030	0.004	0.166	1.730
	Mean	0.483	0.482	0.542	1.623	0.085	0.038	0.220	1.288

Table B3 shows the significance of treatment effects for these asset mispricing variables, as Table 2 did for prices. For readability, we leave out p-values above 10 percent.

Table B3: Significance of treatment effects, mis-pricing measures

RAD (Relative absolute deviation)

Asset H	Bonus	Penalty	p-value	Asset L	Bonus	Penalty	p-value
Absolute	0.116	0.233	***0.016***	Absolute	0.076	0.101	---
Tournament	0.239	0.483	---	Tournament	0.075	0.085	---
p-value	***0.039***	0.095		p-value	---	---	

RD (Relative deviation)

Asset H	Bonus	Penalty	p-value	Asset L	Bonus	Penalty	p-value
Absolute	0.058	0.215	***0.011***	Absolute	−0.009	−0.083	0.065
Tournament	0.230	0.482	---	Tournament	−0.016	0.038	---
p-value	***0.016***	0.079		p-value	---	***0.002***	

PA (Price amplitude)

Asset H	Bonus	Penalty	p-value	Asset L	Bonus	Penalty	p-value
Absolute	0.325	0.600	***0.050***	Absolute	0.251	0.279	0.091
Tournament	0.430	0.542	---	Tournament	0.207	0.220	---
p-value	---	---		p-value	---	---	

TR (Normalised turnover)

Asset H	Bonus	Penalty	p-value	Asset L	Bonus	Penalty	p-value
Absolute	0.997	0.937	---	Absolute	1.019	0.822	---
Tournament	1.724	1.623	---	Tournament	1.256	1.288	---
p-value	***0.038***	***0.009***		p-value	---	0.073	

Note: p-values based on two-tailed robust rank-order tests, session-level data

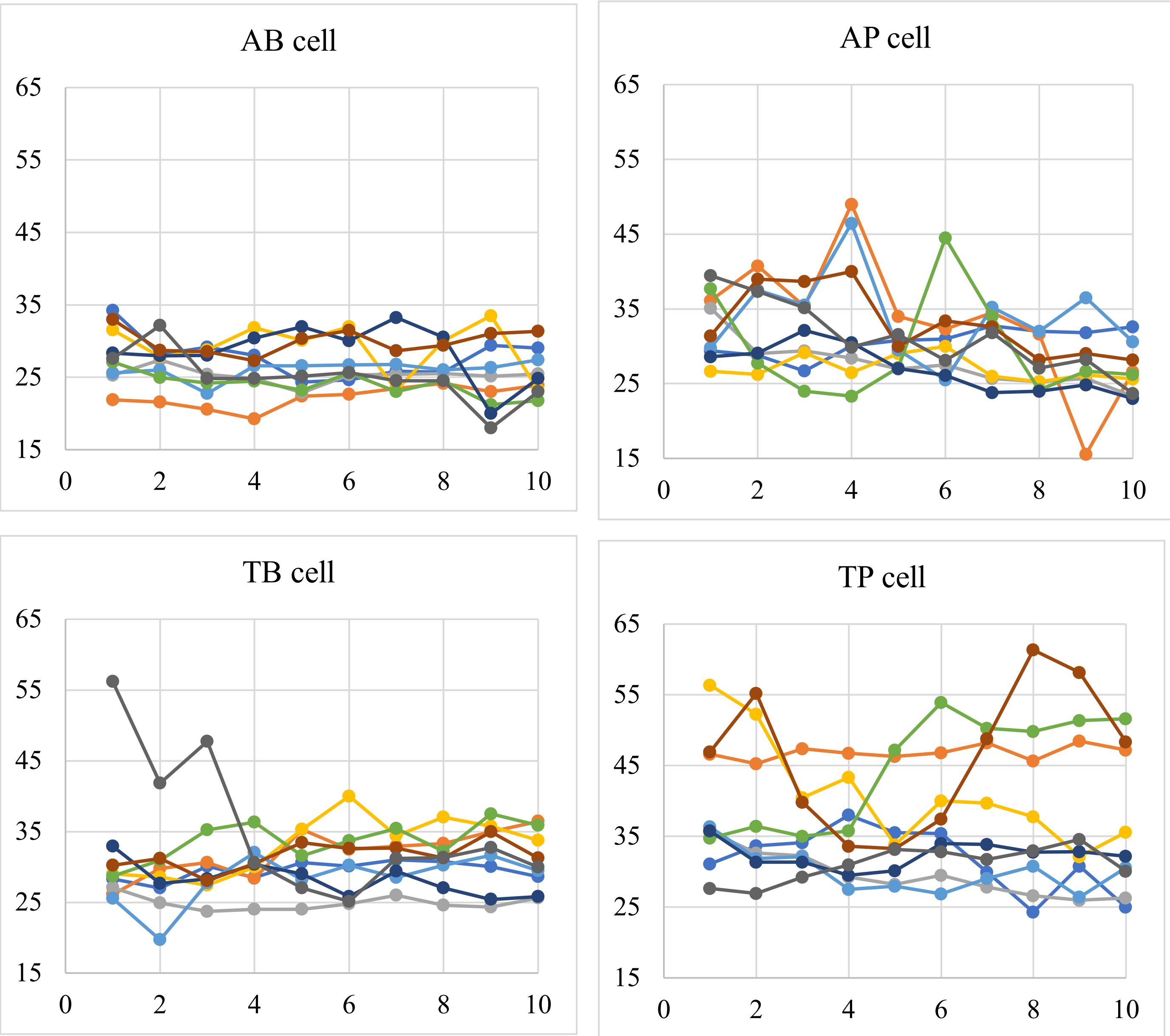
Notice first that the results for RD are identical to those in Table 2 for prices; as noted above, RD is merely an affine transformation of price for each asset (and indeed, since our assets have the same fundamental value, the assets together). Similarly, TR is an affine transformation of the number of transactions, and therefore yields identical results to those in Table 6.

Treatment effects for RAD are nearly identical. The lone exception is in one comparison for the low-risk asset: the effect of tournament incentives in the Penalty treatment is insignificant, where it was significant when RD was used. This is because the asset tended to be over-priced in the TP cell and under-priced in the AP cell. So, RD (which accounts for the sign of mis-pricing) yields a significant difference, while RAD (which ignores sign and only accounts for magnitude) does not.

Treatment effects for PA are similar in some ways to those for RD and RAD, but there are some substantial differences. There is higher PA for the high-risk asset under penalties than under bonuses, as with RD and RAD (though only significant in one of the two comparisons), while there are no systematic differences between absolute-payoff and tournament treatments. The discrepancy between these results and those for RD and RAD is probably due to what each of these variables measures: changes in price over time (for the former) versus differences between the price and fundamental value (for the latter). In particular, if an asset is over-priced throughout a session, it will have a large RD and RAD, but a small PA; examination of Figure B1 below suggests that this sometimes happens in the TP cell, thus under-stating differences between this cell and the other cells. Such persistent over-pricing could be interpreted as a bubble that does not burst during the session, which may be due to our use of an indefinite endpoint, in contrast to some other studies that used fixed endpoints (making it very likely that any bubbles would crash by the final round). This possibility suggests that PA may be a better measure of bubble activity under a fixed-endpoint design, while RAD and (especially) RD may be more appropriate when an indefinite endpoint is used.

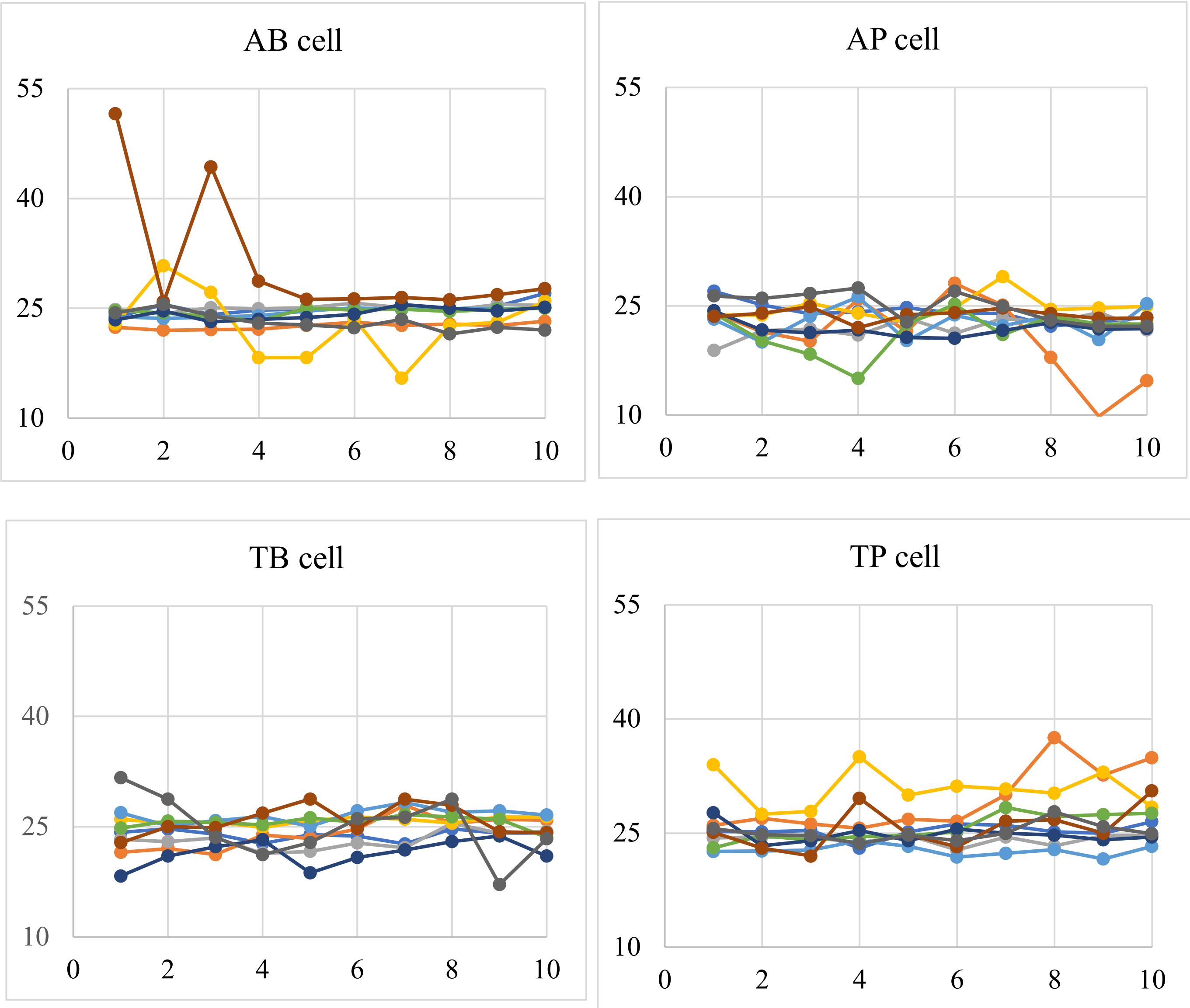
Figures B1 and B2 show time series of asset prices for each session of each cell.

Figure B1: Asset-H prices (in lab-taler, vertical axis) by round number (horizontal axis), indiv. sessions



Note: fundamental value = 25.

Figure B2: Asset-L prices (in lab-taler, vertical axis) by round number (horizontal axis), indiv. sessions



Note: fundamental value = 25.

Table B4 shows some results for the difference in volume between low- and high-risk assets (again, for readability, we leave out p-values above 10 percent).

Table B4: Difference in volume between assets (all rounds), and within- and between-cell significance tests

Asset H transactions minus Asset L transactions	Bonus	Penalty	p-value (Bonus vs. Penalty)
Absolute	−0.46	+2.10**	---
Tournament	+9.36***	+6.69**	---
p-value (Absolute vs. Tournament)	***0.009***	---	

*Notes: ** (***): significantly different from zero at 5-percent (1-percent) level, two-tailed Wilcoxon signed-rank test with session-level data. P-values for treatment differences based on two-tailed robust rank-order tests, session-level data.*

In three of the four cells, there is more trading of the high-risk asset than the low-risk asset. In the remaining cell, AB, trading is approximately the same for both assets. There is some evidence for treatment effects, with the high-minus-low difference significantly smaller in the AB cell than TB. (There is additional suggestive evidence for an effect of tournament incentives, with a larger difference in TP than AP; however this misses significance at conventional levels with a p-value of 0.126.)

Table B5 presents some observed correlations in the data. The top three rows of correlations show that when all cells of the experiment are pooled, there is significant positive association between the price and trading volume of the high-risk asset (the corresponding association for the low-risk asset is also positive, but not significant), as well as a weakly significant association between the asset-H premium and the difference in trading volume between the high- and low-risk assets. However, similar correlations at the level of the individual cell exhibit high variation across cells – indeed, the correlation in one of the four cells is actually negative – and even the positive correlations are not significantly different from zero (suggesting substantial across-session variation within cells). The combination of significant positive correlation in pooled cells and no significant correlation in individual cells suggests that while there are treatment effects on both prices and volumes (Results 1 and 2, and Result 5, respectively), there is no causality between prices and volumes (in either direction).

Table B5: Observed correlations in data

Cells	All (pooled)	AB	AP	TB	TP
Price and volume, session-level data (all rounds)					
Asset H price vs. Asset H volume	+0.422**	−0.294	+0.119	+0.523	+0.331
Asset L price vs. Asset L volume	+0.266	+0.347	+0.353	+0.264	−0.189
Difference in prices vs. difference in volumes	+0.327*	+0.234	−0.127	+0.499	+0.042
Portfolio risk and earnings, subject-level data (round 10)					
Portfolio risk and portfolio balance (lab money)	−0.318***	−0.309***	−0.232**	−0.428***	−0.326***
Portfolio risk and expected earnings (real money)	−0.263***	−0.324***	−0.273***	−0.443***	−0.337***

The last two rows show that at both the pooled-cell and individual-cell level, there is a strong negative relationship between individual subjects' round-10 portfolio risk and their earnings (either in lab money or in expected real money according to the incentive scheme). Thus subjects taking on riskier portfolios (in effect, buying more of the high-risk asset) perform worse in the experiment, as would be expected given the typical over-pricing of the high-risk asset relative to its fundamental value.