Experiments on the determinants of asset price bubbles

Author:
Paul, Debapriya Jojo

Publication Date:
2015

DOI:
https://doi.org/10.26190/unsworks/18477

Experiments on the Determinants of Asset Price Bubbles

Debapriya Jojo Paul

A dissertation submitted in fulfilment of the requirements for the degree of Doctor of Philosophy (Ph.D.)

School of Banking and Finance

UNSW Business School

September 2015

Acknowledgements

This PhD is incomplete without acknowledging the enormity of the contribution that so many have made. First and foremost my supervisors, Associate Professor Julia Henker and Dr. Sian Owen, thank you for your unlimited support, patience, and guidance during my PhD pursuit. It has been a very long journey and your ideas, words of motivation, and encouragement have played a critical role in helping me achieve my goals.

I am thankful to Gavin Northey and Matt Tolhurst at the ASB Experimental Research Laboratory for helping ensure that my experiments ran as smoothly as possible. Partial funding for my experiments was also provided by the ASB Lab, for which I am very grateful. Thank you also to Ben Greiner for helpful advice regarding experimentation, and to Thomas Henker, for nudging me to get that paper done!

To the ASB admin staff, particularly Shirley Webster, Stephanie Osborne, and Ashley Coull, thank you for your assistance on a countless number of administrative tasks. In addition, I am appreciative of the financial support provided to me by the university through the APA Scholarship and PRSS travel funds.

My now former PhD colleagues, Eric Tan and Xiangping You, it was a pleasure to be on this journey with you. Your support and thesis stories have helped me normalise my experiences, and made it all the less stressful. To my dear friends, thank you for bringing balance to my life and for your continuous support, particularly during the final few months of this journey.

To my family, which has been the foundation to my success, motivation, and drive in this project, thank you for your unconditional love, support, and patience. Finally, to my wonderful wife, without your unflinching love and understanding, this journey would never have been made, and as the last 24 hours have so amply demonstrated, you make all the difference. Thank you for all that you do.

Abstract

Price ‘bubbles’, which refer to sustained overvaluation in an asset, represent a serious threat to the stability of markets. This dissertation contributes to the understanding of the determinants of bubbles in experimental asset markets.

The first study investigates whether bubbles emerge in experimental markets because the assets used are endowed rather than earned, thus diminishing their ‘legitimacy’. It takes a new methodological approach to this question by requiring participants in some markets to earn their initial allocation. The results suggest that asset legitimacy is not likely to be a serious threat to the validity of existing results, as the frequency, severity, and duration of bubbles does not noticeably differ between markets where the initial allocation is earned versus endowed.

The second study examines how relative-performance based compensation (‘tournament’ incentives) and its composition impacts price behaviour. Existing studies suggest that tournament incentives exacerbate bubbles, and that this worsens with experience. In contrast to the existing studies, which use single-asset markets, this study implements a two-asset market, allowing for more natural risk-taking. Mispricing in tournaments is found to diminish with experience, while compelling evidence of a difference in price behaviour between tournament and absolute-performance based compensation (‘normal’ incentives) is not detected. This outcome suggests the conclusions of earlier studies are likely driven by the single-asset nature of their markets. Furthermore, in markets containing inexperienced traders, adding penalties for underperformance is associated with less trading activity, but also larger and longer-lived bubbles compared to reward-only tournaments. This result is consistent with herding-driven price behaviour.

The third study explores whether peer effects driven by relative performance feedback explain the price behaviour observed in tournaments. In this study, normally-incentivised participants traded in markets while receiving periodic feedback about the average trader's performance. Results from these markets, assessed in conjunction with the results from the second study indicate that when traders are experienced and compensated under normal incentives, supplying relative performance feedback reduces mispricing. In contrast, introducing tournament compensation when relative feedback is already provided magnifies bubbles, especially under rank-based compensation. These results suggest that information on 'benchmarks' may aid market efficiency.

Contents

List of Tables

List of Figures

CHAPTER 1: Introduction

Almost a decade on, the legacy of the Global Financial Crisis (GFC) can still be felt in many parts of the world today. The liquidity crisis that debilitated financial markets across the world precipitated a deep and long-lasting downturn in global economic activity that was unprecedented in modern times. Of course, the now well-known trigger for the GFC and the ensuing Great Recession was the spectacular implosion of the US housing market in 2006-2007, which resulted in massive declines in household wealth and caused many mortgage-exposed institutions to suffer severe losses. Notably, the dramatic and widespread collapse in housing prices followed years of prodigious growth, particularly in the early part of this century. This narrative of boom and bust is unfortunately something of a recurring theme in asset markets. Similar historical episodes have been identified as far back as the 17th century in the price of Dutch tulip bulbs, while more recent examples include the U.S. stock market in the late 1920s, Japanese real estate and stocks in the late1980s, and internet-related or "dot-com" stocks in the late 1990's[1]. In all of these cases, the bust wiped out most of the gains accumulated during the boom in a relatively short period of time, and much like the recent crisis, real economic activity suffered acutely in the aftermath of the crashes

[1] For a detailed examination of many other historical examples, see Kindleberger and Aliber (2005).

of 1929 and 1990, heralding the arrival of the Great Depression and Japan's "lost decade" respectively.

What explains such a potentially destructive dynamic in asset prices? This dissertation is concerned with one frequently offered response to this question: a "bubble". Although definitions vary in their specificities, a bubble generally refers to a *sustained* period where an asset trades at prices not justified by its fundamental or intrinsic worth. Due to the historical prevalence and drama associated with booms and busts however, the classic characterisation of a bubble is one of persistent and growing *overvaluation*, followed often by a sudden and relatively rapid crash as the bubble 'pops' and prices re-align with fundamentals.

Of course, the concept is not without controversy. Due to its characterisation of long-lasting mispricing, proponents of informationally efficient markets argue that a bubble is an implausible abstraction because the actions of competitive rational arbitrageurs will move prices back to fundamentals quickly. However, it is now well established that arbitrageurs in the real world face numerous constraints – short-selling costs, capital constraints, and limited horizons, amongst others – that hinder their ability and/or willingness to trade against mispricing[2,3]. Indeed in some circumstances, arbitrageurs may prefer to 'ride the bubble' rather than trade against it, exacerbating the initial mispricing (e.g. De Long, Shleifer, Summers and Waldman 1990; Abreu and Brunnermeier 2003).

Nonetheless, while limits to arbitrage may explain the persistence of mispricing, it still leaves unanswered a crucial question: why does a bubble exist in the first place?

---

[2] Limits to arbitrage that restrict short-selling may also explain why bubbles are associated more with overvaluation. Gromb and Vayanos (2010) review the extensive theoretical literature on the limits of arbitrage.

[3] Conversely, Malkiel (2003) argues that the presence of a bubble is *not* inconsistent with the Efficient Market Hypothesis if the mispricing does not produce obvious profitable arbitrage opportunities. By this interpretation of informational efficiency, "no free lunch" takes precedence over "price is right".

This is a question that has attracted considerable attention in the theoretical literature. For example, the conditions under which so-called ‘rational bubbles’ may exist have been studied extensively. These models, in which a bubble grows (in expectation) into perpetuity, lay a basis for the fervent speculation that often accompanies booms, revealing that even if it is common knowledge that prices are too high, fully rational people may still choose to invest in a bubble because prices are expected to rise further; in the absence of common knowledge, bubbles may exist in even finitely-lived assets since investors can hope to profit by selling to an even ‘greater fool’ (Camerer 1989). Other theoretical efforts have demonstrated that bubbles may originate from sources such as agency conflicts (e.g. Allen and Gorton 1993; Allen and Gale 2000), positive feedback trading (De Long et al. 1990), relative wealth concerns (DeMarzo, Kaniel and Kremer 2008), and investor disagreement that potentially originates from behavioural biases (e.g. Scheinkman and Xiong 2003; Hong, Scheinkman and Xiong 2006)[4].

Alongside this theoretical work, the origins of price bubbles have been extensively studied in a parallel literature that draws on observations from the experimental laboratory. The great appeal of the experimental method for this task is that it addresses a key criticism directed at empirical studies that utilise historical data – bubbles in the real world can almost never be identified with complete certainty, even in hindsight, because fundamental values are unobservable. This problem is absent from an experimental setting because experimenters have full control over the fundamental value process. In addition, the ability to exercise control over variables such as the information environment, incentives, market rules, and institutions means that experimental results are less prone to the potential sources of confound that afflict studies involving observational data. While this additional control does come at the

[4] See Scherbina and Schlusche (2014) and Brunnermeier and Oehmke (2012) for reviews of the bubbles literature. Camerer (1989) provides a review of the early theoretical literature.

expense of potentially diminished generalisability, the experimental method nonetheless represents an important tool in understanding the phenomenon of bubbles.

The now canonical experimental framework within which asset price bubbles are studied, developed by Smith, Suchanek and Williams (1988), involves participants trading units of a risky dividend-paying asset in a double-auction market for a fixed number of trading periods, after which the asset expires worthless. The distribution governing dividends and the declining (risk-neutral) fundamental value is known to all traders at all times. Under classical assumptions, the process of backward induction should rule out the existence of a bubble in this environment. However, with inexperienced traders Smith et al. (1988) observe that prices in these markets generally follow a marked bubble-and-crash trajectory – prices begin below fundamental value, rise quickly above fundamental value during the middle periods before crashing back to fundamental value towards the end – which only dissipates in repeated trials with twice-experienced subjects.

As a consequence of the reliability with which bubbles can be produced in this set-up, the Smith et al. (1988) design has become fertile ground for a vast literature that, by modifying elements of the basic design, has sought to understand the factors that drive and diminish bubbles. Examples of this include, but are certainly not limited to, the impact of derivative markets (e.g. Porter and Smith 1995), trading institutions (e.g. Lugovskyy, Puzzello and Tucker 2009), the number of assets (e.g. Fisher and Kelly 2000) market liquidity (e.g. Caginalp, Porter and Smith 1998), short-selling (e.g. Haruvy and Noussair 2006), participant irrationality (e.g. Ackert, Charupat, Deaves and Kluger 2009), rational speculation (e.g. Lei, Noussair and Plott 2000), trader experience (e.g. Dufwenberg, Lindqvist and Moore 2005), compensation incentives (e.g. James and Isaac 2000), asset characteristics (e.g. Noussair and Powell 2010), participant confusion (e.g. Huber and Stöckl 2012), asymmetric information (e.g. Sutter, Huber and Kirchler 2012), and emotions (e.g. Andrade, Odean and Lin 2013)[5].

This dissertation adds to our understanding of the determinants of price bubbles by examining how the design of market experiments, compensation incentives, and the information environment contributes to the formation of bubbles in an experimental setting. While each study is discussed in more detail below, briefly, the first study asks if the act of endowing participants with trading funds at the onset of the experiment unintentionally contributes to the bubbles observed. The second study examines how the structure of tournament compensation contracts, which compensate traders on the basis of relative performance, impacts bubble behaviour in multi-asset markets. The third study investigates whether the price behaviour observed under tournament compensation is driven by monetary incentives or traders' intrinsic desire to outdo others.

1.1 Are the assets used in markets experiments legitimate?

Virtually all market experiments begin with participants being endowed with a trading portfolio. However, this simple and seemingly benign act may fundamentally alter how participants behave during the experiment because, having not been *earned* by participants, the assets in the portfolio may lack 'legitimacy'. Put simply, subjects may perceive the received funds as windfall gains or "other people's money", generating a house-money effect (Thaler and Johnson 1990) in which participants take more risk than they otherwise would with their own earned money. Clearly, this poses a potential problem for the validity of asset market experiments because heightened speculation arising from the illegitimacy of assets may confound the effect of the variable that the

[5] The interested reader may refer to Palan (2013), who provides a comprehensive review of the literature on bubbles in experimental asset markets.

experimenter is truly interested in. Chapter 2 explores this possibility by examining whether prices behave differently when participants are required to trade with earned wealth compared to unearned wealth.

Other studies have produced both affirmative (Ackert, Charupat, Church and Deaves 2006a) and negative (Schwarz and Ang 1989[6]; Ang, Diavatopoulos and Schwarz 2010) results regarding the importance of house-money effects in experimental assets, albeit with designs that are difficult to compare. The study in Chapter 2 adopts a new methodological approach to the problem, following Cherry, Frykblom and Shogren (2002) by implementing an earnings task prior to the market in one treatment, while participants in the other treatment are simply endowed their portfolios. The results do not indicate a significant difference in price behaviour between markets where the initial allocation is 'free' versus markets where it is earned.

These findings contribute to the broader literature on asset legitimacy in economic experiments, which suggests that the importance of asset legitimacy varies according to the type of experiment; behaviour in dictator games is quite sensitive to earned wealth (Cherry, Frykblom and Shogren 2002; Oxoby and Spraggon 2008), while behaviour in public good experiments is not (Clark 2002; Cherry, Kroll and Shogren 2005). The results in Chapter 2 indicate that asset legitimacy may not be a pressing concern for market experiments.

1.2 Tournaments in two-asset markets

A tournament or contest is characterised by compensation that depends on an individual's *relative* performance, and is representative of many of the incentive schemes operating in the finance industry. Since the most lucrative prizes go to winners,

[6] In Porter and Smith (1995)

tournaments entail a large upside, which is often not matched by a corresponding downside; in other words, they are convex incentive schemes. Considerable concerns have been raised, especially in the fall-out from the GFC, that such schemes endanger the stability of markets via the potentially excessive risk-taking that they encourage (see Rajan 2008; Blinder 2009). Despite this, studies of the aggregate impacts of tournament incentives remain sparse; hence, the market-level implications are not well understood.

The few studies that have investigated this issue reach the concerning conclusion that tournament incentives exacerbate bubbles in experimental markets relative to absolute-performance based incentives, and that this effect only worsens with trading experience (James and Isaac 2000; Cheung and Coleman 2014). However, participants in all of these studies are only able to trade a single risky asset, which unduly restricts the risk-taking options available to traders relative to real-world markets. Hence, the generalisability of these results is uncertain. The study in Chapter 3 addresses this issue by examining bubble behaviour under tournament incentives in a market where participants can trade two *differentiated* risky assets, allowing traders to vary their risk-exposure more naturally than in a single-asset market.

If higher prices under tournament incentives arise as a result of risk-taking fuelled by the convexity of tournament payoffs, a possible 'solution' is to introduce a penalty for poor relative performance into the tournament contract. The presence of a penalty induces people to take less risk because it creates a consequence for underperforming (Gilpatric 2009; Qiu 2003; Kempf, Ruenzi and Thiele 2009; Hu, Kale, Pagani and Subramanian 2011). However, it is unclear what the aggregate or pricing implications of this reduced risk-taking would be, since the fear of underperforming may perversely lead to higher prices and larger bubbles if it prompts traders to herd as a form of insurance (Rajan 2006; Dass, Massa and Patgiri 2008). The aggregate impact of penalties in tournament contracts has received very little attention in the existing literature. Chapter 3 fills this gap by studying how the addition of a penalty into a reward-only tournament contract affects price behaviour in a two-asset experimental market.

In contrast to the existing literature, bubbles in the current study dissipate with experience under tournament incentives in two-asset markets. Furthermore, compelling evidence that prices under tournament incentives are significantly different to those under absolute-performance based incentives is not detected in two-asset markets. Hence, these results suggest that a single-asset environment is an important factor behind the results observed in past studies of tournaments. Moreover, Chapter 3 reveals that adding a penalty to a tournament contract in which participants are rewarded for above-average performance results in *lower* trading volumes, but *longer* booms for both assets, and *higher* prices for the riskier of the two assets. Overall, these effects are more consistent with predictions of prices under herding, although they do not survive with experienced participants.

In addition to augmenting the literature on tournament incentives in experimental markets, Chapter 3 also contributes to the strand of the experimental bubbles literature that examines trade in multiple risky assets (Fisher and Kelly 2000; Ackert, Charupat, Church and Deaves 2006b; Childs and Mestelman 2006; Chan, Lei and Vesely 2013) by introducing tournament incentives. Price behaviour under tournaments in these markets broadly resembles the earlier literature – individual assets do bubble and crash, but relative prices between assets generally remain ‘correct’.

1.3 Are tournament effects driven by social competition?

Schoenberg and Haruvy (2012) show that relative performance feedback affects the size of bubbles in experimental markets, even in the absence of tournament incentives. This finding, which is consistent with the idea that people have intrinsic competitive preferences, or a desire to "keep up with the Joneses", potentially suggests that the price behaviour observed under tournament incentives may also be driven by this inherent competitive desire to outdo others, rather than the monetary incentives that define tournaments. This prospect is explored in Chapter 4, which builds on the study in Chapter 3 by examining absolute performance-based markets in which relative performance feedback is provided. By then comparing it to the data from the study in Chapter 3, Chapter 4 is able to isolate the incremental effect of relative performance information from the incremental effect of competitive monetary incentives on prices.

The results of this study suggest that relative concerns induced by the availability of relative performance information may help produce more efficient markets. When traders are compensated for their absolute performance, introducing relative performance feedback has the effect of producing *smaller* bubbles, but only with experienced traders. Adding tournament compensation to a trading environment where relative performance feedback is already provided has the opposite effect, *increasing* the size of bubbles.

1.4 Structure

This dissertation comprises three studies that examine the determinants of asset price bubbles in experimental asset markets. These studies are presented in Chapters 2-4, each of which contains its own introduction, literature review, experimental design, results, and conclusion.

Chapter 2 examines the influence of asset legitimacy on bubbles in experimental markets. Specifically, it studies if prices in experimental markets behave differently when participants have to earn their initial portfolios rather than having it endowed to them.

Chapter 3 is a study on the aggregate impacts of tournament incentives. In particular, it investigates how tournament incentives affect the behaviour of prices relative to normal incentives when experimental markets contain more than one type of risky asset. In addition, the study in Chapter 3 examines how the composition of a tournament contract, specifically the balance between rewards for outperforming others and the penalties for underperformance, affects price behaviour in these markets.

Chapter 4 examines how social comparison influences the behaviour of asset prices in experimental markets. Specifically, it seeks to isolate the price-impact of two different incentives created by tournament schemes – the competitive pecuniary incentives that define tournaments, and the intrinsic competitive incentives generated by the availability of relative performance feedback.

Chapter 5 concludes with a summary of the main findings, a discussion of the limitations and implications of each study, and suggests avenues for future research.

1.5 Presentations and Publications

The research comprising this dissertation has been presented at both domestic and international conferences, as detailed below.

The research comprising Chapter 2 was presented at the 2013 Accounting & Finance Association of Australia and New Zealand Conference (Perth, Australia), the 2013 Behavioural Finance and Capital Markets Conference (Adelaide, Australia), and the 2013 Behavioural Finance Working Group Conference (London, UK). This Chapter is forthcoming in the *Journal of Behavioral Finance*.

Chapter 3 is scheduled to be presented at the 2015 European Financial Management Association Conference (Amsterdam, Netherlands).

CHAPTER 2: Asset Legitimacy in Experimental Markets

2.1 Introduction

Individual behaviour in economic experiments frequently deviates from the predictions of traditional economic theory. Recent evidence suggests that the prevalence of this phenomenon may be explained by the origin of the assets used in these experiments. In particular, if the claims to assets are not *legitimate* in the sense that they are simply endowed to participants – as is predominantly the case – rather than having been earned through effort, then participants may treat those assets as “other people’s money”. Hence, they may exhibit unexpectedly high levels of other-regarding (Cherry et al. 2002; Oxoby and Spraggon 2008) and risk-taking behaviour (Thaler and Johnson 1990; Arkes et al. 1994). We examine whether this issue of asset legitimacy explains one of the long-standing puzzles in experimental economics – the severity and frequency with which asset price bubbles occur in experimental asset markets of the type designed by Smith et al. (1988).

A ‘price bubble’ is defined as a sustained period in which the market price of an asset deviates from (normally exceeding) its intrinsic or fundamental value. They pose serious challenges for investors, policymakers, and regulators alike due to their distortionary effects on the price signal, and the economic disruption inflicted if they burst. Since the seminal study by Smith et al. (1988) that first documented a bubble-and-crash pattern to prices in experimental continuous double-auction asset markets containing inexperienced participants, an extensive literature has emerged that seeks to understand the drivers of bubbles by varying specific participant and/or institutional characteristics of the original experimental design. However, in virtually all such studies participants are initially provided with an allocation of cash and stock, which they use to trade in the experimental market. This failure to legitimise the assets through effort, essentially giving participants ‘free money’, has the potential to elicit a “house-money effect”, whereby participants react to their windfall gain by being more risk-taking than they would naturally be if they were trading with their own (earned) money (Thaler and Johnson 1990). As a result, they may be more willing to overpay for the asset and engage in speculation, leading to the generation (or at least, amplification) of the bubble-and-crash phenomenon typically seen in such studies.

Earlier studies on house-money effects in experimental market settings are sparse and provide conflicting evidence. While Schwarz and Ang (1989, in Porter and Smith 1995) and Ang et al. (2010) do not find a significantly dampened bubble-and-crash pattern to their prices even when participants are required to trade with their own money (which may or may not be earned), Ackert et al. (2006a) report a tendency for prices to be significantly higher (and remain so) in markets in which participants are given larger endowments. However, the comparability of such studies is complicated by the fact that they use significantly different experimental designs.

This study contributes to the literature on house money effects in experimental asset markets by taking an alternative methodological approach. The specific question we seek to answer is whether prices in experimental asset markets behave differently when participants are required to trade over earned wealth compared to unearned wealth. If the associated house-money effects are important in such markets, then bubbles/mispricing should be significantly more common and severe when trade occurs using endowed wealth. By examining a previously untested class of experiments, we also contribute to the on-going debate in the economic literature about the need to legitimise assets with effort in economic experiments in order to conduct valid tests of theory.

A two-treatment experimental design based on Cherry et al. (2002) was implemented to answer our question; one treatment called *Earned* involved participants completing a money-earning task (a GMAT-based quiz) that determined their initial allocation of cash and stocks, while in the second, called *Free*, participants were randomly assigned their initial endowment. Participants in both treatments then traded in an experimental asset market based on the standard double-auction asset market design created by Smith et al. (1988).

Our results do not support the claim that the allocation of ‘free money’ to traders has a significant impact on price behaviour. Markets in both treatments were characterised by the formation of price bubbles with similar frequency, and the size of the bubbles/mispricing is not significantly different in the *Free* treatment and the *Earned* treatment. The use of house-money also did not significantly change the length/duration of any bubbles/mispricing in the *Free* treatment compared to the *Earned* treatment. Hence, issues of asset legitimacy may not be particularly salient for Smith et al. type asset markets.

The remainder of this chapter is organised as follows. The next section reviews the related literature and develops the research hypotheses. Following this, the experimental methodology and results are described in sections 2.3 and 2.4 respectively, while conclusions are presented in section 2.5.

2.2 Literature Review

2.2.1 Endowment origin and asset legitimacy

Traditional (normative) economic theory's contention that the origin of wealth is irrelevant to the decision-making process[7] – that the level of total current wealth is what matters, not how it is obtained, and by extension that incremental costs/benefits are relevant to decisions, whilst historical costs/benefits are not – has been challenged in both the economics and psychology literature. Experimental evidence shows that real peoples' decisions are sensitive to sunk costs (Arkes and Blumer 1985; Garland 1990) and that prior gains or windfall gains increase the propensity for individuals to consume and take risk (Thaler and Johnson 1990; Arkes et al. 1994). Thaler and Johnson named this latter phenomenon a 'house-money' effect, which conveys the intuition that people appear to be more willing to risk losing what they consider 'other people's money' than their own.

Given that the endowments used to initiate economics experiments can themselves be characterised as windfall gains or 'other people's money' for participants, recent attention has been directed towards determining if participants behave differently when required to earn their initial allocations. The most striking effects of endowment origin are found in dictator games, which involve two players – the 'dictator' and 'responder' – in which the dictator must decide how to split a certain sum of money between himself/herself and the responder (the 'responder' is passive and must accept whatever is offered). Although the game-theoretic equilibrium predicts that the dictator

[7] This assumption is known as the fungibility of money/income.

will offer nothing to the receiver, experiments using house-money have consistently shown that dictators exhibit other-regarding behaviour, offering a significant portion of their endowments to the receiver. However, when Cherry et al. (2002) require their dictators to earn their wealth via a money-earning task (a GMAT-based quiz), they find that other-regarding behaviour is virtually eliminated. Oxoby and Spraggon (2008) produce similar results in their dictator game experiments and interpret the process of legitimising claims to the assets of the experiment via the expenditure of effort as akin to the establishment of property rights.

However, asset legitimacy does not have similarly strong effects in all types of economic experiments. For example, the level of free-riding or lack-thereof in public good experiments is not affected by whether participants contribute earned money or house money (Cherry et al. 2005; Clark 2002). Hence, the generalisability of endowment-origin effects is an open empirical question. In this study, we seek to examine whether endowment-origin is a relevant concern for a different class of economic experiments – asset markets.

2.2.2 House-money effects in experimental asset markets

Since their seminal study, the experimental asset market designed by Smith et al. (1988) has provided the most reliable means to study asset price bubbles; unlike real-world data, the experimenter can observe the fundamental value of an experimental asset and control the information environment in which investors (i.e. participants) trade. The now standard/baseline experimental market design comprises a continuous double-auction market in which participants trade for a finite number of periods a homogenous hypothetical asset (a ‘stock’) that pays a stochastic dividend at the end of each trading period whose value is drawn from a probability distribution that is known to all participants (dividend draws are i.i.d.). In such an environment, backward induction under classical assumptions should rule out the existence of trade at values exceeding the risk-neutral fundamental value. Yet, despite participants possessing common knowledge about the dividend generating process, prices in these markets regularly exhibit marked bubble-and-crash patterns, typically starting below but rising rapidly above the fundamental value, before eventually crashing back to the intrinsic value towards the end.

Research following Smith et al. (1998) focused on replicating this general experimental design while modifying specific aspects, for example, by introducing elements such as short-selling, margin buying, brokerage fees, or futures markets, amongst others, to discern which factors helped moderate or eliminate bubbles[8]. The results of these studies indicate that bubbles in experimental markets are robust to the alteration of numerous participant and institutional features. Only common group experience with the experimental market design, at least on the part of a portion of market participants (Porter and Smith 1994; Dufwenberg, Lindqvist, and Moore 2005; Haruvy, Lahav and Noussair 2007) is sufficient to reliably ensure trade at fundamental value. More recent research suggests that the inducement of common expectations through training in the fundamental value process (Cheung et al. 2014), or the reduction of apparent participant confusion (Huber and Kirchler 2012; Kirchler et al. 2012) may also have a similar effect.

A common characteristic of virtually all these bubble experiments is that participants are endowed with a combination of cash and/or stock before the commencement of trade. This act of giving participants ‘free money’ to trade with may contribute to the generation of bubbles by inadvertently influencing participants’ risk

[8] See Palan (2013) for a comprehensive review of the experimental bubbles literature.

appetites via a house-money effect.

House-money effects have previously been studied in an experimental asset market setting by Schwarz and Ang (1989), Ackert et al. (2006a), and Ang et al. (2010). The experiments conducted by Schwarz and Ang (1989) and Ang et al. (2010) required participants to use their own money to trade in some sessions. Prices in their markets however are still prone to bubble and crash, suggesting that any house money effect is negligible. However, a potential complication with this approach is that the origin of the money is not controlled for – that is, the money brought in by participants is earned by presumption only. Corgnet et al. (2014) also point out that asking participants to bring their own money may induce a selection bias, whereby predominantly risk-seeking types who are happy to lose their money self-select into the subject pool. In addition, the results of Ang et al. (2010) are based on a very small sample size (2 sessions), owing to the fact that their examination of house-money effects is a robustness test, rather than the focus of their study.

Instead of requiring participants to use their own money, Ackert et al. (2006a) vary the size of the endowment given to participants, and detect a significant house-money effect on asset prices in their experimental market. They find that participants who are given a larger endowment are willing to bid larger amounts to acquire the asset. As a consequence, market prices are also significantly higher, and remain so for the duration of their market. These seemingly conflicting results are not directly comparable however, since Ackert et al. (2006a) use a markedly different experimental design, namely a Vickery auction in which participants are only able to place bids to purchase new (but identical) assets in each trading period, as opposed to the double auction market used in a typical bubble experiment that allows traders to buy and sell a fixed number of assets. The duration of their market is also significantly shorter, consisting of only 3 trading periods instead of the typical 15. In addition, Ackert et al. do not explicitly examine the issue of price bubbles, making it difficult to draw conclusions regarding the issue.

This study seeks to reconcile the results of the preceding research by taking an alternative approach to examining the house money effect. Specifically, we incorporate an element of the Cherry et al. (2002) methodology – a money-earning stage – into a Smith et al. (1988) type market. This allows for a clearer differentiation between 'house/unearned money' and 'earned money', and hence a more robust test of the effect of house-money/endowment-origin on prices in experimental asset markets.

We note that a concurrent study by Corgnet et al. (2014) exists which also examines the impact of earned and unearned wealth on price bubbles using a money-earning task. However, our study differs from theirs along two key dimensions. First, whereas they employ a "real effort task" that involves participants contributing to the development of a research database by downloading academic research papers, we use a GMAT-based quiz (as in Cherry et al. (2002)). Second, we explicitly map effort to earnings in our task. In contrast, the "real-effort task" employed by Corgnet et al. is characterised by a fixed payment that is the same for all participants regardless of how much effort they actually expend on the task[9]. Hence, a comparison of our result(s) to that of Corgnet et al. is useful in the sense that it reveals the extent to which earned/unearned money effects are sensitive to task-type.

If trading with endowed money does not result in significantly more risk-seeking behaviour on the part of (inexperienced) participants, then one would expect prices in

[9] Another difference between our designs is that Corgnet et al. employ a certain dividend in their market, whereas our dividend is stochastic. Given that dividend certainty does not produce significantly different price behaviour to uncertain dividends (Porter and Smith 1995), we do not expect this to explain any observed differences between the studies.

markets where trade occurs with house-money to behave similarly to prices in markets where participants are trading with earned money. Consequently, we test the following two null hypotheses using numerous measures of the magnitude and duration of bubbles that exist in the experimental asset market literature (defined in section 2.4).

H1: *Mispricing/Overvaluation in markets where participants' initial allocation is endowed is not different in magnitude to markets where participants' initial allocation is earned.*

H2: *Mispricing/Overvaluation in markets where participants' initial allocation is endowed is not longer in duration than in markets where participants' initial allocation is earned.*

2.3 Experimental Design

The experiment consisted of 16 sessions conducted at the ASB Experimental Research Laboratory at the University of New South Wales in August and October 2012, using a student sample recruited through ORSEE (Greiner 2004). A total of 459 students, predominantly undergraduate, participated in the study, none of whom had any prior experience with market experiments[10]. The experiment was computerised using zTree (Fischbacher 2007), with the exception of an end-of-experiment questionnaire, which was completed on paper. All trading was conducted in 'francs' (experimental currency), with earnings converted and paid out in Australian dollars at the end of the experiment at an exchange rate that varied depending on the treatment and the session.

[10] Some participants with multiple ORSEE profiles managed to participate in more than one session of this experiment. Markets which contained these participants are marked with an asterisk in Table 1, and were excluded from the analysis to mitigate any potential confounding of treatment effects.

2.3.1 Treatments

A between-subjects design was implemented, consisting of two treatments – *Earned* and *Free* – which differ only in the way in which the initial allocation of assets and cash in the market was assigned to traders. Prior to their involvement in the market, participants in the *Earned* treatment completed a task that determined their initial allocation. As in Cherry et al. (2002), the task was a timed quiz consisting of questions from the Graduate Management Admissions Test (GMAT)[11]. However unlike their study, in which participants had to answer 17 questions in 40 minutes, subjects in this study were given the task of answering 10 multiple-choice questions (5 numerical reasoning, 5 verbal reasoning) in 20 minutes[12].

To create an incentive for participants to expend effort on our task, the size of the earnings from the task (the initial market allocation) was linked to participants' relative performance. Performance was measured by the number of questions answered correctly. In the event of a tie, the amount of time taken to complete the quiz was considered; the participant who took less time was deemed to have performed better.

The top 50% of performers were allocated to *'high-stakes'* (*HS*) markets, where the initial allocations consisted of twice the amount of cash and assets received by those in *'low stakes' (LS)* markets, to which the bottom 50% were allocated. In effect, participants *earned* their initial allocations in the market. All experiment sessions were designed to consist of two *HS* and two *LS* markets. Once participants were assigned to either the *HS* or *LS* category of markets on the basis of their performance in the task,

[11] Given that participants were university students, the completion of a quiz should have resembled or felt like 'work'.

[12] In addition to multiple-choice numerical and verbal reasoning questions, the GMAT-based quiz used by Cherry et al. (2002) also contained a number of extended response questions. We chose to exclude extended response questions from our quiz due to considerations regarding the length of the experiment, and also to computerise the implementation and grading of the quiz. The 20 minutes given to participants to complete the quiz was selected to give participants more than enough time to attempt all questions, and in fact, 90% of subjects completed the quiz before the allotted time expired.

they were then randomly allocated to one of the two independent *HS*/*LS* markets[13]. That is, their performance in the task played no role in determining which specific *HS*/*LS* market they were allocated to. To the extent that GMAT performance correlates with intelligence (and intelligence correlates with trading behaviour), this was done in order to mitigate the possibility that prices in one *HS*/*LS* market varied systemically from the other because it contained more intelligent participants.

In contrast, participants in the *Free* treatment did not complete a task. Instead, they were randomly assigned to a *HS* or *LS* market. Hence, their initial portfolios were simply endowed.

2.3.2 Market structure

Each experiment session held in August (October) was designed to run 4 separate markets – 2 *HS* and 2 *LS* – of 8 (6) traders each. The market, which was computerised, allowed subjects to trade units of a risky asset called 'X'[14]. The market ran for 10 periods, each lasting 3 minutes in the August sessions, and 2 minutes in the October sessions[15]. Trade occurred according to continuous double auction trading rules (Smith et al. 1988) with an open order book; in each trading period, traders were able to post bids and asks, and/or accept any posted bid or ask, subject to the constraints posed by their holdings of cash and the risky asset[16]. All trades were for single units of the asset. Short selling and buying on margin were not allowed.

[13] Note that *HS* (*LS*) traders only traded with other *HS* (*LS*) traders allocated to the same market.
[14] The parameters of the market were adapted from Dufwenberg et al. (2005), whose design consists of minor variations from the baseline market design of Smith et al. (1988) in relation to market length, trading period length, the number of traders, and the distribution of the dividend. Dufwenberg et al. markets produce price and trading features that qualitatively mirror those of Smith et al. (1988). Our market parameters in the October sessions are identical in all respects to that of Dufwenberg et al. Our August sessions differed from this design in terms of the number of traders (we have 8 vs. 6), and the length of a trading period (3 min. vs. 2 min.).
[15] Trading periods in October are shorter because the conditions of the funding for the October sessions necessitated a shorter experiment length.
[16] The open order book represents another deviation from the Smith et al. (1988) design. The depth of the order book does not significantly affect price behaviour in these markets (Caginalp et al. 2001).

At the end of each period, the asset paid a dividend of 0 or 20 francs with equal probability; all units of the asset paid the same dividend at the end of a given period. Dividends were drawn independently each period by the computer, and the probability distribution governing them was known to all participants. Any non-zero dividends paid were added to the trader's cash balance and their end-of-period portfolio carried over to the next period. Since the average dividend in each period is 10 francs, the risk-neutral fundamental value of the asset is equal to the expected total future dividend stream, or 10 multiplied by the number of remaining trading periods (including the current one). Hence, the fundamental value of the asset in our market declined in steps of 10 from 100 francs in period 1 to 10 francs in period 10, before expiring worthless after the final dividend draw.

Traders commenced the market with an initial allocation of cash and assets that depended on whether they were (a) assigned to a *HS* or *LS* market and (b) assigned trader type 1 or 2.[17] Subjects knew their own initial allocations, but did not know the initial allocations of others in their market. In each *LS* market, half the traders were randomly assigned type 1, and began the market with 6 units of the asset and 200 francs, while the remainder, type 2, were allocated 2 units of the asset and 600 francs[18]. Since the fundamental value of the asset is 100 at the start of the market, all traders in the *LS* market began with a portfolio initially worth (in expectation) 800 francs. Since Type 1 and 2 traders in the *HS* markets were allocated twice the amounts of cash and assets as their analogues in the *LS* market, the initial expected value of the portfolios of all *HS* market traders was twice as much as the *LS* portfolios, or 1600 francs.

The allocations described above determine the initial liquidity of our markets,

[17] Subjects in both treatments did not know which trader type they were. Subjects in the *Free* treatment also did not know if they had been allocated to a high-stakes or low-stakes market.
[18] Markets with an odd number of traders had an extra Type 1/Type 2 trader.

which is measured by the cash-to-assets ratio – the ratio of total cash to total fundamental value of all assets at the beginning of the market. Increasing the initial liquidity in an experimental market has been observed to increase the magnitude of bubbles (Caginalp et al. 1998, 2000, 2001). Hence, to control for this factor, all our markets were designed to have an initial cash-to-assets ratio of 1, provided that markets contained an even number of participants. However, as some sessions ran with fewer than the full complement of participants, the actual number of markets in a session, the number of traders in a market, and consequently the initial cash-to-assets ratio occasionally varied from the intended design. Summary information on all experimental sessions, including the number of traders and the initial cash-to-assets ratio in each market is provided in Table 2.1.

2.3.3 Procedures

Each experimental session corresponded to a single treatment, and sessions (hence, subjects) were randomly assigned to a treatment. *Earned* treatment sessions consisted of three stages – *Task*, *Market*, and *Questionnaire* – while *Free* treatment sessions consisted only of the latter two. The duration of an *Earned* treatment session was approximately 2 hours (1.5 hours in the October sessions), while the *Free* sessions ran for about 1.5 hours (1 hour in the October sessions)[19]. Written instructions were given to participants in all stages, which were also communicated verbally by the experiment administrator[20]. To mitigate potential interaction effects, participants were

[19] All experimental sessions (for both treatments) were advertised as lasting 2 hours (1.5 hours for the October sessions) to ensure that (self-)selection biases induced by the relative attractiveness of the length of a treatment's session did not render one treatment's subject pool systematically different to the other.
[20] The written protocol for the market stage, which can be found in Appendix A2, was adapted from those used by Dufwenberg et al. (2005), Noussair et al. (2001), Noussair and Powell (2010) and Lugovskyy et al. (2009). The experiment administrator read from a script to ensure consistency in the delivery of verbal instructions between sessions and to help mitigate the possibility of experimenter-induced biases. Participants were also given time to read the instructions on their own, and to ask any clarifying questions privately (which were also answered privately).

Table 2.1: Experimental sessions

Panels A and B of this table provide summary information on the experimental sessions of the *Earned* and *Free* treatments respectively. Each session was designed to run two 'High-Stakes' (HS) and two 'Low-Stakes' (LS) markets. The allocation of assets and cash within a market was designed so that the initial ratio of total cash to total asset value (Cash-to-Assets ratio) in a market was equal to 1. The actual number of markets in a session, the number of traders in a market, and the initial Cash-to-Assets ratio occasionally varied from this design due to an insufficient number of participants attending a session. Markets marked with an asterisk are 'contaminated', which means that the market contained a trader who participated in an earlier session of the experiment.

Panel A: Earned Treatment

High Stakes Markets				Low-Stakes Markets			
Session	*Market*	*No. Traders*	*Initial Cash/Assets*	*Session*	*Market*	*No. Traders*	*Initial Cash/Assets*
E1	E_HS1	8	1	**E1**	E_LS1	8	1
	E_HS2	8	1		E_LS2	6	1
E2	E_HS3	8	1	**E2**	E_LS3	8	1
	E_HS4	8	1		E_LS4	8	1
E3	E_HS5	8	1	**E3**	E_LS5	7	0.87
	E_HS6	8	1		E_LS6	8	1
E4	E_HS7	8	1	**E4**	E_LS7	7	0.87
	E_HS8	8	1		E_LS8	7	1.15
E5	E_HS9	8	1	**E5**	E_LS9*	8	1
	E_HS10	8	1		E_LS10	8	1
E6	E_HS11	8	1	**E6**	E_LS11*	8	1
	E_HS12	8	1		E_LS12	8	1
E7	E_HS13	6	1	**E7**	E_LS13	6	1
	E_HS14	6	1		E_LS14	6	1
E8	E_HS15*	6	1	**E8**	E_LS15*	8	1
	E_HS16	6	1				

Panel B: Free Treatment

High Stakes Markets				Low-Stakes Markets			
Session	*Market*	*No. Traders*	*Initial Cash/Assets*	*Session*	*Market*	*No. Traders*	*Initial Cash/Assets*
F1	F_HS1	8	1	**F1**	F_LS1	8	1
	F_HS2	8	1		F_LS2	8	1
F2	F_HS3	8	1	**F2**	F_LS3	8	1
	F_HS4	7	0.87		F_LS4	7	0.87
F3	F_HS5	8	1	**F3**	F_LS5	8	1
	F_HS6	8	1		F_LS6	8	1
F4	F_HS7	8	1	**F4**	F_LS7	7	0.87
	F_HS8	7	0.87		F_LS8	8	1
F5	F_HS9*	8	1	**F5**	F_LS9	8	1
	F_HS10	8	1		F_LS10	8	1
F6	F_HS11	7	0.87	**F6**	F_LS11	8	1
	F_HS12	8	1		F_LS12	8	1
F7	F_HS13	8	1	**F7**	F_LS13	9	0.89
F8	F_HS14	6	1	**F8**	F_LS14	6	1
	F_HS15	6	1		F_LS15	6	1

not allowed to communicate with each other for the duration of the experiment, while the anonymity of their data was ensured by randomly allocating ID numbers to participants before the start of the experiment.

In the *Task* stage, subjects in the *Earned* treatment were given 20 minutes to complete the earnings task on the computer. They were informed that their relative performance in the task would determine their initial allocation of cash and assets in the market, with the top 50% of performers being assigned to markets where the initial portfolios would be twice the size of the initial portfolios in the markets to which the bottom 50% would be allocated. Once all subjects had completed the task, each subject was shown an on-screen summary of their performance including their rank, their allocated market type, and the exact allocation of cash and assets that they would begin the *Market* stage with[21].

The *Market* stage began with participants receiving instructions on how to use the market's trading screen to make and accept bids and offers (5 minutes), followed by 10 minutes where subjects practiced trading using the interface. After the end of the practice period, participants were instructed on the other features of the asset market, following which the market proper began. At the end of each trading period, traders were shown a summary screen of their dividend earnings for that period and their end-of-period cash balance and asset inventory.

Following the end of the market, the final stage involved participants completing an end-of-market questionnaire, which gathered general demographic information about the subject pool, and their experience(s) and thought-processes during the market[22].

[21] Participants were told that they were assigned to market type "A" or "B", rather than "high-stakes" or "low-stakes" markets. The more neutral language of the former is less likely to have an unexpected impact on behaviour.
[22] The questionnaire, which can be found in Appendix A3, is a modified version of the one used by Ackert and Church (2001).

Participants were then called up one-by-one, paid their earnings (in envelopes) and dismissed. Participants' earnings from the experiment consisted of their (converted) market earnings plus a $5 participation fee. As each unit of the risky asset expired worthless at the end of the market, participants' market earnings were equal to the Australian dollar equivalent of their cash balance at the end of the market[23]. Participants' average earnings from the experiment were $30.

2.4 Results

We test our hypotheses by separately comparing *HS*/*LS* markets in the two treatments. Given the aforementioned positive association between bubble behaviour and the initial cash-to-assets ratio, we control for its effect by restricting our analysis to markets with an initial cash-to-assets ratio of 1. This results in the loss of 6 *Free* treatment markets (3 *HS* and 3 *LS*), and 3 *Earned* treatment markets (all *LS)* from the data. To control for the effects of experience, we also exclude any markets that contained subject(s) who had participated in an earlier session of the experiment. This happened with 5 participants who had multiple ORSEE accounts, and results in the loss of 4 *Earned* treatment markets (1 *HS* and 3 *LS*) and 1 *Free* treatment market (*HS*), all of which are identified in Table 2.1. Our analysis is conducted using the remaining 15 *HS* and 9 *LS* markets in the *Earned* treatment, and 11 *HS* and 12 *LS* markets in the *Free* treatment.

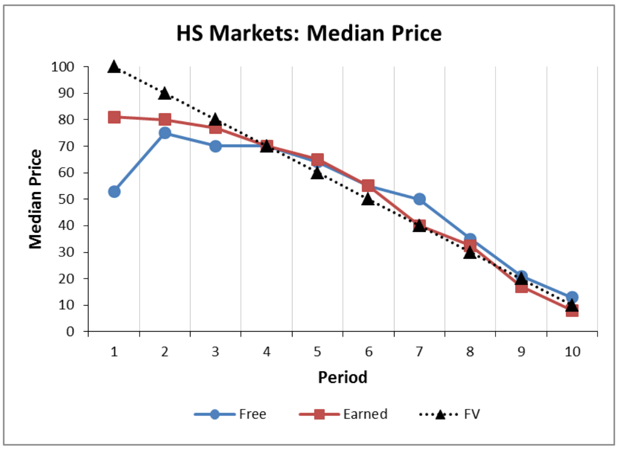
2.4.1 Descriptive summary

Figures 2.1(a) and 2.1(b) chart, respectively, the evolution of the median

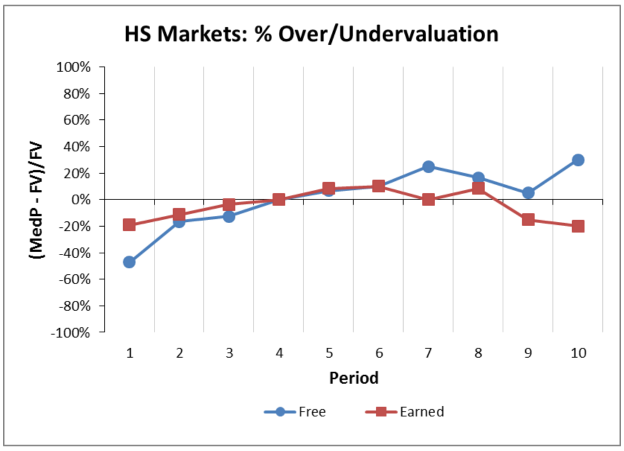
[23] Ending cash balance = initial cash balance + dividend earnings + sales revenue – expenditure on purchases

Figure 2.1: Median prices in *HS* markets

The evolution of median transaction prices in the *HS* markets of each treatment is shown in (a), while the percentage deviation of median price from fundamental value is shown in (b). For each treatment, the median price shown in each period is the median of the median transaction prices of all *HS* markets in that treatment for that period. Only markets that had an initial cash-to-asset ratio of 1 and did not contain participants with prior exposure to the experimental design are included in the analysis.



(a)



(b)

transaction price, and the percentage deviation of the median price from the risk-neutral fundamental value in the *HS* markets of each treatment; the median prices shown are the medians of the median transaction prices of all *HS* markets in each treatment in each period. Median prices in both treatments appear to 'track' fundamental value (FV) remarkably well for Smith et al.-type markets containing inexperienced participants, with no obvious bubble-and-crash phenomenon present. As is typical with inexperienced participants, prices in both treatments begin below fundamental value and remain there for roughly the first third of the market before going above fundamental value (see Porter and Smith, 1995; Palan, 2013). The degree of underpricing in this initial stage of the market appears greater in the *Free* treatment, especially in the first trading period, which potentially indicates better price discovery in the *Earned* treatment. Alternatively, greater underpricing could be the result of more risk-averse trading behaviour in the initial stages in the *Free* treatment, which is inconsistent with the assertion that house-money necessarily encourages greater risk-seeking behaviour. Both points are moot however, as the differences in median prices between treatments in each of the first 3 periods are not statistically significant[24]. For the remainder of the market, prices in both treatments appear to move in tandem, although the *Free* treatment exhibits persistently higher median prices in the final periods, averaging 19% above fundamental value in periods 7-10, compared to 7% undervaluation in the *Earned* treatment over the same period. While this is consistent with the house-money effect story, these differences are only significant in period 10[25].

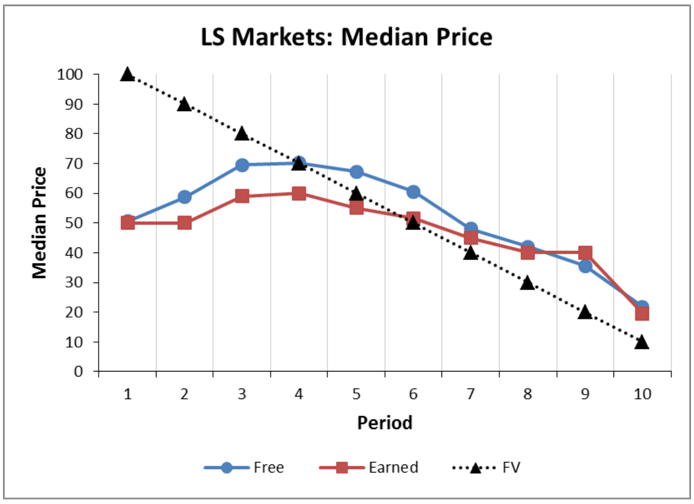
Higher median prices in the *Free* treatment are also evident in the *LS* markets shown in Figures 2.2(a) and 2.2(b). Here, median prices in the *Free* treatment exceed those of the *Earned* treatment in 8 out of 10 periods but, once again, these differences are not statistically significant. The absence of a clear bubble-and-crash pattern to median prices in both treatments is also evident in the *LS* markets. Compared to the *HS* markets, median prices in the *LS* markets of both treatments appear to exhibit more

[24] Testing the significance of the difference in median prices in the *HS* markets of *Free* vs. *Earned* treatments using a Wilcoxon-Mann-Whitney (WMW) U test ($n_E = 15$, $n_F = 11$) returns two-sided p-values of 0.16 (U = 55) , 0.31 (U = 62.5), and 0.38 (U = 65) for periods 1, 2, and 3 respectively.

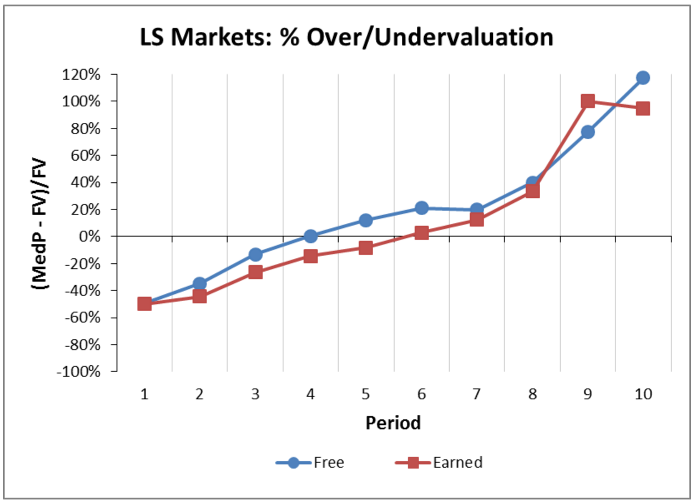
[25] WMW test p-values (two-sided) for periods 7, 8, 9, and 10 are 0.25 (U = 105), 0.62 (U = 92.5), 0.16 (U = 110), and 0.02 (U = 129) respectively.

Figure 2.2: Median prices in *LS* markets

The evolution of median transaction prices in the Low-stakes (*LS*) markets of each treatment is shown in (a), while the percentage deviation of median price from fundamental value is shown in (b). For each treatment, the median price shown in each period is the median of the median transaction prices of all *LS* markets in that treatment for that period. Only markets that had an initial cash-to-asset ratio of 1 and did not contain participants with prior exposure to the experimental design are included in the analysis.



(a)



(b)

severe underpricing in the early stages of the market. This is especially the case for the *Earned* treatment, where the median price stays below fundamental value in the first half of the market, compared to only the first 3 periods in the *HS* markets. Moreover, although prices in both treatments track FV in the sense that they generally decline over the course of the market, they do so more poorly than in the *HS* markets. Specifically, median prices in both treatments fail to fully converge to FV. This is most clearly seen in Figure 2.2(b), where the degree of overvaluation climbs to the region of 100% in both treatments by the end of the market. This lack of convergence is also borne out in the data by the number of overpriced trades, which is defined as a transaction that occurs at a price exceeding the maximum possible dividend earnings from the asset. These trades occurred more commonly, and their effects felt more strongly, in the latter periods of a market, when trading volumes were relatively low. While also present in *HS* markets, they were puzzlingly far more prevalent in the *LS* markets of both treatments. Hence these trades, possibly driven by irrationality (Lei et al. 2001) or speculative interest (or both), appear to be a significant factor behind the lack of convergence to FV late in the *LS* markets.

While the figures described above do not provide any striking evidence that endowed money significantly exacerbates speculative activity in experimental asset markets, they do however mask considerable heterogeneity in price behaviour at the individual market level. Figures A1 and A2 (A3 and A4) in Appendix A1 show the time series of median transaction prices in each *HS* (*LS*) market of the *Earned* and *Free* treatment respectively. While it is difficult to detect any strong difference between treatments by examining these figures, we note that many individual *HS* and *LS* markets in both treatments are in fact characterised by sustained periods of underpricing ("negative bubbles") rather than overpricing. [26] Earning your initial allocation also clearly does not prevent the bubble-and-crash phenomenon – of all the *HS* markets in both treatments, it is the *Earned* treatment that contains the most obvious example of

[26] Markets E_HS6, E_HS10, E_HS12, E_LS3, E_LS6, and E_LS12 in *Earned*, and markets F_HS6, F_HS15, F_LS1, F_LS2, F_LS12, and F_LS14 in *Free*.

that phenomenon (E_HS8). Bubble-and-crash patterns also do not appear to occur significantly more often in *Free* treatment *HS* markets, with arguably only F_HS2 and F_HS3 being candidates. The case for a house-money effect finds a little more encouragement in the *LS* markets, where it is the *Free* treatment with market F_LS3 that has the clearest bubble-and-crash pattern. Markets F_LS8 and F_LS15 also exhibit severe overpricing but without the associated crash. For *Earned*, only market E_LS4 has an obvious bubble-and-crash pattern.

2.4.2 Statistical analysis

2.4.2.1 Bubble measures

To conduct a more rigorous analysis of the individual markets and formally test our hypotheses, we calculate a number of variables commonly used in the literature to measure mispricing/bubbles. These measures can broadly be categorised into two groups. “Bubble magnitude” measures, such as *Price Amplitude, Total Dispersion, Turnover, Normalised Deviation, Haessel-R^2,* and *Average Bias* assess the degree and/or direction of mispricing in a market. “Bubble-length” measures on the other hand, such *Duration, Boom Duration,* or *Bust Duration,* examine how long mispricing lasts in a market.

Price Amplitude measures the extent to which average price in a market changes relative to FV. Haruvy and Noussair (2006) measure it as $\max_t\{(\bar{P}_t - F_t)/F_t\} - min_t\{(\bar{P}_t - F_t)/F_t\}$, where the largest and smallest deviations of average price $\bar{P}_t$ from the fundamental value F_t are normalised by the fundamental value in the respective period. Large values of this measure indicate big swings in price relative to fundamental value and hence the possible existence of a bubble.

Total Dispersion (Haruvy and Noussair 2006) measures the aggregate absolute deviation of median price from fundamental value across all trading periods, and is defined as $\sum_t |MedianP_t - F_t|$. Since it treats both positive and negative deviations from FV identically, it is a measure of mispricing rather than of over or undervaluation, with smaller values indicating a closer correspondence between price and fundamental value. *Turnover*, a normalised measure of trading activity, is used as a measure of magnitude since bubble periods are typically associated with large trading volumes. Due to the difference in trading period lengths between the August and October sessions of our experiment (3 min. vs. 2 min. respectively), we construct a modified version of the variable defined by King et al. (1993), namely $\sum_t V_t/(p \times TSU)$, where V_t, the volume of trade in period *t* is normalised by *TSU*, the total asset supply in the market, and an additional variable *p*, which is the length of a trading period in minutes. *Normalised Deviation*, as measured by Haruvy et al. (2007), combines the preceding two measures to account for both the size of the price deviation and the level of trading activity in a market. We calculate it as $\sum_t V_t |MedianP_t - F_t|/(p \times TSU)$. Large values of this measure may be caused by large absolute deviations of price from fundamental value and/or a high volume of trade, which may suggest heightened speculative or irrational trading behaviour. We can also assess how closely prices tend to track changes in fundamental value in a market by determining its *Haessel-R^2* (Dufwenberg et al. 2005), which is simply the R-squared calculated by regressing average prices on fundamental values. A goodness-of-fit measure, it tells you how much of the variation in average price across periods is explained by changes in fundamental value; values closer to 0 (1) suggest the potential existence (absence) of price bubbles[27].

[27] Stöckl et al. (2010) however, point out that using Haessel-R^2 as a measure of price bubbles is problematic since markets in which prices increase (monotonically) over the life of the asset may also have high R-squared values even though there’s no real fit between FV and price. Indeed, this is evident in two of our markets, F_LS8 and F_LS12. Excluding these markets from our comparison of Haessel R^2 between treatments does not change our qualitative result.

Note that none of the above measures determine whether the asset is generally overvalued or undervalued. To gauge the degree of overpricing/underpricing, we calculate *Average Bias* (Haruvy and Noussair 2006), which measures how far median prices deviate from fundamental value on average over the course of the market, and is calculated as $\frac{1}{N}\sum_{t=1}^{N}(MedianP_t - F_t)$. Large positive (negative) values suggest that prices tend to stay above (below) FV. Values close to zero may suggest that prices stay close to fundamental value or that the asset experiences equal degrees of over and underpricing in the market; assessing the *Average Bias* in conjunction with *Total Dispersion* helps to shed light in this regard, since observing a small (large) *Total Dispersion* at the same time as a near-zero *Average Bias* would imply the former (latter) (Haruvy and Noussair 2006).

The first of the bubble-length measures, *Duration* (Porter and Smith 1995), calculates the maximum number of consecutive periods where average price increases relative to fundamental value, or $max\{m: \bar{P}_t - F_t < \bar{P}_{t+1} - F_{t+1} < \cdots < \bar{P}_{t+m} - F_{t+m}\}$. Larger values of *Duration* point to sustained periods where changes in (average) transaction price across trading periods do not 'adequately' track changes in the FV, potentially indicating the presence of a bubble. *Boom (Bust) Duration* (Haruvy and Noussair 2006) is defined as the maximum number of consecutive periods where median prices stay above (stay below) FV; large values indicate long periods of overvaluation (undervaluation), potentially signalling the presence of a bubble ("negative bubble").

We now re-state our null hypotheses in relation to the above measures. Hypothesis 1, which contends that bubbles/mispricing is equal in magnitude when participants have to trade over earned wealth compared to unearned wealth becomes:

H1: The *Price Amplitude, Total Dispersion, Normalised Deviation, Average Bias* and *Haessel-R^2* measures do not differ between the *Free* and *Earned* treatments.

We test this against a two-sided alternative hypothesis that contends that these measures are significantly different when participants trade with unearned wealth. Since a house-money effect predicts greater risk-taking, speculation, and hence mispricing, we also examine the one-sided alternative hypothesis that the above measures are significantly larger in the *Free* treatment than in the *Earned*. The sole exception to this is *Haessel-R^2*, where the direction of the one-sided alternative is reversed.

Hypothesis 2, which states that trading over earned wealth makes no difference to the length of bubbles/mispricing, becomes:

H2: *Duration, Boom Duration,* and *Bust Duration* do not differ between the *Free* treatment and the *Earned* treatment.

Here, the two-sided alternative once again is that these measures differ significantly between the two treatments. The one-sided alternative for *Duration* and *Boom Duration* is that they are significantly larger in the *Free* treatment than in *Earned*. If house-money effects lead to higher prices and potentially more prolonged overpricing, then periods of sustained underpricing should be shorter in the *Free* markets. Hence, the one-sided alternative for *Bust Duration* is that it is significantly larger in the *Earned* treatment.

2.4.2.2 High-stakes markets

Panels A and B of Table 2.2 contain the values of the bubble measures from each of the *HS* markets in the *Earned* and *Free* treatments respectively. The median value of each measure across all *HS* markets of the respective treatments is shown at the bottom of each panel, along with the associated median absolute deviations[28]. Due to the limited number of observations (each market is a single observation), we examine the statistical significance of the difference between treatments using a Wilcoxon-Mann-Whitney U (WMW) test, which is the non-parametric equivalent of the independent samples t-test.[29] The exact p-values associated with the one-sided and two-sided tests are shown in Panel C.

A cursory glance at the relative median values of the bubble-magnitude variables in the *HS* markets would appear to suggest the presence of house-money effects. The average values of *Amplitude, Total Dispersion, Turnover,* and *Normalised Deviation* are all higher in the *Free* treatment, while *Haessel-*R^2 has a higher median value in the *Earned* treatment. The only exception is *Average Bias*, for which the *Earned* treatment actually has a higher average value. However, the results of the two-sided WMW tests reveal that these differences are in fact not statistically significant. Nonetheless, we find some support for a house-money effect in the one-sided WMW tests, where *Amplitude* and *Normalised deviation* are significant, albeit only at the 10% level (*Amplitude* p-value = 0.09; *Normalised deviation* p-value = 0.099). Even though *Normalised deviation* is marginally significant, its components, *Total Dispersion* and *Turnover* return insignificant results on the one-tailed test (p-values of 0.222 and 0.175).

That *Average Bias* is not significantly larger in the *Free* treatment than in the *Earned* treatment (one-sided p-value = 0.596) is especially problematic for the case of a house-money effect, since it predicts heightened risk-taking/speculative behaviour when

[28] Like the standard deviation, median absolute deviation (MAD) is a measure of the spread of a distribution. MAD is calculated as the median of the absolute deviations of all values from the median value. We report the median and MAD of each measure in preference to the mean and standard deviation due to the small number of observations involved, and their lower sensitivity to outliers.
[29] The WMW test compares the rank-sums of the observations from two independent samples under the null hypothesis that both samples come from the same underlying distribution. The null is rejected if the observed rank-sum for one of the samples is unusually large/small relative to that expected under the null.

Table 2.2: Bubble measures in High-Stakes markets

Panels A and B of the table below show the values of the bubble measures in each High-Stakes (*HS*) market of the *Earned* and *Free* treatments respectively. Only markets that were not 'contaminated' by participants with prior experience of the experimental design, and had an initial cash-to-asset-value ratio of 1 are shown and included in the analysis. The definitions of the bubble measures can be found in section 2.4.2.1. The figures shown in parentheses are the median absolute deviations (MAD) of each measure in each treatment. We test if the bubble measures differ significantly between the *Earned* and *Free* treatments using the non-parametric Wilcoxon-Mann-Whitney (WMW) U test under the null hypothesis that values in both groups come from the same distribution (i.e. no significant difference). The two-sided alternative hypothesis is that values in both groups come from different distributions. The one-sided alternative hypothesis that we test is that bubble measure values are significantly larger in the *Free* treatment than in the *Earned* treatment, except in the case of *Haessel* R^2 and *Bust Duration*, where it is the opposite. Exact p-values associated with these tests are shown in Panel C.

Panel A: Earned treatment

Market	*Amp*	*Tot Disp*	*Avg Bias*	*Haessel* R^2	*Turn*	*Norm Dev*	*Dur*	*Boom Dur*	*Bust Dur*
E_HS1	0.96	127.50	0.65	0.69	1.95	25.24	6	5	3
E_HS2	0.17	8.50	-0.75	1.00	1.10	0.85	1	1	6
E_HS3	0.76	91.50	3.95	0.85	1.02	10.30	6	7	2
E_HS4	0.74	64.50	-0.25	0.89	0.87	5.54	4	5	3
E_HS5	2.03	122.00	6.00	0.53	0.58	6.86	7	6	3
E_HS6	0.82	203.00	-19.70	0.68	0.58	15.27	6	1	9
E_HS7	1.17	154.00	-2.40	0.55	2.26	35.60	7	4	4
E_HS8	3.59	393.50	32.95	0.09	1.51	44.58	6	7	2
E_HS9	0.31	64.00	-6.40	0.95	0.67	4.59	3	0	5
E_HS10	0.65	232.50	-22.65	0.77	0.58	15.06	7	2	7
E_HS11	1.10	181.00	10.60	0.94	0.90	15.00	2	6	4
E_HS12	0.89	270.00	-25.80	0.20	0.99	28.72	8	2	7
E_HS13	0.47	67.00	2.70	0.98	0.71	4.76	1	5	2
E_HS14	1.01	94.50	-7.05	0.87	1.28	11.18	4	2	5
E_HS16	0.98	96.50	5.65	0.94	0.91	6.34	1	5	2
Median:	**0.89**	**122.00**	**-0.25**	**0.85**	**0.91**	**11.18**	**6.00**	**5.00**	**4.00**
	(0.21)	(57.50)	(6.15)	(0.13)	(0.24)	(5.64)	(2.00)	(2.00)	(2.00)

Panel B: Free treatment

Market	*Amp*	*Tot Disp*	*Avg Bias*	*Haessel* R^2	*Turn*	*Norm Dev*	*Dur*	*Boom Dur*	*Bust Dur*
F_HS1	0.42	48.00	4.00	0.95	0.34	1.67	1	5	2
F_HS2	2.39	194.00	13.60	0.55	1.98	32.77	6	6	3
F_HS3	1.65	165.00	8.30	0.51	1.43	20.72	7	6	3
F_HS5	1.34	117.00	11.50	0.90	1.01	9.72	3	8	1
F_HS6	0.68	191.00	-18.50	0.72	1.05	19.13	4	1	6
F_HS7	1.36	192.50	-9.95	0.50	0.76	14.99	8	4	5
F_HS10	1.01	227.00	-4.70	0.60	1.00	28.23	6	4	3
F_HS12	0.44	34.00	-2.70	0.92	0.81	5.68	1	1	1
F_HS13	1.08	187.00	-16.70	0.49	2.35	57.18	6	2	5
F_HS14	1.55	136.50	-0.75	0.64	2.18	30.54	7	6	3
F_HS15	1.55	362.50	-34.75	0.41	0.94	43.25	9	1	9
Median:	**1.34**	**187.00**	**-2.70**	**0.60**	**1.01**	**20.72**	**6.00**	**4.00**	**3.00**
	(0.31)	(40.00)	(11.00)	(0.11)	(0.24)	(10.99)	(2.00)	(2.00)	(2.00)

Panel C: Wilcoxon Mann-Whitney U Test

	Amp	*Tot Disp*	*Avg Bias*	*Haessel* R^2	*Turn*	*Norm Dev*	*Dur*	*Boom Dur*	*Bust Dur*
p-value (1-sided)	0.090	0.222	0.596	0.109	0.175	0.099	0.263	0.480	0.262
p-value (2-sided)	0.180	0.443	0.809	0.217	0.350	0.198	0.524	0.961	0.524

trading with unearned wealth and consequently higher prices. Note also that the median *Average Bias* in both treatments is close to zero while the median values of *Total Dispersion* are considerably greater than zero. This observation suggests that the 'average' *HS* market in both treatments is characterised by periods of overvaluation *and* undervaluation that tend to (mostly) cancel each other out. This tallies with our earlier observations on the evolution of median prices in both treatments, as well as the fact that many individual markets are characterised by prolonged under-pricing rather than overpricing.

Turning to the bubble-length measures, the differences between the *Free* and *Earned* treatments on all three measures are not significant. Hence, the *HS* markets provide no support for the claim that trading with unearned money prolongs the duration of bubbles/mispricing by fuelling the urge to speculate.

2.4.2.3 Low-stakes markets

The values of the bubble measures from each of the *LS* markets in the *Earned* and *Free* treatments are shown in Panels A and B of Table 2.3 respectively. The results of the corresponding WMW tests are shown in Panel C.

Here, only the relative median values of *Total Dispersion* and *Normalised Deviation* in the bubble-magnitude measures correspond to what would be expected if unearned money had an effect on prices. However once again, the differences between the treatments are not statistically significant (two-sided test) on all magnitude measures. Even with the one-sided test, only *Total Dispersion* returns a significant result, but that too only marginally (p-value = 0.082). The story is the same for the bubble-length measures in the *LS* markets, where no significant difference is detected between treatments for *Duration*, *Boom Duration*, or *Bust Duration*. As in the *HS*

Table 2.3: Bubble measures in Low-Stakes markets

Panels A and B of the table below show the values of the bubble measures in each Low-Stakes (*LS*) market of the *Earned* and *Free* treatments respectively. Only markets that were not 'contaminated' by participants with prior experience of the experimental design, and had an initial cash-to-asset-value ratio of 1 are shown and included in the analysis. The definitions of the bubble measures can be found in section 2.4.2.1. The figures shown in parentheses are the median absolute deviations (MAD) of each measure in each treatment. We test if the bubble measures differ significantly between the *Earned* and *Free* treatments using the non-parametric Wilcoxon-Mann-Whitney (WMW) U test under the null hypothesis that values in both groups come from the same distribution (i.e. no significant difference). The two-sided alternative hypothesis is that values in both groups come from different distributions. The one-sided alternative hypothesis that we test is that bubble measure values are significantly larger in the *Free* treatment than in the *Earned* treatment, except in the case of *Haessel* R^2 and *Bust Duration*, where it is the opposite. Exact p-values associated with these tests are shown in Panel C.

Panel A: Earned treatment

Market	*Amp*	*Tot Disp*	*Avg Bias*	*Haessel* R^2	*Turn*	*Norm Dev*	*Dur*	*Boom Dur*	*Bust Dur*
E_LS1	0.39	67.50	0.45	0.91	1.07	12.23	2	5	2
E_LS2	2.79	165.50	12.55	0.91	2.89	48.75	5	7	3
E_LS3	1.65	282.00	-25.10	0.63	2.23	71.59	9	2	8
E_LS4	2.61	232.50	2.65	0.15	2.18	52.68	8	6	4
E_LS6	1.86	328.00	-28.20	0.03	2.76	107.52	9	2	8
E_LS10	1.90	213.50	-5.65	0.28	3.97	90.07	8	5	5
E_LS12	2.42	193.00	-8.30	0.88	2.21	46.35	8	3	6
E_LS13	3.89	257.50	-8.85	0.25	5.31	136.46	9	4	6
E_LS14	0.64	97.50	-1.25	0.83	2.19	28.40	2	7	2
Median:	**1.90**	**213.50**	**-5.65**	**0.63**	**2.23**	**52.68**	**8.00**	**5.00**	**5.00**
	(0.71)	(48.00)	(6.10)	(0.28)	(0.53)	(24.28)	(1.00)	(2.00)	(2.00)

Panel B: Free treatment

Market	*Amp*	*Tot Disp*	*Avg Bias*	*Haessel* R^2	*Turn*	*Norm Dev*	*Dur*	*Boom Dur*	*Bust Dur*
F_LS1	0.56	252.50	-25.25	0.89	1.36	53.60	6	0	10
F_LS2	2.17	299.50	-22.25	0.16	2.89	96.44	8	3	6
F_LS3	2.49	361.00	4.80	0.01	1.50	63.26	5	4	3
F_LS5	0.74	109.50	-0.25	0.76	0.91	13.96	5	6	2
F_LS6	1.34	207.00	19.70	0.62	0.70	12.99	2	8	1
F_LS8	8.65	372.50	25.25	0.90	2.86	101.51	9	7	3
F_LS9	2.84	214.00	-10.40	0.73	2.40	63.81	6	4	6
F_LS10	0.83	120.00	-6.50	0.64	1.54	28.74	5	6	3
F_LS11	1.54	134.00	11.40	0.89	1.55	17.42	2	8	1
F_LS12	4.11	399.50	-24.65	0.88	2.70	134.14	9	4	6
F_LS14	1.45	361.00	-34.30	0.20	3.54	121.06	9	1	8
F_LS15	3.00	585.00	58.50	0.93	3.17	196.25	3	10	0
Median:	**1.86**	**276.00**	**-3.38**	**0.74**	**1.97**	**63.54**	**5.50**	**5.00**	**3.00**
	(1.01)	(90.75)	(20.08)	(0.15)	(0.81)	(42.05)	(2.50)	(2.00)	(2.50)

Panel C: Wilcoxon Mann-Whitney U Test

	Amp	*Tot Disp*	*Avg Bias*	*Haessel* R^2	*Turn*	*Norm Dev*	*Dur*	*Boom Dur*	*Bust Dur*
p-value (1-sided)	0.459	0.082	0.351	0.649	0.827	0.377	0.751	0.319	0.237
p-value (2-sided)	0.917	0.164	0.702	0.702	0.345	0.754	0.515	0.637	0.467

markets, the ‘average’ *LS* market in both treatments exhibits periods of overvaluation nullified by periods of undervaluation of similar magnitude (or vice versa), as suggested by the small median *Average Bias* values and the relatively large *Total Dispersion* medians.

The weight of the statistical evidence from both *HS* and *LS* markets points to a failure to reject hypothesis H1. Bubble/Mispricing magnitude measures are not significantly different between the two treatments when examined with a two-sided WMW test. Although some measures are deemed significantly larger in the *Free* treatment when applying a one-sided test, they are only marginally so. In addition, these variables are not significant in both *HS* and *LS* markets. In regards to hypothesis H2, the lack of significance associated with any of the measures of bubble-length makes the failure to reject it a considerably more straightforward issue.

While this may indeed suggest that trading with house-money has no effect on prices, we note that the failure to reject the null hypotheses may be influenced by other factors, including elements of the experimental design. For example, rewarding the top 50% of performers in the earnings task with a place in a *HS* market may have induced a positive emotional state akin to excitement in these subjects. Given the reported link between induced excitement and larger asset price bubbles in experimental markets (Andrade, Odean, and Lin 2013), this may produce an upward bias to prices in the *HS* markets of the *Earned* treatment, thus acting in the opposite direction to earned money, and confounding its impact[30]. On the other hand, the expected impact, if any, of negative emotions on the *Earned LS* markets is considerably less clear; Andrade et al. do not find a significant link between negative emotions (specifically, fear and sadness)

[30] This ‘hangover’ from the earnings task should be especially strong in the early stages of the market, which may explain the relatively high median prices observed in the initial periods of *Earned HS* markets (albeit not significantly) compared to the *Free HS* markets.

and price bubbles. However, if negative emotions were to produce the opposite effect to positive emotions, then it should in fact *increase* the likelihood of finding a significant difference between *Free* and *Earned* treatments in the *LS* markets.

Another possible source of confound lies in the aforementioned potential positive correlation between performance in the earnings task (GMAT questions), and intelligence[31]. Since subjects in the *Earned* treatment were allocated to *HS* or *LS* markets on the basis of their relative performance in the task, the *HS* (*LS)* markets of the *Earned* treatment could be characterised as being populated with above-average' ('below-average') intelligence traders. Due to random assignment, the *HS* and *LS* markets of the *Free* treatment on the other hand are of 'average' intelligence. If intelligence, of the sort measured by the GMAT, is associated with smaller bubbles[32], then all else being equal, mispricing should be greater (lower) in the *Earned LS* (*HS)* markets than in the *Free LS* (*HS)* markets. Hence, the effect of intelligence may be to increase the likelihood of a non-significant result in *LS* markets, but would actually increase the likelihood of finding a significant difference between treatments in the *HS* markets.

While both of these issues – emotions and intelligence – arise from the nature of the earnings task and represent limitations of our design, it is important to remember that our earnings task is primarily designed to create a credible or 'real' incentive to expend effort. Moreover, these concerns are somewhat mitigated by the fact that our results qualitatively mirror those of Corgnet et al. (2014), who use a different earnings

[31] To mitigate this as much as possible, we selected 10 questions that we deemed low to moderate difficulty, and gave participants 20 minutes to complete the task, which was more than enough time to attempt all questions. While 90% of students completed the task before the allotted time expired, there was still a significant difference in performance (median score *Earned HS* = 8, median score *Earned LS* = 5, WMW p-value (two-sided) < 0.001).

[32] To our knowledge, there is no extant research on the relationship between generalised or specific measures of intelligence and price behaviour in experimental asset markets.

task that crucially does not rank or allocate subjects on the basis of performance, and hence should not be affected by either issue. Another implication of this is that our main results are robust to alternative earnings-task types.

2.4.2.4 Earnings dispersion

One area where our results do not fully coincide with Corgnet et al. (2014) is in their finding that the dispersion of earnings is significantly lower in the earned money treatment. Earnings dispersion – measured by the standard deviation of final earnings (in francs) – in our experiments is summarised in Panels A and B of Table 2.4 for the *HS* and *LS* markets respectively. Looking at the *HS* markets, we observe that earnings dispersion is only significantly different (and larger) in the *Free* treatment at the 10% level (two-sided p-value = 0.097). In the *LS* markets, the median earnings dispersion in *Earned* markets is again lower than the *Free* treatment, but the difference between treatments is not significant.

Corgnet et al. (2014) point to significantly larger trading volumes in their unearned treatment markets as an important factor behind their finding of greater earnings dispersion. We believe the relationship between earnings dispersion and trading volume is indeed a likely reason for the weaker results on earnings dispersion in our study. Whereas the Corgnet et al. experiment is characterised by a significant difference in turnover between treatments, turnover does not differ significantly in ours (*HS* markets p-value (two-sided) = 0.350, *LS* markets p-value (two-sided) = 0.345).

Table 2.4: Earnings dispersion

The standard deviation of participants' earnings in each High-Stakes (Low-Stakes) market is shown in Panel A (B). Markets are categorised according to the treatment in effect – *Earned* or *Free*. Only markets that were not 'contaminated' by participants with prior experience of the experimental design, and had an initial cash-to-asset-value ratio of 1 are shown and included in the analysis. The statistical significance of the difference between treatments is determined using a two-sided Wilcoxon-Mann-Whitney (WMW) U test under the null hypothesis that values in both treatments come from the same distribution (i.e. no significant difference). Exact p-values associated with this test are reported.

Panel A: High-stakes markets

Treatment: Earned		**Treatment: Free**	
Market	*SD Earnings*	*Market*	*SD Earnings*
E_HS1	237.75	F_HS1	617.80
E_HS2	178.70	F_HS2	433.59
E_HS3	295.30	F_HS3	587.62
E_HS4	412.53	F_HS5	228.97
E_HS5	132.62	F_HS6	440.29
E_HS6	471.87	F_HS7	220.11
E_HS7	923.67	F_HS10	1282.52
E_HS8	992.44	F_HS12	440.86
E_HS9	136.72	F_HS13	661.25
E_HS10	193.02	F_HS14	748.71
E_HS11	466.71	F_HS15	329.28
E_HS12	604.25		
E_HS13	76.06		
E_HS14	268.62		
E_HS16	259.17		
Median Earned:	**268.62**	**Median Free:**	**440.86**

WMW p-value (two-sided) = 0.097

Panel B: Low-Stakes markets

Treatment: Earned		**Treatment: Free**	
Market	*SD Earnings*	*Market*	*SD Earnings*
E_LS1	450.16	F_LS1	478.24
E_LS2	226.01	F_LS2	169.76
E_LS3	309.93	F_LS3	537.55
E_LS4	378.70	F_LS5	181.64
E_LS6	250.51	F_LS6	187.10
E_LS10	669.88	F_LS8	280.43
E_LS12	377.46	F_LS9	533.06
E_LS13	686.34	F_LS10	487.00
E_LS14	455.81	F_LS11	321.24
		F_LS12	233.30
		F_LS14	464.06
		F_LS15	834.24
Median Earned:	**378.70**	**Median Free:**	**392.65**

WMW p-value (two-sided) = 0.808

2.4.2.5 The impact of endowment size

The *Free* treatment markets, which are free of the selection issues[33] associated with the *Earned* markets, provide an opportunity to test whether the finding of Ackert et al. (2006a) – that larger (cash) endowments result in higher prices – also applies to Smith et al. (1988) type double auction markets. Recall that participants in *HS* markets were provided with twice the level of cash and assets as those in *LS* markets. Table 2.5 compares the bubble measures in the *HS* (Panel A) and *LS* (Panel B) markets of the *Free* treatment. We compare the null hypothesis of no difference between these two market types against a two-sided alternative that contends that there is a difference, and the one-sided alternative that mispricing/bubbles is more severe in the *HS* markets. The p-values of the respective WMW tests are shown in Panel C.

The one-tailed tests fail to reject the null that mispricing is not greater in the *HS* markets. If anything, most bubble-magnitude measures in the *LS* markets are larger, significantly so in the case of *Normalised Deviation* (two-sided p-value = 0.019). This appears to be driven by both greater mispricing (*Total Dispersion* two-sided p-value = 0.042) and larger trading volumes in the *LS* markets (two-sided p-value of *Turnover* = 0.032). Furthermore, the null hypothesis that *HS* markets do not experience a greater degree of overvaluation cannot be rejected, since *Average Bias* in both *HS* and *LS* markets is close to zero and not significantly different from each other. Moreover, we do not detect a significant different between the two types of markets in any of the bubble-length measures.

The probable root of this apparent contradiction of Ackert et al. (2006a) is in the

[33] That is, participants in the *Free* treatment were randomly assigned to *HS* and *LS* markets (and were not informed on the type of market in which they were trading), whereas those in the *Earned* treatment were selected into one or other based on their task performance. If task performance is correlated with intelligence (or some other factor), then *HS* and *LS* markets in the *Earned* treatment will systematically differ in a factor other than just the endowment level.

Table 2.5: Bubble measures in Free treatment markets

Panels A and B of the table below show, respectively, the values of the bubble measures in each High-stakes (*HS*) and Low-stakes (LS) market of the *Free* treatment. Only markets that were not 'contaminated' by participants with prior experience of the experimental design, and had an initial cash-to-asset-value ratio of 1 are shown and included in the analysis. The definitions of the bubble measures can be found in section 2.4.2.1. The figures shown in parentheses are the median absolute deviations (MAD) of each measure in each treatment. We test if the bubble measures differ significantly between *HS* and *LS* markets of the *Free* treatment using the non-parametric Wilcoxon-Mann-Whitney (WMW) U test under the null hypothesis that values in both groups come from the same distribution (i.e. no significant difference). The two-sided alternative hypothesis is that values in both groups come from different distributions. The one-sided alternative hypothesis that we test is that bubble measure values are significantly larger in *HS* markets than in the *LS* markets, except in the case of *Haessel* R^2 and *Bust Duration*, where it is the opposite. Exact p-values associated with these tests are shown in Panel C.

Panel A: High-Stakes markets of Free treatment

Market	*Amp*	*Tot Disp*	*Avg Bias*	*Haessel* R^2	*Turn*	*Norm Dev*	*Dur*	*Boom Dur*	*Bust Dur*
F_HS1	0.42	48.00	4.00	0.95	0.34	1.67	1	5	2
F_HS2	2.39	194.00	13.60	0.55	1.98	32.77	6	6	3
F_HS3	1.65	165.00	8.30	0.51	1.43	20.72	7	6	3
F_HS5	1.34	117.00	11.50	0.90	1.01	9.72	3	8	1
F_HS6	0.68	191.00	-18.50	0.72	1.05	19.13	4	1	6
F_HS7	1.36	192.50	-9.95	0.50	0.76	14.99	8	4	5
F_HS10	1.01	227.00	-4.70	0.60	1.00	28.23	6	4	3
F_HS12	0.44	34.00	-2.70	0.92	0.81	5.68	1	1	1
F_HS13	1.08	187.00	-16.70	0.49	2.35	57.18	6	2	5
F_HS14	1.55	136.50	-0.75	0.64	2.18	30.54	7	6	3
F_HS15	1.55	362.50	-34.75	0.41	0.94	43.25	9	1	9
Median:	**1.34**	**187.00**	**-2.70**	**0.60**	**1.01**	**20.72**	**6.00**	**4.00**	**3.00**
	(0.31)	(40.00)	(11.00)	(0.11)	(0.24)	(10.99)	(2.00)	(2.00)	(2.00)

Panel B: Low-Stakes markets of Free treatment

Market	*Amp*	*Tot Disp*	*Avg Bias*	*Haessel* R^2	*Turn*	*Norm Dev*	*Dur*	*Boom Dur*	*Bust Dur*
F_LS1	0.56	252.50	-25.25	0.89	1.36	53.60	6	0	10
F_LS2	2.17	299.50	-22.25	0.16	2.89	96.44	8	3	6
F_LS3	2.49	361.00	4.80	0.01	1.50	63.26	5	4	3
F_LS5	0.74	109.50	-0.25	0.76	0.91	13.96	5	6	2
F_LS6	1.34	207.00	19.70	0.62	0.70	12.99	2	8	1
F_LS8	8.65	372.50	25.25	0.90	2.86	101.51	9	7	3
F_LS9	2.84	214.00	-10.40	0.73	2.40	63.81	6	4	6
F_LS10	0.83	120.00	-6.50	0.64	1.54	28.74	5	6	3
F_LS11	1.54	134.00	11.40	0.89	1.55	17.42	2	8	1
F_LS12	4.11	399.50	-24.65	0.88	2.70	134.14	9	4	6
F_LS14	1.45	361.00	-34.30	0.20	3.54	121.06	9	1	8
F_LS15	3.00	585.00	58.50	0.93	3.17	196.25	3	10	0
Median:	**1.86**	**276.00**	**-3.38**	**0.74**	**1.97**	**63.54**	**5.50**	**5.00**	**3.00**
	(1.01)	(90.75)	(20.08)	(0.15)	(0.81)	(42.05)	(2.50)	(2.00)	(2.50)

Panel C: Wilcoxon Mann-Whitney U Test

	Amp	*Tot Disp*	*Avg Bias*	*Haessel* R^2	*Turn*	*Norm Dev*	*Dur*	*Boom Dur*	*Bust Dur*
p-value (1-sided)	0.948	0.979	0.536	0.393	0.984	0.991	0.601	0.793	0.435
p-value (2-sided)	0.104	0.042	0.928	0.786	0.032	0.019	0.797	0.415	0.869

difference between the initial cash-to-assets ratios of their study and ours. The market design of Ackert et al. (2006a) involves participants bidding to buy single units of a fixed supply of new stock in each period using an endowment of only cash. As the supply of stock is the same in both treatments, the initial cash-to-assets ratio in their high-endowment treatment is necessarily larger than in their low-endowment treatment. In contrast, the *HS* and *LS* markets of the current study both have the same initial cash-to-assets ratio of 1, since larger cash endowments are accompanied by an equivalent increase in the asset supply. Hence, the higher prices associated with larger cash endowments observed by Ackert et al. (2006a) appear to be driven by the liquidity of their markets (the cash-to-assets ratio) rather than the size of the of the (cash) endowment per se. This is consistent with the findings of Caginalp et al. (1998; 2000).

2.5. Conclusion

Recent evidence has drawn attention to the potentially distortionary effects that endowing experiment participants with unearned assets has on individual behaviour. Individuals who experience windfall gains are likely to consume more, take more risk, and display greater other-regarding behaviour than they normally would with their own (earned) money. Of particular relevance to experimental asset markets of the type designed by Smith et al. (1988) is the heightened tendency for risk-taking, or 'house-money effect', which could amplify the bubble-and-crash patterns observed in such experiments. We examine whether this is indeed the case using a two-treatment design where participants in one treatment are required to earn their initial wealth via a money-earning task while participants in the other treatment are simply endowed it.

The results suggest that issues regarding asset legitimacy in experimental asset markets may not be especially important and may not confound the results of existing studies. A distinguishable difference in price behaviour between the two treatments is not observed. Hence, legitimising assets with effort does not appear to be necessary for this particular class of experiment.

CHAPTER 3: Tournaments in Two-Asset Markets

3.1 Introduction

Professionals in the intensely competitive world of finance routinely vie for ‘prizes’ such as bonuses, fund flows, and promotions that are tied to their performance relative to others. This gives many of the incentive schemes used in the industry the flavour of a tournament, which is characterised by compensation that depends on an employee’s *relative* rather than absolute performance. The sizeable upside provided by these compensation structures, often not matched by an offsetting downside, creates ‘convex’ incentives. In light of the severe dysfunction that has punctuated financial markets this century – most notably during the collapse of the dot-com bubble and the US sub-prime mortgage crisis – concerns have been raised that such incentives may help precipitate market instability by encouraging excessive risk-taking and short-termism (Rajan 2006; Bebchuk and Spamann 2010; Wagner 2013)[34]. Yet despite the obvious importance of these concerns, the market-level effects of tournament incentives remain relatively unexplored and are not well understood.

On the one hand, that tournaments alter the risk-taking incentives of *individuals* is a well-established point in the literature – since the most lucrative prizes go to

[34] See also Rajan (2008) and Blinder (2009) for treatments of this issue in the news media.

winners, tournament incentives may encourage some contestants, particularly those who trail the leader, to take more risk in an attempt to 'win' (Bronars 1987; Hvide 2002; Cabral 2003; Tsetlin, Gaba and Winkler 2004). However, the implications of such behaviour for market prices have only been examined in a handful of studies, mostly in the confines of the experimental laboratory where traders' incentives and asset fundamentals can be readily manipulated. The experimental studies, all conducted using the continuous double-auction bubble-market design of Smith et al. (1988), support the view of tournaments as a distortionary force in markets. They generally find that tournament incentives exacerbate asset price bubbles – periods of sustained overvaluation – compared to absolute-performance based incentives, an effect that is alarmingly only magnified as participants gain more experience (James and Isaac 2000; Cheung and Coleman 2014).

In this study, we extend the nascent experimental literature on the aggregate impacts of tournament incentives by addressing two issues affecting existing studies that potentially reduce the generalisability and real-world relevance of their results. First, we examine how market prices behave under tournament incentives when subjects can trade more than one type of risky asset. In contrast, existing studies only examine single-asset trading environments, which unduly restrict the risk-taking options available to traders compared to real-world markets. Investors seeking to 'get ahead' in the real world not only have the ability to speculate on a specific security, but also alter the risk profile of their portfolios by shifting into alternative, inherently riskier asset classes or securities. Thus, it stands to reason that price behaviour in single-asset markets may differ from multi-asset markets.

Second, we investigate what impact adding a penalty for underperformance to a tournament contract has on market prices. Despite theoretical and empirical evidence suggesting that penalties, or ‘sticks’ in tournament contracts can curtail risk-taking by contestants (e.g. Gilpatric 2009; Qiu 2003; Kempf, Ruenzi and Thiele 2009; Hu, Kale, Pagani, and Subramanian 2011), the experimental literature pays scant attention to the role played by disincentives, focusing instead on the ‘carrots’, or rewards paid for good relative performance. Thus, rather than strictly being a tournament phenomenon, it may be that the heightened overvaluation seen in existing studies are driven by the absence of consequences attached to poor performance that arises from excessive speculation. However, given that the fear of underperformance potentially encourages traders to herd (Rajan 2006; Dass, Massa and Patgiri 2008), the addition of a penalty may actually result in even higher prices, and hence is an open empirical issue.

To examine these issues, we implemented a between-subjects experimental design featuring three treatments that differ only in the way in which participants were remunerated – a normal incentive *Baseline* treatment and two tournament treatments *Carrot* and *Stick*, where the latter is identical to *Carrot* but includes a penalty for underperformance in the form of a significantly reduced payment (zero). While the compensation contracts in the *Carrot* and *Stick* treatments rewarded/penalised traders on the basis of their performance relative to the ‘average’ trader (as in James and Isaac (2000) and Isaac and James (2003)), we also implemented a set of alternative tournament contracts based on a rank-order tournament (i.e. where rank determines payoff), called *GilCarrot* and *GilStick.* Participants in all treatments traded in a Smith et al. (1988)-type experimental asset market featuring two risky assets – a low-risk asset called X, which paid a modestly sized divided in each period, and a high-risk asset called Y, which paid a lottery-like dividend, thus allowing participants to more naturally vary risk than earlier studies have allowed.

Our first result suggests that the main conclusions in the existing literature are driven by the single-asset nature of their markets. We do not find any compelling evidence that tournament incentives – whether measured using the *Carrot, Stick, GilCarrot,* or *GilStick* treatments – distort prices more than absolute-performance based incentives (*Baseline*), as gauged by the size and duration of mispricing/bubbles in the markets of the respective treatments. Moreover, we find that bubbles under tournament incentives *do* moderate in size and duration as traders gain experience of the experimental design. In fact, evidence of improvement in price behaviour with once-experienced traders is generally weaker in the normal-incentive *Baseline* treatment than it is in the tournament treatments.

On the impact of penalties, this study finds that in markets populated with inexperienced traders, embedding a penalty into a tournament contract that rewards traders for above-average performance (i.e. *Stick)* reduces the amount of trading activity compared to the corresponding reward-only contract (*Carrot*). However, consistent with the herding hypothesis, the trading activity that occurs in *Stick* markets is actually characterised by significantly *longer* booms (periods of overvaluation) in both risky assets, in addition to significantly *higher* prices in the high-risk asset. These differences however disappear with experienced participants. Moreover, we do not detect a significant difference in price behaviour between the rank-order tournament treatments, *GilCarrot* and *GilStick*, with inexperienced or experienced traders.

The remainder of this chapter is structured as follows. In section 3.2, we review the related literature and develop testable hypotheses. Section 3.3 details the experimental design, while section 3.4 describes the results. We present conclusions in section 3.5.

3.2 Literature Review

3.2.1 Tournaments at the individual-level

The effect of tournament incentives on the behaviour of individuals is the subject of an extensive academic literature, the theoretical underpinnings of which have its formal beginnings in optimal labour market contracting under moral hazard. Starting with Lazear and Rosen (1981) and developed further by Green and Stokey (1983), Nalebuff and Stiglitz (1983), O'Keeffe, Viscusi and Zeckhauser (1984), Rosen (1986) and others, much of the early literature, along with the associated empirical work (e.g. Bull, Schotter and Weigelt 1987, Ehrenberg and Bognanno 1990) examined the comparative efficiency and optimality of the incentives provided by tournaments vis-à-vis other incentive structures such as piece rates. Importantly, the variable of interest and the only lever available to agents to affect their chances of winning in these early models is the amount of effort they choose to expend. Of course, players in a tournament can often also vary the amount of risk they take, and thus a sub-strand of the literature has emerged that focuses on risk-taking incentives in tournaments.

Bronars (1987, cited in Hvide 2002, p. 880) was the first to introduce risk-taking as a choice variable into a tournament model, finding that in sequential tournaments, it is optimal for leaders to reduce risk, while followers are inclined to increase risk in order to catch-up. The basic intuition underlying this result arises from the convexity of payoffs produced by tournaments. Faced with a win/lose dichotomy, the consequences of losing by a lot are the same as losing by a small margin[35]. Hence, laggards are better off 'going for broke' in the hope of maximising their chances of securing the larger prize earned by the winner(s), whereas leaders should try and 'lock in' their gains by

[35] This describes most tournament models in the literature, where 2 (or more) players compete over two levels of prizes differentiating winner(s) from loser(s), W_1 and W_2, where $W_1 > W_2$.

playing conservatively. Results in line with this are also reported in a multi-period setting by Tsetlin et al. (2004) and in an infinite-period model by Cabral (2003).

Other models however reveal a more nuanced relationship. Gaba and Kalra (1999) and Gaba, Tsetlin and Winkler (2004) show in a one-period setting that risk-taking incentives are sensitive to the proportion of players deemed ‘winners’/‘losers’. When the proportion of winners is low (specifically, less than 0.5), players have an incentive to ‘break away from the herd’ by increasing risk (as measured by variance). Conversely, when the proportion of winners is high (greater than 0.5), the priority is to avoid an especially poor performance, thus making a low-variance strategy optimal. Nieken and Sliwka (2010) demonstrate using a two-player model that the correlation between the outcomes of contestants’ risky strategies is another important determinant of risk-taking preferences. When risky outcomes are uncorrelated between players – as is typical of most tournament models in the literature – the leader (laggard) prefers to play it safe (take risks), provided the additional expected return from the risky strategy is sufficiently small relative to size of the lead. However, as the correlation increases, it becomes more attractive for the leader to mimic the (anticipated) risky strategy of the trailing agent as a means of maintaining their lead. Of course, the trailing agent is aware of this, hence at high correlations (>0.5), a mixed strategy equilibrium may exist in which the leading player chooses the risky strategy with a higher probability than the trailing player.

While the aforementioned studies of risk-taking in tournaments ignore effort as a choice-variable and consider only the risk-level, Hvide (2002) combines the two by examining a one-period symmetric tournament where players simultaneously choose both the mean (effort) and variance (risk) of their output. In equilibrium, all participants adopt the highest possible level of risk and expend low effort; since expending effort is costly, players have a common incentive to take high risk because it induces noise in the level of output, making differences in effort less important to their chances of winning/losing[36]. However, Kräkel and Sliwka (2004) show that the uniform preference for high risk and low effort does not necessarily hold when contests are asymmetric. They model a two-player tournament where risk-neutral players differ in ability (or equivalently, in their relative starting positions), and choose risk first and effort second. They find that diverse equilibria are possible, with the exact equilibrium depending on the interplay of a number of factors including the magnitude of the difference in abilities, the associated interaction between the effect of risk-taking on effort and on the probability of winning, the shape of the cost-of-effort function, and the prize spread. Although no equilibrium in their model sees the high-ability/leading agent adopt a high-risk strategy whilst the low-ability/trailing agent takes a low-risk strategy, the reverse (i.e. low-risk for high-ability/leaders and high-risk for low-ability/laggards) does not always hold.

Furthermore, in a result that holds particular significance to the current study, Gilpatric (2009) shows that asymmetry in the prize structure of a tournament can also affect the incentive to take risk. Specifically, Gilpatric demonstrates that adding a third payoff level – an explicit penalty for finishing last (a 'stick') – to the customary prizes for the winner (a 'carrot') and the also-rans in a winner-takes-all contest can curb risk-taking by risk-neutral contestants. In the presence of a penalty for severe underperformance, those who trail the leader no longer increase risk (i.e. variance) with impunity, since increasing risk also entails a greater possibility of finishing with an even

[36] In the extreme case where risk is unbounded, players make zero effort and take an infinite amount of risk, causing the tournament compensation scheme to fail. In the case of bounded variance, the prize spread (the difference between the winning and losing prizes) can be adjusted to maintain first-best levels of effort (i.e. an efficient tournament contract) when players are risk-neutral. However, in their model, tournaments will be less efficient than piece rates when agents are risk-averse.

lower payoff. In the model, the precise amount of risk-taking in equilibrium can be controlled by adjusting the relative sizes of the carrot (the additional reward to the winner) and the stick (the penalty for coming last) – the larger the carrot relative to the stick, the greater the incentive to engage in risk-seeking behaviour.

In addition to the theoretical literature, a growing body of research has examined risk-taking in tournaments empirically. In finance, the relevance of tournament theory to the funds management industry has attracted much interest[37]. Brown, Harlow and Starks (1996) argue that mutual fund managers engage in annual contests with each other because their compensation is typically tied to the value of funds under their management, which in turn depends on their recent performance relative to other funds – the best-performing funds receive the largest inflows of new funds, while those performing poorly do not experience similar-scaled outflows (Sirri and Tufano 1998). This convexity in the relationship between relative performance and compensation motivates Brown et al. to hypothesise that fund managers who are 'losing' mid-way through the year will increase the risk of their portfolios more than mid-year 'winners'. Although they find evidence supporting their hypothesis in their sample, subsequent research has provided mixed, often contradictory results[38]. These conflicts can (at least partially) be reconciled by the aforementioned contributions to tournament theory that show that it is not always optimal for laggards (leaders) to be more risk-seeking

[37] The dominance of relative-performance concerns and the presence of highly incentivised contestants mean that the world of professional sport has also attracted significant empirical interest. A number of studies look at risk-taking by contestants in motorsports (Becker and Huselid 1992; Bothner, Kang and Stuart 2007), golf (Brown and Li 2010), basketball (Grund, Höcker, and Zimmerman 2013), weightlifting (Genakos and Pagliero 2012), and high-stakes poker (Lee 2004).

[38] See for example, Chevalier and Ellison (1997), Koski and Pontiff (1999), Busse (2001), Elton, Gruber and Blake (2003), Qiu (2003), Goriaev, Nijman and Werker (2005), Kempf and Ruenzi (2008), Chen and Pennacchi (2009), Kempf, Ruenzi and Thiele (2009), Elton, Gruber, Blake, Krasny and Ozelge (2010), and Hu, Kale, Pagani, and Subramanian (2011).

(conservative)[39]. Moreover, significantly for the current study, a number of empirical studies lend support to the theory posited by Gilpatric (2009) that penalties or disincentives serve to moderate risk-taking – Qiu (2003), Kempf, Ruenzi and Thiele (2009), and Hu, Kale, Pagani, and Subramanian (2011) all observe that greater termination risk (the risk of job-loss) has a negative effect on risk-taking by fund managers.

3.2.2 Tournaments at the market-level

In contrast to the substantive literature on tournament behaviour at the individual-level, the market-level impacts of tournaments have received relatively little attention. To the best of our knowledge, a series of experiments by James and Isaac (2000) and Isaac and James (2003) represent the first and until recently, only attempt by researchers to study the aggregate effects of tournament incentives. Noting that tournaments can alter individuals' risk-taking incentives, the question these studies pose is whether tournament incentives distort market prices by fuelling speculative asset price bubbles. Using a within-subject design and the oft-replicated Smith et al. (1988) double-auction market as a baseline, they examine how prices are affected by the introduction of a tournament condition that rewards traders on the basis of their performance relative to the 'average' trader. Typically, when traders are compensated

[39] Brown et al. (1996) have also inspired a number of theoretical tournament models of the mutual fund industry. Taylor (2003) tackles the issue of risk-taking by leaders/laggards in a two-player mutual fund tournament. In a strategic setting where the risky strategy yields the same return for both players, the (mixed-strategy) equilibrium is characterised by the mid-year winner being more likely adopt a high-risk strategy than the loser. Taylor's model is a special case of the model developed by Nieken and Sliwka (2010), where the correlation between contestants' risky strategies is set at 1. Bagnoli and Watts (2000) consider the risk choices of fund managers in the presence of return-chasing investors (i.e. in a tournament). They show that risk-neutral fund managers will invest in riskier portfolios compared to the case where investors don't chase returns, and this behaviour will be amplified if investors select funds based on rankings rather than performance relative to the average. Acker and Duck (2006) examine the propensity of fund managers to take 'extreme' positions (mostly cash or mostly shares) in a 2-period model where one of the managers is an exogenous passive 'benchmark'. They find that trailing funds are more likely to take extreme positions, especially if they are far behind, or as the end of the tournament approaches.

according to their absolute performance, prices converge quickly to fundamental value in markets consisting of twice-experienced traders. However, this convergence fails to occur once James and Isaac introduce tournament contracts. In fact, repeated exposure to tournament incentives causes prices to deviate further from fundamental value. James and Isaac explain this by showing that it may be mutually advantageous (i.e. rational) for risk-neutral traders to transact at prices above (or below) fundamental value under their tournament contract.

A recent study by Cheung and Coleman (2014) reinforces the main results of James and Isaac (2000). They investigate prices under tournament incentives in both declining (i.e. Smith et al. (1988)) and constant fundamental value markets but use a different tournament compensation contract, based on the mutual fund industry's convex performance-fund flow relationship; they also use a between-subjects design, which is free of the order effects that can afflict within-subject designs[40]. Somewhat contrastingly, Robin, Straznicka, and Villeval (2012) find that long-term and short-term competitive bonuses have differing effects. Their long-term contract pays a bonus at the end of the market based on relative performance over the course of the entire market, whereas the short-term contract awards a bonus at the end of each trading period. They find that their long-term bonus contract produces less price distortion than both short-term bonus and normal incentive contracts, although it is unclear if their results are driven by the incentives, or the changing liquidity of the market that is induced by their payment of bonuses in their short-term contract (Palan 2013).[41]

[40] Specifically, Cheung and Coleman (2014) detect significantly larger bubble Amplitudes and Durations under tournament incentives in inexperienced Smith et al. (1988) markets. Differences become larger in experienced Smith et al. markets, across a wider range of bubble measures. The effect of tournaments in their constant fundamental value markets is milder than in declining fundamental value markets, but nonetheless, it is still sizeable.

[41] Ang, Diavatopoulos and Schwarz (2010) also implement tournament incentives in an experimental market, but only in every alternate trading period as an additional compensation scheme, with the aim of

A common element of the design of these experimental studies is their use of markets featuring only one type of risky asset. This poses a potential problem for the generalisability of their results because unlike real markets, where investors have the opportunity to alter portfolio risk by shifting between a variety of asset classes and securities that are intrinsically more/less risky or speculative in nature, the risk-taking options for traders looking to 'win' in a single-asset environment are extremely limited – they are restricted to simply acquiring more of the same asset by paying higher prices in the hope of selling at a profit or getting lucky with high dividend payments (or both). Therefore, it is uncertain how applicable the behaviour elicited by single-asset markets is to multi-asset environments, or indeed the real world. Thus, by better approximating real-world markets, an experimental market containing more than one type of risky asset should allow for more natural risk-taking behaviour and thus better scope to understand the aggregate impacts of tournament incentives. We fill this gap in the literature by examining if tournament contracts distort prices more/less than absolute-performance-based incentives in experimental asset markets where participants can simultaneously trade *two* differentiated risky assets. Thus the first hypothesis, stated in the null, is:

Hypothesis 1: Price behaviour does not differ between tournament markets and normal-incentive markets.

---

creating a shortened investment horizon for traders. This significant departure in methodology makes it difficult to place their results amongst the other experimental literature. Ang et al. find that the effect of shortened horizons/tournament incentives on bubbles depend on the risk-attitudes of the traders in their markets and whether participants trade with their own money.

We also examine if the reported tendency for bubbles to worsen with trading experience under tournament conditions is sustained in a two-asset environment, leading to the second (null) hypothesis:

Hypothesis 2: Price behaviour under tournament incentives does not differ between markets containing traders who are inexperienced with regards to the experimental design versus once-experienced traders.

Another attribute that the existing experimental studies have in common lies in the design of their tournament compensation contracts – being solely characterised by the payment of additional rewards for good relative performance, their contracts are all 'carrot', no 'stick'. For instance, James and Isaac (2000) pay a flat fee to traders who perform below average, while above-average performers are rewarded with an additional bonus that is proportional to the degree to which they outperform the average. Similarly, Cheung and Coleman (2014) and Robin et al. (2012) periodically award new funds to traders based on their relative performance in the previous sub-period of the market. In doing so, these studies overlook the importance of disincentives or 'sticks' in employment contracts, specifically penalties attached to poor performance, which may serve to moderate risk-taking (Gilpatric 2009). Hence, it may not be tournament incentives driving their results per se, rather the balance, or lack thereof, between 'carrots' and 'sticks' in their contracts.

However, while introducing a penalty into a tournament contract may deter *individuals* from taking risks, the implications for market prices are less clear. On the one hand, penalising underperformance may produce lower prices by discouraging traders from bidding excessively for assets, since doing so makes them more likely to underperform. However, this fear of underperformance may perversely result in *higher* prices. Rajan (2006) argues that relative-performance based compensation encourages investment managers to herd, since herding reduces the chances of underperforming. While this in itself may create a bubble, it also means fund managers may choose to 'ride the bubble' because the alternatives of trading against it or doing nothing expose them to the risk of underperforming if the mispricing persists. As a result, bubbles may 'inflate' further and last longer. In this context, herding becomes the safe strategy while the risky strategy is to deviate (Dass et al. 2008). Of course, all of this crucially relies on there being real consequences for underperforming. Since including a financial penalty in a tournament contract makes the consequences more salient compared to a penalty-free contract, the tendency to herd and thus the severity of asset price bubbles may be greater in the presence of a penalty. Furthermore, Dass et al. argue that bonuses for outperforming the competition ('carrots') may actually help to deflate bubbles by inducing fund managers to try and win the tournament, something that can only be achieved by leaving the safety of the herd. In support of this, they find that during the dot-com bubble, the more highly incentivised fund managers had smaller holdings of so-called 'bubble stocks'.

We extend the experimental literature on tournaments by seeking to resolve the uncertainty surrounding the aggregate-level impacts of penalties. While Isaac and James (2003) also consider a tournament contract with an explicit penalty for severe underperformance, they only run two sessions and unsurprisingly obtain inconclusive results. In contrast, we comprehensively investigate whether including penalties for underperformance affects the severity of mispricing/bubbles compared to tournament contracts that contain no such penalties. Hence the third hypothesis of this study, stated in the null, is:

Hypothesis 3: Prices do not behave differently between 'carrots'-only tournament markets and 'carrots-and-sticks' tournament markets.

In examining the above hypotheses, we also contribute to the literature on bubbles in multi-asset experimental markets, first studied by Fisher and Kelly (2000). Like our study, participants in these experiments typically trade two different assets in a market that mimics the basic Smith et al. (1988) continuous double-auction design. Research following Fisher and Kelly (2000) has examined how prices in these markets behave when assets becomes differentiated by characteristics such the mean and/or the variance of payoffs, or maturity (see Ackert, Charupat, Church and Deaves 2006b; Childs and Mestelman 2006; Chan, Lei, and Vesely 2013). We build on this literature by introducing tournament incentives into a two-asset market, whereas all research has hitherto been based on 'normal', or absolute-performance based incentives. In addition, we examine the effect of trading experience on bubbles in multi-asset markets, which to our knowledge has not been investigated before.

Our study perhaps has most in common with Kleinlercher, Huber and Kirchler (2014), who also examine the effect of different incentive schemes on price behaviour in a two-asset experimental market. Similar to our study, their incentive schemes include an option-like, reward-only "Bonus" contract and a "Penalty" contract. However, the key difference between their study and the current study is that whereas we examine tournament incentives, they do not. Rather, their "Bonus" and "Penalty" treatments are absolute-performance based compensation schemes featuring an *exogenous* benchmark, specifically a pre-defined final cash balance. In contrast, the benchmarks in tournament schemes like ours are *probabilistic*, such as the performance of the average trader[42]. Since optimal risk-taking behaviour may differ for individuals faced with an exogenous benchmark versus a contest scenario (Taylor 2003; Tsetlin et al. 2004), the aggregate implications may also vary.

3.3 Experimental Design

The experiment comprises 35 independent markets carried out across 19 sessions at the ASB Experimental Research Laboratory at UNSW Australia between August and November 2013, with 261 subjects taking part across all treatments. Participants were university students with no prior experience in market experiments, recruited using ORSEE (Greiner 2004)[43]. We begin by describing the parameters of the market institution that were common to all sessions before detailing the specific treatment variables. We finish with an overview of the procedures followed in each session.

3.3.1 Market structure

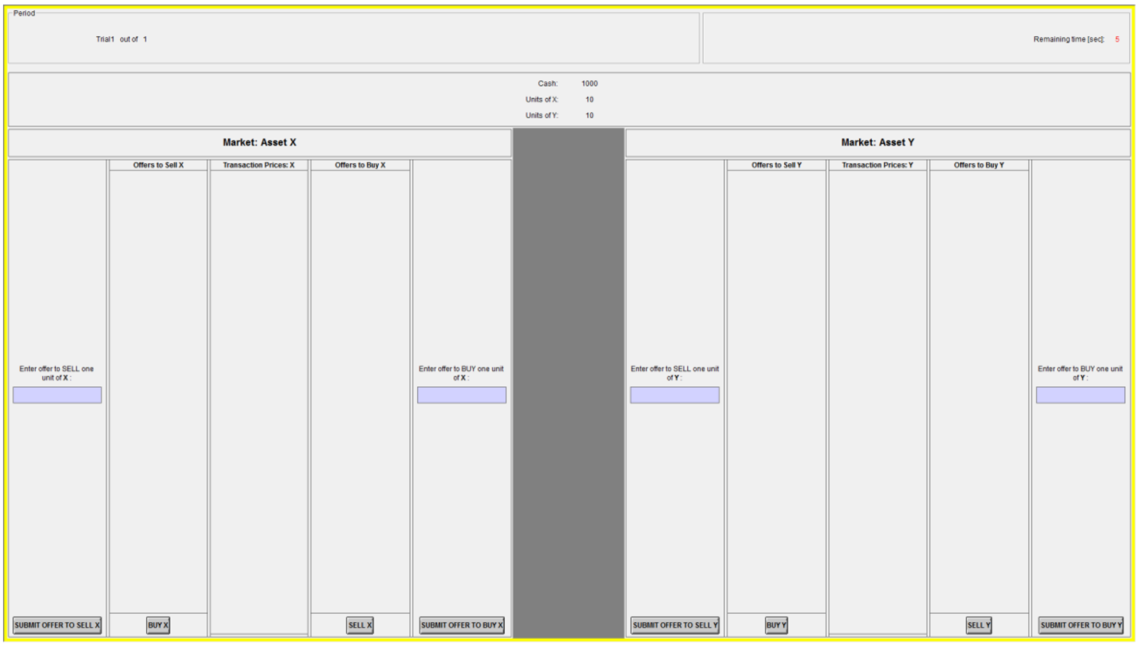
In each session, participants were given the opportunity to trade two types of assets concurrently, one called “X”, the other called “Y”. The market for both assets ran

[42] To further highlight the difference, Kleinlercher et al. (2014) point out that it is possible under their bonus compensation contract for *all* traders to receive a bonus, since with a favourable dividend outcome, all traders could exceed the benchmark-level of cash. In the absence of collusion, this is not possible in tournament schemes where bonuses are paid for above-average performance, since almost certainly *someone* will perform below average.

[43] In total, 38 markets were run. However, some participants with multiple ORSEE profiles managed to ‘slip through’ and participated in more than one session of this experiment. To mitigate the potential confounding of treatment effects, we have excluded from the analysis any data from the 3 markets that contained a subject who had participated in an earlier session.

for 12 periods, each lasting 3 minutes[44]. Trade occurred according to continuous double-auction rules; participants were allowed to post bids and asks for both assets in separate open order books, and accept any posted bid or ask for either asset, subject to the constraints posed by their asset holdings and cash balance. All trade occurred in single units, and short-selling and buying on margin were not permitted. Trade was conducted in experimental currency called ‘francs’, with earnings being paid out at the end of the experiment in Australian dollars at a pre-announced exchange rate of 200 francs to 1 Australian dollar. The market institution was fully computerised using zTree (Fischbacher, 2007) – the trading interface is shown in Figure 3.1[45].

Figure 3.1: Trading interface



[44] While experimental studies of tournament incentives have largely stuck with the parameters in Smith et al. (1988), there is considerable heterogeneity in studies involving multiple assets. The number of trading periods in these studies ranges from 12 (Ackert et al. 2006) to 30 (Chan et al. 2013), while trading period lengths vary between 3 (Chan et al. 2013) and 6 minutes (Fisher and Kelly, 2000). The parameters chosen for our sessions are consistent with the lower end of this range, and represent a suitable compromise given the constraints posed by budgets and time. In particular, we were mindful of avoiding sessions that were ‘too long’ and risked inducing boredom, given the repetitive nature of the market experiments.

[45] Note that given the previously documented tendency for trading activity to be biased in favour of the market that appears on the left-hand side of the screen (see Chan et al. 2013) the market for Asset X was placed on the left for roughly half of the sessions in each treatment, and on the right for the remainder.

All traders began the market with the same initial endowment of assets and cash – 5 units each of X and Y, and 1950 francs. This ensured that the relative position of any trader in the market was not affected by the composition of their initial allocation, and also that the expected earning opportunities for all traders were initially the same. At the end of each trading period, Asset X paid a cash dividend drawn from the distribution {10, 30} with equal probability, while Asset Y paid a dividend from the distribution {0, 100} with respective probabilities (0.8, 0.2)[46]. These distributions were known to all participants. Dividend draws, which were made by the computer, were independent across trading periods and between the two types of assets. Any dividend earnings were added to the trader's cash balance, and their end-of-period portfolio carried over to the next trading period.

Note that the expected dividend paid by both X and Y in each period is 20 francs. Hence, the risk-neutral fundamental value (FV) of both assets is the same, and is equal to the expected total future dividend stream, or 20 multiplied by the number of trading periods remaining (including the current period). [47] As shown by the solid black line in Figure 3.2 below, the resulting risk-neutral FV process of both assets begins at 240 in period 1 and declines in steps of 20 in each period, falling to 20 in period 12 before expiring worthless after the final dividend is drawn at the end of period 12. The FV process represents another difference between our study and Kleinlercher et al. (2014). Whereas we adopt the declining FV environment of Smith et al. (1988) in line

[46] These dividend structures mimic that of Ackert et al. (2006b), who also use a standard/lottery-asset dichotomy, albeit with a much more pronounced difference in potential payoffs between the two types. Their 'standard' asset's dividend distribution is {0.50, 0.90, 1.2} with respective probabilities (0.48, 0.48, 0.04), while their 'lottery' asset pays a dividend from the distribution {0, 18} with associated probabilities (0.96, 0.04). The maximum possible payoff in a period from their lottery asset is 15x the maximum payoff from the standard asset, whereas the corresponding multiple in our study is 3.33x. This is intentional, as we wanted participants to still view Asset Y as a viable "investment" rather than a purely speculative bet.

[47] The expected value of the total future dividend stream was common knowledge, and was communicated to participants in the form of an "average holding value" table contained within the written instructions given to all participants.

with other experimental research on tournament incentives and multi-asset experimental markets, they study experimental assets that have a constant FV. In comparison to declining FV markets, constant FV markets of the type examined by Kleinlercher et al. are less prone to bubble under normal incentives (Smith, van Boening, and Wellford 2000).

Figure 3.2: Fundamental value process, assets X and Y

The solid black line in the graph below depicts the risk-neutral fundamental value process of assets X and Y. Both assets pay an expected dividend of 20 per period. The dashed and solid grey (blue) lines depict the largest and smallest possible cumulative future dividend realisations of asset X (Y) respectively. Asset X pays a minimum of 10 francs in dividends each period, and a maximum of 30 per period. Asset Y pays a minimum of zero every period and a maximum of 100 every period. Hence, the blue dotted line, which is only partially graphed, starts at 1200 in period 1 and falls in steps of 100 in each ensuing period.

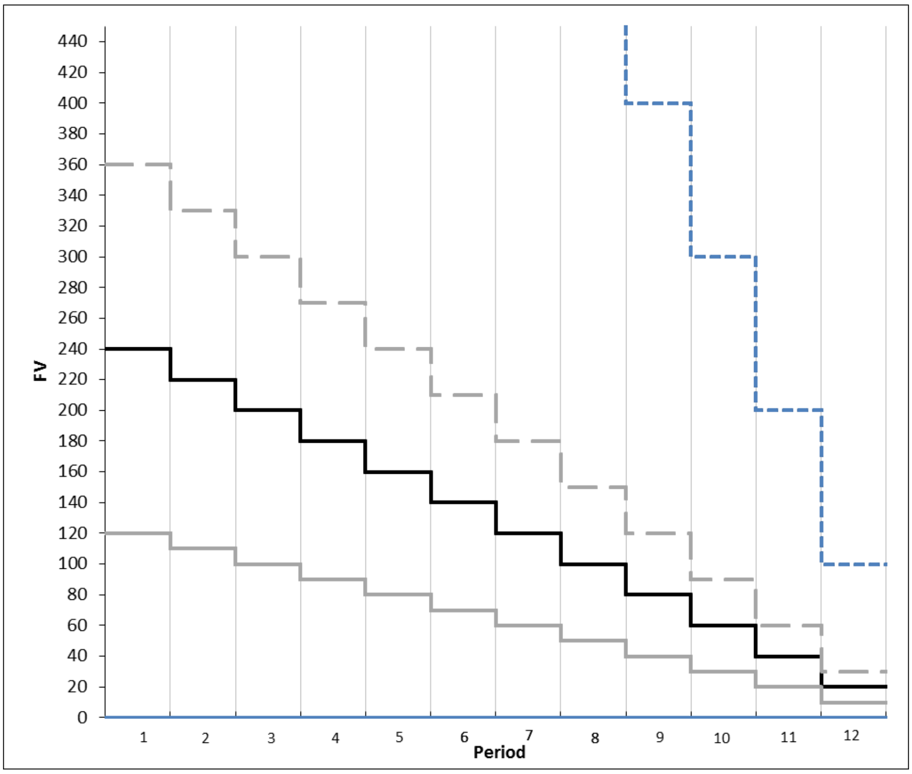


Figure 3.2 also illustrates the largest (dotted line) and smallest (solid line) possible cumulative future dividend realisations of each asset in the experiment (X in grey, Y in blue). To keep the other features of the graph from being obscured, the step

function for the maximum possible dividends from asset Y – which potentially pays 100 francs in each period – is only partially displayed; the blue dotted line begins at 1200 in period 1 and falls in steps of 100 in each ensuing period. In contrast, the minimum possible cumulative dividend payment from asset Y is zero, while asset X pays at least 10 and possibly 30 francs in each period. These step functions serve to demonstrate that although both asset types have the same expected dividend, the variance of Y's dividend payoff is much greater than X's. As asset X always pays at least 10 francs in each period, it represents a 'safe' investment, whereas Y with its lottery-like characteristic is riskier/more speculative. This presence of a second, risky/speculative asset in the market environment provides a more natural and realistic avenue for traders to increase risk in the hope of greater reward than what the single-asset environments of earlier tournament studies provide.

The parameters discussed above determine the initial liquidity of our markets, as measured by the initial cash-to-assets ratio – the ratio of total cash to the total intrinsic value of all assets (X and Y) at the beginning of the market. This ratio was 0.8125 in all sessions, allowing us to control for the effects of liquidity on prices, which is known to be positively associated with the magnitude of bubbles in experimental markets (Caginalp, Porter and Smith 1998, 2000, 2001). While existing experimental studies of tournament incentives and multiple assets have used a variety of initial cash-to-asset ratios, our choice of 0.8125 reflects the cash-to-assets ratio in the most oft-replicated Smith et al. (1988) design, as well as being the initial liquidity used by Cheung and Coleman (2014) in their tournament study.

3.3.2 Treatments

3.3.2.1 James and Isaac tournament contracts

To examine the influence of rewards ('carrots') and penalties ('sticks') in tournament contracts, we implemented a between-subjects design with 3 treatments that differ in the way participants were remunerated for their performance in the market. In the *Baseline* or 'normal'/linear incentives treatment, participants were compensated on the basis of their absolute performance in the market. Since Assets X and Y expired worthless at the end of the market, this means that traders were paid their final cash balance[48].

The remaining two treatments invoke tournament incentives. In both the *Carrot* and *Stick* treatments, we mirrored the approach taken by James and Isaac (2000) and Isaac and James (2003) by compensating traders on the basis of their performance relative to the 'average' trader. The *Carrot* compensation contract rewarded above-average performance with a bonus payment, while paying all other traders a fixed amount, using the following rule:

$$Earnings_i = \begin{cases} 3000 & if \quad C_i < C^* \\ 3000 + 2(C_i - C^*) & if \quad C_i \geq C^* \end{cases}$$

C_i is the final cash balance of trader *i* and C* is the average of the final cash balances of all traders in the market. All units and amounts shown are denominated in francs.

[48] Ending cash balance = initial cash balance + dividend earnings + sales revenue – expenditure on purchases

The compensation contract in the *Stick* treatment introduced an additional component to the contract used in the *Carrot* treatment – a penalty intended to reflect the consequences of a scenario where a trader performs so poorly that they lose their job.

$$Earnings_i = \begin{cases} 0 & if \quad C_i < \frac{1}{2}C^* \\ 3000 & if \quad \frac{1}{2}C^* \leq C_i \leq C^* \\ 3000 + 2(C_i - C^*) & if \quad C_i > C^* \end{cases}$$

While our *Carrot* contract and the "Bonus" contract used by Kleinlercher et al. (2014) have similar, convex functional forms, our *Stick* contract differs markedly from their "Penalty" contract, which deducts a proportional penalty from a fixed payment, effectively placing a cap on traders' earnings[49]. Hence unlike their study, a comparison between our *Carrot* and *Stick* treatments indicates *only* the effect of introducing a penalty for poor performance.

At the end of each trading period, participants in the two tournament treatments were given information on-screen about their relative performance. Specifically, they were informed of the value of their own Account Total and the average Account Total in their market. Based on a measure of the same name used by Schoenberg and Haruvy (2012), Account Total is akin to the market value of a trader's portfolio, and is defined as the sum of a trader's end-of-period cash balance and the value of their end-of-period asset holdings; the end-of-period holdings of X and Y in our study were valued at their respective median traded prices in that period. Like Cheung and Coleman (2014), we

[49] That is, the "Penalty" contract used by Kleinlercher et al. (2014) is a penalty-only contract, whereas our *Stick* contract is a bonus-and-penalty contract.

chose the median price in preference to the final trading price or highest bid (as used by Schoenberg and Haruvy) because it is more difficult for traders to manipulate[50]. Since all assets expired worthless after the final dividend payment, the Account Total at the end of period 12 (i.e. at the end of the market) reverted to the final cash balance[51]. Traders in the *Baseline* treatment were also informed of their own Account Totals at the end of each trading period, but were not told the average in their market.

3.3.2.2 Gilpatric tournament contracts

We also tested two alternative tournament treatments, *GilCarrot* and *GilStick*, which more closely reflect the type of tournament modelled by Gilpatric (2009). Being rank-order tournaments, participants in these treatments were paid a *fixed* amount determined purely by their relative position, specifically their final rank. Our *GilCarrot* contract paid the trader with the largest final cash balance 10,000 francs, while all other traders received the significantly lower payment of 4000.[52] The *GilStick* contract is the same, except the worst performing trader – the trader with the lowest final cash balance – received nothing from the market. Contrast these with the 'James and Isaac' tournaments contracts described above, where payoffs depend not only on being better/worse than average but also the extent to which a trader's absolute performance exceeds the average. By severing any link between absolute performance and compensation, the 'Gilpatric' contracts can be considered 'purer' tournaments, in the Lazear and Rosen (1981) sense, where *only* relative performance matters. Since the

[50] In periods where there was no trade in an asset, the median transaction price was replaced by the median buy offer for that asset in the period. This was done to avoid misleading fluctuations in the Account Total, and participants were made aware of this before the market began.
[51] This small change in the definition of the Account Total for period 12 was necessary, since otherwise, it would create an incentive for participants to arbitrarily bid up the prices of assets X and Y in period 12 in the hope of maximising their Account Totals.
[52] The minimum payment here was set to 4000 francs compared to 3000 francs in the equivalent James and Isaac tournament contract *Carrot* to ensure that the average compensation per trader in real currency, Australian dollars, was roughly equal across treatments, and to also conform to the ASB Lab ethics protocol which specified an average payment range of $15-20 per hour per participant.

appropriate piece of relative-performance information in these treatments is the trader's rank, participants in Gilpatric tournament treatments were informed of their rank at the end of each period (calculated on the basis of Account Total), in addition to the other relative performance information described above.

3.3.3 Procedures

Each experimental session corresponded to a single treatment to which it (and hence, each subject within it) was randomly assigned[53]. Sessions were designed to run two independent market-groups of (up to) 8 traders each and ran for approximately 2.5 hours[54]. To ensure consistency in the delivery of instructions between sessions and reduce experimenter demand effects, all participants received written instructions, which were also communicated verbally by the experiment administrator[55]. Potential interaction effects between participants were mitigated by prohibiting subjects from communicating with each other for the duration of the experiment.

The procedure followed in each session was identical, regardless of the treatment. Sessions began with participants being randomly allocated to a computer/workstation that determined their market-group[56]. They then received training on how to use the trading screen to make and accept bids and offers for each asset (10 minutes), following which they were given 10 minutes to practise trading using the interface. After the practice period, subjects were given further information about the

[53] The only exception to this was a single session where a *Carrot* treatment market ran alongside a *GilStick* market. The instructions and procedures were appropriately modified for this session to prevent contamination of the subject pool.
[54] That is, excluding the practice period, participants only traded with other participants who were in the same market-group. Dividends were also drawn independently for each market-group.
[55] To ensure consistency with the procedures used in the existing literature, the written protocol was adapted from Dufwenberg et al. (2005), Noussair et al. (2001), Noussair and Powell (2010), Lugovskyy et al. (2009), Childs and Mestelman (2006), and Cheung and Coleman (2014). Participants were also given time to read the instructions on their own, and to ask any clarifying questions privately (which were also answered privately). The written protocol can be found in Appendix B3.
[56] The workstation number also served as a participant's ID, thus ensuring the anonymity of their data.

other features of the market environment, including how their earnings would be calculated. After this, the market-proper began. Upon the conclusion of the market, participants were informed that they would be taking part in another 12-period market with the same traders (i.e. market-group). Participants' inventory of assets and cash were reset to their starting levels, and trading commenced for a second round.

After the end of the second round, participants completed an untimed survey consisting of 3 sections[57]. The first section gathered general demographic information about participants and their experiences and thought-processes during the market(s)[58]. The second and third sections, which form part of a related study, comprise the *Cognitive Reflection Test* (CRT) and *Domain-Specific Risk-Taking* (DOSPERT) Scale. The CRT is a measure of cognitive ability developed by Frederick (2005) that consists of 3 problem-solving type questions that assess the ability of respondents to reject an impulsive and intuitive incorrect answer in favour of a correct answer that requires more deliberation. In addition to general measures of cognitive ability, performance in the CRT is correlated with time and risk preferences (Frederick 2005), as well as certain behavioural biases (Oechssler, Roider, and Schmitz 2009). The 30-item DOSPERT Scale, designed by Blais and Weber (2006), is a psychometric scale that measures risk preferences and perceptions across five separate decision-making domains: Financial (split into Investing and Gambling), Health/Safety, Recreational, Ethical, and Social[59]. Respondents use a 7-point scale to rate the likelihood of their participation (Part 1), the perceived riskiness (Part 2), and the benefits expected to accrue (Part 3) from engaging

[57] The survey, which can be found in Appendix B4, was initially paper-based (9 sessions), but was computerised using the *Qualtrics* survey software and administered electronically in the October and November sessions (10 sessions).
[58] This is a modified version of the end-of-experiment questionnaire used by Ackert and Church (2001).
[59] Compared to the original 40-item DOSPERT scale (Weber, Blais and Betz 2002), which was developed for American undergraduate college students, the revised 30-item DOSPERT scale (Blais and Weber 2006) is designed to be more readily applicable to a more diverse range of cultures, age groups, and educational levels.

in 30 different domain-specific risky activities. Of course, administering the DOSPERT Scale after the market stage carries with it the risk that responses may be influenced by participants' experiences during the market. However, given our main objective is to study price behaviour, this is the 'lesser of two evils', as the alternative of implementing the scale before the market could in turn influence participants' trading behaviour. A summary of the demographic characteristics of the subject pool, CRT scores (out of 3), and DOSPERT likelihood/preference scores in the most relevant domain, Financial (ranges from 6 to 42, higher scores indicate greater willingness to take financial risks), is presented in Table 3.1, categorised by treatment.

Once the surveys were completed, participants were called up individually, paid their earnings and dismissed. Participants' total earnings from the experiment were calculated as the sum of their earnings from both rounds of the market, converted to Australian dollars, plus a $5 participation fee. The average payment to participants, inclusive of the participation fee, was $49.

Table 3.1: General demographic information

This table reports general demographic information on the subject pool, categorised by the experimental treatment to which participants were randomly assigned. 'Business student' is defined as someone studying Finance, Economics, Actuarial, Accounting, or "Commerce" (self-reported). In a post-experiment survey, all participants completed the Cognitive Reflection Test (CRT) developed by Frederick (2005), which measures cognitive ability; CRT scores are out of 3 and higher scores indicate better performance. Participants also completed the Domain-Specific Risk-Taking (DOSPERT) Scale (Blais and Weber 2006). The score reported here relates to participants' (self-reported) likelihood of engaging in risky financial activities. Scores range from 6 to 42, with higher scores indicating a greater likelihood of engaging in risky activities.

	Baseline	*Carrot*	*Stick*	*GilCarrot*	*GilStick*
No. markets	7	8	8	6	6
No. subjects	51	58	61	45	46
Average age	22.3	22.4	22.2	22.7	22.6
Male (%)	65	52	52	42	46
Business students (%)	29	40	31	36	43
Avg. CRT score	1.6	1.3	1.5	1.5	1.3
Avg. DOSPERT Fin. score	19.2	19.4	18.5	19.8	18.8

3.4. Results

3.4.1 Inexperienced traders

3.4.1.1 Descriptive summary

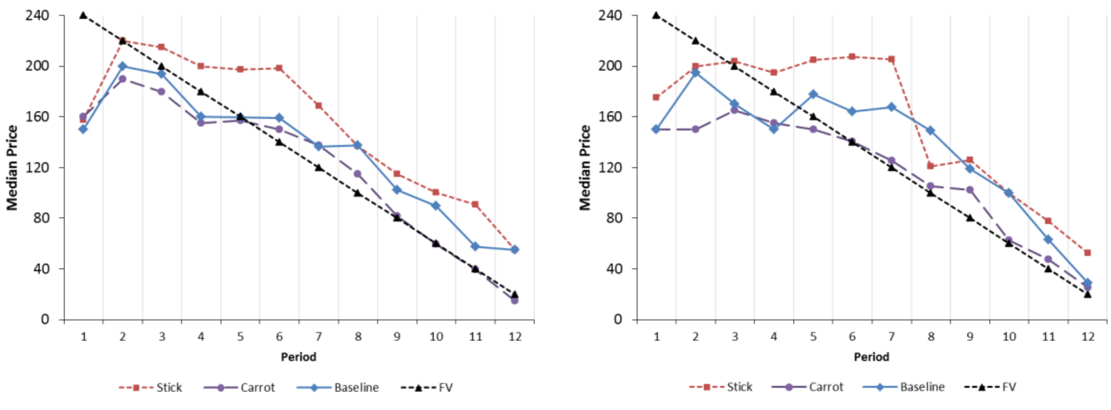
Panels (a) and (b) of Figure 3.3 chart the time-path of the median transaction price of assets X and Y respectively in the *Baseline*, *Carrot*, and *Stick* treatments during the first round of the market; for each treatment, the charted price in each period is the median of the median transaction prices from all markets in that treatment. Median prices in Figure 3.3 for both assets in all treatments broadly follow the pattern associated with Smith et al. (1988)-type markets populated with inexperienced traders – prices start below fundamental value and remain there in the initial periods before rising above fundamental value. However, with the possible exception of asset Y in the *Stick* treatment, which experiences a precipitous fall in median price from periods 7 to 8, the characteristic bubble-and-crash is notably missing. In fact, median prices in the *Carrot* treatment can hardly be described to 'bubble' at all, though it should be noted that these graphs hide considerable heterogeneity at the individual market level. In fact, bubbles-and-crashes were observed in individual markets of all treatments, although they did not occur with the regularity reported in other studies of multi-asset experimental markets such as Fisher and Kelly (2000)[60].

Perhaps the most notable feature of Figure 3.3 is the persistently higher median prices/more pronounced overvaluation exhibited by the *Stick* treatment in comparison to the *Carrot* treatment. In fact, median prices are higher in the *Stick* treatment in every

[60] The evolution of median transaction prices in each individual market of the *Baseline*, *Carrot*, and *Stick* treatment in Round 1 can be found in Appendix B1, Figures B1-B3

Figure 3.3: Median prices in Round 1

Median transaction prices in the *Baseline* (solid blue line), *Carrot* (dashed purple line), and *Stick* (dotted red line) treatments during the first round of the market (i.e. with inexperienced traders) are shown below for the ‘low-risk’ asset X (panel (a)) and ‘high-risk’ asset Y (panel (b)), along with the risk-neutral fundamental value process for each asset (dashed black line). For each treatment, the plotted median price in each period is the median of the median transaction prices from all markets belonging to that treatment. Any markets that were ‘contaminated’ by the presence of subjects who had participated in an earlier session of the experiment are excluded.



(a) Asset X, Round 1 (b) Asset Y, Round 1

trading period for the more speculative asset Y, and in all bar 1 period for asset X. A Wilcoxon Mann-Whitney (WMW) U test – the non-parametric equivalent of the independent samples t-test – reveals that these differences are significant at the 5% level in period 2 though to 6 in asset X, and in periods 2 and 5 for asset Y[61]. Furthermore, it is the *Carrot* treatment where prices appear to most closely conform to fundamental value, even in comparison to the normal-incentive *Baseline* treatment, which can be tentatively described as charting a path in between the two tournament contract treatments, particularly for asset X. These observations run contrary to the notions that tournaments necessarily distort prices, and that rewards (penalties) encourage (discourage) the formation of bubbles. Moreover, they also present a sharp contrast to Kleinlercher et al.

[61] While we do not report full results here, the two-sided p-values in periods 2-6 for asset X are 0.045, 0.049, 0.048, 0.024, and 0.027. For asset Y, the p-values in period 2 and 5 are 0.014 and 0.046 respectively. In addition, median transaction prices for asset Y are higher in the *Stick* treatment than the *Carrot* treatment at the 10% level in periods 1 and 6 (p-value = 0.094 and 0.066 respectively)

(2014), who find that average prices for their "high-risk" asset (equivalent to our asset Y) are highest under their "Bonus" treatment and lowest in their "Penalty" treatment.

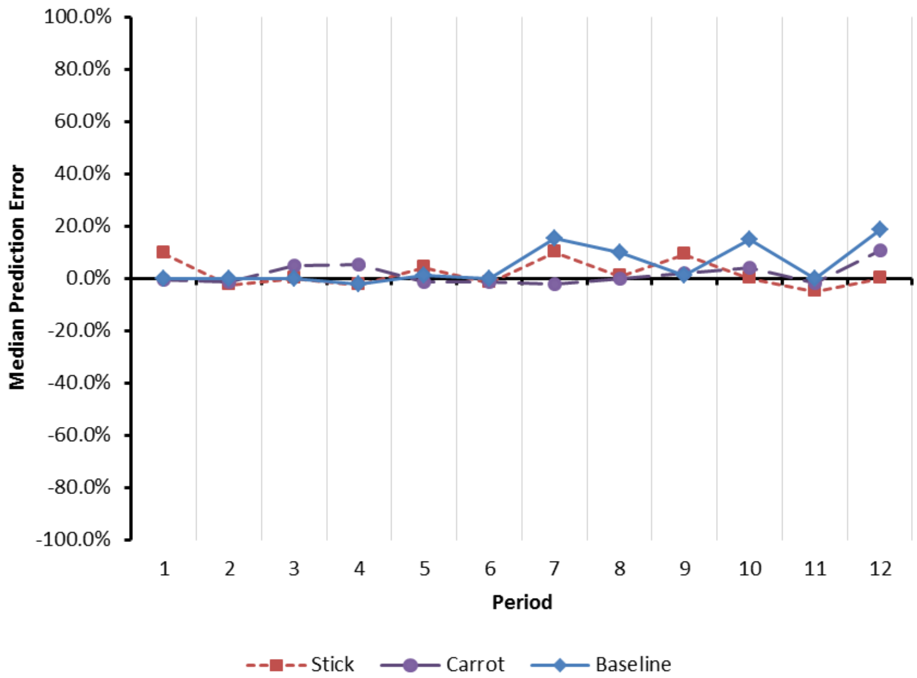
Figure 3.3 also reveals a high degree of correlation in the prices of assets X and Y, which is consistent with behaviour observed in earlier studies of multi-asset markets (e.g. Fisher and Kelly 2000, Childs and Mestelman 2006), where *relative* prices between asset-types tend to remain close to the 'correct' value (i.e. risk-neutral value) even when individual assets exhibit severe mispricing. This is more clearly illustrated by Figure 3.4, which graphs the median *Prediction Error* in each period for each treatment. Like Fisher and Kelly (2000), we define the *Prediction Error* in each period of an individual market as the percentage deviation of the relative price of asset Y (median price of Y divided by the median price of X in that period) from the risk-neutral benchmark (equal to 1 in this study)[62]. More positive (negative) values indicate a greater willingness by market participants to pay a premium to acquire the riskier (less risky) asset Y (X). As Figure 4 illustrates, median prediction errors in all treatments remain relatively close to zero throughout the market[63].

Furthermore, mirroring the approach of Brown et al. (1996) by comparing the first and second half of the market in Figure 3.4 does not indicate the presence of an obvious 'tournament effect' in the two tournament treatments. The effect we seek to detect here is heightened 'risk-seeking' behaviour by traders in the second half of the market, as evidenced by a substantial rise in the price of Y relative to X[64]. Instead, we

[62] Unlike Fisher and Kelly (2000), we report the median of the Prediction Errors across all sessions/markets rather than the average, due to the lower sensitivity of the median to outliers in small samples.
[63] The behaviour of *Prediction Error* in the individual markets of the *Baseline*, *Carrot,* and *Stick* treatments in Round 1 can be seen in Appendix B1, Figures B1-B3.
[64] Of course, Brown et al. (1996) were concerned with adjustments made by individual fund managers in portfolio risk between the two halves of the year rather than an aggregate metric like relative price.

Figure 3.4: Median values of *Prediction Error*, Round 1

The figure below plots the evolution of the median *Prediction Error* in the *Baseline* (solid blue line), *Carrot* (dashed purple line), and *Stick* (dotted red line) treatments during the first round of the market (i.e. with inexperienced traders). For each treatment, the plotted value in each period is the median of the *Prediction Errors* from all markets in that treatment. *Prediction Error* is defined as the percentage difference between the relative price of Y (i.e. median price of asset Y divided by median price of asset X) and the risk-neutral benchmark of 1. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



see that relative prices in *Carrot* and *Stick* behave similarly in both halves – the average of the median *Prediction Errors* in the *Carrot* treatment in periods 1-6 is 1% vs. 2.2% in periods 7-12, while the corresponding values for the *Stick* treatment are 1% and 2.4%. In contrast, the average of the median *Prediction Errors* during the first half of the market in the *Baseline* treatment is -0.2%, compared to 10% in the second half, potentially indicating that participants were willing to pay more to acquire the riskier asset Y in the latter stages of the market. Note however that the desire to move up the leader board is unlikely to be an adequate explanation for this apparent risk-seeking behaviour since relative performance information was not shown to participants in the *Baseline* treatment[65].

3.4.1.2 Statistical analysis

<u>Bubble measures</u>

To conduct a more formal comparison of the treatments, we calculate a number of measures of mispricing/bubbles that are frequently used in the experimental asset market literature. These bubble measures can broadly be categorised into two groups that assess two different dimensions of mispricing – magnitude and length. Readers familiar with the bubble measures used in Chapter 2 may skip the following overview of the various measures without loss of continuity, although note that the definition of *Turnover* is different in this study (see footnote 66 below).

Amplitude (Haruvy and Noussair 2006), the first of the magnitude measures, quantifies the extent to which average prices in a market change relative to FV. It is calculated as $\max_t\{(\bar{P}_t - F_t)/F_t\} - min_t\{(\bar{P}_t - F_t)/F_t\}$, where the largest and smallest deviations of average price P_t from fundamental value F_t are normalised by the FV in the respective period *t*. Large values of this measure indicate big swings in price relative to FV and hence the possible presence of a bubble. *Total Dispersion* (Haruvy and Noussair, 2006) measures the aggregate absolute deviation of median price from FV across all trading periods, and is defined as $\sum_t |MedianP_t - F_t|$. Since it treats both positive and negative deviations from FV identically, it is a measure of aggregate mispricing rather than over or undervaluation, with smaller values indicating a closer correspondence between price and fundamental value. *Turnover*, a normalised measure of trading activity, is used as a measure of magnitude since bubble phases are typically

[65] Having said that, it is possible that even if relative performance information is not provided, participants have an internalised benchmark of what "average" performance looks like, and hence whether they are performing better or worse.

associated with higher trading volumes. We calculate turnover as defined by King, Smith, Williams, and van Boening (1993), namely $\sum_t V_t/(TSU)$, where V_t, the volume of trade in period *t* is normalised by *TSU*, the total number of units of the asset (X or Y) in the market[66]. *Normalised Deviation*, measured by Haruvy, Lahav, and Noussair (2007) as $\sum_t V_t|MedianP_t - F_t|/(TSU)$, combines the preceding two measures to account for both the size of the price deviation and the level of trading activity in a market. To examine how closely prices track changes in FV, we calculate *Haessel-R^2* (Dufwenberg, Lindqvist, and Moore 2005), which is the R-squared from the regression of average prices on fundamental values. Being a goodness-of-fit measure, it conveys how much of the variation in average price across periods is explained by changes in FV; values closer to 0 (1) suggest the potential existence (absence) of price bubbles. Note that none of the aforementioned measures determine whether the asset is generally overvalued or undervalued. To gauge the degree of overpricing/underpricing, we calculate *Average Bias* (Haruvy and Noussair 2006), which measures how far median prices on average deviate from FV over the course of the market, and is calculated as $\frac{1}{N}\sum_{t=1}^{N}(MedianP_t - F_t)$. Large positive (negative) values suggest that prices tend to stay above (below) FV. Values close to zero may suggest that prices stay close to FV or that the asset experiences equal degrees of over and underpricing in the market; assessing the *Average Bias* in conjunction with *Total Dispersion* helps to shed light in this regard, since observing a small (large) *Total Dispersion* at the same time as a near-zero *Average Bias* would imply the former (latter) (Haruvy and Noussair 2006).

The first of the bubble-length measures, *Duration* (Porter and Smith 1995), calculates the maximum number of consecutive periods where average price increases

[66] Note that this differs from the definition of *Turnover* used in Chapter 2, in which the original King et al. (1993) measure was modified by the number of minutes in each trading period. Since all markets in the current study have the same trading period length, we use the original measure.

relative to fundamental value, or $max\{m: \bar{P}_t - F_t < \bar{P}_{t+1} - F_{t+1} < \cdots < \bar{P}_{t+m} - F_{t+m}\}$. Larger values of *Duration* point to sustained periods where changes in (average) transaction prices across trading periods do not 'adequately' track changes in the FV, potentially indicating the presence of a bubble. *Boom* (*Bust*) *Duration* (Haruvy and Noussair 2006) is defined as the maximum number of consecutive periods where median prices stay above (stay below) FV; large values indicate long periods of overvaluation (undervaluation), potentially signalling the presence (absence) of a bubble.

The behaviour of individual assets

Panels A and B of Table 3.2 report the median values of the bubble measures in each treatment for assets X and Y respectively in Round 1, along with the associated median absolute deviations[67]. For each asset-type, each measure produces one observation per market; hence the medians are based on 7 observations in the *Baseline* treatment, and 8 observations each in the *Carrot* and *Stick* treatments[68]. The bottom half of each panel reports two-sided exact p-values from Wilcoxon Mann-Whitney U (WMW) tests of the differences in the measures between treatments, under the null that the values from both treatments come from the same distribution. The WMW test, which is a non-parametric test, is the appropriate statistical test given the small sample size.

We begin by comparing the two tournament treatments, *Carrot* and *Stick* (i.e. Hypothesis 3). In the case of the 'safe' asset X, the differences between the *Carrot* and *Stick* treatments on most bubble measures are not statistically significant. Of the

[67] The median absolute deviation (MAD) is a measure of the spread of a distribution, and is calculated as the median of the absolute deviations of all values in a sample from the median. We report the median value and MAD of each measure in preference to the mean and standard deviation due to the small number of observations involved, and their lower sensitivity to outliers.
[68] The bubble measure values observed in the individual markets of each treatment are tabled in Appendix B2. See Table B1 and B2 for the values from Round 1.

Table 3.2: Summary of bubble measures for assets X and Y in Round 1

This table reports median values of each bubble measure in the *Baseline*, *Carrot*, and *Stick* treatments during Round 1 of the market (i.e. with inexperienced traders); the associated median absolute deviations are displayed in parentheses. Markets contaminated by subjects who had participated in an earlier session are excluded. Panel A (B) reports bubble measure data relating to Asset X (Y). For definitions of the relevant bubble measures, refer to section 3.4.1.2. The statistical significance of the difference between treatments in each measure is assessed using a two-sided Wilcoxon Mann-Whitney U Test, under the null that values from both treatments come from the same distribution. Exact p-values are reported. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively.

Panel A: Asset X, Round 1:

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	3.31	298.00	1.65	0.78	2.45	83.43	5.00	5.00	4.00
	(1.53)	*(181.50)*	*(15.73)*	*(0.16)*	*(0.72)*	*(45.08)*	*(2.00)*	*(2.00)*	*(1.00)*
Carrot [8]	1.51	569.50	-9.42	0.56	2.93	171.09	5.00	3.50	4.50
	(1.06)	*(330.25)*	*(28.77)*	*(0.36)*	*(0.66)*	*(104.79)*	*(2.50)*	*(2.00)*	*(1.50)*
Stick [8]	2.63	607.00	25.63	0.46	2.16	98.89	4.50	10.00	1.50
	(1.55)	*(267.75)*	*(25.56)*	*(0.35)*	*(0.34)*	*(47.69)*	*(1.00)*	*(1.50)*	*(0.50)*
WMW U-Test p-values (two-sided):									
Baseline vs. Carrot	0.867	0.779	0.536	0.536	0.281	0.397	0.799	0.317	0.290
Baseline vs. Stick	0.694	0.281	0.232	0.121	0.779	0.613	0.465	0.081*	0.400
Carrot vs. Stick	0.382	0.878	0.105	0.328	0.065*	0.878	0.576	0.007***	0.039**

Panel B: Asset Y, Round 1

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	1.63	530.50	15.38	0.77	2.03	99.06	4.00	7.00	4.00
	(1.01)	*(150.50)*	*(25.33)*	*(0.05)*	*(0.35)*	*(54.11)*	*(1.00)*	*(1.00)*	*(2.00)*
Carrot [8]	1.76	584.75	-9.96	0.32	2.85	194.11	6.00	4.00	4.00
	(1.20)	*(299.50)*	*(13.67)*	*(0.27)*	*(0.39)*	*(71.02)*	*(1.00)*	*(2.00)*	*(1.00)*
Stick [8]	1.99	529.25	24.85	0.56	1.89	96.18	4.00	8.00	3.00
	(0.96)	*(270.00)*	*(27.04)*	*(0.30)*	*(0.38)*	*(39.99)*	*(1.00)*	*(2.50)*	*(1.00)*
WMW U-Test p-values (two-sided):									
Baseline vs. Carrot	0.955	0.955	0.232	0.463	0.281	0.694	0.421	0.013**	0.405
Baseline vs. Stick	0.779	0.955	0.336	0.613	0.613	0.867	1.000	0.755	0.669
Carrot vs. Stick	0.721	0.798	0.028**	0.721	0.038**	0.279	0.329	0.022**	0.124

magnitude measures, only *Turnover* is marginally significantly lower in the *Stick* treatment compared to *Carrot* (p-value = 0.065), which lends some support to the notion that penalties embedded in tournament contracts inhibit speculation. However, the bubble-length measures present the opposite story, with significantly longer *Boom Durations* (p-value = 0.007) and significantly shorter *Bust Durations* (p-value = 0.039) in the *Stick* treatment indicative of more prolonged periods of overvaluation compared to the *Carrot* treatment; the median market in the *Stick* (*Carrot*) treatment experiences 10 (3.5) consecutive periods where the median price of X exceeds fundamental value, and only 1.5 (4.5) consecutive periods below fundamental value.

For the riskier asset Y, the degree of mispricing is comparable to asset X, as suggested by the similarity in the median bubble measure values between the two asset-types in each tournament treatment. Like asset X, *Turnover* for asset Y is significantly higher in the *Carrot* treatment (p-value = 0.038), while *Boom Duration* is again significantly longer in the *Stick* treatment (p-value = 0.022). In addition, prices for Asset Y are also significantly lower in the *Carrot* treatment according to the *Average Bias* measure (p-value = 0.028), which shows that Asset Y is on average overvalued by 25 francs in each period under *Stick* incentives, whereas *Carrot* incentives are associated with asset Y being *undervalued* on average by almost 10 francs per period. This difference in *Average Bias* between the two tournament treatments is also mirrored in asset X, however the failure to attain statistical significance there is due to greater noise in the *Carrot* treatment.

Taken as a whole, the bubble measures are consistent with our observations from Figure 3.3. They reveal that when penalties are embedded into tournament contracts that reward participants for beating the ‘market’, inexperienced traders trade less compared to reward-only contracts. However, the trade that does occur actually happens at higher prices, and periods of overvaluation last longer, especially in the riskier asset. Thus, rather than curtailing the impetus to speculate on riskier ventures, our findings mostly suggest that the addition of a ‘stick’ achieves the opposite result. While we do not examine individual-level behaviour in this study, these results are consistent with pricing expected under the herding hypothesis (Rajan 2006; Dass et al. 2008).

Like Kleinlercher et al. (2014), our results appear to be driven by the attendant incentives and not by differences between the treatments in participants’ inherent risk attitudes – average DOSPERT scores in the Financial domain (or its subsets), though collected after the market stage, do not differ significantly between *Carrot* and *Stick* markets (WMW test p-value (two-sided) = 0.529, $n_{Carrot} = n_{Stick} = 8$); nor are they driven by differences in cognitive ability, as measured by CRT scores (WMW test p-value (two-sided) = 0.6, $n_{Carrot} = n_{Stick} = 8$)[69]. However, while our findings here coincide with Kleinlercher et al. regarding trading volumes, they contrast strongly with respect to the degree of mispricing observed in the riskier asset; overvaluation (calculated similarly to *Average Bias)* in their “high-risk” asset is greatest in their “Bonus” treatment and lowest in their “Penalty” treatment[70]. Since the absence of relative performance evaluation reduces the inclination to herd, a possible explanation for this stark difference between our studies lies in the non-competitive incentives faced by their traders. This, combined with the penalty-only nature/framing of their “Penalty” contract, may focus participants’

[69] Debate surrounding the collection of the DOSPERT/CRT data before or after the market is in some ways moot, since random assignment of participants to treatments should ensure that the treatment groups are on average ‘equivalent’ at the outset of the experiment. Nonetheless, we tested for differences as an additional safety measure. Though we do not report the full results here (available upon request), we do not find a significant difference between any of the treatments in the CRT scores (market average or individual) or DOSPERT scores (market average or individual).

[70] For their “low-risk” asset (equivalent to asset X in our study), Kleinlercher et al. (2014) report significantly lower (higher) average prices in their ‘Penalty’ (‘Linear’ incentives) treatment compared to other treatments, although the differences are not economically significant – the price paths in all their treatments for the low-risk asset are very similar.

thoughts on avoiding the uncertainty associated with the riskier asset, as there is no competition to beat, or 'reward' to be gained. Moreover, while we do not speculate on the precise mechanism, the constant FV process used by Kleinlercher et al. may also play a part.

Comparing the tournament treatments to the normal incentive *Baseline* treatments (Hypothesis 1) in Table 3.2 is also revealing. For both assets X and Y, we fail to find a significant difference between *Baseline* and the two tournament treatments on any of the bubble-magnitude measures. Of the bubble-length measures, *Boom Duration* is smaller in the *Baseline* treatment than the *Stick* treatment for asset X, but only marginally so (p-value = 0.081), while being significantly larger in the *Baseline* treatment compared to the *Carrot* treatment for asset Y (p-value = 0.013). This mostly runs contrary to much of the evidence from single-asset experimental studies going back to James and Isaac (2000) that find tournament incentives to be associated with significantly larger bubbles. Hence, our results suggest that the findings of these earlier studies may be an artefact of speculation in a single-asset environment. When inexperienced traders are given the ability to bet on a higher payoff from an alternate, risky asset, tournament incentives do not distort prices any more than normal incentives.

Relative prices

Turning to relative prices, Panel A of Table 3.3 shows the median value of the *Average Prediction Error* in each treatment, along with the associated median absolute deviations. *Average Prediction Error* is identical to the 'overall normalised exchange rate deviation' measure used by Fisher and Kelly (2000), and is calculated by averaging the *Prediction Errors* (defined above) in all periods of a session/market. Like the bubble measures, this yields one observation per market. The table reports median *Average*

Table 3.3: Average Prediction Errors

Median values of the *Average Prediction Error* in the *Baseline, Carrot, and Stick* treatments in Round 1 and 2 are shown below in Panels A and B respectively, with the associated median absolute deviations in parentheses. Markets contaminated by subjects who had participated in an earlier session are excluded. *Average Prediction Error* is calculated using all periods in a market, the first 6 periods, and the final 6 periods in *Avg PredErr*, *AvgPredErr_p1to6*, and *AvgPredErr_p7to12* respectively. The statistical significance of the individual measures is assessed using a (two-sided) one-sample Wilcoxon Signed-rank test, under the null that the median is equal to zero. The statistical significance of the difference between treatments is assessed using the Wilcoxon Mann-Whitney U Test under the null that values from both treatments come from the same distribution. The statistical significance of the difference between *AvgPredErr_p1to6* and *AvgPredErr_p7to12* within each treatment is assessed using a (paired-sample) Wilcoxon Signed-rank test, against the one-sided alternative hypothesis that *AvgPredErr_p7to12 > AvgPredErr_p1to6*. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively.

Panel A: Round 1

Treatment [N]	*Avg PredErr (%)*	*AvgPredErr_ p1to6 (%)*	*AvgPredErr_ p7to12 (%)*	*Signed-rank p-value (1-sided) 1-6 vs. 7-12*
Baseline [7]	3.78	-0.32	7.88	0.064*
	(9.46)	*(3.36)*	*(9.26)*	
Carrot [8]	3.17	-1.04	4.56	0.242
	(4.43)	*(7.20)*	*(8.01)*	
Stick [8]	3.23	0.14	4.03	0.200
	(4.08)	*(6.85)*	*(1.37)*	
WMW U-Test p-values (2-sided):				
Baseline vs. Carrot	0.694	0.867	0.463	
Baseline vs. Stick	0.955	0.955	0.463	
Carrot vs. Stick	0.878	1.000	0.959	

Panel B: Round 2

Treatment [N]	*Avg PredErr (%)*	*AvgPredErr_ p1to6 (%)*	*AvgPredErr_ p7to12 (%)*	*Signed-rank p-value (1-sided) 1-6 vs. 7-12*
Baseline [7]	2.72	-0.12	5.23	0.032**
	(11.33)	*(2.27)*	*(21.08)*	
Carrot [8]	2.59	-12.70**	15.04*	0.013**
	(14.92)	*(8.01)*	*(24.94)*	
Stick [8]	-5.00	-3.68	-6.62	0.556
	(8.41)	*(6.76)*	*(7.08)*	
WMW U-Test p-values (2-sided):				
Baseline vs. Carrot	1.000	0.189	0.779	
Baseline vs. Stick	0.397	0.281	0.397	
Carrot vs. Stick	0.279	0.505	0.105	

Prediction Errors based on the entire duration of a market ("Avg PredErr"), as well as in each half of the market ("Avg PredErr_p1to6" and "Avg PredErr_p7to12).

Looking at the whole-of-market measure ("Avg PredErr"), we see that the median values are very similar – around 3% – in all three treatments. In fact, the median values are not significantly different from zero in any of the treatments, which is consistent with other multi-asset studies that find relative prices do not significantly deviate from the risk-neutral theoretical value when assets are differentiated by the variance of payoffs alone (Ackert et al. 2006, Childs and Mestelman 2006)[71,72]. However, this contrasts again with Kleinlercher et al. (2014), whose high-risk asset sells at a significant premium to the low-risk asset in their "Bonus" and the normal-incentive "Linear" treatments, while the opposite holds true in their "Penalty" treatment.

To examine if relative prices behave differently between the first and second half of the market – specifically due to heightened speculation on the risky asset in the second half – we compare the measures corresponding to each half *within* treatments. We use a Wilcoxon Signed-rank test to examine the one-sided alternative hypothesis that *Average Prediction Errors* in the second half of the market are higher than in the first half; the corresponding p-values (one-sided) are reported in the right-most column of Panel A[73]. The results do not provide compelling evidence of a 'tournament effect' in relative prices in Round 1. Although median *Average Prediction Errors* are larger in all treatments in the second half, these differences are only statistically significant in the

---

[71] The statistical significance of the median *Average Prediction Error* in each treatment is assessed using the one-sample Wilcoxon Signed-rank test, under the null that the median is equal to zero. It is the non-parametric equivalent of the one-sample t-test.

[72] However, this result contrasts with Fisher and Kelly (2000) who report that their riskier asset sells at a slight premium to the safer asset, although they do not obtain enough observations to make a formal statistical comparison. In addition, their results are potentially confounded by the differing levels of experience of some traders in their markets. Ackert et al. (2006) find a preference for assets with lottery-like payoffs only when trade occurs with borrowed money.

[73] The (paired-sample) Wilcoxon Signed-rank test is the non-parametric equivalent of the paired t-test. The null hypothesis is that values from both groups come from the same distribution.

normal-incentive *Baseline* treatment (median *Average Prediction Error* is -0.32% in 1st half vs. 7.88% in 2nd median), but only at the 10% level (p-value (one-sided) = 0.064).

The bottom of Panel A reports exact p-values (two-sided) from WMW tests comparing the measures between treatment-pairs. The null hypothesis is that values in both treatments come from the same distribution. The failure to achieve statistical significance on any of the tests means that we do not find support for the conjecture that tournament incentives have a significant impact on relative prices compared to normal incentives (Hypothesis 1), or that relative prices in *Carrot* markets behave differently to *Stick* markets (Hypothesis 3).

3.4.2 The effect of experience

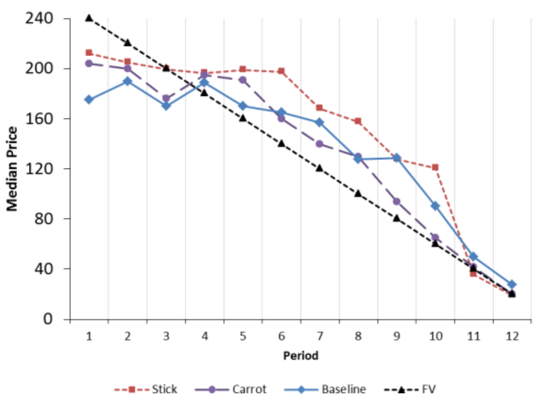
<u>Individual assets</u>

Figure 3.5 depicts the evolution of median transaction prices for assets X and Y (panels (a) and (b) respectively) in the *Baseline, Carrot*, and *Stick* treatments during the second round of trading[74]. With once-experienced traders, we see that the difference in price behaviour, in particular between the two tournament treatments is much less obvious compared to Round 1 (cf. Figure 3.3). The convergence between two is especially pronounced in the case of the risky asset Y, where in a reversal of Round 1, the *Stick* treatments appears to conform more closely to fundamental value than *Carrot.* Although median prices for asset X are again higher in the *Stick* treatment compared to the *Carrot* treatment in most trading periods, the differences do not appear to be as large as in Round 1, especially in the early and latter stages of the market. In fact, in contrast to Round 1, the differences in median price between the *Carrot* and *Stick* treatment are

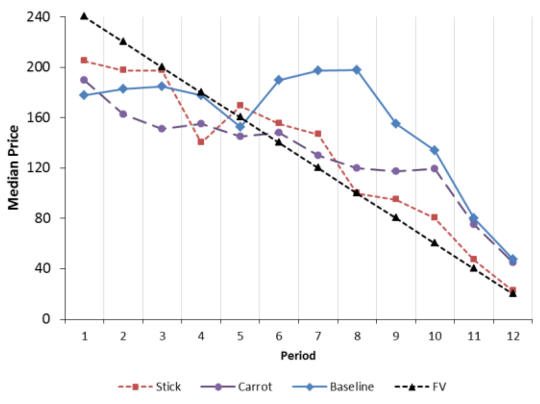
[74] Refer to Appendix B1, Figures B6-B8 to see the evolution of median prices in individual markets of the *Baseline*, *Carrot*, and *Stick* treatments in Round 2.

Figure 3.5: Median prices in Round 2

Median transaction prices in the *Baseline* (solid blue line), *Carrot* (dashed purple line), and *Stick* (dotted red line) treatments during the second round of the market (i.e. with experienced traders) are shown below for the 'low-risk' asset X (panel (a)) and 'high-risk' asset Y (panel (b)), along with the risk-neutral fundamental value process for each asset (black dotted line). For each treatment, the plotted median price in each period is the median of the median transaction prices from all markets belonging to that treatment. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



(a) Asset X, Round 2



(b) Asset Y, Round 2

not statistically significant in any trading period, for both assets (unreported WMW tests). Furthermore, notwithstanding the considerable heterogeneity in price behaviour at the individual market level, it is the *Baseline* treatment where median prices exhibit the most obvious bubble – Asset Y – resulting in median prices for Y that are significantly higher than the *Carrot* treatment at the peak of the bubble in periods 7 and 8[75].

The bubble measures corroborate these observations. Median values of the bubble measures in each treatment in Round 2, along with exact p-values (two-sided) from the associated WMW tests are detailed in Table 3.4[76]. For asset X (Panel A), relative median values on most bubble measures point to greater mispricing and

[75] WMW test p-values are 0.045, and 0.048 in periods 7 and 8 respectively. The *Baseline* treatment also registers a significantly higher median price than *Carrot* in asset Y in period 3 (p-value = 0.049)
[76] Refer to Table B3 and B4 in Appendix B2 for the values of the bubble measures in each individual market of these treatments.

overvaluation in the *Stick* treatment compared to the *Carrot* treatment, whereas the opposite holds true for asset Y (Panel B). However, on most measures, these differences in median values between the two treatments are smaller in Round 2 compared to Round 1. More importantly, in contrast to Round 1, we fail to reject the null of no difference between the *Carrot* and *Stick* treatments (Hypothesis 3) on any of the measures for either asset X or Y.

Comparing the *Baseline* treatment against the tournament treatments (Hypothesis 1), we do not find a significant difference in Round 2 between the *Baseline* treatment and the two tournament treatments in any of the bubble measures for asset X. For the riskier asset Y, all of the median bubble-measure values, with the exception of *Turnover*, indicate *greater* mispricing/bubble behaviour in the *Baseline* treatment than in either the *Carrot* or *Stick* treatments. Of these however, the only difference that is significant at the 5% level is in *Boom Duration*, which is smaller in the *Stick* treatment ($\text{median}_{Baseline}$ = 7 vs. median_{Stick} = 3, p-value = 0.035). *Average Bias* is also higher in *Baseline* than in the *Carrot* treatment, but is only marginally significant ($\text{median}_{Baseline}$ = 24.33 vs. median_{Carrot} = 1.08, p-value = 0.094).

Hence, aggregate-level differences between the incentive schemes seem to largely dissipate when traders are experienced in relation to the experimental design (and their trading cohort). To help understand what drives this, Figures 3.6 and 3.7 compare, by treatment, the evolution of median prices in the two rounds for assets X and Y respectively. While it is difficult to make strong conclusions based on these figures, it is notable that the most striking change between rounds occurs in the *Stick* treatment for asset Y, where median prices are lower in most periods and adhere much more closely to FV in Round 2. Improved adherence to fundamental value in Round 2,

Table 3.4: Summary of bubble measures for assets X and Y in Round 2

This table reports median values of each bubble measure in the *Baseline*, *Carrot*, and *Stick* treatments during Round 2 (i.e. with experienced traders); median absolute deviations are displayed in parentheses. Markets contaminated by subjects who had participated in an earlier session are excluded. Panel A (B) reports bubble measure data relating to Asset X (Y). For definitions of the relevant bubble measures, refer to section 3.4.1.2. The statistical significance of the difference between treatments in each measure is assessed using a two-sided Wilcoxon Mann-Whitney U Test, under the null that values from both treatments come from the same distribution. Exact p-values are reported. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively.

Panel A: Asset X, Round 2

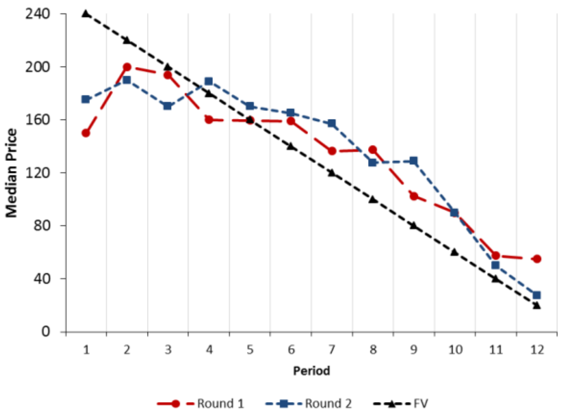
Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	1.15	323.50	10.96	0.82	1.60	39.45	3.00	8.00	3.00
	(0.79)	*(247.00)*	*(13.55)*	*(0.18)*	*(0.69)*	*(32.05)*	*(1.00)*	*(1.00)*	*(1.00)*
Carrot [8]	0.80	302.00	2.94	0.85	2.04	64.30	4.00	5.50	3.00
	(0.35)	*(151.50)*	*(11.25)*	*(0.10)*	*(1.00)*	*(52.34)*	*(2.00)*	*(2.50)*	*(2.00)*
Stick [8]	1.97	689.00	40.17	0.77	1.66	101.15	5.00	8.00	3.00
	(1.59)	*(503.50)*	*(47.38)*	*(0.22)*	*(0.45)*	*(66.21)*	*(1.50)*	*(2.50)*	*(1.50)*
WMW U-Test p-values (two-sided):									
Baseline vs. Carrot	0.463	0.694	0.336	0.779	0.779	0.694	0.411	0.271	0.540
Baseline vs. Stick	0.955	0.867	1.000	0.779	0.536	0.955	0.797	1.000	0.717
Carrot vs. Stick	0.645	0.505	0.279	1.000	0.382	0.645	0.345	0.456	0.917

Panel B: Asset Y, Round 2

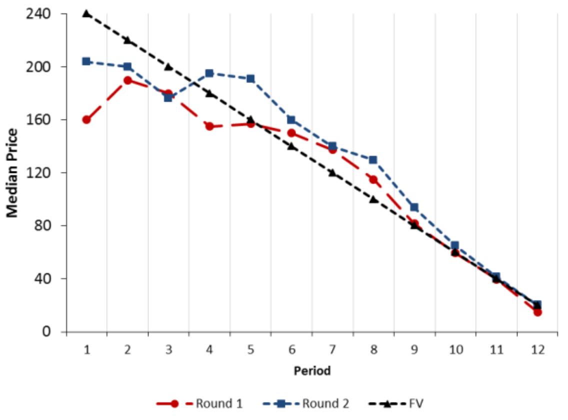
Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	2.16	626.00	24.33	0.52	1.50	127.50	5.00	7.00	3.00
	(0.72)	*(492.50)*	*(68.88)*	*(0.13)*	*(0.47)*	*(64.98)*	*(3.00)*	*(4.00)*	*(2.00)*
Carrot [8]	1.91	391.75	1.08	0.78	1.89	71.52	4.50	5.50	5.00
	(0.83)	*(56.25)*	*(12.58)*	*(0.12)*	*(0.63)*	*(48.75)*	*(1.00)*	*(1.50)*	*(1.50)*
Stick [8]	1.23	481.50	11.25	0.84	1.26	60.95	3.00	3.00	3.50
	(0.86)	*(216.75)*	*(47.93)*	*(0.13)*	*(0.18)*	*(40.14)*	*(1.00)*	*(3.00)*	*(2.00)*
WMW U-Test p-values (two-sided):									
Baseline vs. Carrot	0.867	0.121	0.094*	0.867	1.000	0.779	0.715	0.183	0.184
Baseline vs. Stick	0.613	0.281	0.536	0.613	0.463	0.281	0.282	0.035**	0.378
Carrot vs. Stick	0.574	0.505	0.721	0.574	0.487	0.798	0.209	0.197	0.530

Figure 3.6: Median prices for asset X, Round 1 vs. Round 2

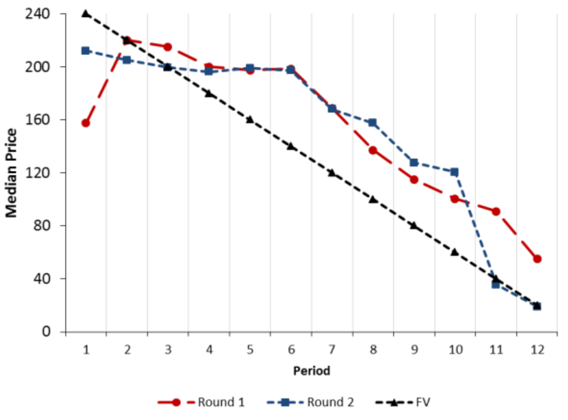
This figure compares the median-price behaviour of the 'low-risk' asset X between the two rounds of the market in the *Baseline* (panel (a)), *Carrot* (panel (b), and *Stick* (panel (c)) treatment. The red dashed line depicts Round 1 prices; the blue dashed line depicts Round 2 prices, while the black dotted line is the fundamental value process. The plotted median price in each period is the median of the median transaction prices from all markets in that treatment. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



(a) Baseline, Asset X



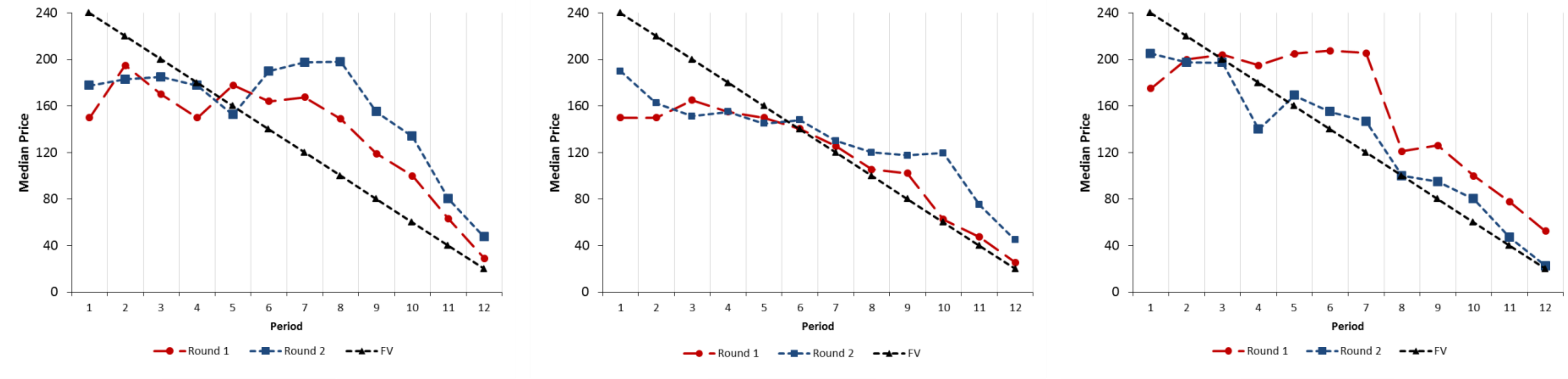
(b) Carrot, Asset X



(c) Stick, Asset X

Figure 3.7: Median prices for asset Y, Round 1 vs. Round 2

This figure compares the median-price behaviour of the 'high-risk' asset Y between the two rounds of the market in the *Baseline* (panel (a)), *Carrot* (panel (b), and *Stick* (panel (c)) treatment. The red dashed line depicts Round 1 prices; the blue dashed line depicts Round 2 prices, while the black dotted line is the fundamental value process. The plotted median price in each period is the median of the median transaction prices from all markets in that treatment. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



(a) Baseline, Asset Y

(b) Carrot, Asset Y

(c) Stick, Asset Y

particularly in the early stages of the market, is also evident for asset X in both the *Carrot* and *Stick* treatments. Median prices for asset X in the latter treatment also appear to adjust more quickly and successfully to fundamental value towards the end of the market. In contrast, median prices for asset Y in the *Baseline* and *Carrot* treatments seem to conform less well to FV with experienced traders, especially in the second half of the market.

We formally assess if price behaviour changes significantly between rounds within each treatment (i.e. Hypothesis 2) by conducting Wilcoxon Signed-rank tests on the various bubble measures; the null hypothesis is that there is no difference in the bubble measure between rounds. The two-sided p-values from these tests are shown in Table 3.5. Beginning with Panel A, which corresponds to the 'safe' asset X, we see that the *Baseline* treatment shows no significant change in behaviour between the two rounds on any of the measures. In contrast, the results point to an 'improvement' in price behaviour in the two tournament treatments, with a number of bubble measures indicative of significantly reduced mispricing/bubble behaviour. This is especially the case in the *Stick* treatment, where *Boom Duration* and *Turnover* are both significantly smaller in Round 2 (p-value = 0.014 and 0.03 respectively), *Haessel-R^2* is significantly larger (p-value = 0.036), while *Normalised Deviation* and *Bust Duration* show improvements that are marginally significant (p-value = 0.093 and 0.078 respectively)[77]. These improvements help drive the trend to insignificance between the *Stick* and *Carrot* treatments in Round 2 for asset X. Even in the *Carrot* treatment, where median prices conform relatively well to FV in the first round of trading and hence the scope for 'improvement' is more limited, we see that *Turnover* and *Normalised Deviation* are

[77] Even though the median value of *Normalised Deviation* for asset X in the *Stick* treatment is higher in Round 2 than in Round 1, the signed-rank test nonetheless reveals a marginally significant *improvement* because the Round 2 value of this measure is actually *lower* than the corresponding Round 1 value in 6 out of 8 markets.

Table 3.5: Comparing bubble measures between rounds

This table reports the results of within-treatment comparisons of the bubble measures between market rounds in the *Baseline*, *Carrot*, and *Stick* treatments. Markets contaminated by subjects who had participated in an earlier session are excluded. The values shown below are p-values from a two-sided Wilcoxon signed-rank test of the null hypothesis that bubble measure values do not differ significantly between rounds 1 and 2. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively

Panel A: Asset X, Round 1 vs. Round 2:

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	1.000	0.499	0.128	0.499	0.176	0.866	0.317	0.230	0.333
Carrot [8]	0.575	0.327	0.575	0.674	0.012**	0.012**	0.160	0.323	0.256
Stick [8]	0.674	0.674	0.779	0.036**	0.030**	0.093*	0.574	0.014**	0.078*

Panel B: Asset Y, Round 1 vs. Round 2:

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	0.735	0.091*	0.128	0.612	0.018**	0.866	0.475	0.932	0.795
Carrot [8]	0.779	0.093*	0.779	0.401	0.012**	0.012**	0.013**	0.477	0.725
Stick [8]	0.208	0.401	0.161	0.124	0.036**	0.124	0.031**	0.011**	0.019**

both significantly lower in Round 2 for asset X (p-value = 0.012 for both measures). The significance of the latter measure appears to be driven by the decline in *Turnover* however, since *Total Dispersion* is not significantly different between the two rounds.

The moderating effect of experience on mispricing/bubbles under tournament incentives is also seen in the riskier asset, Y (Panel B). They confirm what Figure 3.7 strongly suggests – that the price behaviour of asset Y in the *Stick* treatment shows marked improvement in its adherence to FV in Round 2. We see improvement in all of the bubble measures over the two rounds, significantly in the case of *Turnover* (p-value = 0.036) and the bubble-length measures, *Duration* (p-value = 0.031), *Boom Duration* (p-value = 0.011), and *Bust Duration* (p-value = 0.019). The adjustment is particularly large in the case of *Boom Duration*, where the median value falls from 8 to 3 periods. This primarily drives the shift to insignificance (significance) for the *Stick* treatment on this measure with respect to the *Carrot* (*Baseline*) treatment, where the same measure does not change significantly between rounds.

Despite the impression created by Figure 3.7 that prices for asset Y in the *Carrot* treatment are distorted more by experience, we do not find any evidence supporting this in the bubble measures. In fact, the median values of most bubble measures suggest *less* distortion in Round 2, significantly in the case of *Duration* (p-value = 0.013) and *Turnover* (p-value = 0.013), and marginally significantly for *Total Dispersion* (p-value = 0.093). The composite measure of *Turnover* and *Total Dispersion, Normalised Deviation*, is also significantly smaller with experienced traders (p-value = 0.012).

Once again, the effect of experience is least pronounced in the *Baseline* treatment. Indeed, consistent with Figure 3.7, most bubble measures for asset Y in this treatment actually ‘worsen’ in Round 2, although the deterioration is only significant in one measure – *Total Dispersion* – and that too only marginally so (p-value = 0.091). The only measure to achieve significance at the 5% level is *Turnover*, which like the two tournament treatments, is actually significantly reduced by experience (p-value = 0.018).

Thus, the trend towards convergence in price behaviour between the treatments as participants gain experience is primarily driven by the reduction in mispricing/bubbles in the two tournament treatments, especially in the *Stick* treatment. This result contrasts strongly with James and Isaac (2000) and Cheung and Coleman (2014), who find that prices under tournament incentives diverge more from fundamental value (i.e. bubbles become larger) as traders gain experience. The most likely explanation for this discrepancy lies in a crucial difference between these earlier studies and ours – they examine single-asset environments, whereas participants in our markets trade two differentiated risky assets. Hence, our result suggests that the number of assets available for trade also plays an important role in determining how trading experience interacts with tournament incentives in affecting prices.

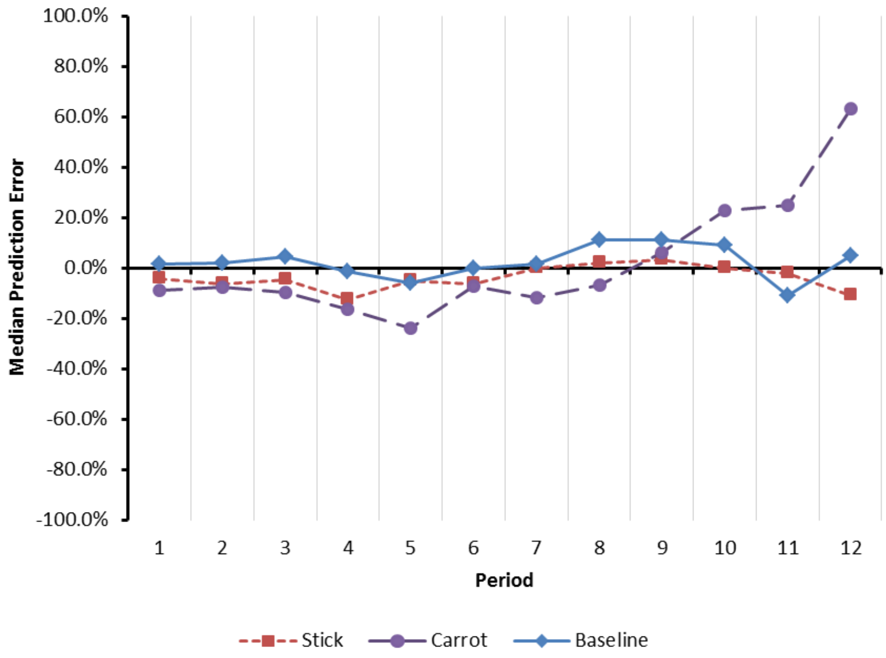
<u>Relative prices</u>

Having examined how experience affects the prices of individual assets, we now turn to its impact on relative prices. Figure 3.8 shows the median *Prediction Error* in each period for each treatment in Round 2[78]. The most interesting aspect of this chart and the most obvious change from Round 1 (cf. Fig. 3.4) is that the *Carrot* treatment exhibits a pronounced upward trajectory in the second half of the market, particularly in the last 3 periods. Participants in the median *Carrot* market were willing to pay a 63% premium to acquire asset Y relative to the price paid for X in the final period.

[78] To see how *Prediction Error* behaves in the individual markets of the *Baseline*, *Carrot*, and *Stick* treatments in Round 2, refer to Figures B6-B8 in Appendix B1.

Figure 3.8: Median values of *Prediction Error*, Round 2

The figure below plots the evolution of the median *Prediction Error* in the *Baseline* (solid blue line), *Carrot* (dashed purple line), and *Stick* (dashed red line) treatment during the second round of the market (i.e. with experienced traders). For each treatment, the plotted value in each period is the median of the *Prediction Errors* from all markets in that treatment. *Prediction Error* is defined as the percentage difference between the relative price of Y (i.e. median price of asset Y divided by median price of asset X) and the risk-neutral benchmark of 1. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



Furthermore, the average of the median *Prediction Errors* in the *Carrot* treatment in the first 6 periods is -12.2% compared to 16.4% in the final six. This is consistent with tournament-induced risk-seeking by traders who are hoping to improve their rankings as the end of the market approaches by betting on receiving the relatively large dividend that asset Y provides. In contrast, relative prices do not seem to behave in an overtly similar manner in the *Baseline* and *Stick* treatments in Figure 3.8, with median *Prediction Errors* for both treatments staying in the region of zero in all periods.

Panel B of Table 3.3, which reports the results of statistical tests on the *Average Prediction Errors* in Round 2, confirms the changing behaviour of relative prices within the *Carrot* treatment – the median *Average Prediction Error* in the second half of the market (15.04%) is significantly larger (one-sided p-value = 0.013) than the corresponding value for the first half of the market (-12.70%). The fact that we also observe a similar effect in the *Baseline* treatment (one-sided p-value = 0.032), albeit one that is smaller – median *Average Prediction Error* in periods 1-6 is -0.12% vs. 5.23% in periods 7-12 – indicates that this not a purely 'tournament' phenomenon. However, the larger magnitude of the difference in the *Carrot* treatment suggests that competitive incentives may play an amplifying role. As for why such an effect appears with experienced traders even though it is missing in Round 1, we posit that a possible explanation could be that traders become more aware of the strategic use of the riskier asset Y as they become more familiar with the trading environment and dividend structures of the two assets[79].

Unlike the two other treatments, the *Stick* treatment does not show a significant difference in relative price behaviour between the two halves of the market. Moreover, when relative prices are examined over the course of the entire market, the whole-of-market *Average Prediction Error* ('Avg PredErr') in all three treatments is not significantly different from zero (i.e. relative prices conform to the theoretical value, 'on average'). For the *Carrot* treatment, this result arises because the statistically significant and large relative discount for asset Y in the first half of the market (one-sample Wilcoxon Signed-rank p-value = 0.0499) is offset by the (marginally) significant and large premium in the second (one-sample Wilcoxon Signed-rank p-value = 0.093). Furthermore, we do not detect a significant difference between any of the treatments on the whole-of-market measure, and we also fail to reject the null of no difference between the treatments in the measures corresponding to each half of the market.

[79] Of course, it is possible that there *is* a similar difference in the behaviour of relative prices in Round 1, but our test lacks the power to detect it.

Hence, like Round 1, relative prices in our treatments do not appear to behave significantly differently from each other.

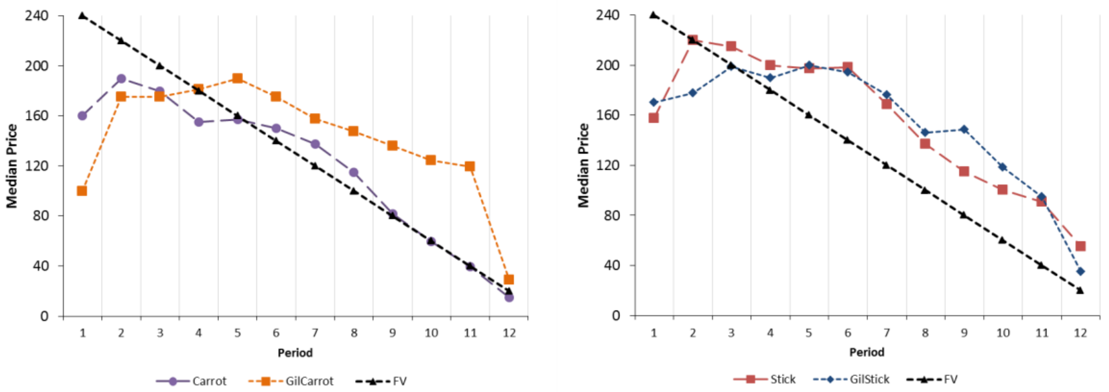
3.4.3 Rank-order contracts

<u>Individual assets</u>

Panels A and B of Figure 3.9 (3.10) compare the price behaviour of asset X (Y) in the *Carrot* and *Stick* treatments against their respective rank-order tournament equivalents, *GilCarrot* and *GilStick* in Round 1. These figures reveal that median prices in the rank-order tournaments and their James & Isaac tournament counterparts are generally closely associated, especially in the case of the penalty-based contracts (*Stick* and *GilStick*). However, the relationship does not appear to be as close between the

Figure 3.9: Median prices in J&I tournament vs. Gilpatric tournament, asset X

The median-price behaviour of the 'low-risk' asset X in Round 1 (i.e. inexperienced traders) is compared between James and Isaac (2000)-based tournament treatments and the corresponding Gilpatric (2009)-based rank-order tournament treatment below. Panel (a) depicts median prices in the *Carrot* treatment (purple dashed line) and the *GilCarrot* treatment (orange dotted line), while Panel (b) shows medians prices in the *Stick* treatment (red dashed line) and the *GilStick* (blue dotted line) treatment. Also shown is the risk-neutral fundamental value process (black dashed line). For each treatment, the plotted median price in each period is the median of the median transaction prices from all markets belonging to that treatment. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.

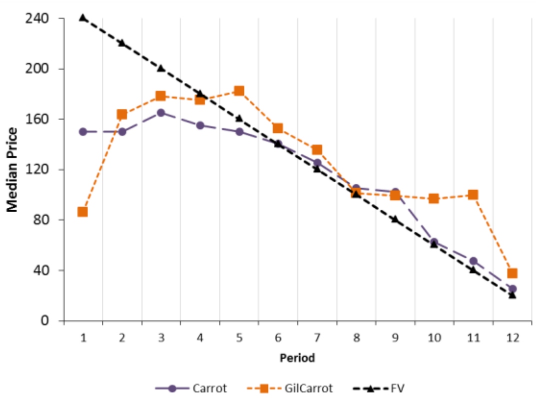


(a) Asset X, Round 1 - Carrot vs GilCarrot

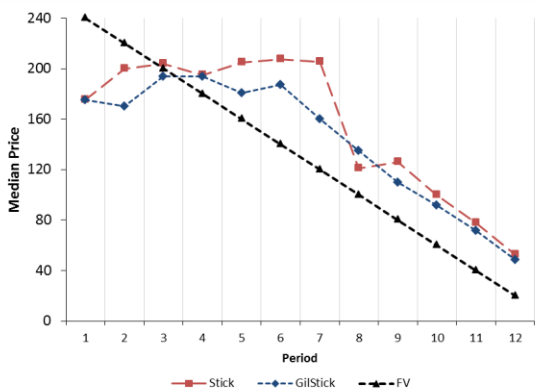
(b) Asset X, Round 1 - Stick vs GilStick reward-only treatments, where *GilCarrot* produces higher median prices than the *Carrot* treatment in most periods. Except for the first period, these differences are generally small or negligible in asset Y but are larger and more persistent in asset X[80]. As a consequence, the noticeable difference in median prices that exists between the *Carrot* and *Stick* treatments in inexperienced markets (see Fig. 3.3) is greatly diminished in the case of *GilCarrot* and *GilStick*, as shown in panels A and B of Figure 3.11[81].

Figure 3.10: Median prices in J&I tournament vs. Gilpatric tournament, asset Y

The median-price behaviour of the 'high-risk' asset Y in Round 1 (i.e. inexperienced traders) is compared between James and Isaac (2000)-based tournament treatments and the corresponding Gilpatric (2009)-based rank-order tournament treatment below. Panel (a) depicts median prices in the *Carrot* treatment (purple dashed line) and the *GilCarrot* treatment (orange dotted line), while Panel (b) shows medians prices in the *Stick* treatment (red dashed line) and the *GilStick* (blue dotted line) treatment. Also shown is the risk-neutral fundamental value process (black dashed line). For each treatment, the plotted median price in each period is the median of the median transaction prices from all markets belonging to that treatment. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



(a) Asset Y, Round 1 - Carrot vs GilCarrot



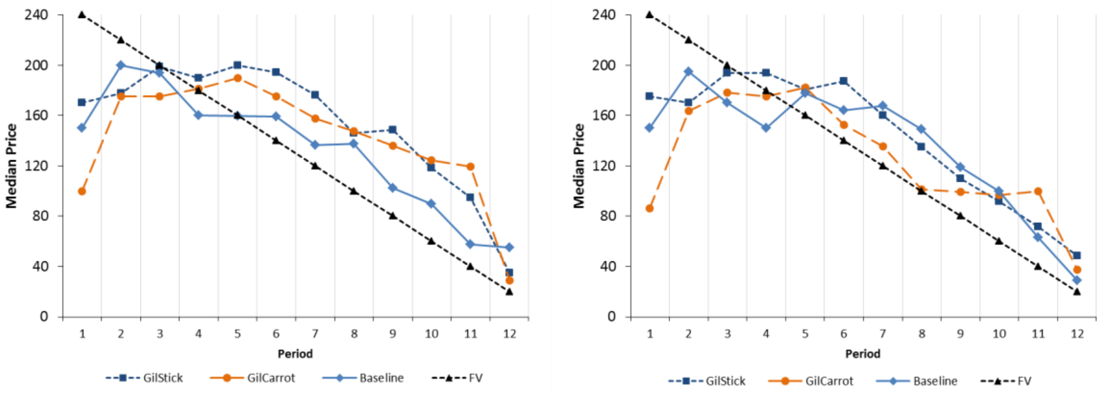
(b) Asset Y, Round 1 - Stick vs GilStick

80 Median prices for asset X in the *GilCarrot* treatment are significantly higher than *Carrot* in periods 9 and 12, but only at the 10% level (WMW p-value = 0.07 and 0.092 respectively). Unreported WMW tests comparing bubble measures between the *Carrot* and *GilCarrot* treatments show that *GilCarrot* has a longer *Boom Duration* that is marginally significant (p-value = 0.064) for asset X; all other measures for asset X return insignificant differences, while there are no significant differences between *Carrot* and *GilCarrot* on any measures for asset Y. Similarly, *Stick* and *GilStick* do not show significant differences on any of the bubble measures for either asset, except the trading activity measure *Turnover*, which is marginally significantly higher in the *GilStick* treatment for asset X (p-value = 0.053).

81 The evolution of Round 1 median prices in each individual market of the *GilCarrot* and *GilStick* treatment is exhibited in Figures B4 and B5 respectively in Appendix B1.

Figure 3.11: Median prices in rank-order tournaments, Round 1

Median transaction prices in the *Baseline* (solid blue line), *GilCarrot* (dashed orange line), and *GilStick* (blue dotted line) treatments during the first round of the market (i.e. with inexperienced traders) are shown below for the 'low-risk' asset X (panel (a)) and 'high-risk' asset Y (panel (b)), along with the risk-neutral fundamental value process (dashed black line). For each treatment, the plotted median price in each period is the median of the median transaction prices from all markets belonging to that treatment. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



(a) Asset X, Round 1

(b) Asset Y, Round 1

The bubble measures from the rank-order tournaments in Round 1 are summarised in Table 3.6, along with exact p-values (two-sided) from the corresponding WMW tests[82]. Consistent with the visual data, and in contrast to the James & Isaac tournament contracts (cf. Table 3.2), we do not detect, for either asset-type, a significant difference between *GilCarrot* and *GilStick* on any of the bubble measures (Hypothesis 3). Relative to the James & Isaac tournaments, Table 3.6 is also somewhat more supportive of the argument that tournament contracts distort prices more than normal incentives (Hypothesis 1) – *Haessel-*R^2 is higher, and *Turnover* and *Normalised Deviation* are both significantly lower in the *Baseline* treatment than in the *GilCarrot* treatment, albeit only marginally (p-values of 0.051, 0.073, and 0.073 respectively), and only for asset X. Given that we fail to find a significant difference in *Total Dispersion* between *Baseline* and *GilCarrot*, it is also likely that the difference in *Normalised*

[82] Refer to Table B1 and B2 in Appendix B2 for the values of the bubble measures in the individual markets of these treatments. Round 2 equivalents can be found in Tables B3 and B4.

Table 3.6: Summary of bubble measures in Round 1 using rank-order tournaments

This table reports median values of each bubble measure in the *Baseline*, *GilCarrot*, and *GilStick* treatments during Round 1; median absolute deviations are displayed in parentheses. Markets contaminated by subjects who had participated in an earlier session are excluded. Panel A (B) reports bubble measure data relating to Asset X (Y). For definitions of the relevant bubble measures, refer to section 3.4.1.2. The statistical significance of the difference between treatments in each measure is assessed using a two-sided Wilcoxon Mann-Whitney U Test, under the null that values from both treatments come from the same distribution. Exact p-values are reported. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively.

Panel A: Asset X, Round 1:

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	3.31	298.00	1.65	0.78	2.45	83.43	5.00	5.00	4.00
	(1.53)	*(181.50)*	*(15.73)*	*(0.16)*	*(0.72)*	*(45.08)*	*(2.00)*	*(2.00)*	*(1.00)*
GilCarrot [6]	3.05	758.50	23.42	0.46	3.91	262.55	5.00	8.00	3.50
	(1.71)	*(202.25)*	*(19.52)*	*(0.11)*	*(0.85)*	*(62.06)*	*(2.50)*	*(1.00)*	*(1.00)*
GilStick [6]	1.99	614.50	28.85	0.63	2.79	156.44	6.00	8.00	2.50
	(0.58)	*(222.00)*	*(34.25)*	*(0.08)*	*(0.29)*	*(64.60)*	*(1.00)*	*(2.00)*	*(1.50)*
WMW U-Test p-values (two-sided):									
Baseline vs. GilCarrot	0.945	0.366	0.836	0.051*	0.073*	0.073*	0.736	0.178	0.950
Baseline vs. GilStick	0.731	0.445	0.181	0.234	0.509	0.234	0.457	0.229	0.668
GilCarrot vs. GilStick	0.589	0.937	0.589	0.240	0.167	0.485	0.558	0.864	0.381

Panel B: Asset Y, Round 1

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	1.63	530.50	15.38	0.77	2.03	99.06	4.00	7.00	4.00
	(1.01)	*(150.50)*	*(25.33)*	*(0.05)*	*(0.35)*	*(54.11)*	*(1.00)*	*(1.00)*	*(2.00)*
GilCarrot [6]	3.44	676.75	-5.05	0.35	2.79	169.82	4.00	5.00	3.50
	(1.99)	*(217.50)*	*(31.55)*	*(0.25)*	*(0.44)*	*(46.82)*	*(2.00)*	*(1.00)*	*(1.50)*
GilStick [6]	2.52	518.50	12.20	0.61	2.62	117.21	5.50	6.00	2.50
	(1.01)	*(103.25)*	*(32.79)*	*(0.10)*	*(0.93)*	*(53.74)*	*(1.00)*	*(3.50)*	*(1.00)*
WMW U-Test p-values (two-sided):									
Baseline vs. GilCarrot	0.731	0.836	0.366	0.731	0.445	0.534	0.871	0.508	0.530
Baseline vs. GilStick	0.731	1.000	0.836	0.945	0.731	0.731	0.508	0.458	0.751
GilCarrot vs. GilStick	0.818	0.699	0.699	0.699	0.937	0.937	0.675	0.894	0.374

Deviation simply reflects the same effect as *Turnover*.

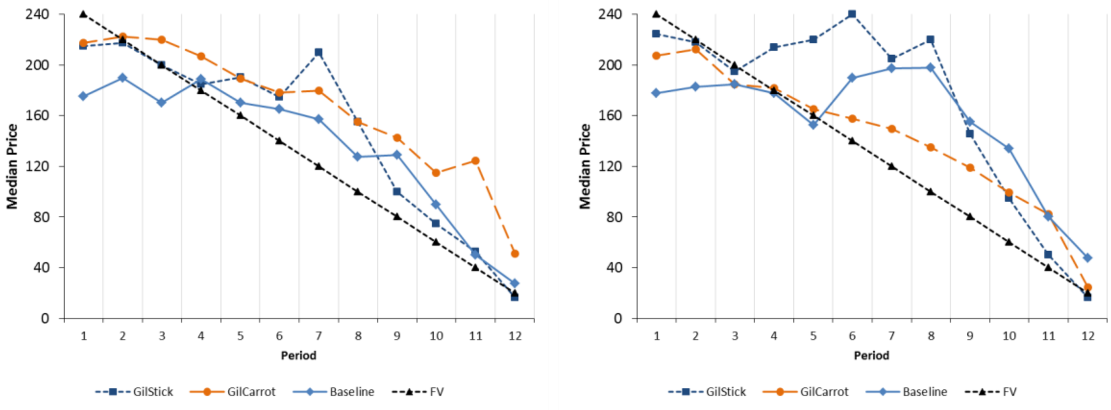
In markets with experienced traders (Round 2), we see in Figure 3.12 that differences in median price between *GilCarrot* and *GilStick* appear to be greater for asset Y than X[83]. Indeed, median prices in the *GilStick* treatment exhibit a sizeable bubble in asset Y. However as Table 3.7 reveals, like Round 1, none of the bubble measures for asset Y in Round 2 differ significantly between the two rank-order tournament treatments. Also, with the exception of *Amplitude*, which is marginally significantly greater in the *GilCarrot* treatment than in *GilStick* (p-value = 0.065), all other bubble measures for asset X return insignificant differences. In regards to the *Baseline* treatment, we fail to find a significant difference on any of the bubble measures with respect to the *GilCarrot* or *GilStick* treatments for either asset.

As with the James & Isaac tournaments, we find that greater trading experience is associated with smaller bubbles under rank-order tournament conditions, especially in the *GilStick* treatment (Hypothesis 2). The results (two-sided p-values) of Signed-rank tests carried out on the bubble measures of each rank-order tournament treatment are shown in Table 3.8. The table reveals that for asset X (Panel A) in the *GilStick* treatment, *Amplitude* and *Total Dispersion* are significantly smaller in Round 2 (p-value = 0.028 and 0.046 respectively), while *Normalised Deviation*, *Duration*, and *Boom Duration* are also smaller but only at the 10% level (p-value = 0.075, 0.091, and 0.058 respectively). For asset Y (Panel B) in the *GilStick* treatment, *Turnover* is significantly reduced by experience (p-value = 0.046), while smaller *Amplitudes* and *Durations* in Round 2 are marginally significant (p-value = 0.075 and 0.058 respectively). On the

[83] Median prices from Round 2 in each individual market of the *GilCarrot* and *GilStick* treatment are charted in Figures B9 and B10 respectively in Appendix B1.

Figure 3.12: Median prices in rank-order tournaments, Round 2

Median transaction prices in the *Baseline* (solid blue line), *GilCarrot* (dashed orange line), and *GilStick* (blue dotted line) treatments during the second round of the market (i.e. with experienced traders) are shown below for the 'low-risk' asset X (panel (a)) and 'high-risk' asset Y (panel (b)), along with the risk-neutral fundamental value process (dashed black line). For each treatment, the plotted median price in each period is the median of the median transaction prices from all markets belonging to that treatment. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



(a) Asset X, Round 2

(b) Asset Y, Round 2

other hand, the evidence that experience reduces mispricing/bubbles is weaker in the *GilCarrot* treatment, where the only measure that changes significantly between rounds is the trading activity measure, *Turnover*, which is smaller in Round 2 for both assets (p-value = 0.028 in both cases). Importantly however, none of the bubble measures in either rank-order tournament treatment indicate significantly *more* mispricing as participants gain experience.

<u>Relative prices</u>

Relative prices in the Gilpatric tournaments in Round 1 behave in a qualitatively similar manner to the James and Isaac tournaments. That is, on 'average', relative prices conform to the theoretical value and do not differ significantly between the two treatments, *GilCarrot* and *GilStick*. This can be seen in Figure 3.13(a), where the

Table 3.7: Summary of bubble measures in Round 2 using rank-order tournaments

This table reports median values of each bubble measure in the *Baseline*, *GilCarrot*, and *GilStick* treatments during Round 2; median absolute deviations are displayed in parentheses. Markets contaminated by subjects who had participated in an earlier session are excluded. Panel A (B) reports bubble measure data relating to Asset X (Y). For definitions of the relevant bubble measures, see section 3.4.1.2. The statistical significance of the difference between treatments in each measure is assessed using a two-sided Wilcoxon Mann-Whitney U Test, under the null that values from both treatments come from the same distribution. Exact p-values are reported. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively.

Panel A: Asset X, Round 2:

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	1.15	323.50	10.96	0.82	1.60	39.45	3.00	8.00	3.00
	(0.79)	*(247.00)*	*(13.55)*	*(0.18)*	*(0.69)*	*(32.05)*	*(1.00)*	*(1.00)*	*(1.00)*
GilCarrot [6]	2.33	457.75	36.81	0.87	2.55	115.18	5.00	7.50	2.00
	(1.12)	*(119.75)*	*(22.92)*	*(0.02)*	*(0.48)*	*(45.08)*	*(0.50)*	*(2.50)*	*(1.00)*
GilStick [6]	0.92	446.25	25.54	0.85	2.59	81.13	3.50	6.00	3.00
	(0.49)	*(234.25)*	*(30.01)*	*(0.12)*	*(0.86)*	*(59.78)*	*(1.00)*	*(2.00)*	*(0.50)*
WMW U-Test p-values (two-sided):									
Baseline vs. Gil-Carrot	0.366	0.628	0.366	0.445	0.421	0.628	0.864	0.810	0.804
Baseline vs. Gil-Stick	0.731	0.836	0.945	0.731	0.219	0.945	0.493	0.650	0.935
Gil-Carrot vs. Gil-Stick	0.065*	0.485	0.485	0.818	0.784	0.394	0.210	0.303	0.498

Panel B: Asset Y, Round 2

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	2.16	626.00	24.33	0.52	1.50	127.50	5.00	7.00	3.00
	(0.72)	*(492.50)*	*(68.88)*	*(0.13)*	*(0.47)*	*(64.98)*	*(3.00)*	*(4.00)*	*(2.00)*
GilCarrot [6]	1.99	462.50	23.10	0.84	1.74	76.56	6.00	7.00	4.00
	(0.86)	*(143.25)*	*(18.19)*	*(0.02)*	*(0.26)*	*(28.16)*	*(2.00)*	*(2.00)*	*(1.50)*
GilStick [6]	1.75	694.25	52.35	0.57	2.25	96.66	2.50	7.00	1.50
	(0.60)	*(228.25)*	*(18.20)*	*(0.23)*	*(0.94)*	*(39.36)*	*(1.00)*	*(4.00)*	*(0.50)*
WMW U-Test p-values (two-sided):									
Baseline vs. GilCarrot	0.945	0.295	0.628	0.181	0.628	0.628	0.386	0.833	0.705
Baseline vs. GilStick	0.628	0.534	0.731	1.000	0.313	0.836	0.422	0.756	0.755
GilCarrot vs. GilStick	0.818	0.589	0.310	0.240	0.589	0.589	0.106	0.985	0.284

Table 3.8: Comparing rank-order tournament bubble measures between rounds

This table reports the results of within-treatment comparisons of bubble measures between market rounds in the *GilCarrot* and *GilStick* treatments. Markets contaminated by subjects who had participated in an earlier session are excluded. Panel A (B) reports for asset X (Y). The values shown below are p-values from a two-sided Wilcoxon signed-rank test of the null hypothesis that bubble measure values do not differ significantly between rounds 1 and 2. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively

Panel A: Asset X, Round 1 vs. Round 2:

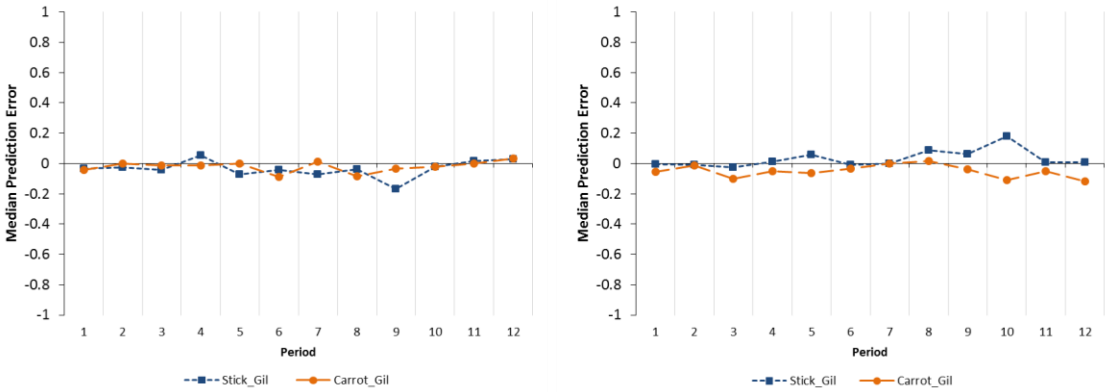
Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
GilCarrot [6]	0.917	0.917	0.116	0.173	0.028**	0.463	0.674	0.916	0.190
GilStick [6]	0.028**	0.046**	0.249	0.173	0.463	0.075*	0.091*	0.058*	0.593

Panel B: Asset Y, Round 1 vs. Round 2:

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
GilCarrot [6]	0.917	0.917	0.463	0.116	0.028**	0.173	0.461	0.597	0.665
GilStick [6]	0.075*	0.600	0.173	0.917	0.046**	0.249	0.058*	0.525	0.597

Figure 3.13: Median values of *Prediction Error*, rank-order tournaments

The evolution of the median *Prediction Error* in the *GilCarrot* (dashed orange line), and *GilStick* (dashed blue line) treatment is shown below for Round 1 of the market in panel (a) and Round 2 in panel (b). For each treatment, the plotted value in each period is the median of the *Prediction Errors* from all markets in that treatment. *Prediction Error* is defined as the percentage difference between the relative price of Y (i.e. median price of asset Y divided by median price of asset X) and the risk-neutral benchmark of 1. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



(a) Round 1 (b) Round 2

median *Prediction Error* in both treatments is close to zero in most periods[84]. Consistent with this, in both treatments we fail to reject the null hypothesis that the median *Average Prediction Error*, summarised in Panel A of Table 3.9, is equal to zero. This is also the case for the two half-market measures. Moreover, within each rank-order tournament treatment, we do not detect a significant difference in relative-price behaviour between the first and second halves of the market. Furthermore, the differences between *GilCarrot* and *GilStick*, assessed using the WMW test, are not statistically significant on any of the measures, nor do the rank-order tournaments differ significantly from the *Baseline* condition.

The results in Round 2 are similar to Round 1. In both rank-order tournaments, median *Prediction Errors* stay in the region of zero, as shown in Figure 3.13(b), while

[84] *Prediction Error* in the individual markets of the *GilCarrot* and *GilStick* treatments in Round 1 is charted in Figures B4 and B5 respectively in Appendix B1. Round 2 analogues can be found in Figures B9 and B10.

Table 3.9: Average Prediction Errors – rank-order tournaments

Median values of the *Average Prediction Error* in the *Baseline, Carrot, and Stick* treatments in Round 1 and 2 are shown below in Panels A and B respectively, with the associated median absolute deviations in parentheses. Markets contaminated by subjects who had participated in an earlier session are excluded. *Average Prediction Error* is calculated using all periods in a market, the first 6 periods, and the final 6 periods in *Avg PredErr*, *AvgPredErr_p1to6*, and *AvgPredErr_p7to12* respectively. The statistical significance of the individual measures is assessed using a (two-sided) one-sample Wilcoxon Signed-rank test, under the null that the median is equal to zero. The statistical significance of the difference between treatments is assessed using the Wilcoxon Mann-Whitney U Test under the null that values from both treatments come from the same distribution. The statistical significance of the difference between *AvgPredErr_p1to6* and *AvgPredErr_p7to12* within each treatment is assessed using a (paired-sample) Wilcoxon Signed-rank test, against the one-sided alternative hypothesis that *AvgPredErr_p7to12 > AvgPredErr_p1to6*. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively.

Panel A: Round 1

Treatment [N]	*Avg PredErr (%)*	*AvgPredErr_ p1to6 (%)*	*AvgPredErr_ p7to12 (%)*	*Signed-rank p-value (1-sided) 1-6 vs. 7-12*
Baseline [7]	3.78	-0.32	7.88	0.064*
	(9.46)	*(3.36)*	*(9.26)*	
GilCarrot [6]	-3.22	-3.21	-3.09	0.377
	(8.79)	*(2.78)*	*(15.48)*	
GilStick [6]	2.05	-0.92	6.00	0.377
	(7.17)	*(4.64)*	*(9.49)*	
WMW U-Test p-values (2-sided):				
Baseline vs. Gil-Carrot	0.295	0.836	0.295	
Baseline vs. Gil-Stick	0.295	0.945	0.534	
Gil-Carrot vs. Gil-Stick	0.937	0.699	0.937	

Panel B: Round 2

Treatment [N]	*Avg PredErr (%)*	*AvgPredErr_ p1to6 (%)*	*AvgPredErr_ p7to12 (%)*	*Signed-rank p-value (1-sided) 1-6 vs. 7-12*
Baseline [7]	2.72	-0.12	5.23	0.032**
	(11.33)	*(2.27)*	*(21.08)*	
GilCarrot [6]	-4.51	-4.14*	-3.63	0.377
	(4.95)	*(4.27)*	*(6.68)*	
GilStick [6]	3.80	3.46	3.85	0.058*
	(9.64)	*(4.29)*	*(8.95)*	
WMW U-Test p-values (2-sided):				
Baseline vs. Gil-Carrot	0.295	0.181	0.366	
Baseline vs. Gil-Stick	0.836	0.836	1.000	
Gil-Carrot vs. Gil-Stick	0.180	0.093*	0.180	

Average Prediction Errors (for the whole market and in each half) are generally not significantly different from zero (see Panel B of Table 3.9). The only exception to this is the *GilCarrot* treatment, in which according to the *Average Prediction Error* measure, asset Y sells at a statistically significant discount to asset X of around 4% in the first half of the median market, but only at the 10% level (one-sample Signed-rank test p-value = 0.075). In addition, while *Average Prediction Error* is higher in the second half of the market in the *GilStick* treatment compared to the first half, the statistical significance of the difference is only marginal (one-sided p-value = 0.058) and the economic significance even less so (3.46% in the first half vs. 3.85% in the second). Notably, relative prices in the *GilCarrot* treatment display none of the signs of heightened speculation in asset Y that is evident in the *Carrot* treatment in Round 2 (cf. Fig. 3.8). Furthermore, comparing the *GilCarrot* and *GilStick* treatments, we see in Panel B of Table 3.9 that differences in *Average Prediction Error* between the two treatments generally fail to attain statistical significance. The first half of the market again presents the exception; there is some evidence that the relative price of asset Y is higher in *GilStick* than in *GilCarrot* during this period, although significance here is only at the 10% level (p-value = 0.093) and the difference is economically quite small (-4.14% in *GilCarrot* vs. 3.46% in *GilStick).*

3.5. Conclusion

Tournament incentives have been accused in the experimental literature of distorting the efficient functioning of markets by exacerbating asset price bubbles. While this narrative tallies with mooted concerns regarding the link between market instability and the proliferation of convex incentive structures in the financial industry, the real-world relevance of these experimental results is limited by their examination of single-asset markets that preclude the ability to trade in securities with dissimilar risk characteristics, an option that is available to investors in real markets. Moreover, the reward-centric focus of existing studies means the role of penalties for poor performance in tournament contracts have been largely ignored, despite the fact that they may help to moderate risk-taking behaviour. We address these gaps in the literature by examining how rewards ('carrots') and penalties ('sticks') embedded in tournament contracts affect price behaviour in experimental asset markets where participants can trade in two differentiated assets. Each asset has the same risk-neutral fundamental value, but one asset is intrinsically riskier by virtue of a lottery-like dividend structure that generates potentially higher payoffs, thus allowing traders to naturally vary their risk by shifting in/out of the asset.

Our results challenge the main conclusions of the existing literature. In two-asset experimental markets, we do not find any compelling evidence to suggest that asset price bubbles are larger under tournament incentives than normal, absolute-performance based incentives. Moreover, unlike earlier studies, bubbles under tournament incentives in our markets do dissipate with experienced traders. Hence, the results of earlier studies appear to be driven by the single-asset nature of their markets.

Furthermore, penalties embedded into tournament contracts that reward traders for 'beating the market' reduce the volume of trading activity in inexperienced markets compared to reward-only contracts. However, the trade that does occur happens at higher prices, and periods of overvaluation last longer, especially in the case of the riskier asset. Thus, in markets with inexperienced traders, 'sticks' or penalties for underperformance are associated with *greater* mispricing, not less. While this may seem a counterintuitive and surprising result, it is consistent with price behaviour under tournament incentives when traders are prone to herd; the inclusion of a penalty for underperformance makes traders more likely to herd as a way to minimise the risk of being an underperformer, thus perversely exacerbating and prolonging mispricing. However, this effect does not appear to survive in our markets when participants are once-experienced. Moreover, we do not observe a significant difference in price behaviour between carrot-only and carrot-and-stick contracts when we implement a rank-order tournament, either with inexperienced or experienced traders.

In light of the on-going debate surrounding compensation practices in the financial industry, our results are particularly relevant to policymakers and regulators. Our findings suggest that, at the aggregate-level, tournament incentives may not be as disruptive a force as earlier studies indicate. Furthermore, regulatory initiatives such as placing caps on finance professionals' bonuses may be misplaced – shifting the balance between carrots and sticks further towards the stick-end may increase the incentive to herd, thereby having the perverse effect of fuelling the instability that such actions seek to prevent.

Although we make important contributions towards better understanding the aggregate effects of tournament incentives, the laboratory environment in which we conduct our study is obviously considerably less complex than real markets and the real world. As such, our study is subject to the limitations of experimentation as a methodology. Foremost amongst these is the 'penalty' in our tournament contracts – a zero payment, which some may reasonably protest is not a 'real' penalty since it does not impose actual losses on traders. Whilst true, a zero payment represents the most an experimenter can penalise experiment participants, given that ethical considerations preclude experimenters from enforcing financial losses/liabilities on subjects. Even if it were possible, potential selection biases make it undesirable, since only certain types of subjects may volunteer for the experiment. Moreover, the gap between our non-penalty payment and the zero payment still represents a sizeable disincentive for university student participants, given the time commitment made. (2.5 hours). Hence, while the impact of 'sticks' may be diminished in an experimental setting compared to the real world – where professionals face the more severe risks of job termination and/or reputational damage – this suggests the differences that we do observe with inexperienced participants are likely to be *underestimated*.

The fact that we consider markets with only two risky assets, whereas real-world markets are characterised by a myriad of potential investments, may be considered another limitation of our study. Furthermore, unlike our treatment groups, market participants in the real world do not all trade under the same incentives. These represent simplifications of the real world required to build a workable experimental design and isolate the effects of different incentive schemes. Thus the extents to which our results can be generalised when these restrictions are relaxed is an open question, and as such, represent potential avenues for future research.

"There is nothing so disturbing to one's well-being

and judgment as to see a friend get rich"

Charles Kindleberger

CHAPTER 4: Peer Effects in Experimental Markets

4.1 Introduction

The allegation that tournament incentives foment destabilising asset price bubbles by virtue of the convex payoffs they generate (Rajan 2006; Dass, Massa and Patgiri 2008) has found its strongest support in the experimental laboratory. Compared to compensation schemes based on absolute performance, larger price bubbles – periods of sustained overvaluation – have been observed in experimental asset markets when traders are subject to tournament compensation, where pay depends on relative performance (James and Isaac 2000; Cheung and Coleman 2014). However, a recent study by Schoenberg and Haruvy (2012) shows that, even in the absence of relative-performance based pay, simply providing periodic feedback to traders about their relative performance can produce price-patterns that are similar to those observed in tournaments. This result, which tallies with theoretical observations linking relative wealth concerns to financial bubbles (DeMarzo, Kaniel and Kremer 2008), raises an intriguing question: is the price behaviour seen in tournament studies actually driven by traders' innate desire to outperform others, rather than the extrinsic monetary incentives associated with tournaments?

The answer to this question has particular import to the on-going debate surrounding compensation practices in the financial industry, as it potentially sheds light on the efficacy of imposing restrictions on pay versus the information environment as a means to achieve market stability (Dijk, Holmen and Kirchler 2014). In this study, we seek an answer to the question posed, by determining the separate impacts of relative performance information and tournament compensation on price behaviour in experimental asset markets. While a number of other experimental studies have sought to unravel these effects in relation to the performance and risk-taking behaviour of *individuals* (Hannan, Krishnan and Newman 2008; Dijk et al. 2014), to our knowledge, our study is the first to investigate the issue at the aggregate or price-level.

Following Dijk et al. (2014), we implemented a between-subjects experimental design comprising three treatments: *Baseline, RelInfo,* and *Tournament.* In both the *Baseline* and *RelInfo* treatments participants were compensated according to their absolute-performance, but while traders in the *RelInfo* treatment were given periodic feedback about their relative performance, those in the *Baseline* treatment were not. On the other hand, participants in the *Tournament* treatment were compensated according to a tournament compensation scheme, while also receiving periodic feedback on relative performance. We followed the canonical tournament studies of James and Isaac (2000) and Isaac and James (2003) in the *Tournament* treatment by remunerating traders on the basis of their performance relative to the 'average' trader. Additionally, an alternative tournament treatment, *GilTournament,* was also implemented which ran a rank-order tournament, where final rank determined the payoff. Participants in all treatments traded in an experimental asset market based on the Smith et al. (1988) continuous double-auction bubble-market design, featuring two risky-assets – a low-risk asset called X, which paid a modestly sized dividend in each trading period, and a high-risk asset called

Y that paid a lottery-like dividend – which allowed participants to naturally vary the risk of their holdings.

Overall, our results suggest that relative performance information and competitive monetary incentives have opposing incremental effects on asset prices. While we do not detect any significant differences in price behaviour between the *Baseline*, *RelInfo*, and *Tournament* treatments with inexperienced traders, differences do emerge once traders gain common group and design-experience. Providing relative performance information to experienced traders who are compensated for their absolute performance serves to actually *reduce* mispricing – price bubbles in the *RelInfo* treatment are smaller than in the *Baseline* treatment, especially for the high-risk asset Y. In contrast, tournament incentives, when introduced into an environment where relative performance feedback is already provided, has the opposite effect, *increasing* mispricing relative to the *RelInfo* treatment. This effect is relatively weak for the *Tournament* treatment, but is much stronger for the *GilTournament* treatment, where price bubbles are larger than the *RelInfo* treatment with both inexperienced and experienced traders.

The remainder of this chapter is structured as follows. In section 4.2, we review the related literature and develop our hypotheses. Section 4.3 details the experimental design, while section 4.4 describes the results. Finally, we present our conclusions in section 4.5.

4.2. Literature Review

Although the standard assumption in neoclassical economics holds that your utility depends only on your *own* wealth, economists going back to Adam Smith (1759, in Tran and Zeckhauser 2012) and Veblen (1899) have long recognised that individuals' happiness may also be influenced by social comparisons[85]. One's relative position or performance may, amongst other things, confer status, affect motivation and self-esteem, and provide signals to aid learning about unobservables such as ability. Studies examining subjective well-being underscore the importance of relative wealth concerns, largely confirming the intuition that individuals' levels of self-reported happiness reflect a desire to "keep up with the Joneses" (see Clark, Frijters and Shields (2008) and Frey and Stutzer (2002) for reviews of the literature; also Clark and Oswald 1996; Luttmer 2005; Boyce, Brown and Moore 2010)[86].

The effect these desires have on individuals' behaviour has given rise to the notion of 'conspicuous consumption' (Veblen 1899) – the consumption or accumulation of luxury and positional goods as a means to advertise social status – and the 'relative income hypothesis' of Duesenberry (1949), which contends that consumption and savings patterns are determined by an individual's relative position in the income distribution. More recently, Dupor and Liu (2003) show that consumption externalities produced by social comparison may lead to overconsumption in equilibrium if individuals are prone to 'jealousy', a term coined to describe the disutility experienced by individuals when the aggregate consumption of others rises. On the subject of risk-

[85] The psychology literature has also examined the importance of relative concerns, beginning with the theory of social comparison developed by Festinger (1954). See Hannan et al. (2008) and Buunk and Gibbons (2007) for an overview of this literature.

[86] In contrast, results from the experimental laboratory often indicate the presence of pro-social or other-regarding preferences (e.g. see Fehr and Schmidt 1999; Charness and Grosskopf 2001; Charness and Rabin 2002). However, given the experimental environment may make such behaviours more likely to be observed than in the real world (e.g. via perceived pressure to behave in a certain or 'correct' manner), the generalisability of these results is uncertain (see Levitt and List (2007) for a discussion of these issues).

taking, theoretical contributions include Robson (1992), who shows that status-seeking preferences can induce convex utility (i.e. risk-seeking) over ranges of wealth where opportunity for rapid upward mobility in status is possible, thus formalising the concave-convex-concave utility forwarded by Friedman and Savage (1948) as an explanation for why individuals simultaneously purchase both insurance and lottery tickets. Becker, Murphy and Werning (2005) demonstrate that when people can participate in a market for status, status-seeking preferences may induce demand for risk-taking if higher status is associated with a higher marginal utility of income. In addition, Roussanov (2010) shows theoretically that the desire for high status may lead investors to underdiversify. In contrast, Bakshi and Chen (1996) argue that concerns for status may prompt investors to be more cautious in their consumption patterns and risk-taking,

Empirical studies that examine the effect of social comparison on risk-taking provide mixed support for the connection to risk-seeking behaviour. A number have examined the issue in the context of reference-dependent decision-making models such as Prospect Theory (Kahneman and Tversky 1979). When peer income acts as the (social) reference point against which (social) gains and losses are measured, these studies find little support for the 'reflection effect', which refers to the theory's prediction that people are risk averse in the domain of gains and risk-seeking in the domain of losses. Using field data, Vendrik and Woltjer (2007) find that self-reported happiness is concave for both positive and negative relative incomes in their sample, implying risk aversion in both domains. Consistent with this, Linde and Sonnemans (2012) observe that individuals in their experiment tend to be risk averse in a lottery-choice task in both social gain and loss domains (especially in the latter) when informed of a referent participant's payoff. Another cornerstone of Prospect Theory, loss aversion, which contends that the pain of a loss is greater than the happiness derived from an equivalent gain, fares better in the presence of a social reference point (Vendrik and Woltjer 2007; Schwerter 2013), although some evidence to the contrary is reported by Bault, Coricelli and Rustichini (2008), who observe that participants exhibit stronger emotional responses for social gains than social losses in a lottery-choice task.

Dijk et al. (2014) present evidence consistent with a competitive desire for high rank in an experimental setting. They find that ranking information influences the composition of investors' portfolios even when that information has no relevance to payoffs. Specifically, underperforming individuals in a portfolio-choice task exhibit a strong inclination to hold positively skewed assets that entail a small probability of a large payment, while outperformers prefer to hold negatively skewed assets that pay a modest amount on most occasions. Interestingly, Baghestanian, Gortner and Van der Weele (2015) present evidence from an experimental asset market that the type of relative performance information affects risk-taking. Highlighting the worst performer leads traders in their market to increase diversification and reduce aggregate risk-taking, while highlighting the best performer has the opposite effect.

More generally, the empirical literature has shown that a number of other decisions are subject to the influence of peers, whether due to concerns about relative wealth or other reasons. So-called 'peer effects' have been observed in stock market participation (Hong, Kubik and Stein 2004; Brown, Ivković, Smith and Weisbenner 2008; Kaustia and Knüpfer 2012), stock selection (Hong, Kubik and Stein 2005; Ivković and Weisbenner 2007; Bursztyn, Ederer, Ferman and Yuchtman 2014), and executive compensation and corporate investment (Shue 2013). Choice behaviour aside, peer effects have also been observed in fundamental attitudes such as risk-aversion (Ahern, Duchin and Shumway 2014).

The influence of social comparison on effort and productivity has also attracted considerable interest in the literature. Lab and field experiments largely point to the existence of a positive peer effect under flat or absolute performance based compensation schemes; feedback on relative performance, received either explicitly or by observing peers, improves individuals' performance under such schemes (Falk and Ichino 2006; Mas and Moretti 2009; Kuhnen and Tymula 2011; Hannan et al. 2008; Blanes i Vidal and Nossol 2011; Azmat and Iriberri 2010; Tran and Zeckhauser 2012). In contrast, introducing relative performance feedback under tournament compensation does not improve performance (Eriksson, Poulsen and Villeval 2009; Hannan et al. 2008)[87] and may in fact cause performance to deteriorate if the feedback is precise (Hannan et al. 2008)[88].

The effect on performance seen under non-competitive remuneration schemes is particularly interesting because it again hints at fundamental competitive desires for status or rank as potential drivers of peer effects. The mechanisms that give rise to peer effects are the subject of a number of recent neuroeconomic (Bault, Joffily, Rustichini and Coricelli 2011; Tomlin, Nedic, Prentice, Holmes and Cohen 2013; Frydman 2015), laboratory (Lahno and Serra-Garcia 2015), and field (Bursztyn et al. 2014) experiments. These studies confirm the importance of concerns for relative payoff or "keeping up with the Joneses" as a channel through which peer effects operate, but also identify social learning (learning from others) and preference for conformity as other drivers[89].

---

[87] Eriksson et al. (2009) find that feedback does not improve performance under piece-rates or tournament pay schemes in their experiment, although in the case of piece-rates, this may be due to participants exerting maximum effort without the feedback.

[88] Precision here refers to how detailed the feedback is about relative position. Hannan et al. (2008) implement two types of feedback: "coarse" and "fine". Under coarse feedback, participants are only told whether their performance is better or worse than the median. The "fine" feedback, which is more precise, informs participants about the decile their performance falls into.

[89] Learning from observing the behaviour of others is a feature of the sizeable literature on informational cascades and herding in finance (e.g. Bikhchandani, Hirshleifer and Welch 1992, 1998). As the focus of

In contrast to the individual-level literature, the effects of "keeping up with the Joneses" preferences on asset prices are comparatively less well studied. While these preferences have been explicitly incorporated into some asset pricing models (e.g. Abel 1990; Gali 1994; Bakshi and Chen 1996; Gómez, Priestley, and Zapatero 2009), most relevant to our study is a model by DeMarzo et al. (2008), where such preferences arise endogenously. In their finite-horizon overlapping generations framework, relative wealth concerns result from competition amongst rational investors for a scarce good, the future affordability of which depends on relative wealth. The fear of relative poverty in the future induces investors to herd as a form of insurance, which in turn generates price bubbles in the asset that the herd invests in.

Support for a link between relative wealth concerns and asset price bubbles is provided by a recent experimental study by Schoenberg and Haruvy (2012), who find that the size and duration of price bubbles in experimental asset markets are sensitive to the type of relative performance information provided to traders, despite that information being irrelevant to their compensation. Specifically, they examine price behaviour in the canonical Smith et al. (1988) bubble-market environment and find that asset price bubbles are significantly more pronounced when traders receive information on the value of the best performing trader's portfolio at the end of each trading period compared to when they receive information about the worst performing trader.

This result is especially interesting because the price behaviour observed in their markets, particularly in their 'upward-reference' markets, compares to the heightened mispricing seen in studies of experimental asset markets under tournament incentives, where relative performance feedback is customarily provided (e.g. James and Isaac

our study is on "keeping up with the Joneses" preferences, we do not review the literature on social learning here.

2000; Isaac and James 2003; Robin, Straznicka and Villeval 2012; Cheung and Coleman 2014). Hence, this potentially suggests that the driving force behind the price dynamics observed in these studies of tournament incentives may in fact be the intrinsic desire of traders to "keep up with the Joneses" rather than the extrinsic monetary incentives that tournaments provide. However as yet, no study has attempted to separate the impact(s) of the two on market prices, and as such, the relative importance of each is an open empirical question.

We bridge this gap by isolating the effect of intrinsic competitive incentives on experimental asset prices by examining price behaviour in absolute performance-based ('normal' incentive) markets with and without the supply of relative performance information. The null hypothesis examined is:

Hypothesis 1: Prices do not behave significantly differently in normal incentive markets when relative performance information is present vs. absent.

We extract the incremental effect of tournament monetary incentives by comparing normal incentive markets where relative performance information is provided against tournament incentive markets, under the null hypothesis:

Hypothesis 2: Price behaviour does not significantly differ between tournament incentive markets and normal-incentive-plus-relative-performance-information markets.

Our approach mirrors Dijk et al. (2014), who also attempt to separate the effect of relative performance information from tournament incentives in an experimental setting. They find that behaviour in their portfolio-choice task is driven almost entirely by 'social competition' or the intrinsic desire for rank. The key distinction between Dijk et al. and the current study is that they examine the issue at the investor-level whereas we focus on impacts at the aggregate/market level.

In doing so, we not only contribute to the experimental literature on tournament incentives, but also to the empirical literature on the aggregate-level impacts of relative wealth concerns, which consists of only the study by Schoenberg and Haruvy (2012). Our study differs from Schoenberg and Haruvy on three counts. First, we examine the impact of a different type of relative performance information – the performance of the average trader, rather than the best or worst performer. Second, instead of a single-asset experimental market, we study an asset market containing two differentiated dividend-paying risky assets, allowing us to assess the impact of competitive preferences within an environment that provides a better approximation of real-world markets. In addition, whereas Schoenberg and Haruvy only study the effect of relative performance feedback on inexperienced traders, by repeating our markets, we also examine how common group and design experience mediates the relationship between feedback and prices.

4.3. Experimental Design

The experiment comprises 43 independent markets across 23 sessions conducted at the ASB Experimental Research Laboratory at UNSW Australia between August and November 2013, with 320 subjects taking part across all treatments. Participants were university students with no prior experience in market experiments, recruited using

ORSEE (Greiner 2004) [90]. As this study was conducted in conjunction with the study detailed in Chapter 3, the experimental designs are identical between the two studies with respect to the structure of the market and experimental procedures. Readers familiar with these elements of the design may skip sections 4.3.1 and 4.3.3 without loss of continuity.

4.3.1 Market structure

In each session, participants were given the opportunity to trade two types of assets concurrently, one called “X”, the other called “Y”. The market for both assets, grounded in the classic Smith et al. (1988) design, ran for 12 periods, each lasting 3 minutes[91]. Trade occurred according to continuous double-auction rules; participants were allowed to post bids and asks for both assets in separate open order books, and accept any posted bid or ask for either asset, subject to the constraints posed by their asset holdings and cash balance. All trade occurred in single units, and short-selling and buying on margin were not permitted. Trade was conducted in experimental currency called ‘francs’, with earnings being paid out at the end of the experiment in Australian dollars at a pre-announced exchange rate of 200 francs to 1 Australian dollar. The market institution was fully computerised using zTree (Fischbacher 2007)[92].

---

[90] In total, 46 markets were run. However, some participants with multiple ORSEE profiles managed to participate in more than one session. To mitigate any potential confounding of treatment effects, we excluded from our analysis any markets that contained subjects who had participated in an earlier session.

[91] While experimental studies of tournament incentives have largely stuck with the parameters in Smith et al. (1988), there is considerable heterogeneity in studies involving multiple assets. The number of trading periods in these studies ranges from 12 (Ackert, Charupat, Church and Deaves 2006) to 30 (Chan, Lei and Vesely 2013), while trading period lengths vary between 3 (Chan et al. 2013) and 6 minutes (Fisher and Kelly, 2000). The parameters chosen for our sessions are consistent with the lower end of this range, and represent a suitable compromise given the constraints posed by budgets and time. In particular, we were mindful of avoiding sessions that were “too long” and risked inducing boredom in participants, given the repetitive nature of market experiments.

[92] The trading interface is shown in Figure 3.1 in Chapter 3. Given the previously documented tendency for trading activity to be biased in favour of the market that appears on the left-hand side of the screen (see Chan et al. 2013), the market for Asset X was placed on the left for roughly half of the sessions in each treatment, and on the right for the remainder.

All traders began the market with the same initial endowment of assets and cash – 5 units each of X and Y, and 1950 francs. This ensured that the relative position of any trader in the market was not affected by the composition of their initial allocation, and also the expected earning opportunities for each trader was initially the same. At the end of each trading period, Asset X paid a cash dividend drawn from the distribution {10, 30} with equal probability, while Asset Y paid a dividend from the distribution {0, 100} with respective probabilities (0.8, 0.2)[93]. These distributions were known to all participants. Dividend draws, which were made by the computer, were independent across trading periods and between the two types of assets. Any dividend earnings were added to the trader's cash balance, and their end-of-period portfolio carried over to the next trading period.

Note that the expected dividend paid by both X and Y in each period is 20 francs. Consequently, the process of backward induction implies that the risk-neutral fundamental value (FV) of both assets is the same, equal to the expected total future dividend stream, or 20 multiplied by the number of remaining trading periods (including the current one). Hence the risk-neutral fundamental values of both assets, shown by the solid black line in Figure 4.1, declines in steps of 20 in each period, beginning at 240 in period 1 and falling to 20 in period 12, before expiring worthless after the final dividend draw at the end of the period[94].

[93] These dividend structures mimic Ackert et al. (2006b), who also use a standard/lottery-asset dichotomy, albeit with a much more pronounced difference in potential payoffs between the two types. Their 'standard' asset's dividend distribution is {0.50, 0.90, 1.2} with respective probabilities (0.48, 0.48, 0.04), while their 'lottery' asset pays a dividend from the distribution {0, 18} with associated probabilities (0.96, 0.04). The maximum possible payoff in a period from their lottery asset is 15x the maximum payoff from the standard asset, whereas the corresponding multiple in our study is only 3.33x, as the intention was to have participants still view Asset Y as a viable "investment" rather than a purely speculative bet.

[94] The expected value of the total future dividend stream was common knowledge, and was communicated to participants in the form of an "average holding value" table contained within the written instructions given to all participants.

Figure 4.1: Fundamental value process of assets X and Y

The solid black line in the graph below depicts the risk-neutral fundamental value process of assets X and Y. Both assets pay an expected dividend of 20 per period. The dashed and solid grey (blue) lines depict the largest and smallest possible cumulative future dividend realisations of asset X (Y) respectively. Asset X pays a minimum of 10 francs in dividends each period, and a maximum of 30 per period. Asset Y pays a minimum of zero every period and a maximum of 100 every period. Hence, the blue dotted line, which is only partially graphed, starts at 1200 in period 1 and falls in steps of 100 in each ensuing period.

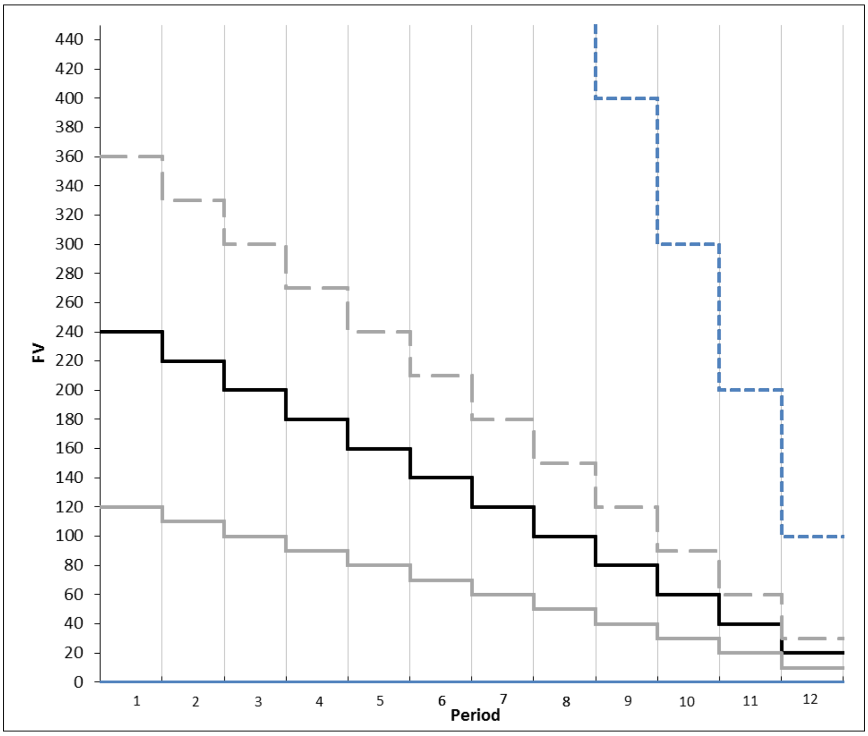


Figure 4.1 also illustrates the largest (dotted line) and smallest (solid line) possible cumulative future dividend realisations of each asset (X in grey, Y in blue). To keep the other features of the graph from being obscured, the step function for the maximum possible dividends from asset Y is only partially displayed; the blue dotted line for Y, which can potentially pay 100 francs in every period, begins at 1200 in period 1 and falls in steps of 100 in each ensuing period. In contrast, the minimum cumulative dividend payment from asset Y is zero, while asset X pays at least 10 and possibly 30 francs in each period. These step functions serve to demonstrate that

although both assets have the same expected dividend, the variance of Y's dividend payoff is much greater than X's. Hence, asset X represents a 'safe' investment, whereas Y with its lottery-like characteristic is riskier/more speculative. The presence of a second, risky/speculative asset in the market provides a more natural and realistic avenue for traders to take risk in the hope of reaping greater rewards than what single-asset environments provide.

The parameters discussed above determine the initial liquidity of our markets, as measured by the initial cash-to-assets ratio – the ratio of total cash to the total intrinsic value of all assets (X and Y) at the beginning of the market. This ratio was 0.8125 in all sessions, allowing us to control for the effects of liquidity on prices, which is known to be positively associated with the magnitude of bubbles in experimental markets (Caginalp, Porter and Smith 1998, 2000, 2001). While existing experimental studies of tournament incentives and multiple assets have used a variety of initial cash-to-asset ratios, our choice of 0.8125 reflects the cash-to-assets ratio in the most oft-replicated Smith et al. (1988) design, as well as being the initial liquidity used by both Schoenberg and Haruvy (2012) and Cheung and Coleman (2014) in their studies of social competition and tournament incentives respectively.

4.3.2 Treatments

To determine the separate effects of relative performance feedback and competitive monetary incentives on price behaviour, we implemented a 3-treatment between-subjects design[95]. In the first two treatments, *Baseline* and *RelInfo*, participants

[95] This is an incomplete 2 x 2 factorial design, where the 2 factors are the type of incentive contract and the information environment. The two levels of the incentive factor are linear incentives and tournament incentives, while the two levels of the information factor are 'with relative performance information' and 'without relative performance information'. Our design does not interact tournament incentives with 'without relative performance information'

were compensated under 'normal'/linear incentives, that is, on the basis of their absolute performance; since both assets X and Y expired worthless at the end of the market, this means that traders were paid their final cash balance[96]. The only difference between the two treatments is that whereas traders in the *RelInfo* treatment were privately informed (on-screen) of their relative performance at the end of each trading period, those in the *Baseline* treatment were not. Specifically, *RelInfo* treatment traders were shown the average Account Total in their market, in addition to their own. *Baseline* participants were only informed of their own Account Totals.

Account Total is akin to the market value of a trader's portfolio. It is based on a measure of the same name used by Schoenberg and Haruvy (2012), and is defined as the sum of a trader's end-of-period cash balance and the value of their end-of-period asset holdings; the end-of-period holdings of X and Y in our study were valued at their respective median traded prices in that period. Like Cheung and Coleman (2014), we chose the median price in preference to the final trading price or highest bid (as used by Schoenberg and Haruvy) because it is more difficult for traders to manipulate[97]. Since all assets expired worthless after the final dividend payment, the Account Total at the end of period 12 (i.e. at the end of the market) reverted to the final cash balance[98].

Since traders' relative performance in the *RelInfo* treatment was private information and experimental subjects were prohibited from communicating with each other, any peer effects in our experiment generated by a desire to "keep up with the Joneses" is more likely to be due to intrinsically-motivated competitive preferences (e.g.

---

[96] Ending cash balance = initial cash balance + dividend earnings + sales revenue – expenditure on purchases

[97] In periods where there was no trade in an asset, the median transaction price was replaced by the median buy offer for that asset in the period. This was done to avoid misleading fluctuations in the Account Total, and participants were made aware of this before the market began.

[98] This small change in the definition of the Account Total for period 12 was necessary, since otherwise, it would create an incentive for participants to arbitrarily bid up the prices of assets X and Y in period 12 in the hope of maximising their Account Totals.

for self-esteem) than status concerns, as the latter is predicated on public recognition. Status concerns are made all the more unlikely by the fact that random assignment of participants to markets reduces the probability that market-groups contained participants who knew each other. Furthermore, since participants were only informed of the market value of the average portfolio but not its composition, or indeed the composition of any other trader's portfolio, the potential for social learning to explain any peer effects in the *RelInfo* treatment is limited. However, it is possible that relative performance feedback may help traders to learn about their own abilities (Festinger 1954), which in turn affects their trading behaviour and thus prices. While disentangling the contributions of these drivers is beyond the scope of this study, we do note that any incremental effect induced by the introduction of relative performance information, though most likely driven by innate competitive preferences, may act in concert with social learning.

In the third treatment, *Tournament*, participants traded under a tournament compensation scheme that paid traders on the basis of their performance relative to the 'average' trader. This treatment in fact comprises two 'sub-treatments' that are the subject of the study in Chapter 3. In the first sub-treatment, called 'Carrot', we followed James and Isaac (2000) and rewarded those who performed better-than-average with higher payments while all others were paid a fixed amount, according to the following rule:

$$Earnings_i = \begin{cases} 3000 & if \quad C_i < C^* \\ 3000 + 2(C_i - C^*) & if \quad C_i \geq C^* \end{cases}$$

C_i is the final cash balance of trader *i* and C* is the average of the final cash balances of all traders in the market. All units and amount shown are denominated in francs.

The second sub-treatment, 'Stick', is the same as 'Carrot' except for the inclusion of a penalty for especially poor relative performance, as in Isaac and James (2003):

$$Earnings_i = \begin{cases} 0 & if \quad C_i < \frac{1}{2}C^* \\ 3000 & if \quad \frac{1}{2}C^* \leq C_i \leq C^* \\ 3000 + 2(C_i - C^*) & if \quad C_i > C^* \end{cases}$$

The data from the two sub-tournaments are pooled together to form the *Tournament* treatment. Like the *RelInfo* treatment, all *Tournament* traders were privately informed of their own Account Totals and the average Account Total in the market at the end of each trading period. Hence, the only difference between the two treatments is the type of compensation contract.

A potential problem that the pooling of the sub-treatments presents is that it may mask any differences between specific sub-treatments and the other treatments, especially if price behaviour varies systematically between the two sub-treatments. While we detect some significant differences between the two with inexperienced traders in Chapter 3, these differences do not survive with experienced participants. In the discussion of the results below, we note any situations where the individual sub-treatments tell a different story to the pooled treatment.

Note that in the *Tournament* treatment described above, payoffs depend not only on being better/worse than average but also the extent to which a trader's *absolute* performance differs from the average. To examine a 'purer' tournament contract, in the Lazear and Rosen (1981) sense, where *only* relative performance matters, we also tested an alternative tournament treatment named *GilTournament*. Based on the rank-order tournaments modelled by Gilpatric (2009), participants in this treatment were paid fixed amounts that were determined purely by their final rank. Like *Tournament*, *GilTournament* consists of two sub-treatments pooled together; sub-treatment 'GilCarrot' paid the trader with the largest final cash balance 10,000 francs, while all other traders received 4000, whereas 'GilStick' retained the same compensation structure but paid the worst-performing trader nothing [99]. Concerns surrounding pooling are less serious here, as we do not detect a significant difference in price behaviour between 'GilCarrot' and 'GilStick' with either experienced or inexperienced participants in Chapter 3. Since the payoff-relevant piece of relative-performance information in these tournaments is the trader's rank, participants in this treatment were informed of their rank at the end of each period (calculated on the basis of Account Total), in addition to the other relative performance information described above.

4.3.3. Procedures

Each experimental session corresponded to a single (sub-)treatment to which it (and hence, each subject within it) was randomly assigned[100]. Sessions were designed to run two independent market-groups of (up to) 8 traders each and ran for approximately

[99] Note also that the minimum payment in 'GilCarrot' was set to 4000 francs compared to 3000 francs in the equivalent *Tournament* treatment contract 'Carrot' to ensure that the average compensation per trader in real currency, Australian dollars, was roughly equal across contract-types, and to also conform to the ASB Lab ethics protocol which specified an average payment range of $15-20 per hour per participant.

[100] The only exception to this was a single session where a 'Carrot' market ran alongside a 'GilStick' market. The instructions and procedures were appropriately modified for this session to prevent contamination of the subject pool.

2.5 hours[101]. To ensure consistency in the delivery of instructions between sessions and reduce experimenter demand effects, all participants received written instructions, which were also communicated verbally by the experiment administrator[102]. As mentioned above, potential interaction effects between participants were mitigated by prohibiting subjects from communicating with each other for the duration of the experiment.

The procedure followed in each session was identical, regardless of the treatment. Sessions began with participants being randomly allocated to a computer/workstation that determined their market-group[103]. They then received training on how to use the trading screen to make and accept bids and offers for each asset (10 minutes), following which they were given 10 minutes to practise trading using the interface. After the practice period, subjects were given further information about the other features of the market environment, including how their earnings would be calculated. After this, the market-proper began. Upon the conclusion of the market, participants were informed that they would be taking part in another 12-period market with the same traders (i.e. market-group). Participants' inventory of assets and cash were reset to their starting levels, and trading commenced for a second round.

After the end of the second round, participants completed an untimed survey consisting of 3 sections[104]. The first section gathered general demographic information

---

[101] That is, excluding the practice period, participants only traded with other participants who were in the same market-group. Dividends were also drawn independently for each market-group.
[102] To ensure consistency with the procedures used in the experimental asset market literature, the written protocol was adapted from Dufwenberg, Lindqvist and Moore (2005), Noussair, Robin and Ruffieux (2001), Noussair and Powell (2010), Lugovskyy, Puzzello and Tucker (2009), Childs and Mestelman (2006), and Cheung and Coleman (2014). Participants were also given time to read the instructions on their own, and to ask any clarifying questions privately (which were also answered privately). The written protocol can be found in Appendix B3
[103] The workstation number also served as a participant's ID, thus ensuring the anonymity of their data.
[104] The survey, which can be found in Appendix B4, was initially paper-based (12 sessions), but was computerised using the *Qualtrics* survey software and administered electronically in the October and November sessions (11 sessions).

about participants and their experiences and thought-processes during the market(s)[105]. The second and third sections, which form part of a related study, comprise the *Cognitive Reflection Test* (CRT) and *Domain-Specific Risk-Taking* (DOSPERT) Scale. The CRT is a measure of cognitive ability developed by Frederick (2005) that consists of 3 problem-solving type questions that assess the ability of respondents to reject an impulsive and intuitive incorrect answer in favour of a correct answer that requires more deliberation. In addition to general measures of cognitive ability, performance in the CRT is correlated with time and risk preferences (Frederick 2005), as well as certain behavioural biases (Oechssler, Roider, and Schmitz 2009). The 30-item DOSPERT Scale, designed by Blais and Weber (2006), is a psychometric scale that measures risk preferences and perceptions across five separate decision-making domains: Financial (split into Investing and Gambling), Health/Safety, Recreational, Ethical, and Social[106]. Respondents use a 7-point scale to rate the likelihood of their participation (Part 1), the perceived riskiness (Part 2), and the benefits expected to accrue (Part 3) from engaging in 30 different domain-specific risky activities. Of course, administering the DOSPERT Scale after the market stage carries with it the risk that responses may be influenced by participants' experiences during the market. However, given our main objective is to study price behaviour, this is the 'lesser of two evils', as the alternative of implementing the scale before the market could in turn influence participants' trading behaviour. A summary of the demographic characteristics of the subject pool, CRT scores (out of 3), and DOSPERT likelihood/preference scores in the most relevant domain, Financial (ranges from 6 to 42, higher scores indicate greater willingness to take financial risks), is presented in Table 4.1, categorised by treatment.

---

[105] This is a modified version of the end-of-experiment questionnaire used by Ackert and Church (2001).

[106] Compared to the original 40-item DOSPERT scale (Weber, Blais and Betz 2002), which was developed for American undergraduate college students, the revised 30-item DOSPERT scale (Blais and Weber 2006) was chosen because it is designed to be more readily applicable to a more diverse range of cultures, age groups, and educational levels.

Table 4.1: General demographic information

This table reports general demographic information on the subject pool, categorised by the experimental treatment to which participants were randomly assigned. Note that the treatment *Tournament* (*GilTournament*) comprises the pooled 'Carrot' and 'Stick' ('Gil-Carrot' and 'Gil-Stick') treatments from Chapter 3. ‘Business student’ is defined as someone studying Finance, Economics, Actuarial, Accounting, or "Commerce" (self-reported). In a post-experiment survey, all participants completed the Cognitive Reflection Test (CRT) developed by Frederick (2005), which measures cognitive ability; CRT scores are out of 3 and higher scores indicate better performance. Participants also completed the Domain-Specific Risk-Taking (DOSPERT) Scale (Blais and Weber 2006). The score reported here relates to participants’ (self-reported) likelihood of engaging in risky financial activities. Scores range from 6 to 42, with higher scores indicating a greater likelihood of engaging in risky activities.

	Baseline	*RelInfo*	*Tournament*	*GilTournament*
No. markets	7	8	16	12
No. subjects	51	59	119	91
Average age	22.3	21.9	22.3	22.7
Male (%)	65	59	52	44
Business students (%)	29	34	35	40
Avg. CRT score	1.6	1.5	1.4	1.4
Avg. DOSPERT Fin. score	19.2	19.5	19.0	19.3

Once the surveys were completed, participants were called up individually, paid their earnings (in envelopes) and dismissed. Participants’ total earnings from the experiment were calculated as the sum of their earnings from both rounds of the market, converted to Australian dollars, plus a $5 participation fee. The average payment to participants, inclusive of the participation fee, was $50.

4.4. Results

We examine our research hypotheses as follows. To determine how relative performance information impacts prices, independent of any monetary incentives attached to relative performance (Hypothesis 1), we compare price/bubble behaviour in the *Baseline* treatment to the *RelInfo* treatment. To then gauge the incremental effect of explicit monetary incentives associated with relative performance (Hypothesis 2), we compare the *RelInfo* treatment against the *Tournament*/*GilTournament* treatment.

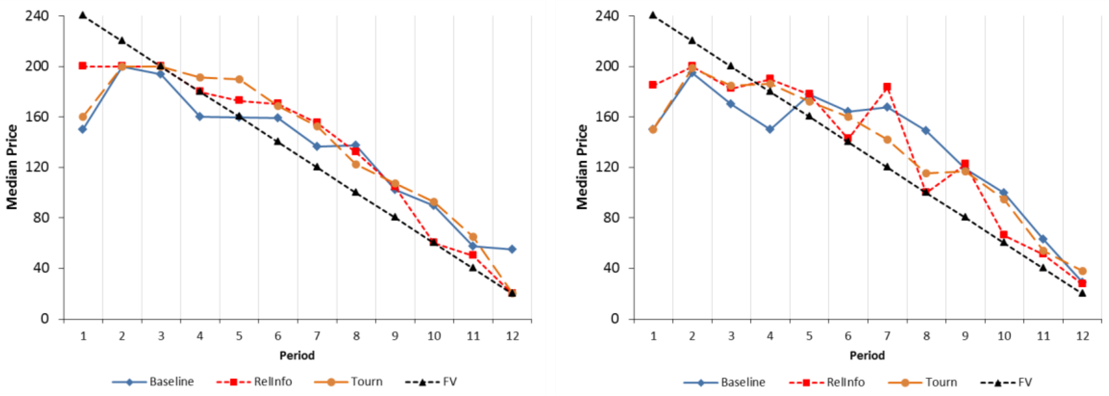
4.4.1 Inexperienced traders

4.4.1.1 Descriptive summary

Panels (a) and (b) of Figure 4.2 chart the time-path of the median transaction price of assets X and Y respectively in the *Baseline*, *RelInfo*, and *Tournament* treatments during the first round of the market; for each treatment, the charted price in each period is the median of the median transaction prices from all markets in that treatment. The price behaviour shown in these graphs is somewhat atypical of Smith et al. (1988)-type markets populated with inexperienced participants; missing here, most surprisingly in the *Baseline* treatment – essentially a replication of earlier two-asset studies such as Fisher and Kelly (2000) – is the characteristic bubble-and-crash pattern to prices. Both assets in all three treatments exhibit only a moderate degree of

Figure 4.2: Median prices in Round 1

Median transaction prices in the *Baseline* (solid blue line), *RelInfo* (dotted red line), and *Tournament* (dashed orange line) treatments during the first round of the market (i.e. with inexperienced traders) are shown below for the 'low-risk' asset X (panel (a)) and 'high-risk' asset Y (panel (b)), along with the risk-neutral fundamental value process for each asset (dotted black line). For each treatment, the plotted median price in each period is the median of the median transaction prices from all markets belonging to that treatment. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



(a) Asset X, Round 1 (b) Asset Y, Round 1 overvaluation in the middle periods, and there is no crash to fundamental value late in the market. However, it should be noted that the median prices shown in Fig. 4.2 conceal substantial heterogeneity within each treatment – bubbles followed by crashes were indeed observed in *individual* markets in all treatments[107].

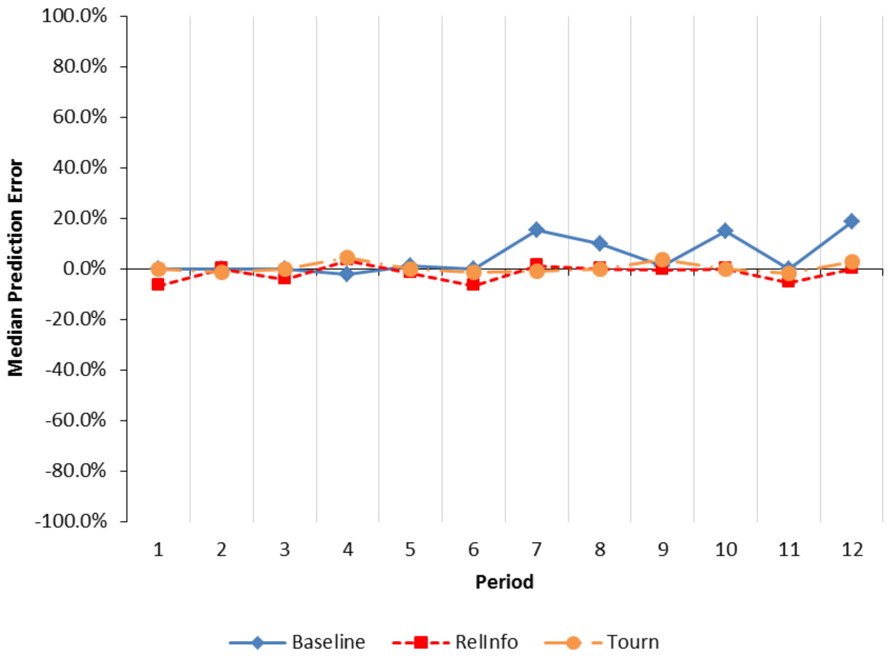
With respect to the research hypotheses of this study, Figure 4.2 does not support the idea that relative performance information or tournament incentives have an impact on price behaviour. There is no obvious difference in median price behaviour between the three treatments in either asset. Indeed, the differences in median price between the treatments are not statistically significant at the 5% level in any period, as ascertained using the non-parametric equivalent of the independent samples t-test, the Wilcoxon Mann-Whitney U test.

A similar picture emerges from the behaviour of the *relative* price between the two asset-types. Here again, the three treatments do not appear to differ noticeably, which can be seen in the behaviour of median *Prediction Error*s charted in Figure 4.3. Following Fisher and Kelly (2000), we define the *Prediction Error* in each period of an individual market as the percentage deviation of the relative price of asset Y (median price of Y divided by the median price of X in that period) from the risk-neutral benchmark (equal to 1 in this study)[108]. More positive (negative) values indicate a greater willingness by market participants to pay a premium to acquire the riskier (less risky) asset Y (X). As Fig. 4.3 shows, the median *Prediction Error* in every treatment remains relatively close to zero throughout the market. This observation appears consistent with other multi-asset studies, which report that relative prices remain

[107] The evolution of median prices in the individual markets of each treatment comprising this study is charted in Figures B1-B12 in Appendix B1.
[108] Unlike Fisher and Kelly (2000), we report the median of the *Prediction Errors* across all sessions/markets rather than the average, due to the lower sensitivity of the median to outliers in small samples.

Figure 4.3: Median values of *Prediction Error*, Round 1

The figure below plots the evolution of the median *Prediction Error* in the *Baseline* (solid blue line), *RelInfo* (dotted red line), and *Tournament* (dashed orange line) treatments during the first round of the market (i.e. with inexperienced traders). For each treatment, the plotted value in each period is the median of the *Prediction Errors* from all markets in that treatment. *Prediction Error* is defined as the percentage deviation of the relative price of Y (i.e. median price of asset Y divided by median price of asset X) from the risk-neutral benchmark of 1. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



'correct' even when/if individual assets are over/under-priced (e.g. Fisher and Kelly 2000, Childs and Mestelman 2006).

Although relative prices in the three treatments appear to closely correspond to each other in general, Fig. 4.3 potentially indicates that this correspondence is weaker during the second half of the market. Specifically, while the average of the median *Prediction Errors* in the second half of the market is 0.9% and 0.7% in *RelInfo* and *Tournament* respectively, it is 10% in the *Baseline* treatment. As we investigate more formally below, this modest premium paid for the riskier asset Y in the *Baseline* treatment during the latter stages of the market may indicate the presence of risk-seeking behaviour. Hence, its absence in the *RelInfo* treatment potentially suggests that providing relative performance feedback discourages such behaviour.

4.4.1.2 Statistical analysis

The visual data thus seems to suggest that in markets with inexperienced participants, relative performance information and competitive monetary incentives have a negligible impact on price behaviour. To see if a similar picture emerges from a formal statistical analysis of treatment differences, a number of measures of bubbles and mispricing commonly employed in the experimental asset market literature are constructed. Readers familiar with the definitions of the bubble measures used in Chapter 3 can skip the following sub-section entitled "Bubble Measures" without loss of continuity.

<u>Bubble measures</u>

The bubble measures examined can be broadly categorised into two groups that assess two different dimensions of mispricing – magnitude and length.

The first of the magnitude measures, *Amplitude* (Haruvy and Noussair 2006), quantifies the extent to which the average price in a market changes relative to FV. It is calculated as $\max_t\{(\bar{P}_t - F_t)/F_t\} - min_t\{(\bar{P}_t - F_t)/F_t\}$, where the largest and smallest deviations of average price P_t from the fundamental value F_t are normalised by the fundamental value in the respective period *t*. Large values of this measure indicate bigger swings in price relative to FV and hence the more likely presence of a bubble. *Total Dispersion* (Haruvy and Noussair, 2006) measures the aggregate absolute deviation of median price from FV across all trading periods, and is defined as $\sum_t |MedianP_t - F_t|$. Since it treats both positive and negative deviations from FV identically, it is a measure of mispricing rather than over or undervaluation, with smaller values indicating a closer correspondence between price and FV. *Turnover*, a normalised measure of trading activity, is used as a measure of magnitude since bubble periods are typically associated with high trading volumes. It is calculated as defined by King, Smith, Williams, and van Boening (1993), namely $\sum_t V_t/TSU$, where V_t, the volume of trade in period *t* is normalised by *TSU*, the total number of units of the asset (X or Y) in the market. *Normalised Deviation*, measured by Haruvy, Lahav and Noussair (2007) as $\sum_t V_t|MedianP_t - F_t|/(TSU)$, combines the preceding two measures to account for both the size of the price deviation and the level of trading activity in a market. *Haessel-R^2* (Dufwenberg et al. 2005) indicates how closely prices track changes in fundamental value. It is the R-squared from the regression of average prices on fundamental values. Being a goodness-of-fit measure, it tells you how much of the variation in average price across periods is explained by changes in FV; values closer to 0 (1) suggest the potential existence (absence) of price bubbles. *Average Bias* (Haruvy and Noussair 2006) is used to gauge the degree of overpricing/underpricing in a market. It measures how far median prices deviate from FV on average over the course of the market, and is calculated as $\frac{1}{N}\sum_{t=1}^{N}(MedianP_t - F_t)$. Large positive (negative) values suggest that prices tend to stay above (below) FV. Values close to zero may suggest that prices stay close to FV or that the asset experiences equal degrees of over and underpricing in the market; assessing the *Average Bias* in conjunction with *Total Dispersion* helps to shed light in this regard, since observing a small (large) *Total Dispersion* at the same time as a near-zero Average Bias would imply the former (latter) (Haruvy and Noussair 2006).

The first of the bubble-length measures, *Duration* (Porter and Smith 1995), calculates the maximum number of consecutive periods where average price increases relative to FV, or $max\{m: \bar{P}_t - F_t < \bar{P}_{t+1} - F_{t+1} < \cdots < \bar{P}_{t+m} - F_{t+m}\}$. Larger values of *Duration* point to sustained periods where changes in (average) transaction price across trading periods do not 'adequately' track changes in the FV, potentially indicating the presence of a bubble. *Boom* (*Bust*) *Duration* (Haruvy and Noussair 2006) is defined as the maximum number of consecutive periods where median prices stay above (stay below) FV; large values indicate long periods of overvaluation (undervaluation), potentially signalling the presence (absence) of a bubble.

The behaviour of individual assets

Panels A and B of Table 4.2 report the median values of the bubble measures in each treatment for assets X and Y respectively in Round 1, along with the associated median absolute deviations (MAD)[109]. Since each measure produces one observation per market for each asset-type, the medians are based on 7 observations in the *Baseline* treatment, 8 in the *RelInfo* treatment, and 16 in the *Tournament* treatment[110]. The bottom half of each panel reports two-sided exact p-values from Wilcoxon Mann-Whitney U (WMW) tests of the differences in the measures between treatments, under the null that the values from both treatments come from the same distribution[111]. The WMW test, which is the non-parametric equivalent of the independent samples t-test, is the appropriate statistical test given the small sample sizes involved.

[109] The median absolute deviation (MAD) is a measure of the spread of a distribution, and is calculated as the median of the absolute deviations of all values in a sample from the median. We report the median value and MAD of each measure in preference to the mean and standard deviation due to the small number of observations involved, and their lower sensitivity to outliers.

[110] The higher number of observations in the latter treatment arises from the fact that *Tournament* comprises markets from both James and Isaac (2000) tournament contracts – *Carrot* and *Stick* – examined in Chapter 2.

[111] The WMW test is the non-parametric equivalent of the independent samples t-test, which is the appropriate test given the small sample size.

Table 4.2: Summary of bubble measures for assets X and Y in Round 1

This table reports median values of each bubble measure in the *Baseline*, *RelInfo*, and *Tournament* treatments during Round 1 of the market; median absolute deviations are displayed in parentheses. Markets contaminated by subjects who had participated in an earlier session are excluded. Panel A (B) reports bubble measure data relating to Asset X (Y). For definitions of the relevant bubble measures, see section 4.4.1.2. The statistical significance of the difference between treatments in each measure is assessed using a two-sided Wilcoxon Mann-Whitney U Test, under the null that values from both treatments come from the same distribution. Exact p-values are reported. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively.

Panel A: Asset X, Round 1:

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel* R^2	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	3.31	298.00	1.65	0.78	2.45	83.43	5.00	5.00	4.00
	(1.53)	*(181.50)*	*(15.73)*	*(0.16)*	*(0.72)*	*(45.08)*	*(2.00)*	*(2.00)*	*(1.00)*
RelInfo [8]	1.03	457.25	11.08	0.57	2.25	79.20	6.50	5.50	2.00
	(0.44)	*(284.25)*	*(29.48)*	*(0.18)*	*(0.66)*	*(37.14)*	*(2.50)*	*(2.00)*	*(1.00)*
Tournament [16]	2.23	569.50	9.21	0.46	2.59	108.04	4.50	8.00	3.00
	(1.70)	*(312.75)*	*(26.06)*	*(0.36)*	*(0.59)*	*(71.52)*	*(1.50)*	*(3.00)*	*(2.00)*
WMW U-Test p-values (two-sided):									
Baseline vs. RelInfo	0.189	0.955	1.000	0.536	0.779	0.955	0.956	0.845	0.977
Baseline vs. Tournament	0.922	0.413	0.769	0.198	0.671	0.413	0.559	0.684	0.906
RelInfo vs. Tournament	0.120	0.461	0.787	0.569	0.928	0.528	0.592	0.323	0.940

Panel B: Asset Y, Round 1

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel* R^2	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	1.63	530.50	15.38	0.77	2.03	99.06	4.00	7.00	4.00
	(1.01)	*(150.50)*	*(25.33)*	*(0.05)*	*(0.35)*	*(54.11)*	*(1.00)*	*(1.00)*	*(2.00)*
RelInfo [8]	1.25	527.00	11.04	0.59	1.90	80.51	4.00	4.00	2.50
	(0.75)	*(221.25)*	*(26.94)*	*(0.18)*	*(0.39)*	*(23.66)*	*(1.00)*	*(1.00)*	*(0.50)*
Tournament [16]	1.79	572.25	2.48	0.38	2.53	120.11	5.00	6.00	3.00
	(0.87)	*(312.00)*	*(16.42)*	*(0.32)*	*(0.64)*	*(78.07)*	*(1.50)*	*(2.50)*	*(1.00)*
WMW U-Test p-values (two-sided):									
Baseline vs. RelInfo	0.463	0.955	0.867	0.694	0.867	0.779	0.943	0.174	0.844
Baseline vs. Tournament	0.820	0.922	0.922	0.452	0.769	0.922	0.636	0.251	0.829
RelInfo vs. Tournament	0.291	0.569	0.976	0.383	0.697	0.528	1.000	0.557	0.635

The results in Table 4.2 are consistent with the visual data. For both asset-types, a significant difference between the *Baseline* treatment and the *RelInfo* treatment (Hypothesis 1) is not detected in any bubble measure. While Schoenberg and Haruvy (2012) also report similar findings in a single-asset setting[112], this contrasts somewhat with Dijk et al. (2014), who find significant differences between their corresponding treatments, albeit in *individual*-level behaviour. Although a direct comparison between Dijk et al. and our study is confounded by considerable differences in intent and experimental design, it potentially serves to underscore that behaviour at the individual and aggregate levels do not necessarily coincide. Where the aggregate-level findings do tally with the individual-level findings of Dijk et al. is in the failure to find, for either asset-type, a significant difference between the *RelInfo* treatment and the *Tournament* treatment on any of the bubble measures (Hypothesis 2)[113].

Relative prices

Similar results are obtained from the analysis of relative prices, which is summarised in Panel A of Table 4.3. The variable of interest here, *Average Prediction Error,* is identical to the 'overall normalised exchange rate deviation' measure used by Fisher and Kelly (2000), and is calculated by averaging the *Prediction Errors* (defined above) in all periods of a market, yielding one observation per market. The median *Average Prediction Error* ("Avg PredErr") and associated MAD in each treatment is reported in the top-half of Panel A. Also reported are median values of *Average*

---

[112] Which is perhaps unsurprising, given they use a t-test with only 7 observations.

[113] While full results are not reported here, a comparison of the bubble measures of the *RelInfo* treatment against the constituent elements of the *Tournament* treatment shows that the bubble measures in *RelInfo* are not statistically significantly different from those in the *Carrot* sub-treatment for asset X or Y. Regarding the *Stick* contract, we do not detect a statistically significant difference in any of the bubble measures for asset Y, while most bubble measures do not differ in a statistically significant manner for asset X, the exceptions being *Boom Duration*, which is significantly higher in *Stick* (10 vs. 5.5, p-value = 0.01), and *Amplitude*, again higher in *Stick* (2.63 vs. 1.03) but only marginally significantly (p-value = 0.07).

Table 4.3: Average Prediction Errors

Median values of the *Average Prediction Error* in the *Baseline*, *RelInfo*, *and Tournament* treatments in Round 1 and 2 are shown below in Panels A and B respectively, with the associated median absolute deviations in parentheses. Markets contaminated by subjects who had participated in an earlier session are excluded. *Average Prediction Error* is calculated using all periods in a market, the first 6 periods, and the final 6 periods in *Avg PredErr*, *AvgPredErr_p1to6*, and *AvgPredErr_p7to12* respectively. The statistical significance of the individual measures is assessed using a (two-sided) one-sample Wilcoxon Signed-rank test, under the null that the median is equal to zero. The statistical significance of the difference between treatments is assessed using the Wilcoxon Mann-Whitney U Test under the null that values from both treatments come from the same distribution. The statistical significance of the difference between *AvgPredErr_p1to6* and *AvgPredErr_p7to12* within each treatment is assessed using a (paired-sample) Wilcoxon Signed-rank test, against the one-sided alternative hypothesis that *AvgPredErr_p7to12* > *AvgPredErr_p1to6*. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively.

Panel A: Round 1

Treatment [N]	*Avg PredErr (%)*	*AvgPredErr_ p1to6 (%)*	*AvgPredErr_ p7to12 (%)*	*Signed-rank p-value (1-sided) 1-6 vs. 7-12*
Baseline [7]	3.78	-0.32	7.88	0.064*
	(9.46)	*(3.36)*	*(9.26)*	
RelInfo [8]	-0.77	-1.44	-1.89	0.200
	(5.16)	*(3.45)*	*(10.36)*	
Tournament [16]	3.23	0.14	4.03	0.151
	(4.43)	*(6.85)*	*(3.85)*	
WMW U-Test p-values (2-sided):				
Baseline vs. RelInfo	0.397	0.536	0.336	
Baseline vs. Tournament	0.769	0.974	0.376	
RelInfo vs. Tournament	0.490	0.787	0.610	

Panel B: Round 2

Treatment [N]	*Avg PredErr (%)*	*AvgPredErr_ p1to6 (%)*	*AvgPredErr_ p7to12 (%)*	*Signed-rank p-value (1-sided) 1-6 vs. 7-12*
Baseline [7]	2.72	-0.12	5.23	0.032**
	(11.33)	*(2.27)*	*(21.08)*	
RelInfo [8]	-5.60*	-5.68	-4.94*	0.663
	(5.74)	*(4.92)*	*(3.46)*	
Tournament [16]	-0.51	-7.45**	1.87	0.025**
	(8.50)	*(10.87)*	*(17.93)*	
WMW U-Test p-values (2-sided):				
Baseline vs. RelInfo	0.152	0.232	0.121	
Baseline vs. Tournament	0.624	0.154	0.769	
RelInfo vs. Tournament	0.320	0.569	0.214	

Prediction Error calculated using only the periods in the first ("Avg PredErr_p1to6") or second ("Avg PredErr_p7to12) half of the market.

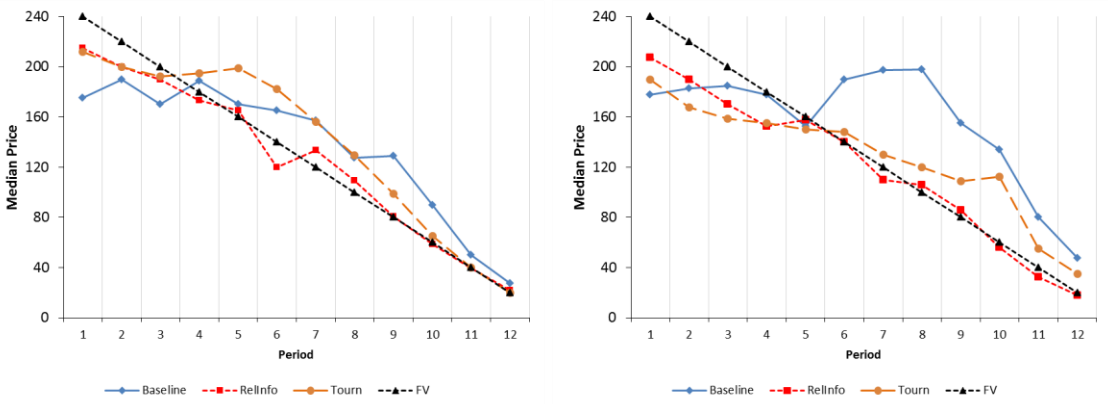
Similar to the analysis of risk-taking in tournaments conducted by Brown, Harlow and Starks (1996), the primary purpose of dividing the measure into two is to see if the behaviour of relative prices *within* treatments differs between the two halves of the market; specifically, whether there is heightened speculation in the 'risky' asset Y in the second half. Consequently, a Wilcoxon Signed-rank test is used to examine the one-sided alternative hypothesis that *Average Prediction Errors* in the second half of the market are higher than in the first half. The corresponding p-values (one-sided) are reported in the right-most column of Panel A[114]. In addition, the two-sided exact p-values reported at the bottom of Panel A correspond to WMW tests comparing *Average Prediction Error* between treatment-pairs in Round 1, under the null that values from both treatments come from the same distribution.

In line with Fig. 4.3, relative prices in Round 1 conform on 'average' to theoretical expectations; the median of the whole-of-market *Average Prediction Error* measure "Avg PredErr" is very small in all three treatments (less than 4% in absolute terms), and not significantly different from zero[115]. Signed-rank tests conducted within treatments do not indicate the presence of heightened speculative behaviour in the second half of the market in either the *RelInfo* or *Tournament* treatment, where competitive pressures may be expected to generate such behaviour. Indeed in all three treatments, we fail to reject the null of no difference in *Average Prediction Error* between the two halves at the 5% level, although it is marginally significantly higher in the second half in the *Baseline* treatment (one-sided p-value = 0.064). More pertinently

[114] The (paired-sample) Wilcoxon Signed-rank test is the non-parametric equivalent of the paired t-test. The null hypothesis is that values from both groups come from the same distribution.
[115] This is assessed using a one-sample Wilcoxon Signed-rank test, under the null that the median value is equal to zero. This is the non-parametric equivalent of the one-sample t-test.

Figure 4.4: Median prices, Round 2

Median transaction prices in the *Baseline* (solid blue line), *RelInfo* (dotted red line), and *Tournament* (dashed orange line) treatments during the second round of the market (i.e. with experienced traders) are shown below for the 'low-risk' asset X (panel (a)) and 'high-risk' asset Y (panel (b)), along with the risk-neutral fundamental value process for each asset (dotted black line). For each treatment, the plotted median price in each period is the median of the median transaction prices from all markets belonging to that treatment. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



(a) Asset X, Round 2

(b) Asset Y, Round 2

for the research hypotheses, a statistically significant difference is not detected between any of the treatments on any of the measures. Hence, as with the behaviour of the individual assets, relative performance information and competitive monetary incentives do not have a discernible impact on relative prices with inexperienced traders.

4.4.2 Experienced traders

<u>Individual assets</u>

Common group and design experience changes the story. Figures 4.4(a) and (b) show how median prices evolve during the second round of trading for asset X and Y respectively. Although definitive judgements from these graphs are difficult to make due to variability within treatments, the differentiation between the treatments in median-price behaviour appears more pronounced here compared to Round 1 (cf. Fig 4.2). Of all the conditions, it is the *RelInfo* treatment where median prices adhere to FV best, not appearing to bubble at all for either asset-type. In contrast, asset X in both the *Baseline* and *Tournament* treatments exhibits moderate overvaluation of roughly equivalent duration, with the 'bubble' peaking earlier in the latter treatment. For asset Y, median prices in the three treatments diverge during the second half of the market, with the *Baseline* treatment exhibiting a pronounced bubble in median prices that does not fully crash back to fundamental value by the end. The *Tournament* treatment on the other hand takes a path in between the *Baseline* and the *RelInfo* treatments.

These observations are mirrored by the bubble-measure data from Round 2, which are reported with the associated WMW test results in Table 4.4. They indicate that the *RelInfo* treatment is associated with smaller and shorter-lived bubbles than the *Baseline* treatment (Hypothesis 1). For the low-risk asset X (Panel A), the relative median values of all bar two of the measures (*Duration* and *Bust Duration)* are consistent with this, although the differences are significant at the 5% level only in the case of *Boom Duration* (p-value = 0.029), where the period of overvaluation lasts more than 3 times as long in the median *Baseline* market than in the corresponding *RelInfo* market ($\text{median}_{Baseline}$ = 8 vs. $\text{median}_{RelInfo}$ = 2.5 periods). A marginally significant difference is also observed in the case of *Haessel-R^2* (p-value = 0.094), where changes in FV explain 94% of the variation in average prices in the median *RelInfo* market, compared to 83% in the *Baseline* treatment.

Somewhat stronger evidence of a difference between the two conditions is found in the 'risky' asset Y (Panel B), where all of the median bubble measure values are consistent with less severe and prolonged overvaluation in the *RelInfo* treatment compared to the *Baseline* treatment. The goodness-of-fit measure *Haessel-R^2* is again significantly larger in the *RelInfo* treatment, this time at the 5% level (p-value = 0.029), with FV explaining 92% of the variation in average price across periods in the median

Table 4. 4: Summary of bubbles measures for assets X and Y in Round 2

This table reports median values of each bubble measure in the *Baseline*, *RelInfo*, and *Tournament* treatments during Round 2 of the market; median absolute deviations are displayed in parentheses. Markets contaminated by subjects who had participated in an earlier session are excluded. Panel A (B) reports bubble measure data relating to Asset X (Y). For definitions of the relevant bubble measures, see section 4.4.1.2. The statistical significance of the difference between treatments in each measure is assessed using a two-sided Wilcoxon Mann-Whitney U Test, under the null that values from both treatments come from the same distribution. Exact p-values are reported. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively.

Panel A: Asset X, Round 2:

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	1.15	323.50	10.96	0.82	1.60	39.45	3.00	8.00	3.00
	(0.79)	*(247.00)*	*(13.55)*	*(0.18)*	*(0.69)*	*(32.05)*	*(1.00)*	*(1.00)*	*(1.00)*
RelInfo [8]	0.55	237.75	1.54	0.94	1.50	22.93	3.50	2.50	3.00
	(0.12)	*(82.25)*	*(10.82)*	*(0.02)*	*(0.47)*	*(3.29)*	*(1.50)*	*(2.50)*	*(2.00)*
Tournament [16]	0.80	367.50	7.33	0.83	1.85	74.44	5.00	7.00	3.00
	(0.42)	*(244.00)*	*(18.89)*	*(0.15)*	*(0.57)*	*(60.92)*	*(2.00)*	*(3.00)*	*(2.00)*
WMW U-Test p-values (two-sided):									
Baseline vs. RelInfo	0.232	0.336	0.281	0.094*	0.482	0.121	0.618	0.029**	0.707
Baseline vs. Tournament	0.720	0.922	0.579	0.720	0.871	0.769	0.749	0.523	0.559
RelInfo vs. Tournament	0.320	0.383	0.320	0.172	0.742	0.192	0.655	0.099*	0.958

Panel B: Asset Y, Round 2

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	2.16	626.00	24.33	0.52	1.50	127.50	5.00	7.00	3.00
	(0.72)	*(492.50)*	*(68.88)*	*(0.13)*	*(0.47)*	*(64.98)*	*(3.00)*	*(4.00)*	*(2.00)*
RelInfo [8]	0.70	310.75	-6.86	0.92	1.23	32.16	3.00	2.50	4.50
	(0.29)	*(143.75)*	*(15.06)*	*(0.03)*	*(0.27)*	*(11.73)*	*(1.50)*	*(2.00)*	*(1.50)*
Tournament [16]	1.53	413.50	1.16	0.78	1.43	71.52	3.50	4.00	4.50
	(1.02)	*(99.75)*	*(22.28)*	*(0.15)*	*(0.39)*	*(40.44)*	*(1.50)*	*(3.00)*	*(1.50)*
WMW U-Test p-values (two-sided):									
Baseline vs. RelInfo	0.094*	0.054*	0.072*	0.029**	0.635	0.152	0.831	0.077*	0.258
Baseline vs. Tournament	0.671	0.118	0.175	0.671	0.660	0.413	0.400	0.046**	0.196
RelInfo vs. Tournament	0.120	0.264	0.383	0.093*	0.869	0.192	0.795	0.555	0.820

RelInfo market versus only 52% in *Baseline*. The bubble-magnitude measures *Amplitude, Total Dispersion,* and *Average Bias* are also significantly smaller in the *RelInfo* treatment at the 10% level (p-value = 0.094, 0.054, and 0.072 respectively), while the bubble-length measure *Boom Duration* is smaller in *RelInfo* by a similar margin to what is observed with asset X, but this time attaining only marginal significance (p-value = 0.077).

In a further contrast to Round 1, a comparison of the *RelInfo* treatment against the *Tournament* treatment (Hypothesis 2) in Table 4.4 provides some evidence that bubbles are larger in the latter when participants are experienced. For asset X, the relative median values of all of the bubble measures (except *Bust Duration*) suggest greater mispricing in the *Tournament* treatment, although a significant difference is detected in only one measure, that too at only the 10% level – *Boom Duration* (p-value = 0.099), which is 7 periods long in the median *Tournament* market compared to 2.5 in *RelInfo.* Similarly for Asset Y, where even though most measures again point to higher prices and larger bubbles in the median market of the *Tournament* treatment (*Bust Duration* again being the exception), just one of the measures reports a (marginally) significant difference – *Haessel* R^2 is higher in the *RelInfo* treatment ($\text{median}_{RelInfo}$ = 92% vs. $\text{median}_{Tournament}$ = 78%; p-value = 0.093)[116].

Although the random assignment of participants to treatments should ensure that the treatment groups are probabilistically equivalent at the outset of the experiment, we nonetheless confirm that the observed differences in price behaviour between treatments

[116] Breaking the *Tournament* treatment down to its constituent contracts, *Carrot* and *Stick*, and comparing them individually to the *RelInfo* treatment produces qualitatively similar results (unreported). For asset X, none of the bubble measures in either *Carrot* or *Stick* are statistically significantly different from those of *RelInfo*. For asset Y, we do not detect a statistically significant difference between *RelInfo* and the *Stick* contract on any bubble measure, however we do find that *Haessel* R^2 is marginally significantly lower (p-value = 0.08), and *Amplitude* is marginally significantly higher (p-value = 0.08) in *Carrot* markets compared to *RelInfo.*

are not driven by disparities between them in participants' risk-attitudes or cognitive abilities. Notwithstanding the sourcing of the data post-market, the averages of the DOSPERT and CRT scores reported in Table 4.1 attest to their likely equivalence between treatments. In unreported WMW tests, we also fail to find a significant difference between any of the treatments in CRT or DOSPERT scores. In contrast, Baghestanian et al. (2015) find that providing information about peers impacts participants' risk attitudes; peer information in their experimental market is associated with a reduction in participants' willingness to take risk in a post-market risk elicitation task (the 'bomb risk elicitation task'). However, numerous differences in the design of our respective studies make it difficult to identify the source(s) of this discrepancy with any confidence. It may lie in the type of relative information used (participants in their "Info" treatment can observe others' portfolios), but could also be due to considerable disparities in the market design, and/or the type of risk elicitation task itself.

Summarising the results from the second round of trading, the evidence suggests that with once-experienced traders, the incremental effect of providing participants with relative-performance information is to actually *reduce* the size and duration of bubbles compared to normal-incentive markets in which this information is not provided. On the other hand, introducing competitive monetary incentives into a trading environment where relative-performance information is already provided has the *opposite* effect, though its effect is relatively weak. Although Dijk et al. (2014) only examine an inexperienced cohort, a parallel found here with their individual-level results is that we also find stronger evidence of the impact made by the social comparison element of tournaments than the attendant monetary incentives. Moreover, these effects appear to largely offset each other in our sample, as suggested by the lack of statistical significance achieved on all bar one bubble measure when comparing the *Baseline* treatment against the *Tournament* treatment – *Boom Duration* remains significantly lower in the *Tournament* treatment for asset Y (p-value = 0.046).

To understand why it is that aggregate-level differences between the treatments emerge with experienced participants, refer to Table 4.5. For both assets X (Panel A) and Y (Panel B), it reports two-sided p-values from Wilcoxon Signed-rank tests conducted on each bubble measure within each treatment, under the null hypothesis that there is no difference between Rounds 1 and 2. The results show that whereas mispricing/bubbles dissipate between rounds in both the *RelInfo* and *Tournament* treatments, no such 'improvement' appears to occur in the *Baseline* treatment. In fact, the only variable that experiences a statistically significant improvement in the *Baseline* treatment is the indirect measure of bubble-magnitude, *Turnover,* in asset Y (p-value = 0.018). The degree of mispricing on the same asset, as measured by *Total Dispersion,* is actually *greater* in Round 2 in the *Baseline* treatment, albeit only marginally significantly (p-value = 0.091). In contrast, the statistical significance of *Total Dispersion*, *Haessel-R^2*, *Turnover,* and *Normalised Deviation* in both assets lays testament to the diminished magnitude of bubbles in the *RelInfo* treatment in Round 2 (p-values$_X$ = 0.017, 0.017, 0.012, and 0.012 respectively; p-values$_Y$ = 0.0499, 0.017, 0.012, and 0.017 respectively). In addition, *Duration* is also marginally significantly smaller in Round 2 for asset Y in *RelInfo* (p-value = 0.085). While the degree of 'improvement' in bubble measures from Round 1 to 2 is generally greater in the *RelInfo* treatment than in the *Tournament* treatment (judged from their median values), the latter also exhibits similar declines in the size of its bubbles. In *Tournament,* statistically significant improvements in *Turnover*, *Normalised Deviation,* and *Haessel-R^2* are reported in asset X, marginally so for the latter measure (p-values = 0.001, 0.003, and

Table 4.5: Comparing bubble measures between rounds

This table reports the results of within-treatment comparisons of the bubble measures between market rounds in the *Baseline*, *RelInfo*, and *Tournament* treatments. Markets contaminated by subjects who had participated in an earlier session are excluded. The values shown below are p-values from a two-sided Wilcoxon signed-rank test of the null hypothesis that bubble measure values do not differ significantly between rounds 1 and 2. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively

Panel A: Asset X, Round 1 vs. Round 2:

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	1.000	0.499	0.128	0.499	0.176	0.866	0.317	0.230	0.333
RelInfo [8]	0.327	0.017**	0.401	0.017**	0.012**	0.012**	0.226	0.177	1.000
Tournament [16]	0.469	0.679	0.501	0.070*	0.001***	0.003***	0.534	0.287	0.774

Panel B: Asset Y, Round 1 vs. Round 2:

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	0.735	0.091*	0.128	0.612	0.018**	0.866	0.475	0.932	0.795
RelInfo [8]	0.208	0.050**	0.327	0.017**	0.012**	0.017**	0.085*	0.320	0.320
Tournament [16]	0.438	0.079*	0.469	0.044**	0.001***	0.004***	0.001***	0.046**	0.100

0.07 respectively). Furthermore, asset Y in the same treatment exhibits significant improvements in *Haessel-R^2*, *Turnover*, *Normalised Deviation*, *Duration*, *Boom Duration,* and a marginally significant improvement in *Total Dispersion* (p-values = 0.044, 0.001, 0.004, 0.001, 0.046, and 0.079 respectively).

Thus the driving force behind the emergence of peer effects in Round 2 appears to be the ‘improvement’ in price behaviour in the *RelInfo* treatment, or equally the absence of a discernible improvement in the *Baseline* treatment. Given that the bubble-reducing effect of trading experience under normal incentives is an empirical regularity in single-asset markets (e.g. Smith et al. 1988; King et al. 1993; Dufwenberg et al. 2005; Haruvy et al. 2007), the lack of obvious improvement in the *Baseline* treatment – particularly for asset Y, where most bubble measure medians actually worsen – is a decidedly surprising outcome[117]. It is especially so since the normal incentive-based *RelInfo* treatment *does* exhibit diminished bubbles in Round 2. Hence, how experience affects price behaviour in multi-asset markets – something that, to our knowledge, this study is the first to examine – is possibly an issue that warrants more attention in the literature.

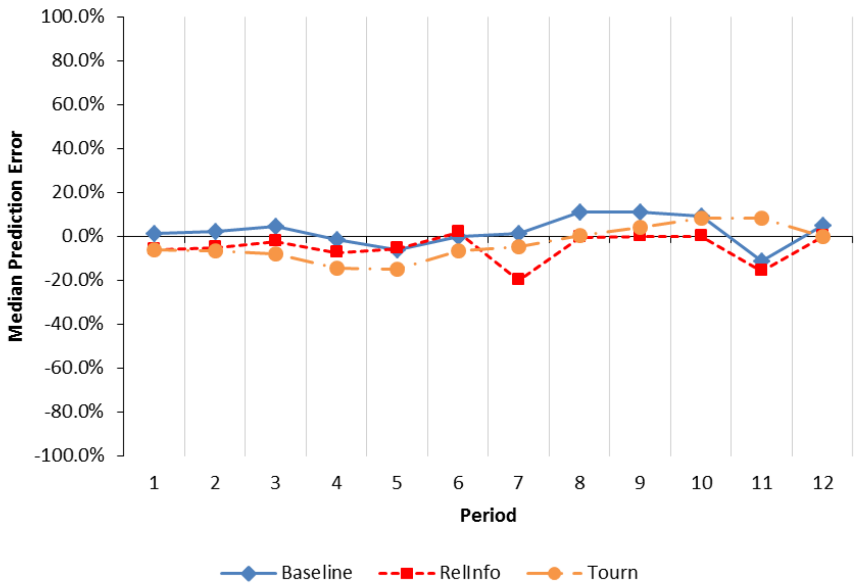
Relative prices

Figure 4.5 plots the evolution of the median *Prediction Error* in each treatment in Round 2. While the median *Prediction Error* generally remains close to the risk-neutral benchmark of zero in all treatments, the graph appears to suggest that asset Y sells at a persistent discount to asset X during the first half of the market in the *Tournament* treatment – the average of the median *Prediction Errors* in the *Tournament*

[117] Under normal incentives, bubbles in Smith et al. (1988) markets generally disappear by the third round of the market, i.e. with twice-experienced traders. In this study, experienced traders were only once-experienced. While this may possibly explain the lack of significant improvement in price behaviour, Haruvy et al. (2007) note that bubbles dissipate steadily across rounds, with generally smaller bubbles that peak earlier being observed in round 2. In this regard, the behaviour of asset Y in the *Baseline* treatment remains puzzling.

Figure 4.5: Median values of *Prediction Error*, Round 2

The figure below plots the evolution of the median *Prediction Error* in the *Baseline* (solid blue line), *RelInfo* (dotted red line), and *Tournament* (dashed orange line) treatments during the second round of the market (i.e. with experienced traders). For each treatment, the plotted value in each period is the median of the *Prediction Errors* from all markets in that treatment. *Prediction Error* is defined as the percentage deviation of the relative price of Y (i.e. median price of asset Y divided by the median price of asset X) from the risk-neutral benchmark of 1. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



treatment in rounds 1-6 is –9.5%. Confirmatory evidence is provided by the results of a one-sample Signed-rank test, which finds that median *Prediction Errors* in the *Tournament* treatment are significantly less than zero in periods 2 through to 6 (p-value = 0.013, 0.007, 0.007, 0.039, 0.031, and 0.088 respectively). However, importantly this result does not translate into statistically significant differences between the *Tournament* treatment and the *RelInfo* treatment, although WMW tests do reveal that the *Baseline* treatment witnessed marginally significantly higher *Prediction Errors* than the *Tournament* treatment in periods 2, 3 and 6 (p-value = 0.084, 0.071, and 0.075 respectively).[118]

Average Prediction Errors in Round 2, which are summarised in Panel B of Table 4.3, reflect a similar picture. Like Round 1, the median of the whole-of-market measure, "Avg PredErr" is small in all treatments. It is not statistically significantly different from zero in any treatment, except *RelInfo*, where Y sells at a significant discount to X. Although only modest in size (5.6%) and of marginal statistical significance (one-sample Wilcoxon Signed-rank test p-value = 0.093), this discount is consistent with the smaller bubbles seen in the *RelInfo* treatment; both of these observations indicate the predominance of risk-averse behaviour in this treatment, which is consistent with past studies that find that relative comparisons are associated with risk-aversion (e.g. Linde and Sonnemans 2012; Vendrik and Woltjer 2007; Baghestanian et al. 2015). Notwithstanding, WMW tests comparing the treatments do not reveal a statistically significant difference between any of the treatments on the whole-of-market *Average Prediction Error* measure, or its two components.

Like Round 1, relative prices in the *Baseline* treatment display a significant change in behaviour from the first half of the market to the second, with *Average Prediction Error* in the median *Baseline* market rising from -0.12% in the first half to 5.23% in the second (one-sided p-value = 0.032). *Tournament* traders exhibit a similar propensity to pay more in the second half of the market to acquire the riskier asset Y relative to the price paid for X (one-sided p-value = 0.025); the *Average Prediction Error* in the median *Tournament* market goes from a significant risk-averse discount of -7.45% (one-sample Wilcoxon Signed-rank test p-value = 0.015) in the first half to a

[118]A marginally significant difference is also detected between *Baseline* and *RelInfo*, but only in period 2 (p-value = 0.085)

very small premium of 1.87% in the second[119]. However, given the modest and statistically insignificant nature of the premiums paid for asset Y in the second half of the market in both the *Baseline* and *Tournament* treatments, it is difficult to interpret these results as evidence of *risk-seeking* behaviour per se, although it does suggest that traders became less risk-averse in the *Baseline* and *Tournament* treatments as the market progressed. In contrast, the *RelInfo* treatment is characterised by the absence of such an effect – if anything, the second half of the market in *RelInfo* sees Y trading at a marginally significant discount to X of approximately 5% (one-sample Wilcoxon Signed-rank test p-value = 0.069) – which could be interpreted as additional evidence that relative performance feedback is associated with greater risk aversion.

4.4.3 Rank-order tournaments

The discussion below relates to our alternative tournament contract, the rank-order based *GilTournament* treatment. As the impact of relative performance information has been covered in the preceding analysis, we restrict the following discussion to the incremental effect of tournament incentives, as represented by *GilTournament.* In doing so, we note that the analysis below is subject to an important caveat arising from the fact that the *GilTournament* treatment differs from the *RelInfo* treatment by not only the type of compensation scheme (tournament vs. normal), but also the fact that subjects in the *GilTournament* treatment were provided with an additional piece of relative performance information – rank. Therefore, while we may refer to any differences between *GilTournament* and *RelInfo* as tournament

[119] This is driven by the behaviour of relative prices in the ‘Carrot’ sub-treatment, where the median *Average Prediction Error* of 15.04% in the second half is significantly higher than the corresponding -12.70% in the first (one-sided p-value = 0.013). In comparison, the ‘Stick’ sub-treatment does not exhibit a significantly higher *Average Prediction Error* in the second half of the market; its first half median is -3.68% vs. -6.62% in the second.

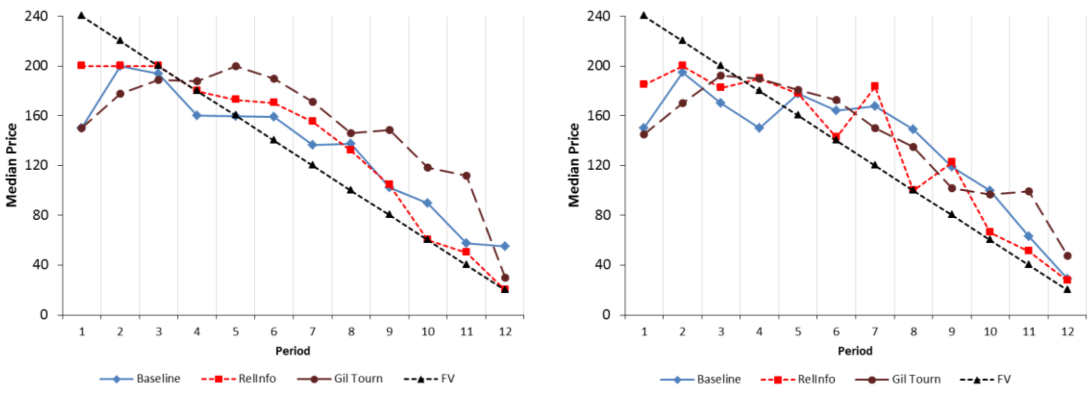
compensation-driven, we cannot definitively conclude that any effects observed are solely due to the compensation scheme.

<u>Individual assets</u>

The median-price behaviour of assets X and Y in the inexperienced markets of the *GilTournament* treatment is illustrated in Panels (a) and (b) of Figure 4.6 respectively, alongside the *Baseline* and *RelInfo* treatments. The graphs reveal that while the behaviour of the 'risky' asset Y in the *GilTournament* treatment does not appear to differ noticeably from the other treatments, the 'safe' asset X in the *GilTournament* treatment exhibits persistently higher median prices and more overvaluation than both the *Baseline* and *RelInfo* treatments, specifically in the middle

Figure 4.6: Median prices, *GilTournament*, Round 1

Median transaction prices in the rank-order tournament treatment *GilTournament* (dashed brown line) are compared to the *Baseline* (solid blue line) and *RelInfo* (dotted red line) treatments during the first round of the market (i.e. with inexperienced traders) below. Panel (a) reports the behaviour of the 'low-risk' asset X, while Panel (b) covers the 'high-risk' asset Y. The risk-neutral fundamental value process for each asset (dotted black line) is shown in both (a) and (b). For each treatment, the plotted median price in each period is the median of the median transaction prices from all markets belonging to that treatment. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



(a) Asset X, Round 1

(b) Asset Y, Round 1

and latter periods of the market.

The associated bubble measure data, summarised in Table 4.6, support these observations. For all bubble measures in asset Y (Panel B), the two-sided WMW tests fail to reject the null hypothesis of no difference between *GilTournament* and the *RelInfo* treatment (or *Baseline).* In contrast, compared to the *RelInfo* treatment, asset X (Panel A) in the *GilTournament* treatment experiences significantly larger *Amplitudes* ($median_{GilTournament}$ = 2.51 vs. $median_{RelInfo}$ = 1.03; p-value = 0.047) and significantly longer *Boom Durations* ($median_{GilTournament}$ = 8 vs. $median_{RelInfo}$ = 5.5; p-value = 0.049). The combined mispricing and trading activity measure *Normalised Deviation* is also significantly higher in the *GilTournament* treatment for X, but only at the 10% level ($median_{GilTournament}$ = 213.54 vs. $median_{RelInfo}$ = 79.20; p-value = 0.069)[120]. The data also provide some evidence that mispricing in asset X is greater in the *GilTournament* treatment than the *Baseline* treatment with inexperienced traders. A smaller *Normalised Deviation* ($median_{GilTournament}$ = 213.54 vs. $median_{Baseline}$ = 83.43) and higher *Haessel-*R^2 ($median_{GilTournament}$ = 56% vs. $median_{Baseline}$ = 78%) is observed in the latter treatment, though both differences are significant at only the 10% level (p-value = 0.068, p-value = 0.056).

In markets comprising experienced traders, the results in both assets suggest a strengthening of the effect of monetary incentives attached to tournaments. Price behaviour in the second round of trading is summarised graphically in Figure 4.7 and via bubble measures in Table 4.7; in both, Panel A (B) relates to asset X (Y). For asset

[120] Compared to the *GilTournament* treatment, evidence of differences between its sub-treatments and the *RelInfo* treatment is relatively weak. This is unsurprising, since the power of statistical tests diminish as sample sizes become smaller. While full results are not reported here, we detect a marginally significant difference between 'GilStick' and *RelInfo* on one measure for asset X, *Amplitude* (p-value = 0.081). We do not detect a significant difference between 'GilCarrot' and *RelInfo* in asset X. For asset Y, like the *GilTournament* treatment, we do not detect any significant differences between its sub-treatments and the *RelInfo* treatment.

Table 4.6: Summary of bubble measures in Round 1 using *GilTournament*

This table reports median values of each bubble measure in the *Baseline*, *RelInfo*, and *GilTournament* treatments during Round 1; median absolute deviations are displayed in parentheses. Markets contaminated by subjects who had participated in an earlier session are excluded. Panel A (B) reports bubble measure data relating to Asset X (Y). For definitions of the relevant bubble measures, see section 4.4.1.2. The statistical significance of the difference between treatments in each measure is assessed using a two-sided Wilcoxon Mann-Whitney U Test, under the null that values from both treatments come from the same distribution. Exact p-values are reported. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively.

Panel A: Asset X, Round 1:

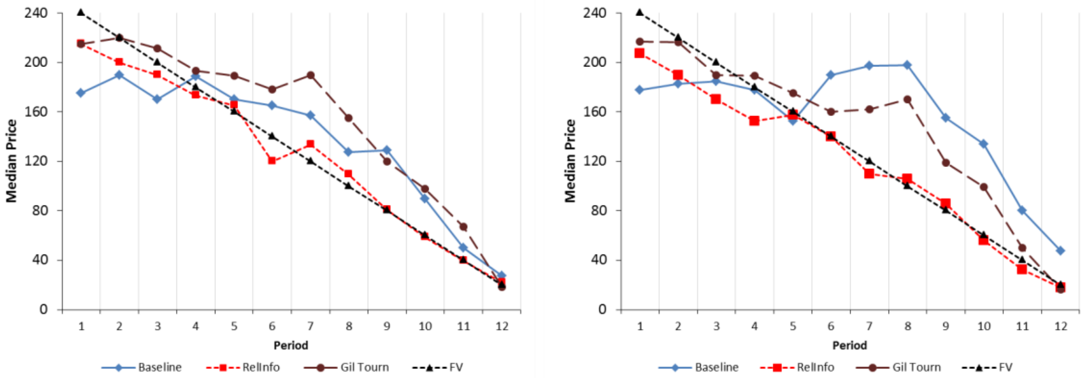
Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	3.31	298.00	1.65	0.78	2.45	83.43	5.00	5.00	4.00
	(1.53)	*(181.50)*	*(15.73)*	*(0.16)*	*(0.72)*	*(45.08)*	*(2.00)*	*(2.00)*	*(1.00)*
RelInfo [8]	1.03	457.25	11.08	0.57	2.25	79.20	6.50	5.50	2.00
	(0.44)	*(284.25)*	*(29.48)*	*(0.18)*	*(0.66)*	*(37.14)*	*(2.50)*	*(2.00)*	*(1.00)*
GilTournament [12]	2.51	714.75	28.67	0.56	2.93	213.54	6.00	8.00	3.00
	(1.24)	*(229.00)*	*(21.13)*	*(0.17)*	*(0.70)*	*(95.98)*	*(1.50)*	*(1.50)*	*(1.00)*
WMW U-Test p-values (two-sided):									
Baseline vs. RelInfo	0.189	0.955	1.000	0.536	0.779	0.955	0.956	0.845	0.977
Baseline vs. GilTournament	0.902	0.299	0.340	0.056*	0.137	0.068*	0.821	0.122	0.854
RelInfo vs. GilTournament	0.047**	0.343	0.395	0.521	0.262	0.069*	0.748	0.049**	0.632

Panel B: Asset Y, Round 1

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	1.63	530.50	15.38	0.77	2.03	99.06	4.00	7.00	4.00
	(1.01)	*(150.50)*	*(25.33)*	*(0.05)*	*(0.35)*	*(54.11)*	*(1.00)*	*(1.00)*	*(2.00)*
RelInfo [8]	1.25	527.00	11.04	0.59	1.90	80.51	4.00	4.00	2.50
	(0.75)	*(221.25)*	*(26.94)*	*(0.18)*	*(0.39)*	*(23.66)*	*(1.00)*	*(1.00)*	*(0.50)*
GilTournament [12]	2.97	575.50	2.62	0.59	2.70	140.37	5.00	5.50	3.00
	(1.45)	*(181.75)*	*(33.88)*	*(0.28)*	*(0.68)*	*(65.40)*	*(2.00)*	*(2.50)*	*(1.00)*
WMW U-Test p-values (two-sided):									
Baseline vs. RelInfo	0.463	0.955	0.867	0.694	0.867	0.779	0.943	0.174	0.844
Baseline vs. GilTournament	0.650	0.902	0.482	0.773	0.482	0.536	0.784	0.386	0.852
RelInfo vs. GilTournament	0.208	0.792	0.910	0.624	0.521	0.181	0.930	0.523	0.923

Figure 4.7: Median prices, *GilTournament*, Round 2

Median transaction prices in the *Baseline* (solid blue line), *RelInfo* (dotted red line), and *GilTournament* (dashed brown line) treatments during the second round of the market (i.e. with experienced traders) are shown below for the 'low-risk' asset X (panel (a)) and 'high-risk' asset Y (panel (b)), along with the risk-neutral fundamental value process for each asset (dotted black line). For each treatment, the plotted median price in each period is the median of the median transaction prices from all markets belonging to that treatment. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



(a) Asset X, Round 2

(b) Asset Y, Round 2

X, the median price in Fig. 4.7(a) is higher in the *GilTournament* treatment than the *RelInfo* treatment in all but the first and last trading periods, with the largest differences occurring mid-market, which is when the *GilTournament* bubble peaks. Consistent with this, virtually all of the bubble-magnitude measures for asset X point to significantly greater overvaluation in *GilTournament* than *RelInfo* in Round 2, most notably *Average Bias*, which indicates that the median price for asset X in the median *GilTournament* market is on average about 37 francs above FV in each period compared to only 2 francs (approx.) in *RelInfo*, a difference that is significant at the 1% level (p-value = 0.007). The *GilTournament* treatment also exhibits a significantly larger *Amplitude* ($\text{median}_{GilTournament}$ = 1.35 vs. $\text{median}_{RelInfo}$ = 0.55; p-value = 0.031), *Normalised Deviation* ($\text{median}_{GilTournament}$ = 108.35 vs. $\text{median}_{RelInfo}$ = 22.93; p-value = 0.02), and *Turnover* ($\text{median}_{GilTournament}$ = 2.55 vs. $\text{median}_{RelInfo}$ = 1.5; p-value = 0.03) than *RelInfo*, while *Total Dispersion* is also higher in *GilTournament* but only at the 10% level (median$_{GilTournament}$ = 457.45 vs. median$_{RelInfo}$ = 237.75; p-value = 0.082). Furthermore, the greater persistence of overvaluation in *GilTournament* that is evident in Fig.4.7(a) is reflected in the significant difference in *Boom Duration* (median$_{GilTournament}$ = 7.5 vs. median$_{RelInfo}$ = 2.5; p-value = 0.025)[121].

Asset Y in Round 2 also tells a similar story. The median price in Fig. 4.7(b) is higher in the *GilTournament* treatment than the *RelInfo* treatment in every period except the last. This is reflected in the bubble measures (Table 4.7, Panel B) via a significantly higher *Average Bias* (median$_{GilTournament}$ = 32.38 vs. median$_{RelInfo}$ = -6.86; p-value = 0.025), larger *Amplitude* (median$_{GilTournament}$ = 1.75 vs. median$_{RelInfo}$ = 0.7; p-value = 0.025), and longer *Boom Duration* (median$_{GilTournament}$ = 7 vs. median$_{RelInfo}$ = 2.5; p-value = 0.023) in the *GilTournament* treatment, which also reports a marginally significantly lower *Haessel-R^2* (median$_{GilTournament}$ = 82% vs. median$_{RelInfo}$ = 92%; p-value = 0.069)[122].

What drives the broader statistical significance of the 'tournament effect' in Round 2 compared to Round 1 under a rank-order tournament? The answer, a comparison of Figures 4.6 and 4.7 strongly suggests, appears to be a reduction in mispricing between rounds in the *RelInfo* treatment (see above, and Table 4.5) that is generally not matched by the *GilTournament* treatment. Table 4.8, which reports the results of two-tailed Wilcoxon Signed-rank tests comparing the round-to-round change in each bubble measure in the *GilTournament* treatment, shows that for Asset X, only *Turnover* experiences an 'improvement' between rounds that is significant at the 5%

[121] The primary driver of these differences is the 'GilCarrot' sub-treatment, where we find that differences in *Average Bias* and *Amplitude* are significant at the 1% level, *Normalised Deviation* and *Boom Duration* are significant at the 5% level, and *Total Dispersion*, *Haessel R^2*, and *Turnover* are significant at the 10% level. The only significant difference between 'GilStick' and *RelInfo* is in *Turnover* at the 10% level.
[122] For asset Y, we find significant differences between 'GilCarrot' and *RelInfo* in *Amplitude* and *Boom Duration* at the 5% level, and *Average Bias* at the 10% level. *Average Bias* is also significantly different between 'GilStick' and *RelInfo* at the 10% level, as is *Boom Duration* and *Bust Duration.*

<u>Table 4.7: Summary of bubble measures in Round 2 using *GilTournament*</u>

This table reports median values of each bubble measure in the *Baseline*, *RelInfo*, and *GilTournament* treatments during Round 2; median absolute deviations are displayed in parentheses. Markets contaminated by subjects who had participated in an earlier session are excluded. Panel A (B) reports bubble measure data relating to Asset X (Y). For definitions of the relevant bubble measures, see section 4.4.1.2. The statistical significance of the difference between treatments in each measure is assessed using a two-sided Wilcoxon Mann-Whitney U Test, under the null that values from both treatments come from the same distribution. Exact p-values are reported. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively.

<u>Panel A: Asset X, Round 2:</u>

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	1.15	323.50	10.96	0.82	1.60	39.45	3.00	8.00	3.00
	(0.79)	*(247.00)*	*(13.55)*	*(0.18)*	*(0.69)*	*(32.05)*	*(1.00)*	*(1.00)*	*(1.00)*
RelInfo [8]	0.55	237.75	1.54	0.94	1.50	22.93	3.50	2.50	3.00
	(0.12)	*(82.25)*	*(10.82)*	*(0.02)*	*(0.47)*	*(3.29)*	*(1.50)*	*(2.50)*	*(2.00)*
GilTournament [12]	1.35	457.75	36.60	0.87	2.55	108.35	4.00	7.50	2.50
	(0.64)	*(203.75)*	*(23.72)*	*(0.08)*	*(0.80)*	*(64.33)*	*(1.00)*	*(2.50)*	*(1.00)*
<u>WMW U-Test p-values (two-sided):</u>									
Baseline vs. RelInfo	0.232	0.336	0.281	0.094*	0.482	0.121	0.618	0.029**	0.707
Baseline vs. GilTournament	0.773	0.902	0.650	0.482	0.218	0.837	0.785	0.918	0.819
RelInfo vs. GilTournament	0.031**	0.082*	0.007***	0.115	0.030**	0.020**	0.773	0.025**	0.524

<u>Panel B: Asset Y, Round 2</u>

Treatment [N]	*Amplitude*	*Total Dispersion*	*Average Bias*	*Haessel R^2*	*Turnover*	*Normalised Deviation*	*Duration*	*Boom Duration*	*Bust Duration*
Baseline [7]	2.16	626.00	24.33	0.52	1.50	127.50	5.00	7.00	3.00
	(0.72)	*(492.50)*	*(68.88)*	*(0.13)*	*(0.47)*	*(64.98)*	*(3.00)*	*(4.00)*	*(2.00)*
RelInfo [8]	0.70	310.75	-6.86	0.92	1.23	32.16	3.00	2.50	4.50
	(0.29)	*(143.75)*	*(15.06)*	*(0.03)*	*(0.27)*	*(11.73)*	*(1.50)*	*(2.00)*	*(1.50)*
GilTournament [12]	1.75	532.50	32.38	0.82	1.80	80.15	5.00	7.00	2.50
	(0.71)	*(243.00)*	*(29.39)*	*(0.13)*	*(0.50)*	*(35.87)*	*(2.50)*	*(2.50)*	*(1.50)*
<u>WMW U-Test p-values (two-sided):</u>									
Baseline vs. RelInfo	0.094*	0.054*	0.072*	0.029**	0.635	0.152	0.831	0.077*	0.258
Baseline vs. GilTournament	0.837	0.299	0.592	0.432	0.351	0.650	0.985	0.721	0.970
RelInfo vs. GilTournament	0.025**	0.135	0.025**	0.069*	0.156	0.115	0.833	0.023**	0.157

Table 4.8: Comparing *GilTournament* bubble measures between rounds

This table reports the results of a within-treatment comparison of the bubble measures of the *GilTournament* treatment between Round 1 and 2. Markets contaminated by subjects who had participated in an earlier session are excluded. The values shown below are p-values from a two-sided Wilcoxon signed-rank test of the null hypothesis that bubble measure values do not differ significantly between rounds 1 and 2. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively

Asset	*Amp*	*Tot Disp*	*Avg Bias*	*H-R*^{2}	*Turn*	*Norm Dev*	*Dur*	*Boom Dur*	*Bust Dur*
Asset X	0.071*	0.158	0.638	0.060*	0.050**	0.084*	0.124	0.133	0.473
Asset Y	0.638	0.695	0.136	0.239	0.003***	0.158	0.752	0.406	0.551

level (p-value = 0.0499). Yet even here, the improvement is modest, with the median falling from 2.93 in Round 1 to 2.55 in Round 2, compared to the corresponding change in the *RelInfo* treatment from 2.25 to 1.55. Other measures for asset X in the *GilTournament* treatment such as *Amplitude, Haessel-R*^{2}, and *Normalised Deviation* also show improvement, but only at the 10% significance level (p-values = 0.071, 0.06, and 0.084 respectively). Furthermore, in asset Y, the only variable to show a statistically significant difference between rounds is *Turnover* (p-value = 0.003).

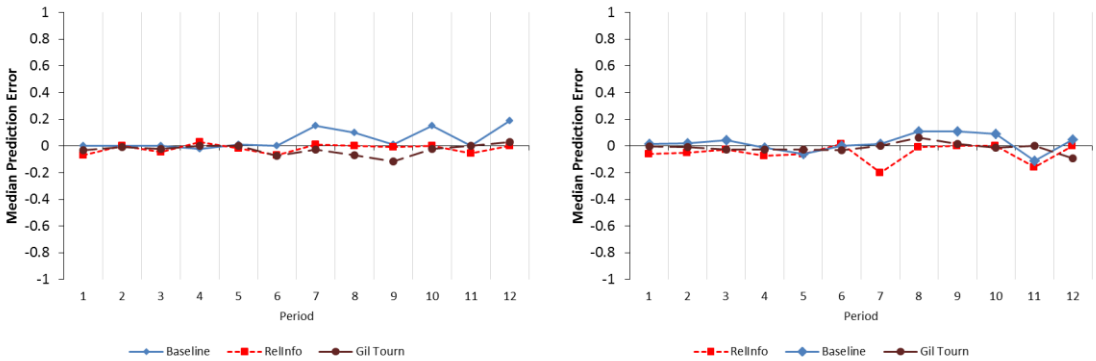
In summary, our findings from the *GilTournament* treatment suggest that introducing competitive monetary incentives based on rank has the effect of exacerbating asset price bubbles, both in terms of size and duration. While this is true for only the low-risk asset in markets containing inexperienced traders, it is more broadly the case when traders are experienced. Of course, these findings are subject to the aforementioned caveat that the additional ranking information in *GilTournament* potentially confounds the results. Nonetheless, *if* ranking information has the same aggregate impact as information about the average trader – that is, to reduce mispricing – then it is possible that these results understate the bubble-magnifying role of rank-order based compensation. Moreover, these results are also much stronger than those obtained with the *Tournament* treatment, where ‘beat-the-market' monetary incentives appear to have little effect on prices with inexperienced traders, and a relatively weak effect with experienced traders, albeit one that works in the same direction as *GilTournament.* Consequently, our results potentially suggest that the type of tournament compensation scheme plays an important role in determining how competitive monetary incentives impact prices above and beyond the effect of relative performance information.

<u>Relative prices</u>

In contrast to the price-behaviour of the individual assets, introducing rank-order based compensation to the market does not appear to systematically affect relative prices. This can be seen in Figures 4.8(a) and (b), which show the evolution of the median *Prediction Error* in the *GilTournament* treatment in Rounds 1 and 2 respectively, along with the *Baseline* and *RelInfo* treatments. The graphs show that

Figure 4.8: Median values of *Prediction Error*, *GilTournament*, Rounds 1 and 2

The evolution of the median *Prediction Error* in the *Baseline* (solid blue line), *RelInfo* (dotted red line), and *GilTournament* (dashed brown line) treatment is shown below for Round 1of the market in panel (a) and Round 2 in panel (b). For each treatment, the plotted value in each period is the median of the *Prediction Errors* from all markets in that treatment. *Prediction Error* is defined as the percentage difference between the relative price of Y (i.e. median price of asset Y divided by median price of asset X) and the risk-neutral benchmark of 1. Any markets that were 'contaminated' by the presence of subjects who had participated in an earlier session of the experiment are excluded.



(a) Round 1 (b) Round 2

relative prices in the *GilTournament* treatment largely conform to theoretical expectations in both rounds of the market. Furthermore, in both rounds of the market, significant differences in *Prediction Error* are not detected between *GilTournament* and the *Baseline* or *RelInfo* treatments in most trading periods.

Average Prediction Errors, which are summarised in Table 4.9, paint the same picture. The median values of all of the measures are very small in the *GilTournament* treatment in both rounds (in all cases, less than 3.5% in absolute terms), and none are found to be significantly different from zero (unreported). In addition, using the WMW test we cannot reject the null of no difference between *GilTournament* and the other two treatments for any of the measures. Furthermore, like the *Tournament* treatment, *Average Prediction Errors* in the *GilTournament* treatment in Round 2 are statistically significantly higher in the second half of the market compared to the first half, but this time at the 10% level (one-sided p-value = 0.091)[123]. However once again, the practical significance of this difference is negligible, as the median only shifts from –2.43% to –0.58%.

4.5 Conclusion

Recent evidence suggests that the severe bubbles exhibited by prices in experimental asset markets under tournament conditions may be driven by traders' intrinsic desire for rank, fuelled by the availability of relative performance information. This study investigates this possibility by isolating the incremental impacts of relative performance information and tournament-related monetary

[123] This is mostly driven by the 'GilStick' sub-treatment, which returns a significant result at the 10% level on the same test, while we fail to reject the null no difference between Round 1 and 2 in 'GilCarrot'.

Table 4.9: Average Prediction Errors – *GilTournament*

Median values of the *Average Prediction Error* in the *Baseline*, *RelInfo*, and *GilTournament* treatments in Round 1 and 2 are shown below in Panels A and B respectively, with the associated median absolute deviations shown in parentheses. Markets contaminated by subjects who had participated in an earlier session are excluded. *Average Prediction Error* is calculated using all periods in a market, the first 6 periods, and the final 6 periods in *Avg PredErr*, *AvgPredErr_p1to6*, and *AvgPredErr_p7to12* respectively. The statistical significance of the individual measures is assessed using a (two-sided) one-sample Wilcoxon Signed-rank test, under the null that the median is equal to zero. The statistical significance of the difference between treatments is assessed using the Wilcoxon Mann-Whitney U Test under the null that values from both treatments come from the same distribution. The statistical significance of the difference between *AvgPredErr_p1to6* and *AvgPredErr_p7to12* within each treatment is assessed using a (paired-sample) Wilcoxon Signed-rank test, against the one-sided alternative hypothesis that *AvgPredErr_p7to12* > *AvgPredErr_p1to6*. Differences that are significant at the 10%, 5% and 1% level are denoted by *, **, and ***, respectively.

Panel A: Round 1

Treatment [N]	*Avg PredErr (%)*	*AvgPredErr_ p1to6 (%)*	*AvgPredErr_ p7to12 (%)*	*Signed-rank p-value (1-sided) 1-6 vs. 7-12*
Baseline [7]	3.78	-0.32	7.88	0.064*
	(9.46)	*(3.36)*	*(9.26)*	
RelInfo [8]	-0.77	-1.44	-1.89	0.200
	(5.16)	*(3.45)*	*(10.36)*	
GilTournament [12]	-0.21	-3.21	1.02	0.377
	(9.38)	*(5.40)*	*(14.35)*	
WMW U-Test p-values (2-sided):				
Baseline vs. RelInfo	0.397	0.536	0.336	
Baseline vs. GilTournament	0.196	0.837	0.299	
RelInfo vs. GilTournament	0.792	0.970	0.734	

Panel B: Round 2

Treatment [N]	*Avg PredErr (%)*	*AvgPredErr_ p1to6 (%)*	*AvgPredErr_ p7to12 (%)*	*Signed-rank p-value (1-sided) 1-6 vs. 7-12*
Baseline [7]	2.72	-0.12	5.23	0.032**
	(11.33)	*(2.27)*	*(21.08)*	
RelInfo [8]	-5.60*	-5.68	-4.94*	0.663
	(5.74)	*(4.92)*	*(3.46)*	
GilTournament [12]	-0.69	-2.43	-0.58	0.091*
	(6.63)	*(5.89)*	*(6.90)*	
WMW U-Test p-values (2-sided):				
Baseline vs. RelInfo	0.152	0.232	0.121	
Baseline vs. GilTournament	0.650	0.536	0.592	
RelInfo vs. GilTournament	0.238	0.384	0.208	

payoffs on asset prices in an experimental market where subjects can trade two differentiated assets. The results suggest that relative performance information and tournament payoffs have distinct impacts on asset prices, but generally only when traders are experienced. When experienced traders are compensated according to their absolute performance (‘normal’ incentives), supplying information about the performance of the ‘average’ trader – akin to informing traders of their performance relative to the market index – serves to *reduce* the size and duration of price bubbles compared to markets where this information is not available, especially for relatively high-risk assets. In contrast, adding tournament compensation to markets where relative performance information is provided results in *larger* bubbles compared to relative performance information-only markets. This effect is relatively weak when the tournament compensation contract is based on a ‘beat-the-market’ scheme, but much stronger when traders are compensated according to rank. Moreover, the offsetting nature of the effects of relative performance information and tournament compensation means that prices in tournament-based markets generally behave similarly to normal incentive markets where no relative information in provided.

These results underscore the importance of social comparison and ‘peer effects’ as a determinant of aggregate behaviour. In doing so, they also suggest potential directions for future research. First, given the aforementioned possible source of confound introduced by the presence of ranking information in the rank-order tournament treatment, future studies could illuminate whether relative performance feedback in the form of rank affects price behaviour differently to ‘market index’ information. This is particularly important because rankings information in the form of league-tables is an especially common method of communicating relative performance information in the real world, particularly in the funds management industry.

Second, the experimental design in this study is an incomplete factorial design, in that it does not interact ‘no relative performance information’ with ‘tournament compensation’. Further research that examines how prices behave in tournaments with no relative performance feedback may provide additional insights into the effects of tournament compensation, as well as the role of relative performance information *within* tournaments.

Third, the private nature of the relative feedback in this study leads to the inference that the most likely channel through which peer effects operate in our asset market is intrinsically-held competitive preferences. However, our experimental design is not equipped to examine the specific underlying mechanism(s) that drive peer effects in the market. Hence, investigating these mechanisms in market environments, potentially using the methods and insights offered by the burgeoning field of neurofinance, may represents a potentially exciting avenue for future research.

CHAPTER 5: Conclusion

Asset prices bubbles present a serious threat to the stability of financial markets and the wider economy, making it imperative to understand what drives them. This dissertation contributes to that cause by examining the determinants of asset price bubbles in an environment where they can be reliably measured and observed – the experimental laboratory.

The first study investigates whether asset price bubbles in experimental markets are fuelled by the seemingly benign act of allocating traders with assets to participate in the market with. Since these assets are not earned, participants may treat it as ‘other people’s money’, hence taking more risk than they otherwise would, thus generating the bubble. We examine this concern by observing price behaviour in markets where participants earn their initial allocation. The results of the study suggest that asset legitimacy is not likely to be a serious threat to the validity of existing results obtained from asset market experiments; price behaviour does not vary significantly between markets where the initial allocation is earned versus markets where they are endowed. However, a potential caveat applies here in that a bubble-and-crash phenomenon is not especially prevalent in unearned markets to begin with, which limits the scope for earned money to dampen bubbles. Another potential limitation to our study is that our earnings task, a GMAT quiz, is correlated with intelligence, which may impact trading behaviour. Moreover, success in the earnings task may generate an affective (emotional) response that could impact trading behaviour. It is difficult to separate these effects from asset legitimacy in the current experimental design. This may be a potentially fruitful avenue for future research to pursue.

The second study examines how prices behave under tournament incentives when participants can trade in a multi-asset market. Existing studies in single-asset environments report that tournaments exacerbate bubbles, and that this effect worsens with experience. However, our study finds that bubbles under tournament incentives *do* dissipate with experience when it is possible to trade in more than one type of risky asset. Moreover, prices under tournament incentives do not appear to behave differently to normal incentives in a two-asset environment. Hence, our results suggest that the findings of past studies are driven by the single-asset nature of the market. Furthermore, we find that penalties embedded into tournament contracts can actually *worsen* bubbles compared to contracts that only reward good relative performance, implying price behaviour that is consistent with investor herding. Of course, experiments are limited in their ability to 'penalise' participants. Hence, future studies may examine how prices behave under more salient penalties than a zero-payoff.

The third study investigates how relative performance feedback affects price and bubble behaviour. Despite suggestions that relative performance feedback may inflame bubbles by inciting a competitive desire within people to "get ahead of the Joneses", we find that prices actually conform to fundamental value better under normal incentives when relative performance feedback is provided. Hence, it may be that relative performance feedback aids market efficiency rather than hindering it. A possible avenue for future research may be to examine how different types of relative performance feedback affect price behaviour.

APPENDICES

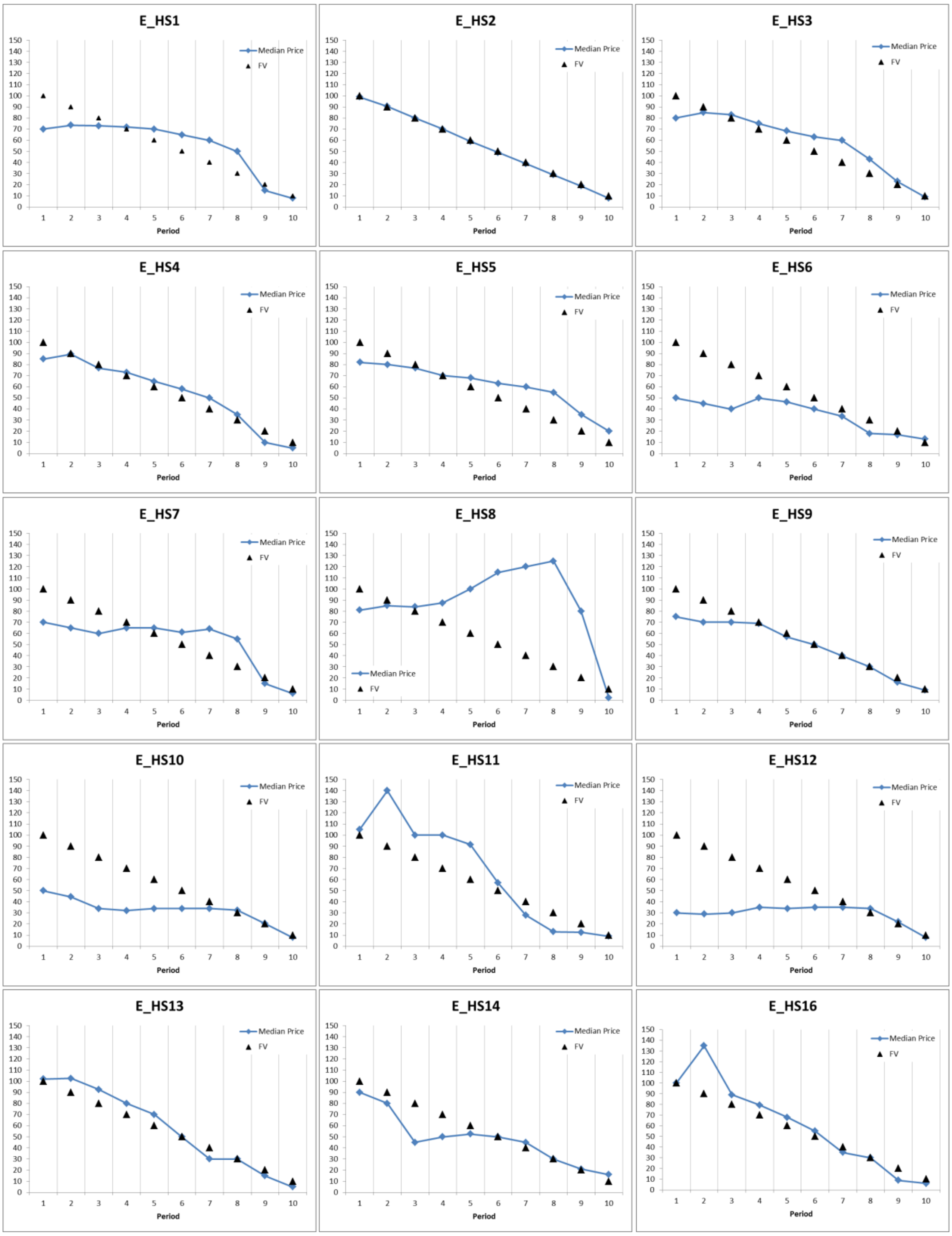
Appendix A – Asset Legitimacy Experiment

Appendix A1 – Additional Figures

Appendix A2 - Participant instructions

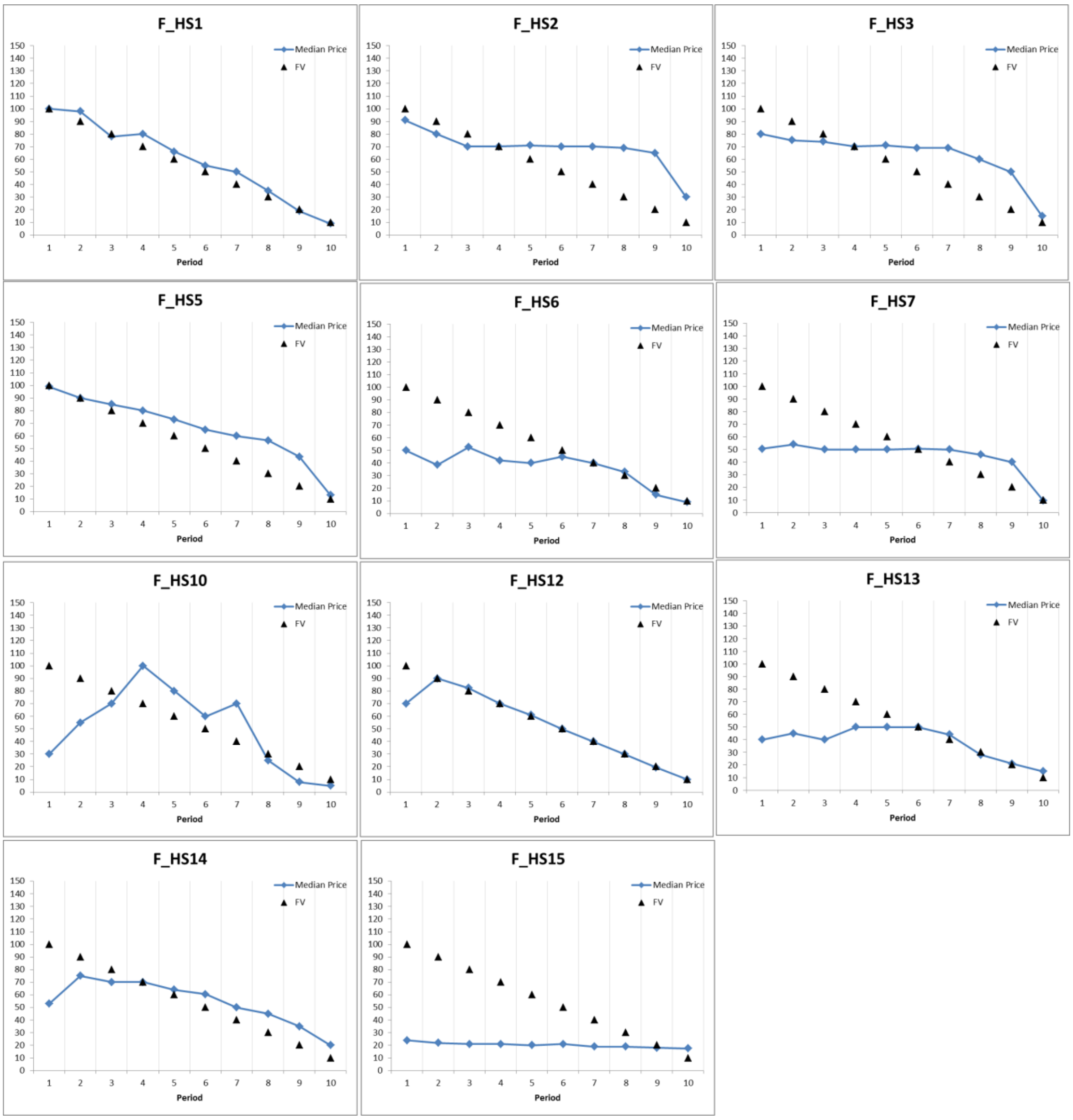
Appendix A1: Additional Figures

Figure A1: Median prices in individual *Earned* HS markets



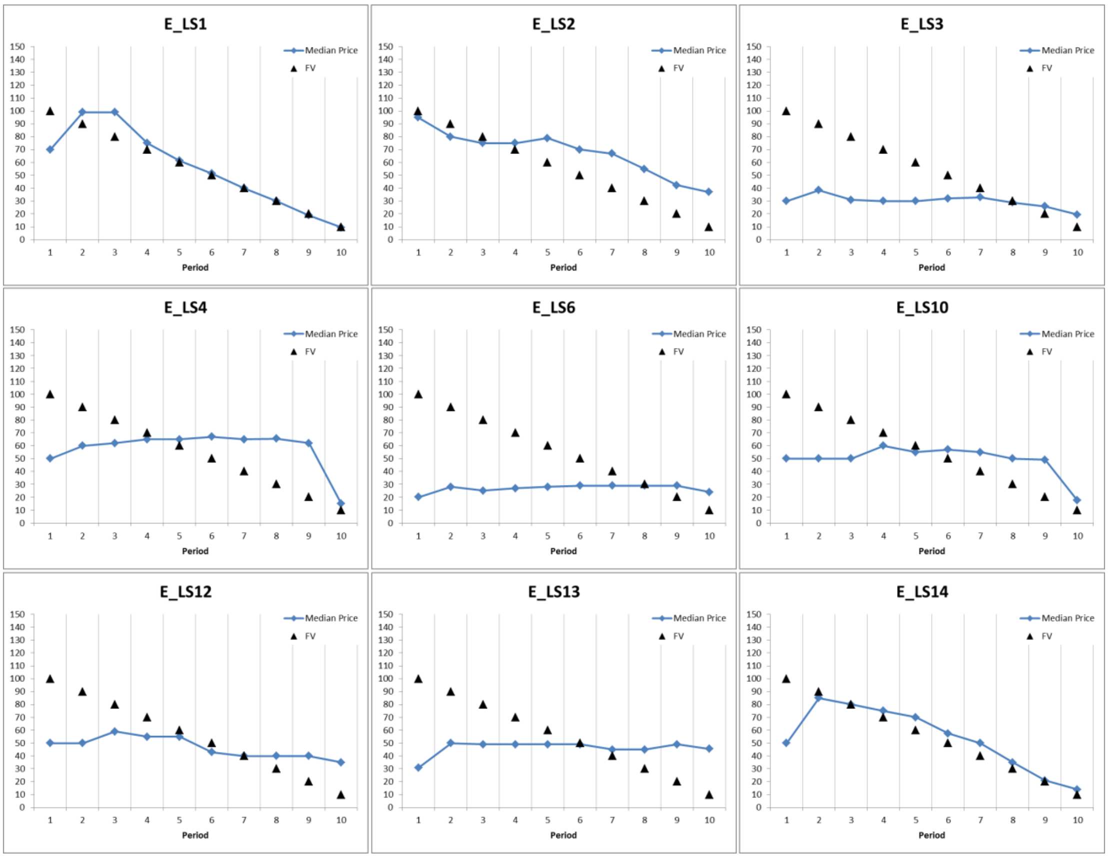
Note: Only *Earned* High-stakes (HS) markets that had an initial cash-to-asset ratio of 1 and did not contain any participants who had prior exposure to the experimental design are shown.

Figure A2: Median prices in individual *Free* HS markets



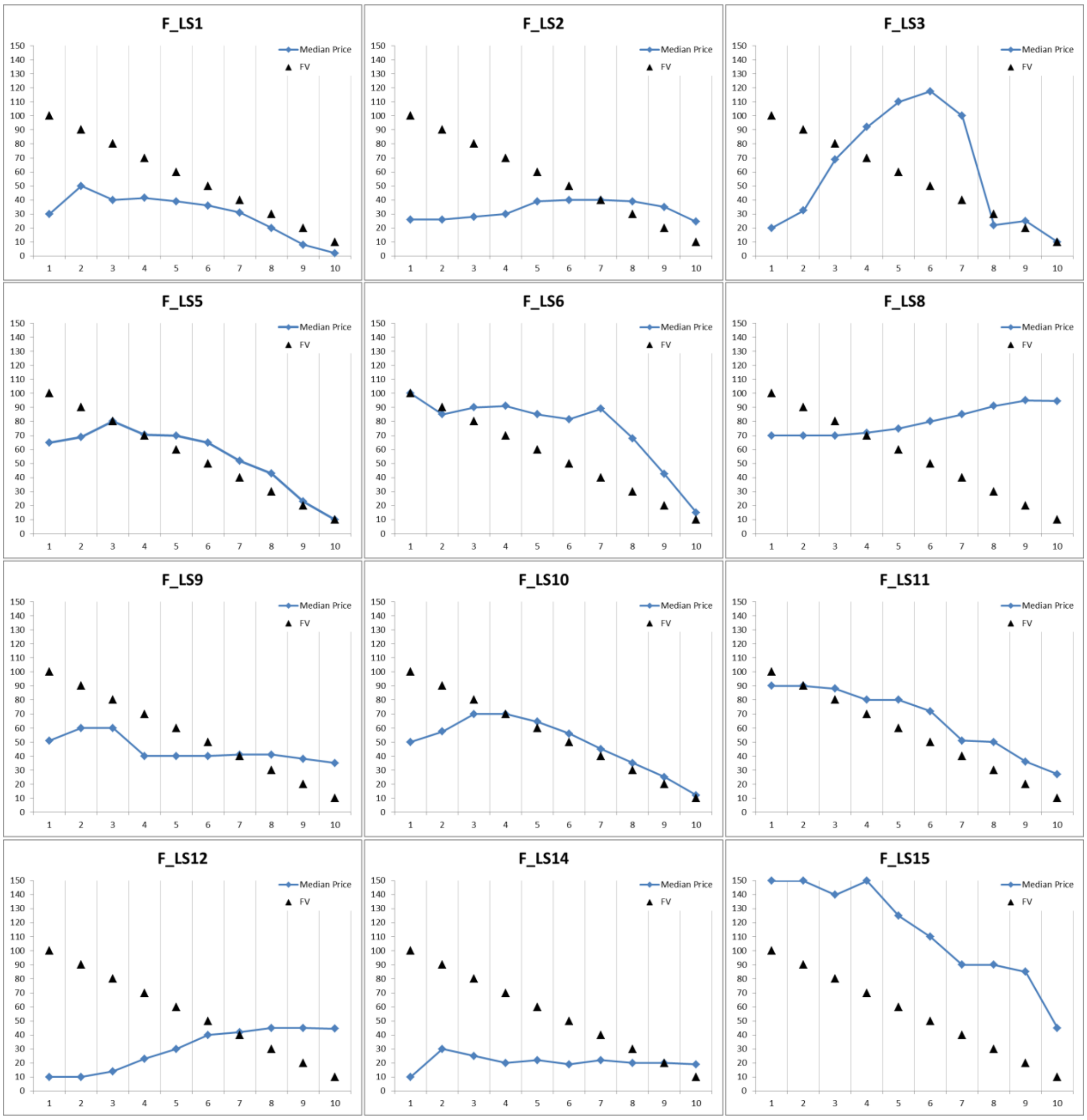
Note: Only *Free* High-stakes (HS) markets that had an initial cash-to-asset ratio of 1 and did not contain any participants who had prior exposure to the experimental design are shown.

Figure A3: Median prices in individual *Earned* LS markets



Note: Only *Earned* Low-stakes (LS) markets that had an initial cash-to-asset ratio of 1 and did not contain any participants who had prior exposure to the experimental design are shown.

Figure A4: Median prices in individual *Free* LS markets



Note: Only *Free* Low-stakes (LS) markets that had an initial cash-to-asset ratio of 1 and did not contain any participants who had prior exposure to the experimental design are shown.

Appendix A2: Instructions

The written instructions provided to participants of the experiment pertaining to the study in Chapter 2 are shown below. The parts of the instructions that are unique to the Earned treatment are bolded, italicised and bracketed in red font ***[like this]****. The content of the instructions for October session participant is identical, except for a shorter experiment length and trading period length.*

General instructions

This is an experiment in the economics of market decision-making. The instructions are simple and if you follow them carefully and make good decisions, you may earn a considerable amount of money, which will be paid to you, in cash, at the end of the experiment.

The experiment consists ***[of two stages. The first stage involves the completion of a task, through which you will earn the money that you will begin the second stage of the experiment with. The second stage will consist]*** of a sequence of trading periods in which you will have the opportunity to buy and sell in a market. All trading will be in terms of *francs*. The cash payment to you at the end of the experiment will be in Australian dollars, rounded up to the nearest 5 dollars. The conversion rate is ____ francs to 1 dollar.

The experiment will last no more than 1.5/***[2]*** hours, and will include up to 30 minutes of instructions and practice. Please do not speak with any other participants during this experiment. Please also remember to switch off your mobile phone. Failure to comply with these rules will result in your exclusion from the experiment and the forfeiture of all payments.

[Stage One

You will have 20 minutes to complete a quiz on the computer. The quiz consists of 10 multiple-choice questions taken from the numerical and verbal reasoning sections of the Graduate Management Admission Test (GMAT).

How you perform in the quiz relative to other participants determines your quiz earnings – better performers earn more. Performance is measured by the number of questions answered correctly. Where two or more participants are tied for the number of correct answers, the amount of time taken to complete the quiz is taken into consideration; the participant who has taken less time is deemed to have performed better.

Quiz earnings are awarded in the form of a portfolio of cash and goods that you will begin the second (i.e. next) stage of the experiment with. The top 50% of performers will be assigned to market type A, while the bottom 50% will be allocated to market type B. The initial portfolios in type A markets consist of twice the amount of cash and goods as the initial portfolios in Type B. As a result, they are worth twice as much.

An introduction screen will shortly appear on your computer, detailing the instructions for the quiz. Please read them carefully before clicking "Start Quiz" to begin the quiz. You may use the supplied calculator and working paper to help you answer the questions. Please also note that the usual rules of a test apply.

Once you and all other participants have completed the quiz, the market type (A or B) that you have been assigned to and your quiz earnings will be communicated to you on-screen. This is your private information – do not reveal it to other participants.]

How to use the Computerised Market

Before proceeding ***[to Stage 2]***, we introduce the market interface that you will be using for the remainder of the experiment. Please note that any actions you take during this demonstration period will <u>not</u> count towards your earnings or influence your position later in the experiment.

In the top right hand corner of the screen you see how much time, in seconds, is left in the current trading Period. The good that can be bought and sold in the market is called X. In the centre of your screen you see how many units of X you currently have and the amount of Cash that you have available. Note that the amounts shown on your screen are for demonstration purposes only and have no relation to ***[your performance in the task in Stage 1 (and hence your initial portfolio in Stage 2)]*** what you will begin the actual market with.

When you would like to offer to sell a unit of X, use the text area entitled "Enter offer to sell" in the first column. In that text area you can enter the price at which you are offering to sell a unit of X, and then select "Submit Offer To Sell". Please do so now. Type in a number in the appropriate space, and then click on the field labelled "Submit Offer To Sell". You will notice that 8 numbers, one submitted by each participant in your market, now appear in the second column from the left, entitled "Offers To Sell". Your offer is listed in blue. You can submit multiple offers; new offers will be added to the list, but will not replace your previous offer(s).

The lowest offer-to-sell price will always be on the top of that list and will, by default, be selected. You can select a different offer by clicking on it. It will then be highlighted. If you select "Buy", the button at the bottom of this column, you will buy one unit of X for the currently selected sell price. Please purchase a unit now by selecting an offer and clicking the "Buy" button. Since each of you had offered to sell a unit of X and attempted to buy a unit of X, if all were successful, you all have the same number of units of X you started out with. This is because you bought one unit of X and sold one unit of X.

When you buy a unit of X, your Cash balance decreases by the price of the purchase, and any existing offers-to-buy submitted by you are cancelled. When you sell a unit of X your Cash balance increases by the price of the sale, and any existing offers-to-sell submitted by you are cancelled. You may make an offer to buy a unit by selecting "Submit offer to buy." Please do so now. Type a number in the text area "Enter offer to buy", then press the red button labelled "Submit Offer To Buy". The highest offer-to-buy price will always be on top of that list and will, by default, be selected. You can accept any of the offers-to-buy by selecting the offer and then clicking on the "Sell" button. Please do so now. In the middle column, labelled "Transaction Prices", you can see the prices at which X has been bought and sold in this period. The most recent transaction will be listed at the top.

You will now have about 10 minutes to buy and sell X. This is a practice period. Your actions in the practice period do <u>not</u> count toward your earnings and do not influence your position later in the experiment. The only goal of the practice period is to master the use of the interface. Please be sure that you have successfully submitted offers to buy and offers to sell. Also be sure that you have accepted buy and sell offers. If you have any questions, please raise your hand and the experimenter will come by and assist you.

[Stage 2 -] **Specific Instructions for the Market**

The market consists of you and 7 other traders ***[who were assigned to the same market type as you (i.e. A's only trade with other A's, B's only trade with other B's)]***. At the beginning of the market, you will be endowed with/***[have been allocated]*** a portfolio of goods (called 'X') and Cash ***[earned by participants of your market-type in Stage 1]***. Other traders in your market may have a different distribution of cash and goods in their initial portfolio to you.

The market has 10 periods, each lasting 3 minutes. In each period, you may buy and/or sell units of the good called X. X can be considered an asset with a life of 10 periods, and your inventory of X carries over from one trading period to the next. Note that your cash balance and asset inventory cannot fall below zero.

At the end of each trading period, each unit of X pays a dividend, which is randomly determined by the computer. The possible dividend values and the associated likelihoods are shown below:

Dividend	→	0	20
Likelihood	→	½	½

Since each dividend is equally likely, the average dividend per period is 10 francs. The dividend draws in each period are independent. This means that the likelihood of a particular dividend in a period is not affected by the dividend in previous periods. After the final dividend is paid at the end of period 10, each unit of X expires worthless.

Average Holding Value Table

You can use the table at the end of this document to help you make decisions. It calculates the average amount of dividends you will receive if you hold a unit of X in your inventory for the rest of the market, or equivalently, how much in dividends you give up, on average, when you sell a unit of X at any time. Each of the 5 columns of the table is described below:

1. *Ending Period*: indicates the last trading period of the market, period 10.
2. *Current Period*: indicates the period during which the average holding value is being calculated.
3. *Number of holding periods:* This is equivalent to the number of times a dividend can be received if a unit of X is held in your inventory from the current period to the end of the market.
4. *Average Dividend Per Period*: gives the average amount that the dividend will be in each period for each unit of X held in your inventory.
5. *Average Holding Value Per Unit of Inventory*: gives the expected total dividend for the remainder of the experiment for each unit of X that is held in your inventory for

the rest of the market. That is, for each unit you hold in your inventory for the remainder of the market, you will receive on average the amount listed in column 5 in dividends. Equivalently, it tells you how much in future dividends you give up on average when you sell a unit in the current period. The number in column 5 is calculated by multiplying the numbers in columns 3 and 4.

Example: Suppose that there are 4 periods remaining. Since the dividend paid on a unit of X has a 50% chance of being 0 and a 50% chance of being 20, the dividend is in expectation 10 per period for each unit of X. If you hold a unit of X for 4 periods, the total dividend paid on the unit over the 4 periods is in expectation 4×10 = 40.

Calculating Your Earnings

Your dividend earnings in each period depends on the number of units of X in your inventory at the **end** of the period, and is calculated as follows:

PERIOD DIVIDEND EARNINGS = END-OF-PERIOD INVENTORY UNITS x DIVIDEND PER UNIT FOR THAT PERIOD

Dividend earnings are added to your cash balance at the end of each period.

When you spend money to buy unit(s) of X, the total amount of cash that you have is reduced by the amount of the purchase. If you sell unit(s) of X, the total amount of cash you have increases by the amount of the sale. Your end-of-period cash balance is calculated as follows:

END OF PERIOD CASH = BEGINNING OF PERIOD CASH + PERIOD DIVIDEND EARNINGS

+ (SALES REVENUE – EXPENDITURES ON PURCHASES)

Since each unit of X expires worthless after the final dividend payment, your earnings from the experiment will equal the balance of your cash account at the end of the market/experiment. Note that you do not have to calculate your earnings by yourself. The computer does all the work.

There will also be a show up fee of $5 (non-tradable) to all participants.

An earnings report will appear on-screen at the end of each period. After seeing your earnings, press the “Continue” button to go to the next period. The next period will begin once all of you press the “Continue” button.

Average Value Holding Table

Ending Period	Current Period	Number of Holding Periods	×	Average Dividend Per Period	=	Average Holding Value Per Unit in Inventory
10	1	10		10		100
10	2	9		10		90
10	3	8		10		80
10	4	7		10		70
10	5	6		10		60
10	6	5		10		50
10	7	4		10		40
10	8	3		10		30
10	9	2		10		20
10	10	1		10		10

Appendix A3: Survey

The following questionnaire, a modified version of the one used by Ackert and Church (2001), was completed by participants at the end of the experiment.

Post-Market Questionnaire

1. What year are you in university? ________

2. What department/school are you in at university (e.g., finance, economics)? ____________

3. What is your sex (tick one) **male ______ female ______**

4. What is your age? _______ years

5. How interesting did you find this experiment? (circle the appropriate number)

Not Very Interesting 1- - - - -2- - - - -3- - - - -4- - - - -5- - - - -6- - - - -7 **Very Interesting**

6. Have you ever traded securities for yourself or others? (tick one) **yes _______ no _______**

7. Have you ever participated in the management of an investment portfolio? (tick one)
yes ________ no ________

8. Compared to the amount of money available to you from alternative sources, how would you characterize the amount of money earned for participating in this experiment? (circle the appropriate number)

Nominal Amount 1- - - - -2- - - - -3- - - - -4- - - - -5- - - - -6- - - - -7 **Considerable Amount**

9. How would you characterize your attitude toward risk while participating in the market? (circle the appropriate number)

Very Risk Averse 1- - - - - 2- - - - -3- - - - -4- - - - -5- - - - - 6- - - - -7 **Very Risk-Taking**

10. Describe as best you can the trading/investment strategy you followed in the market, including any changes in strategy as the market evolved.

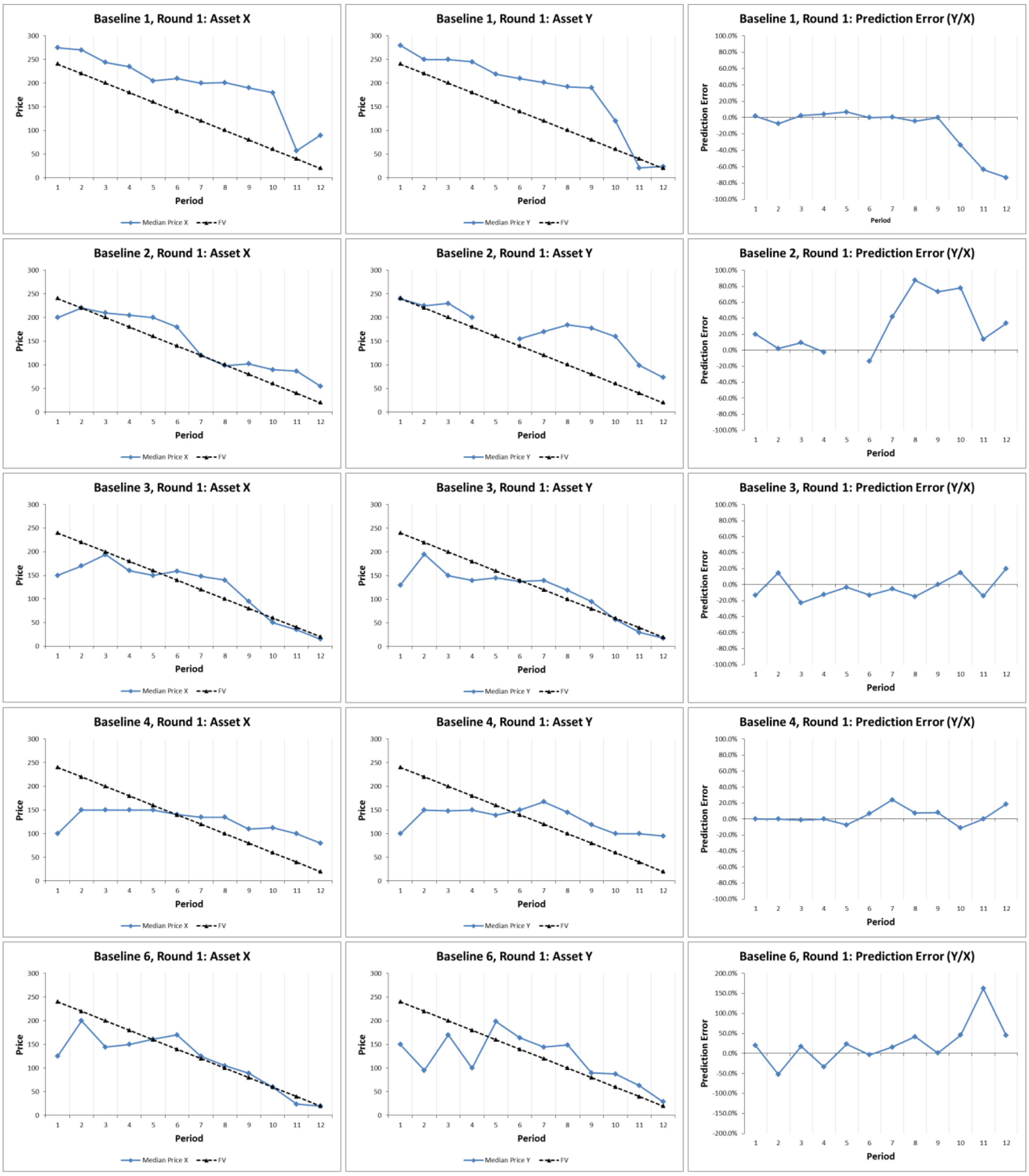
If you wish to leave any feedback for the experimenters regarding this experiment (e.g. the instructions), please do so below.

Appendix B – Experiments relating to Chapters 3 and 4

Appendix B1: Additional Figures

Figure B1: Median prices in individual *Baseline* markets in Round 1

The left-most column of the figure below shows median transaction prices for asset X in each *Baseline* market in Round 1. The middle column shows median prices for asset Y in the corresponding market, while the right-most column reports the resulting *Prediction Error*. *Prediction Error* is the percentage deviation of the relative median price of Y (i.e. median price of Y divided by the median price of X) from the risk-neutral benchmark of 1. Note that *Baseline* market no. 5 is excluded because it contained a subject who had participated in an earlier session.



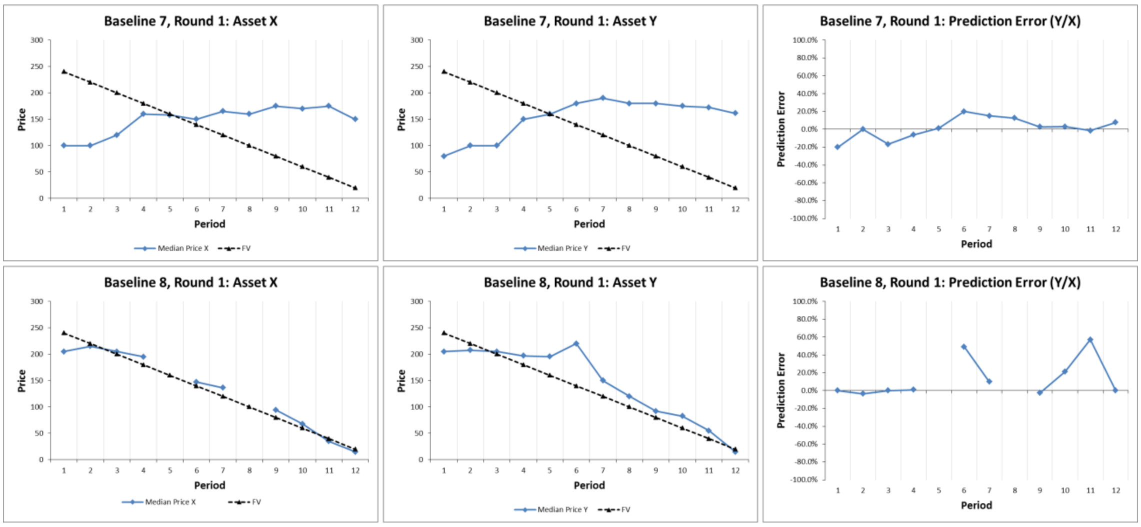
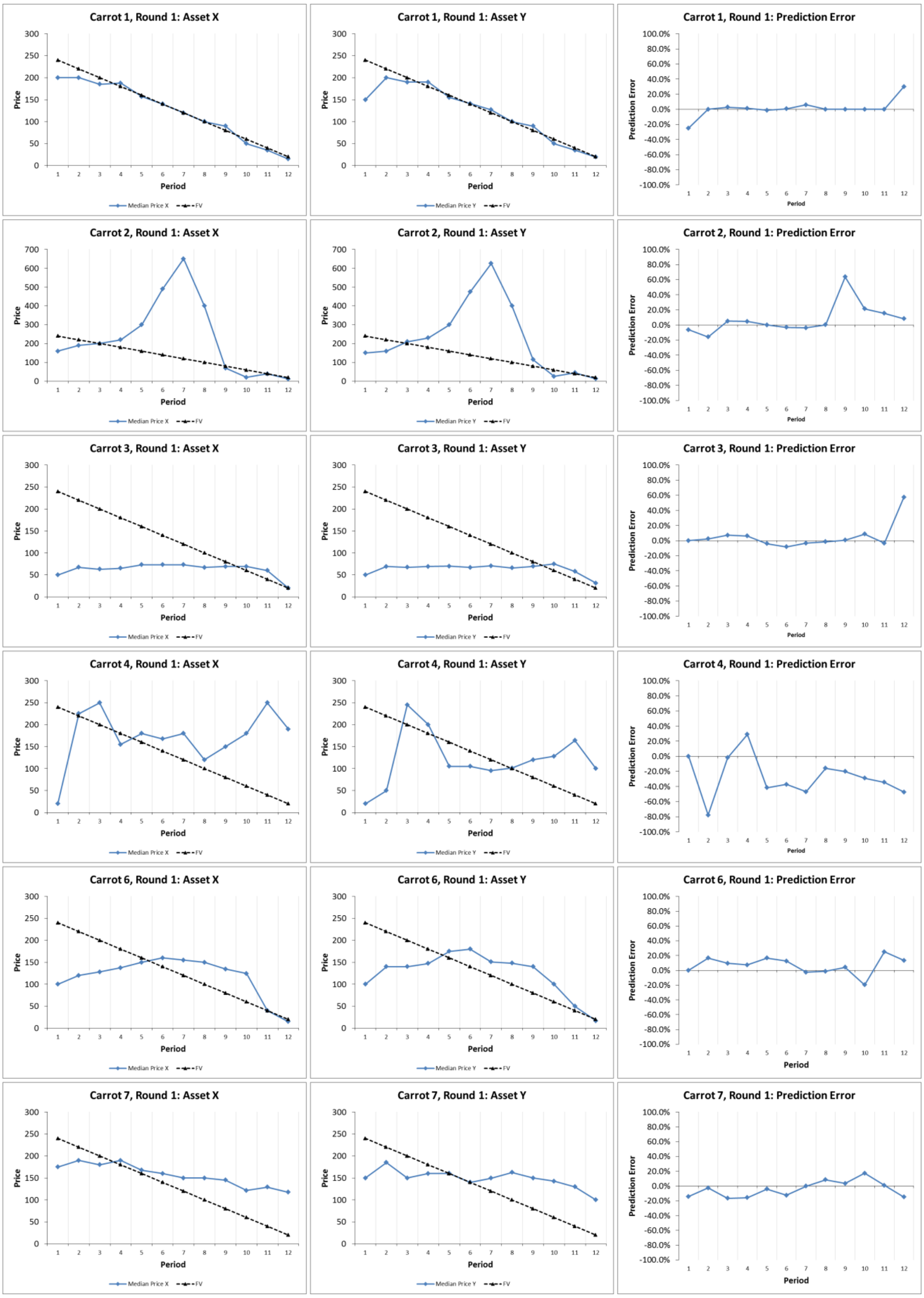

Baseline 7, Round 1: Asset X
Price
Period
Median Price X
FV
Baseline 7, Round 1: Asset Y
Price
Period
Median Price Y
FV
Baseline 7, Round 1: Prediction Error (Y/X)
Prediction Error
Period
Baseline 8, Round 1: Asset X
Price
Period
Median Price X
FV
Baseline 8, Round 1: Asset Y
Price
Period
Median Price Y
FV
Baseline 8, Round 1: Prediction Error (Y/X)
Prediction Error
Period

Figure B2: Median prices in individual *Carrot* markets in Round 1

The left-most column of the figure below shows median transaction prices for asset X in each *Carrot* market in Round 1. The middle column shows median prices for asset Y in the corresponding market, while the right-most column reports the resulting *Prediction Error*. *Prediction Error* is the percentage deviation of the relative median price of Y (i.e. median price of Y divided by the median price of X) from the risk-neutral benchmark of 1. Note that *Carrot* market no. 5 is excluded because it contained a subject who had participated in an earlier session.



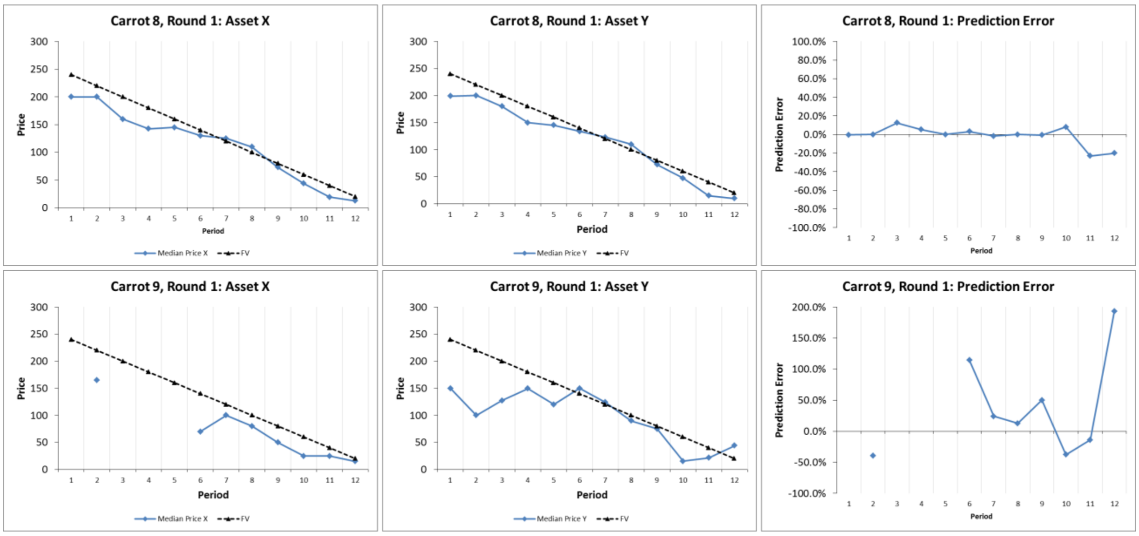
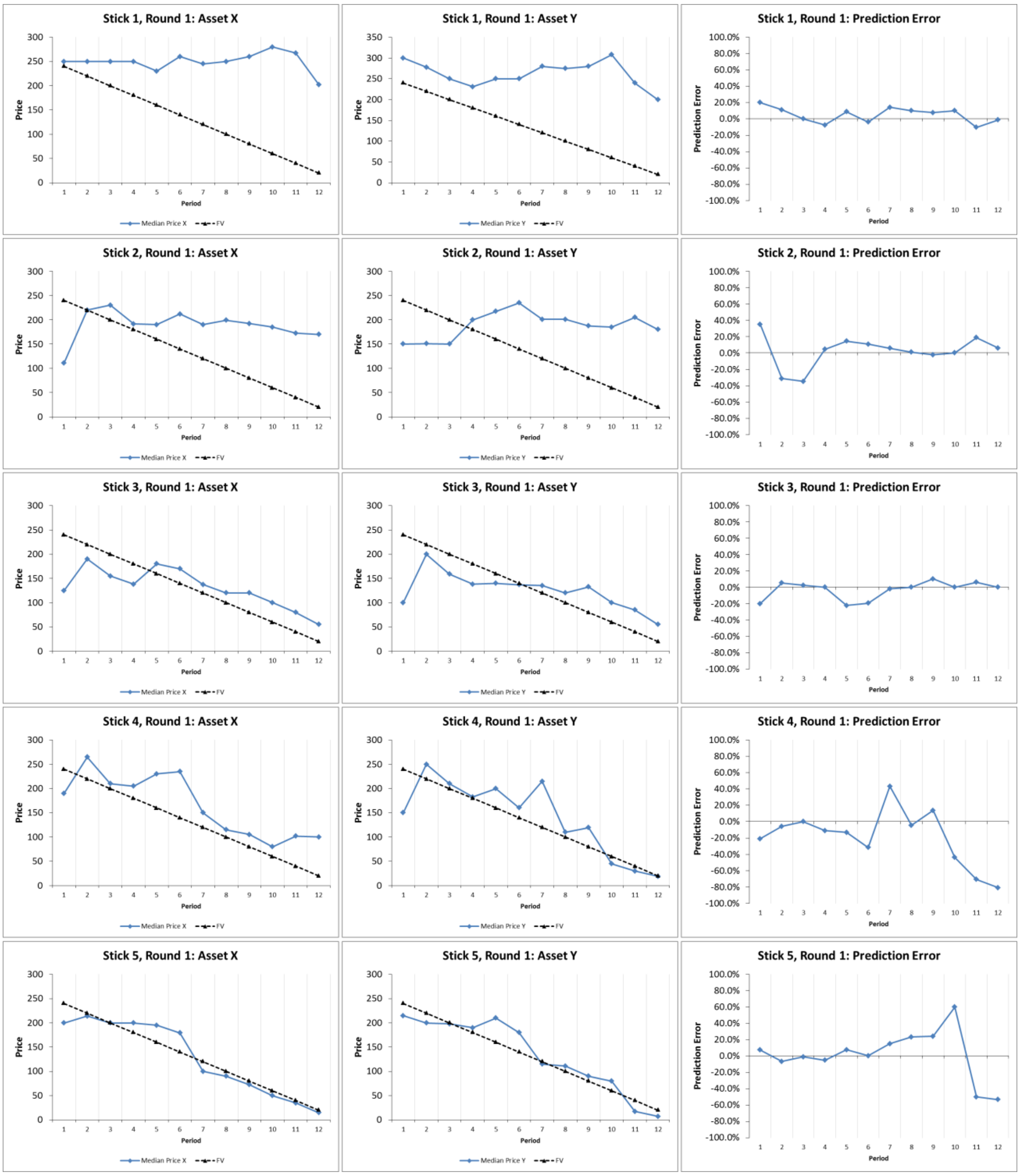

Carrot 8, Round 1: Asset X
Carrot 8, Round 1: Asset Y
Carrot 8, Round 1: Prediction Error
Carrot 9, Round 1: Asset X
Carrot 9, Round 1: Asset Y
Carrot 9, Round 1: Prediction Error
Price
Prediction Error
Period
Median Price X
Median Price Y
FV

Figure B3: Median prices in individual *Stick* markets in Round 1

The left-most column of the figure below shows median transaction prices for asset X in each *Stick* market in Round 1. The middle column shows median prices for asset Y in the corresponding market, while the right-most column reports the resulting *Prediction Error*. *Prediction Error* is the percentage deviation of the relative median price of Y (i.e. median price of Y divided by the median price of X) from the risk-neutral benchmark of 1. Note that the *Prediction Error* in Period 1 of *Stick* market number 8 is 6566%, and hence is not plotted in order to preserve the other features of the graph.



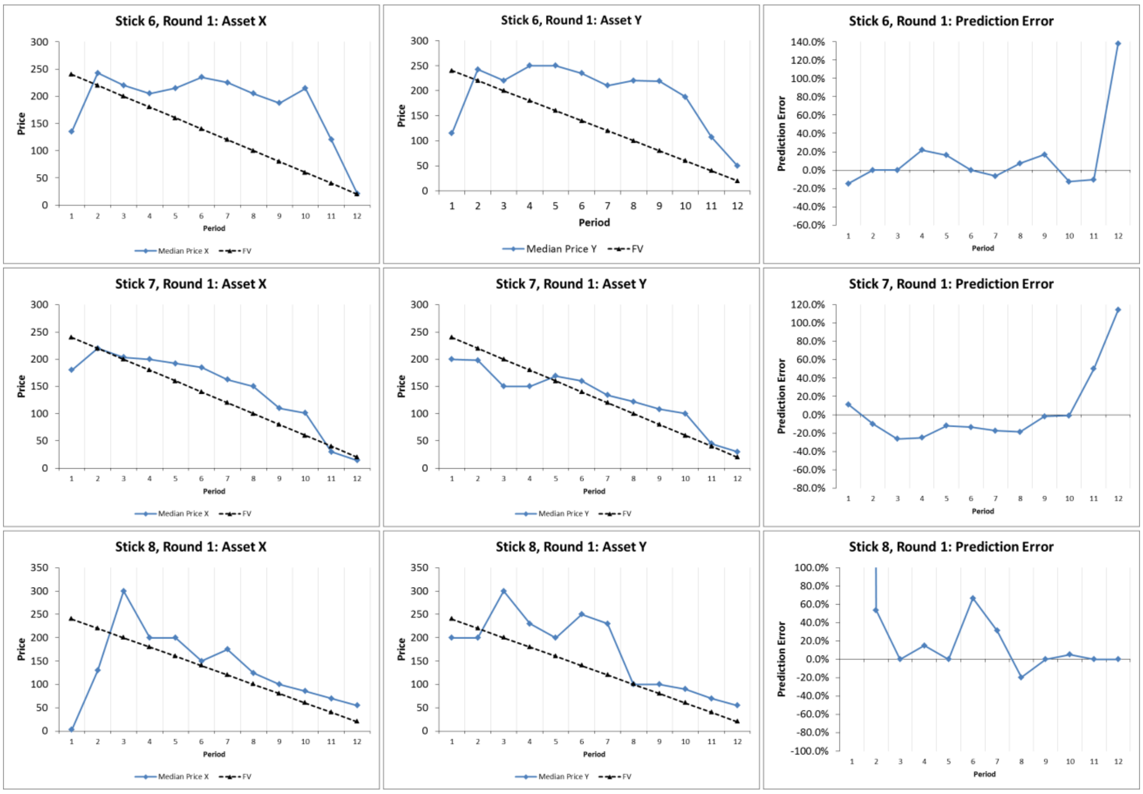

Stick 6, Round 1: Asset X
Stick 6, Round 1: Asset Y
Stick 6, Round 1: Prediction Error
Stick 7, Round 1: Asset X
Stick 7, Round 1: Asset Y
Stick 7, Round 1: Prediction Error
Stick 8, Round 1: Asset X
Stick 8, Round 1: Asset Y
Stick 8, Round 1: Prediction Error
Price
Prediction Error
Period
Median Price X
Median Price Y
FV

Figure B4: Median prices in individual *GilCarrot* markets in Round 1

The left-most column of the figure below shows median transaction prices for asset X in each *GilCarrot* market in Round 1. The middle column shows median prices for asset Y in the corresponding market, while the right-most column reports the resulting *Prediction Error*. *Prediction Error* is the percentage deviation of the relative median price of Y (i.e. median price of Y divided by the median price of X) from the risk-neutral benchmark of 1.

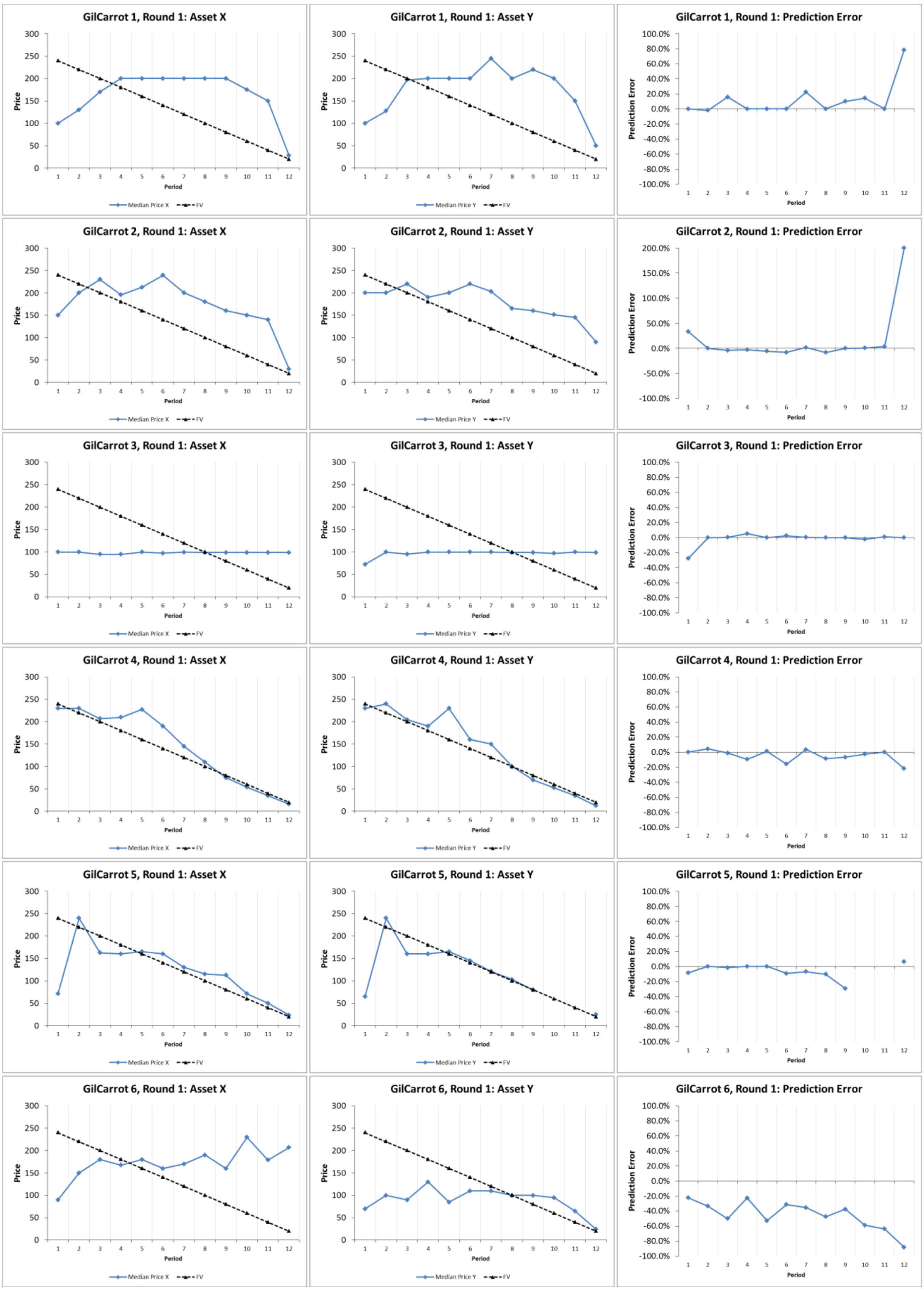


Figure B5: Median prices in individual *GilStick* markets in Round 1

The left-most column of the figure below shows median transaction prices for asset X in each *GilStick* market in Round 1. The middle column shows median prices for asset Y in the corresponding market, while the right-most column reports the resulting *Prediction Error*. *Prediction Error* is the percentage deviation of the relative median price of Y (i.e. median price of Y divided by the median price of X) from the risk-neutral benchmark of 1. Note that *GilStick* market no. 6 is excluded because it contained a subject who had participated in an earlier session.

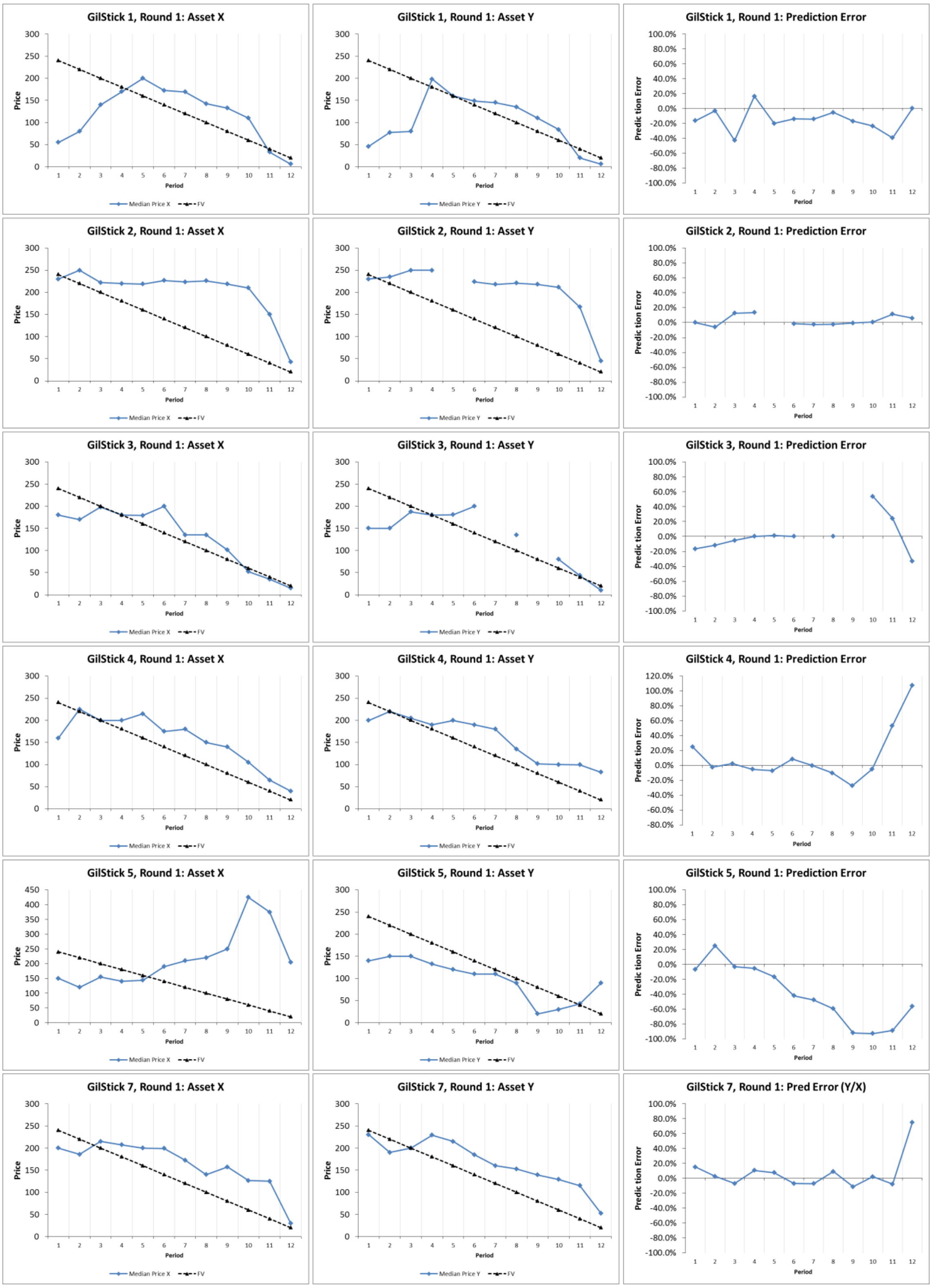
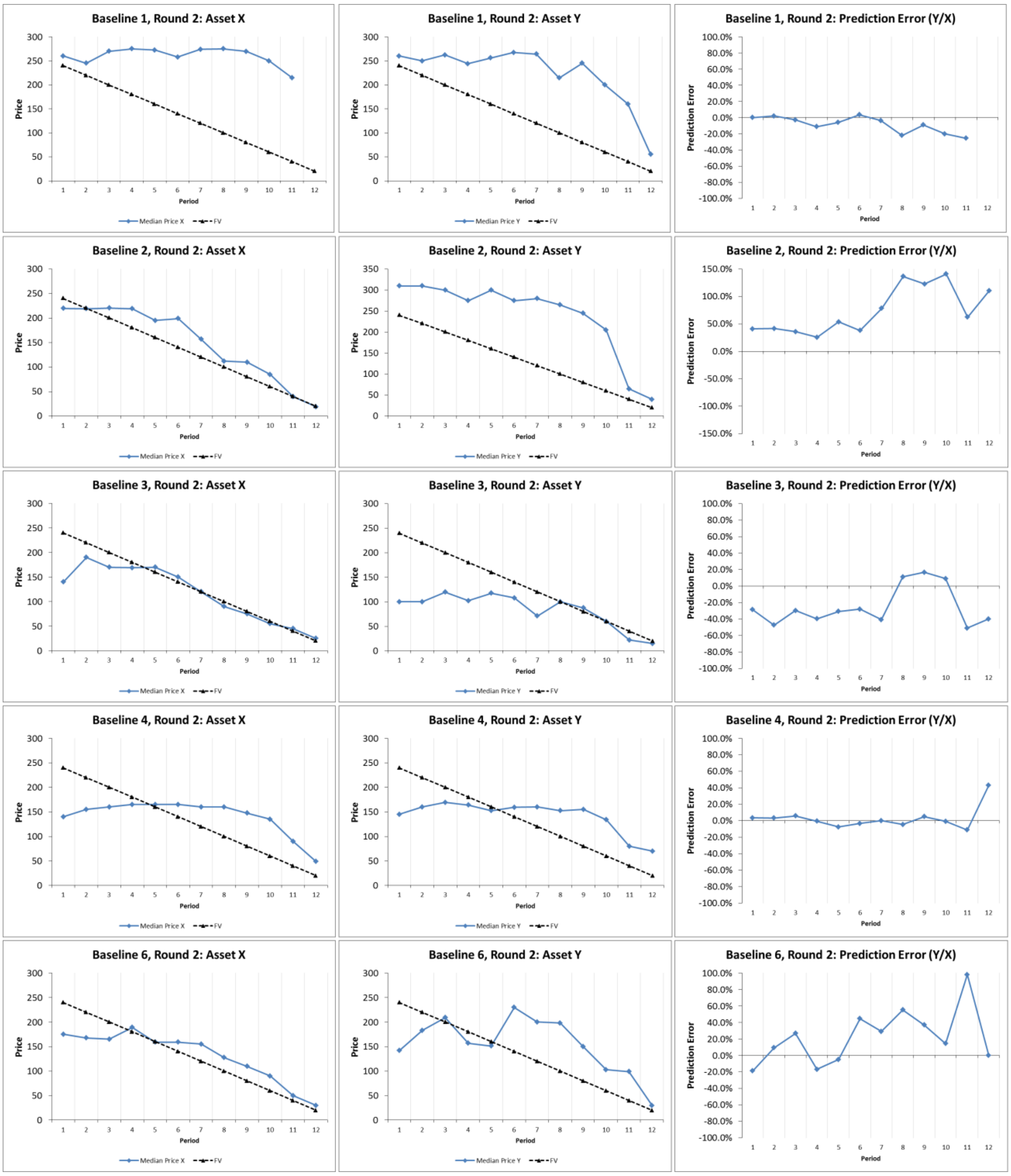


Figure B6: Median prices in individual *Baseline* markets in Round 2

The left-most column of the figure below shows median transaction prices for asset X in each *Baseline* market in Round 2. The middle column shows median prices for asset Y in the corresponding market, while the right-most column reports the resulting *Prediction Error*. *Prediction Error* is the percentage deviation of the relative median price of Y (i.e. median price of Y divided by the median price of X) from the risk-neutral benchmark of 1. Note that *Baseline* market no. 5 is excluded because it contained a subject who had participated in an earlier session. Note also that the median price for asset X (Y) in periods 11 and 12 of *Baseline* market no. 7 is 400 and 2000 (825 and 2100) respectively. To avoid obscuring the other facets of the graph, these points have not been graphed.



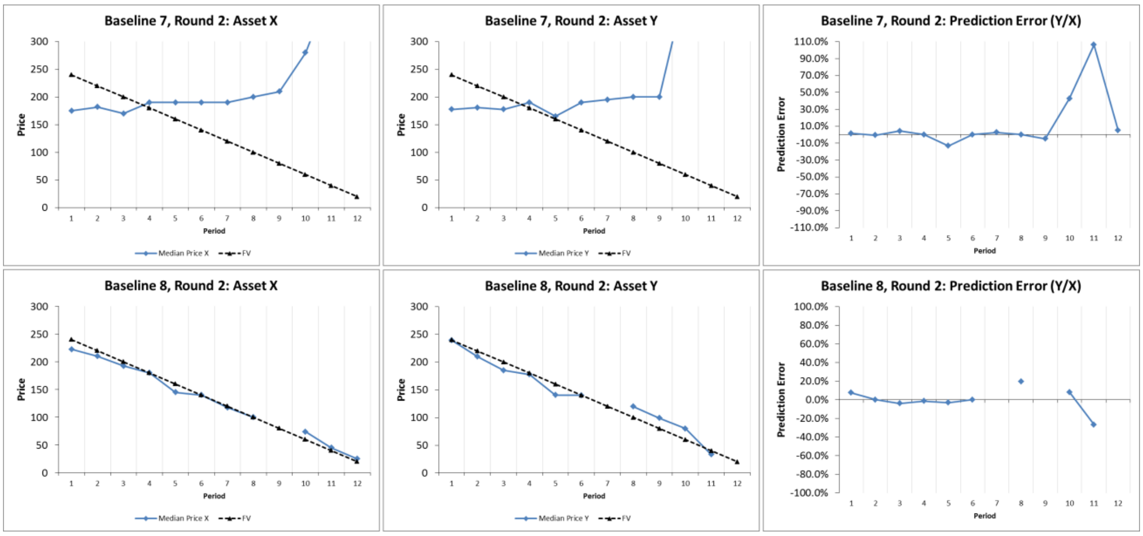
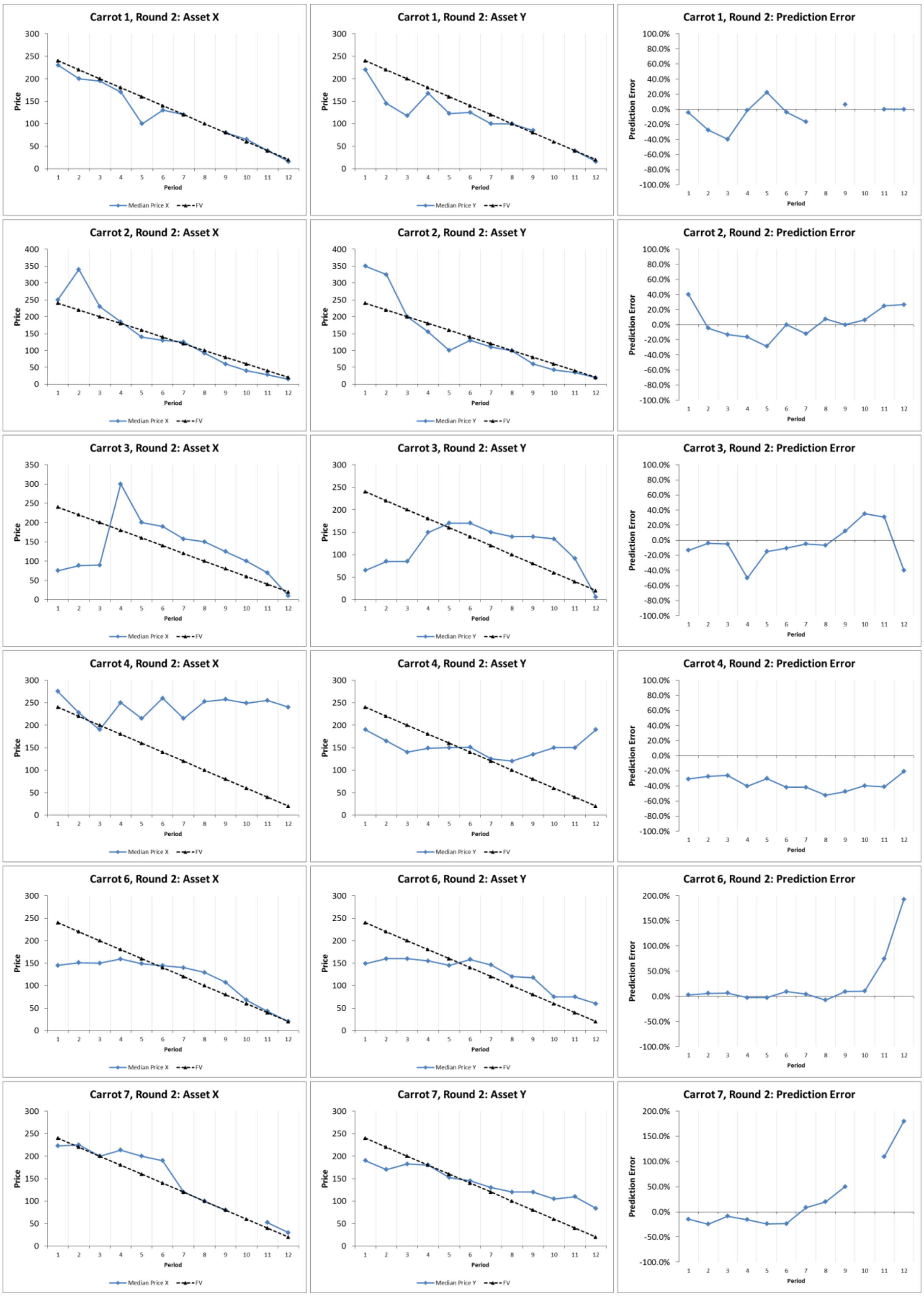
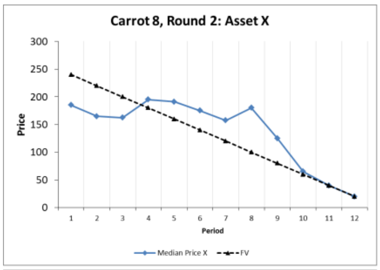
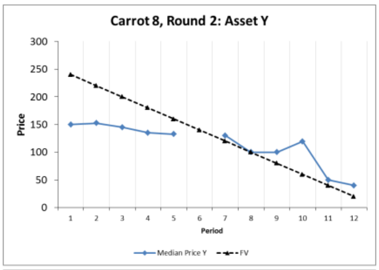

Baseline 7, Round 2: Asset X
Price
Period
Median Price X
FV
Baseline 7, Round 2: Asset Y
Price
Period
Median Price Y
FV
Baseline 7, Round 2: Prediction Error (Y/X)
Prediction Error
Period
Baseline 8, Round 2: Asset X
Price
Period
Median Price X
FV
Baseline 8, Round 2: Asset Y
Price
Period
Median Price Y
FV
Baseline 8, Round 2: Prediction Error (Y/X)
Prediction Error
Period

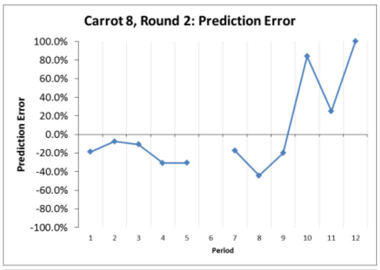
Figure B7: Median prices in individual *Carrot* markets in Round 2

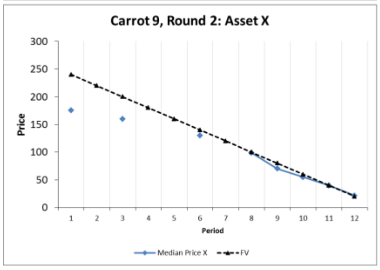
The left-most column of the figure below shows median transaction prices for asset X in each *Carrot* market in Round 2. The middle column shows median prices for asset Y in the corresponding market, while the right-most column reports the resulting *Prediction Error*. *Prediction Error* is the percentage deviation of the relative median price of Y (i.e. median price of Y divided by the median price of X) from the risk-neutral benchmark of 1. Note that *Carrot* market no. 5 is excluded because it contained a subject who had participated in an earlier session.

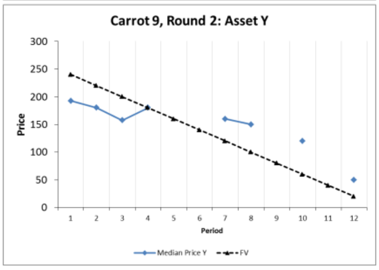



Carrot 8, Round 2: Asset X
Price
Period
Median Price X
FV


Carrot 8, Round 2: Asset Y
Price
Period
Median Price Y
FV


Carrot 8, Round 2: Prediction Error
Prediction Error
Period


Carrot 9, Round 2: Asset X
Price
Period
Median Price X
FV


Carrot 9, Round 2: Asset Y
Price
Period
Median Price Y
FV

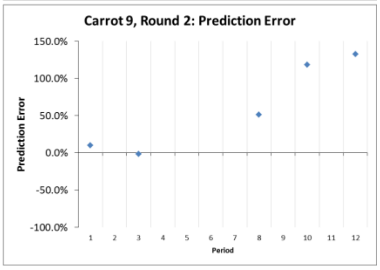
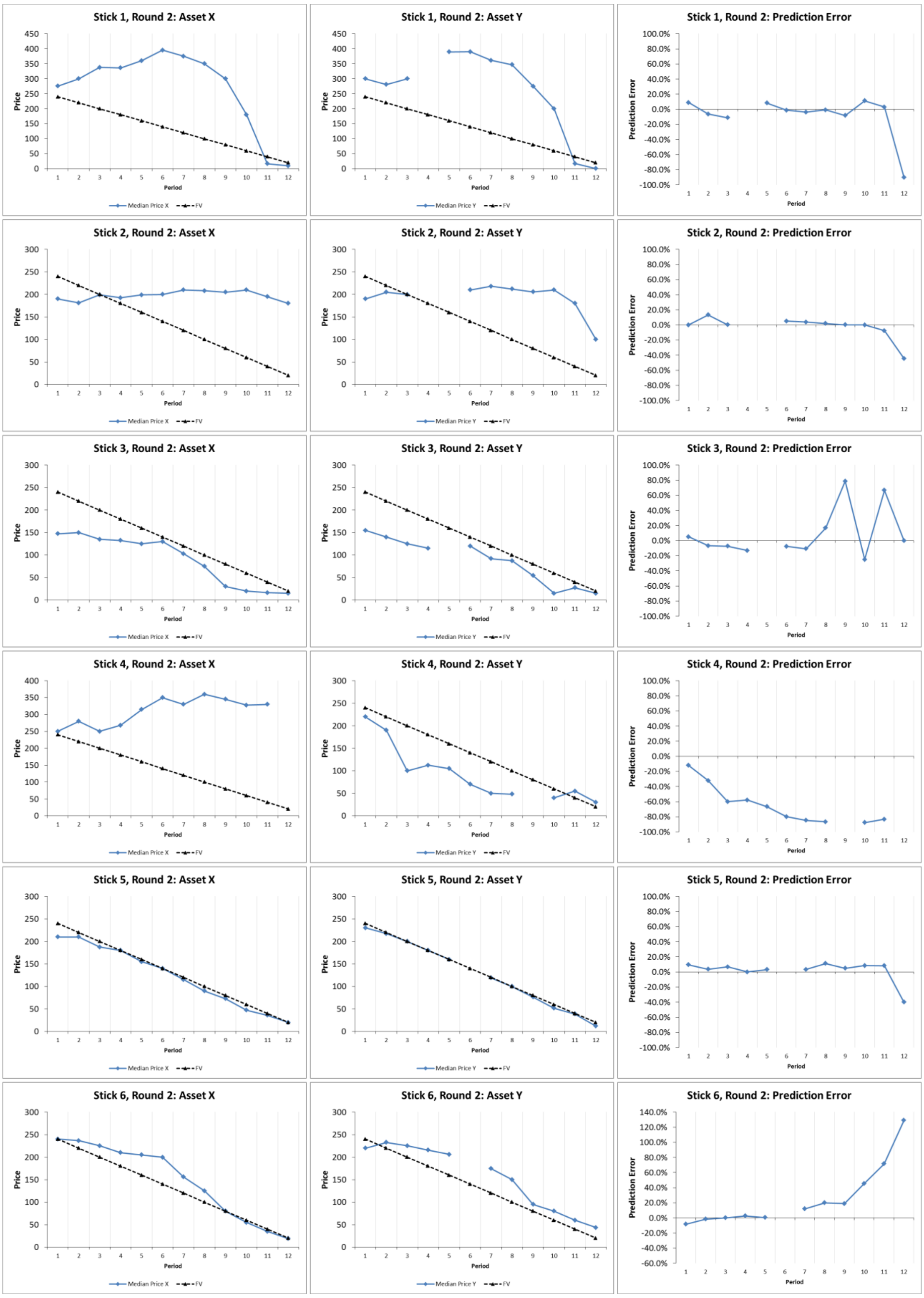

Carrot 9, Round 2: Prediction Error
Prediction Error
Period

Figure B8: Median prices in individual *Stick* markets in Round 2

The left-most column of the figure below shows median transaction prices for asset X in each *Stick* market in Round 2. The middle column shows median prices for asset Y in the corresponding market, while the right-most column reports the resulting *Prediction Error*. *Prediction Error* is the percentage deviation of the relative median price of Y (i.e. median price of Y divided by the median price of X) from the risk-neutral benchmark of 1.



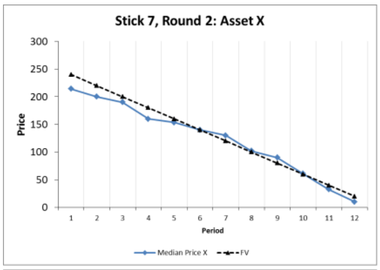
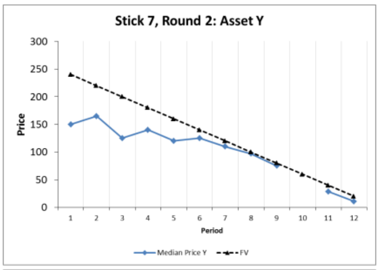
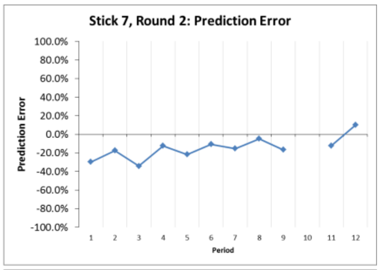
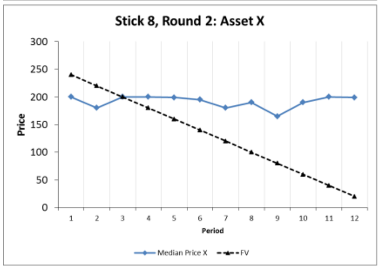
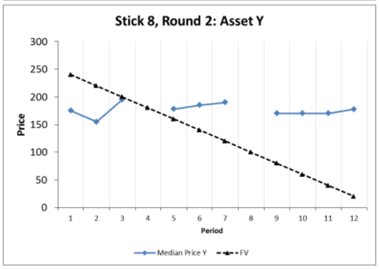
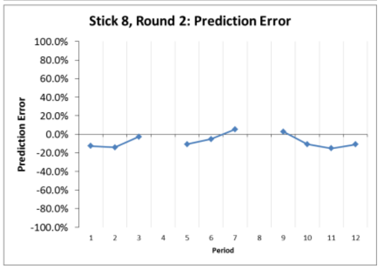

Stick 7, Round 2: Asset X
Price
Period
Median Price X
FV

Stick 7, Round 2: Asset Y
Price
Period
Median Price Y
FV

Stick 7, Round 2: Prediction Error
Prediction Error
Period

Stick 8, Round 2: Asset X
Price
Period
Median Price X
FV

Stick 8, Round 2: Asset Y
Price
Period
Median Price Y
FV

Stick 8, Round 2: Prediction Error
Prediction Error
Period

Figure B9: Median prices in individual *GilCarrot* markets in Round 2

The left-most column of the figure below shows median transaction prices for asset X in each *GilCarrot* market in Round 2. The middle column shows median prices for asset Y in the corresponding market, while the right-most column reports the resulting *Prediction Error*. *Prediction Error* is the percentage deviation of the relative median price of Y (i.e. median price of Y divided by the median price of X) from the risk-neutral benchmark of 1.

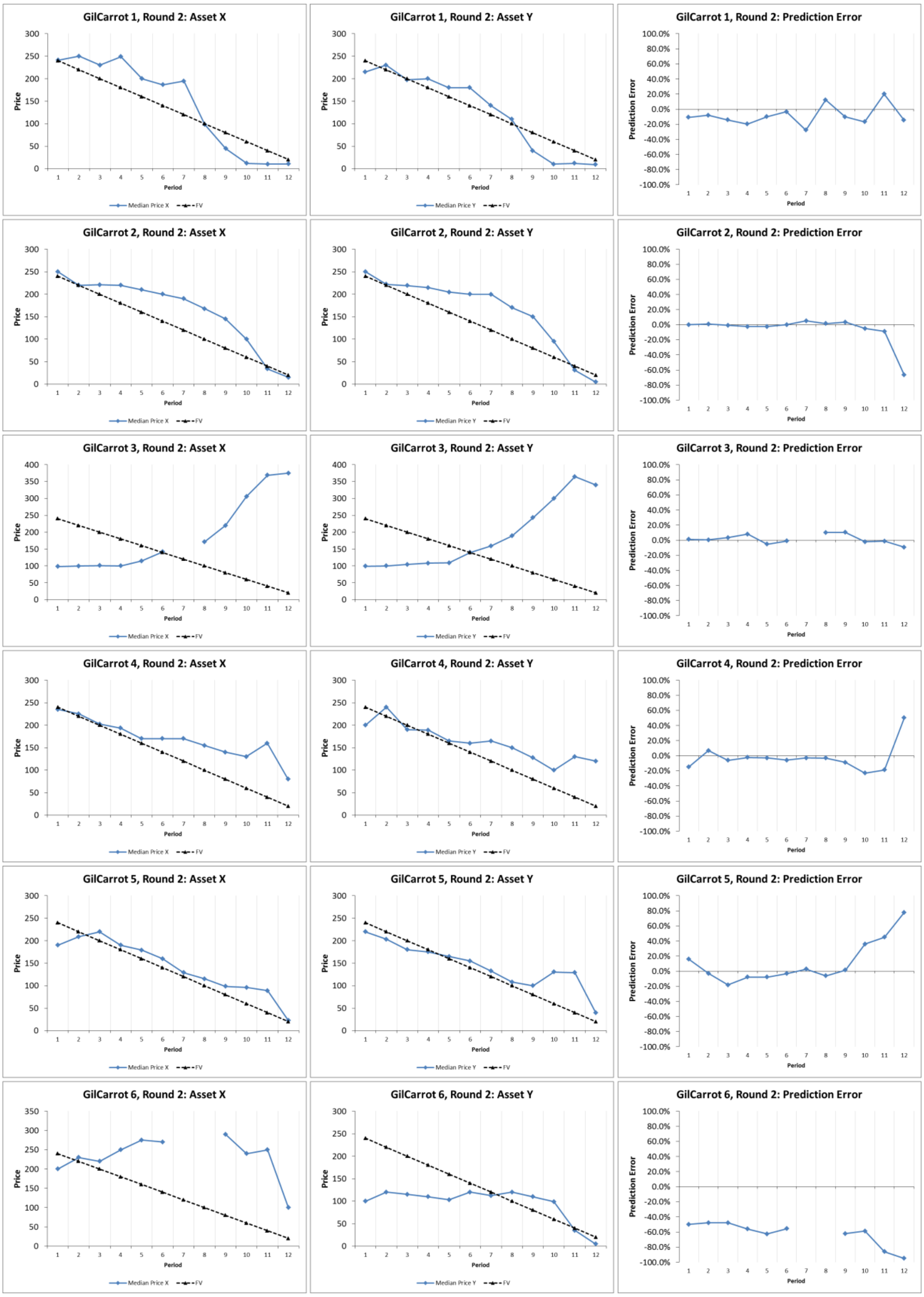


Figure B10: Median prices in individual *GilStick* markets in Round 2

The left-most column of the figure below shows median transaction prices for asset X in each *GilStick* market in Round 2. The middle column shows median prices for asset Y in the corresponding market, while the right-most column reports the resulting *Prediction Error*. *Prediction Error* is the percentage deviation of the relative median price of Y (i.e. median price of Y divided by the median price of X) from the risk-neutral benchmark of 1. Note that *GilStick* market no.6 is excluded because it contained a subject who had participated in an earlier session.

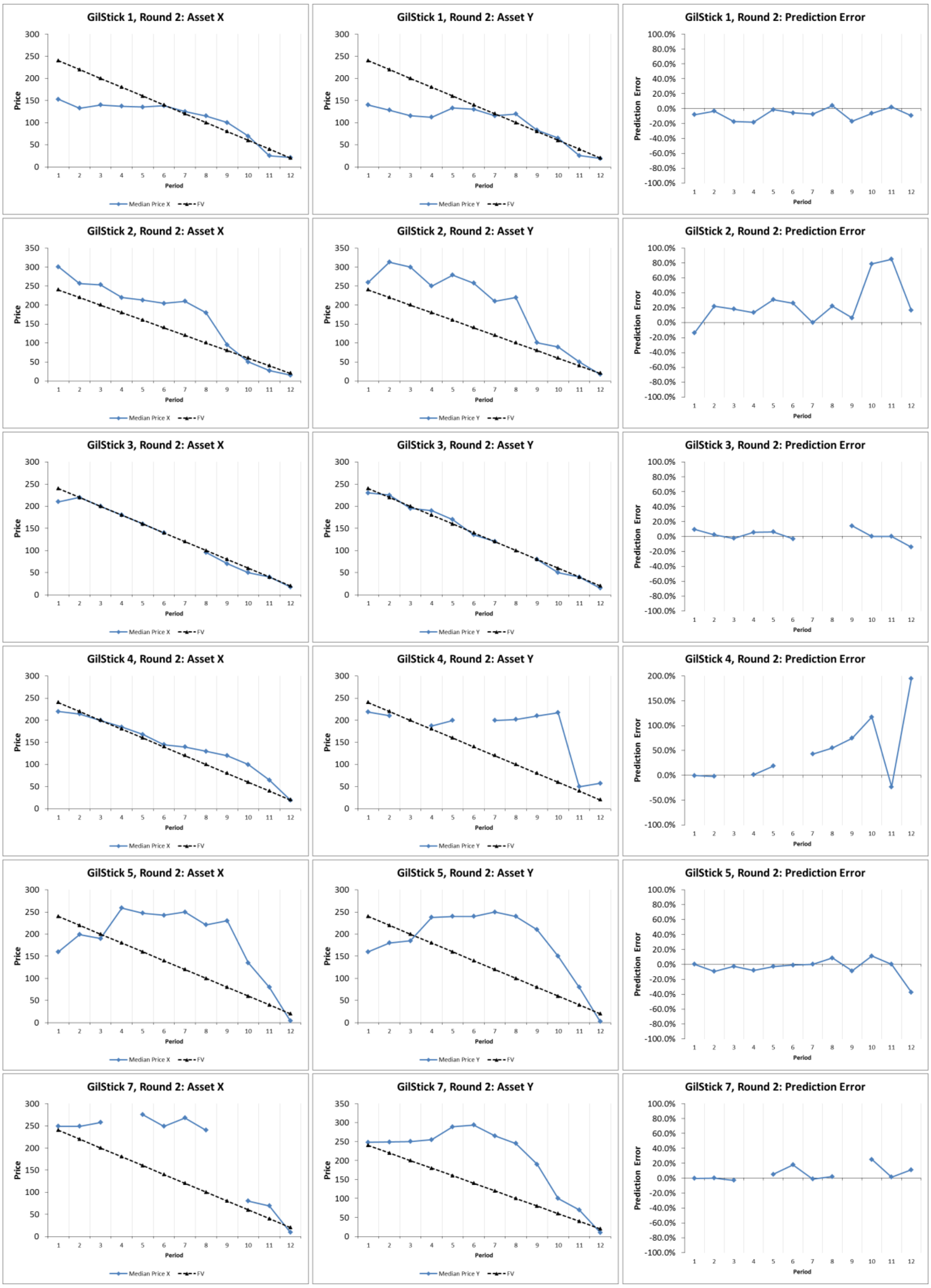
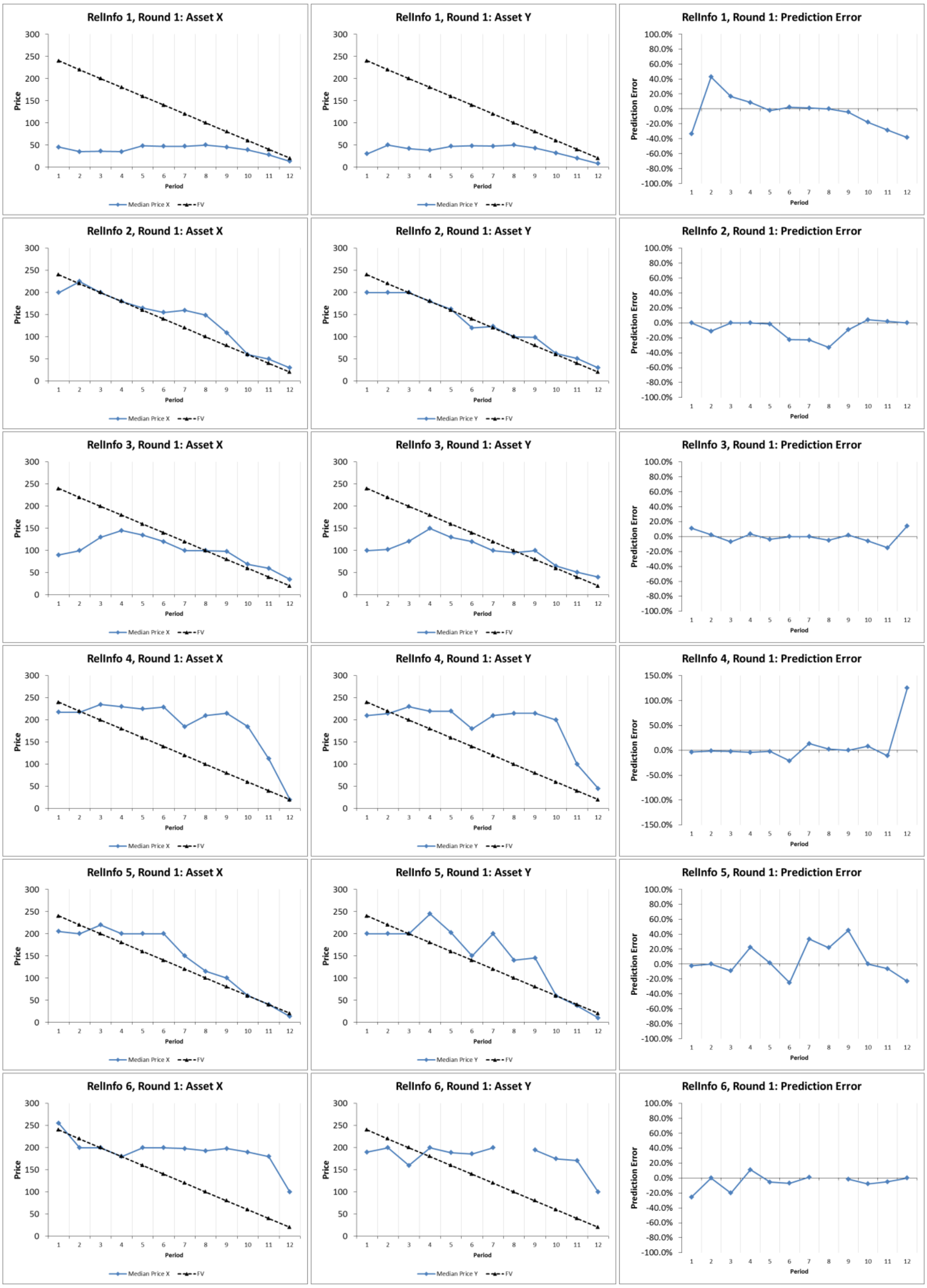
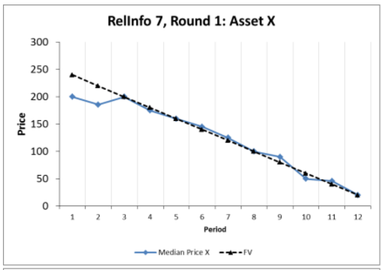
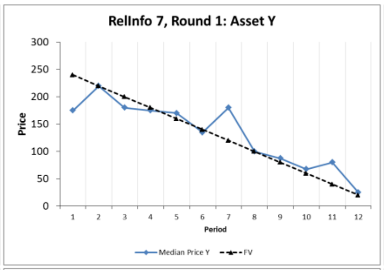


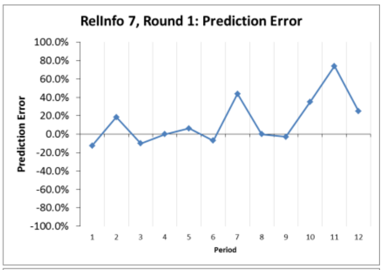
Figure B11: Median prices in individual *RelInfo* markets in Round 1

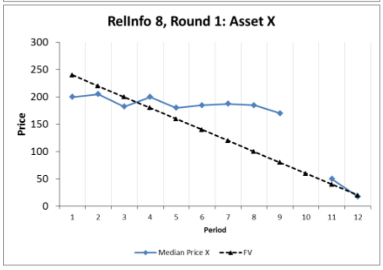
The left-most column of the figure below shows median transaction prices for asset X in each *RelInfo* market in Round 1. The middle column shows median prices for asset Y in the corresponding market, while the right-most column reports the resulting *Prediction Error*. *Prediction Error* is the percentage deviation of the relative median price of Y (i.e. median price of Y divided by the median price of X) from the risk-neutral benchmark of 1.

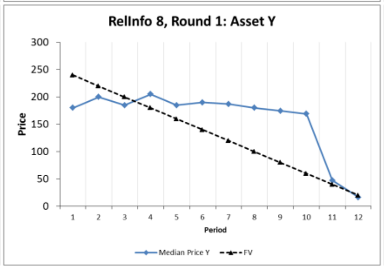



RelInfo 7, Round 1: Asset X
Price
Period
Median Price X
FV


RelInfo 7, Round 1: Asset Y
Price
Period
Median Price Y
FV


RelInfo 7, Round 1: Prediction Error
Prediction Error
Period


RelInfo 8, Round 1: Asset X
Price
Period
Median Price X
FV


RelInfo 8, Round 1: Asset Y
Price
Period
Median Price Y
FV

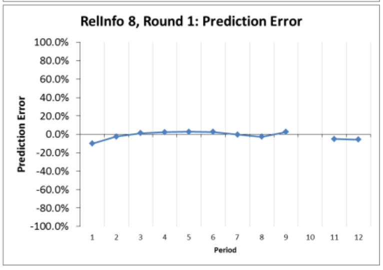
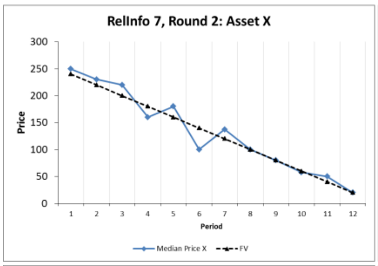
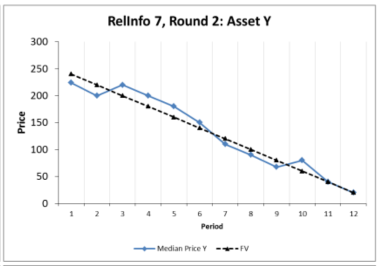
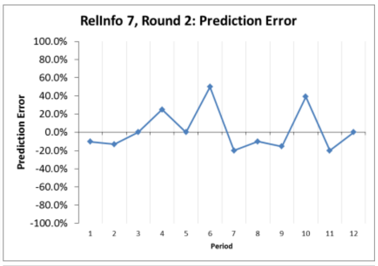
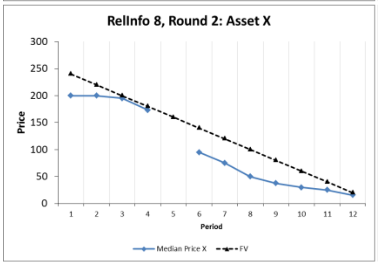
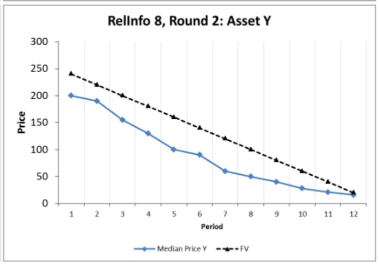
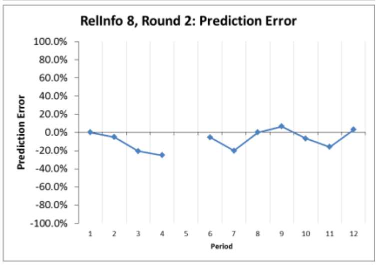

RelInfo 8, Round 1: Prediction Error
Prediction Error
Period

Figure B12: Median prices in individual *RelInfo* markets in Round 2

The left-most column of the figure below shows median transaction prices for asset X in each *RelInfo* market in Round 2. The middle column shows median prices for asset Y in the corresponding market, while the right-most column reports the resulting *Prediction Error*. *Prediction Error* is the percentage deviation of the relative median price of Y (i.e. median price of Y divided by the median price of X) from the risk-neutral benchmark of 1.


RelInfo 7, Round 2: Asset X
Price
Period
Median Price X
FV

RelInfo 7, Round 2: Asset Y
Price
Period
Median Price Y
FV

RelInfo 7, Round 2: Prediction Error
Prediction Error
Period

RelInfo 8, Round 2: Asset X
Price
Period
Median Price X
FV

RelInfo 8, Round 2: Asset Y
Price
Period
Median Price Y
FV

RelInfo 8, Round 2: Prediction Error
Prediction Error
Period

Appendix B2: Additional Tables

Tables B1-B4 below display the individual bubble measure values from each market of each treatment of the study conducted in Chapter 3. Table B1 (B2) reports for asset X (Y) in Round 1 of the market. Table B3 (B4) reports for asset X (Y) in Round 2. The relevant bubble measures are defined in section 3.4.1.2.

Table B1: Bubble measures for asset X in Round 1

Panel A: Baseline

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*$H\text{-}R^2$*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
Baseline	B1	3.36	797.50	66.46	0.78	1.91	128.51	5	12	0
	B2	1.79	291.50	17.38	0.94	2.45	53.45	5	5	1
	B3	3.31	298.00	-7.83	0.56	1.53	43.43	4	4	5
	B4	3.42	553.00	-3.92	0.47	3.33	167.78	11	7	5
	B6	0.82	286.50	-15.54	0.81	2.93	83.43	3	5	4
	B7	7.30	947.00	18.58	0.68	3.17	263.97	11	7	5
	B8	0.40	116.50	1.65	0.94	1.26	17.40	3	2	2
	Median:	**3.31**	**298.00**	**1.65**	**0.78**	**2.45**	**83.43**	**5.00**	**5.00**	**4.00**

Panel B: Carrot

	Market	*Amp*	*Tot. Disp.*	*Avg. Bias*	*$H\text{-}R^2$*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
Carrot	C1	0.41	115.50	-6.71	0.89	2.93	41.53	3	1	3
	C2	4.82	1527.50	99.29	0.09	4.00	331.88	6	5	4
	C3	1.50	868.50	-67.54	0.15	3.80	265.09	10	2	9
	C4	9.02	997.50	42.29	0.20	2.94	267.43	3	8	1
	C6	1.53	593.50	-12.13	0.23	4.40	238.53	9	5	5
	C7	5.13	545.50	26.29	0.92	2.49	103.66	11	9	3
	C8	0.55	228.50	-16.54	0.96	2.69	57.86	4	2	6
	C9	0.50	250.00	-31.25	0.91	0.63	20.67	2	0	7
	Median:	**1.51**	**569.50**	**-9.42**	**0.56**	**2.93**	**171.09**	**5.00**	**3.50**	**4.50**

Panel C: Stick

	Market	*Amp*	*Tot. Disp.*	*Avg. Bias*	*$H\text{-}R^2$*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
Stick	S1	9.58	1435.00	119.58	0.02	3.74	458.57	10	12	0
	S2	7.89	961.00	58.58	0.01	1.88	154.08	5	10	1
	S3	2.19	474.50	0.88	0.66	2.38	112.43	3	8	4
	S4	4.33	527.00	35.58	0.65	2.13	85.35	3	11	1
	S5	0.57	198.00	-0.75	0.91	1.78	23.73	5	3	6
	S6	2.97	875.50	55.46	0.27	1.28	78.68	4	11	1
	S7	1.22	340.00	15.67	0.81	2.20	58.69	5	8	2
	S8	2.28	687.00	2.75	0.12	2.86	184.57	2	10	2
	Median:	**2.63**	**607.00**	**25.63**	**0.46**	**2.16**	**98.89**	**4.50**	**10.00**	**1.50**

Table B1 cont.

Panel D: GilCarrot

	Market	*Amp*	*Tot. Disp.*	*Avg. Bias*	*H-R²*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
GilCarrot	GC1	3.30	913.00	32.75	0.08	3.83	284.38	7	9	3
	GC2	2.81	747.00	43.92	0.39	4.95	349.18	2	10	2
	GC3	4.47	770.00	-31.50	0.33	4.58	300.05	11	4	8
	GC4	0.63	229.00	14.08	0.91	4.00	54.13	2	7	4
	GC5	1.04	353.50	-8.29	0.55	1.63	101.17	6	8	2
	GC6	10.45	1008.50	41.96	0.52	2.87	240.73	4	8	4
	Median:	**3.05**	**758.50**	**23.42**	**0.46**	**3.91**	**262.55**	**5.00**	**8.00**	**3.50**

Panel E: GilStick

	Market	*Amp*	*Tot. Disp.*	*Avg. Bias*	*H-R²*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
GilStick	GS1	1.49	682.50	-12.46	0.09	2.75	225.20	6	6	4
	GS2	2.65	899.00	73.25	0.60	2.43	201.89	7	11	1
	GS3	0.79	280.00	1.67	0.74	2.83	67.68	5	5	3
	GS4	1.62	455.00	24.58	0.75	3.00	96.00	4	9	1
	GS5	9.71	1606.00	85.33	0.57	3.85	302.95	9	7	5
	GS7	2.37	546.50	33.13	0.67	2.43	111.00	6	10	2
	Median:	**1.99**	**614.50**	**28.85**	**0.63**	**2.79**	**156.44**	**6.00**	**8.00**	**2.50**

Note: Market B5 in the *Baseline* treatment, C5 in the *Carrot* treatment, and GS6 in the *GilStick* treatment are excluded because they contain subjects who participated in an earlier session of the experiment.

Table B2: Bubble measures for asset Y in Round 1

Panel A: Baseline

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*Haessel R^2*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
Baseline	B1	1.63	681.00	53.58	0.78	2.06	99.06	4	10	1
	B2	2.66	514.50	46.77	0.82	1.75	77.20	3	7	0
	B3	0.62	310.50	-16.88	0.77	1.55	44.95	4	3	6
	B4	4.25	629.50	0.29	0.05	4.68	240.83	6	7	5
	B6	1.12	530.50	-9.96	0.53	2.03	109.70	2	8	4
	B7	7.35	1089.00	22.42	0.54	3.29	316.63	11	7	4
	B8	0.81	289.50	15.38	0.81	1.69	38.23	5	9	2
	Median:	**1.63**	**530.50**	**15.38**	**0.77**	**2.03**	**99.06**	**4.00**	**7.00**	**4.00**

Panel B: Carrot

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R^2*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
Carrot	C1	0.47	168.50	-9.38	0.88	2.80	60.85	3	2	3
	C2	4.60	1570.50	98.96	0.05	2.85	265.10	6	7	2
	C3	1.61	886.00	-66.42	0.12	3.60	265.17	10	3	9
	C4	4.61	882.50	-10.54	0.05	2.66	224.34	6	5	3
	C6	1.27	559.50	-5.96	0.32	4.45	241.93	5	7	4
	C7	4.54	610.00	18.33	0.33	2.86	163.89	7	6	4
	C8	0.66	200.00	-14.50	0.86	3.43	60.57	7	2	6
	C9	1.92	469.00	-32.75	0.66	0.87	28.57	2	2	5
	Median:	**1.76**	**584.75**	**-9.96**	**0.32**	**2.85**	**194.11**	**6.00**	**4.00**	**4.00**

Panel C: Stick

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R^2*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
Stick	S1	5.74	1582.00	131.83	0.10	1.74	209.09	6	12	0
	S2	7.49	1121.00	58.58	0.02	1.98	211.43	5	9	3
	S3	1.66	473.50	-4.88	0.74	2.98	132.93	4	6	6
	S4	1.05	363.50	10.96	0.70	2.40	85.05	3	8	3
	S5	0.93	228.50	4.46	0.88	1.00	16.63	4	3	3
	S6	2.66	996.50	62.21	0.25	1.33	107.30	8	11	1
	S7	1.00	290.00	0.50	0.85	2.14	52.94	3	8	4
	S8	2.32	585.00	38.75	0.43	1.80	75.86	4	5	2
	Median:	**1.99**	**529.25**	**24.85**	**0.56**	**1.89**	**96.18**	**4.00**	**8.00**	**3.00**

Table B2 cont.

Panel D: GilCarrot

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R²*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
GilCarrot	GC1	2.91	1000.50	44.13	0.06	2.65	216.85	9	9	3
	GC2	3.96	704.00	48.67	0.49	3.93	216.43	2	10	2
	GC3	4.68	788.00	-33.17	0.14	2.93	163.28	11	4	8
	GC4	0.71	194.50	9.54	0.92	3.18	51.63	2	6	4
	GC5	0.43	274.50	-19.65	0.93	1.63	101.57	3	4	2
	GC6	21.86	649.50	-40.04	0.21	2.30	176.37	5	4	7
	Median:	**3.44**	**676.75**	**-5.05**	**0.35**	**2.79**	**169.82**	**4.00**	**5.00**	**3.50**

Panel E: GilStick

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R²*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
GilStick	GS1	1.34	631.00	-29.17	0.07	3.58	235.20	8	5	3
	GS2	3.02	889.50	79.05	0.60	1.71	116.97	4	7	1
	GS3	1.00	322.00	-4.30	0.63	1.45	41.33	5	2	3
	GS4	3.36	424.50	28.71	0.82	2.75	85.63	6	10	1
	GS5	4.24	520.00	-31.33	0.59	4.38	220.95	6	2	10
	GS7	2.03	517.00	36.42	0.78	2.49	117.46	3	9	2
	Median:	**2.52**	**518.50**	**12.20**	**0.61**	**2.62**	**117.21**	**5.50**	**6.00**	**2.50**

Note: Market B5 in the *Baseline* treatment, C5 in the *Carrot* treatment, and GS6 in the *GilStick* treatment are excluded because they contain subjects who participated in an earlier session of the experiment.

Table B3: Bubble measures for asset X in Round 2

Panel A: Baseline

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R^2*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
Baseline	B1	4.76	1324.00	120.36	0.02	2.29	261.09	7	11	0
	B2	0.53	279.50	19.63	0.93	1.50	31.80	3	8	2
	B3	0.62	221.00	-13.42	0.82	1.60	39.45	2	2	4
	B4	2.03	571.50	10.96	0.46	3.05	156.18	9	8	4
	B6	1.15	323.50	1.38	0.82	1.43	39.03	3	7	3
	B7	92.26	3083.00	234.75	0.32	3.00	510.40	11	9	3
	B8	0.35	76.50	-2.59	0.99	0.89	7.40	3	3	3
	Median:	**1.15**	**323.50**	**10.96**	**0.82**	**1.60**	**39.45**	**3.00**	**8.00**	**3.00**

Panel B: Carrot

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R^2*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
Carrot	C1	0.48	124.00	-10.36	0.92	0.75	10.60	2	1	6
	C2	0.83	265.00	6.25	0.93	3.65	71.55	2	4	5
	C3	1.95	829.50	-0.38	0.08	2.74	200.63	3	8	3
	C4	11.63	1346.50	110.54	0.05	2.00	228.09	5	9	1
	C6	0.76	339.00	-12.75	0.78	3.40	115.28	6	7	5
	C7	0.41	168.50	12.14	0.94	1.83	28.23	5	3	1
	C8	1.05	396.00	8.42	0.67	2.09	57.06	7	7	3
	C9	0.35	132.50	-16.19	0.98	0.47	6.70	2	1	3
	Median:	**0.80**	**302.00**	**2.94**	**0.85**	**2.04**	**64.30**	**4.00**	**5.50**	**3.00**

Panel C: Stick

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R^2*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
Stick	S1	3.35	1742.00	139.67	0.48	2.40	288.69	6	10	2
	S2	8.07	989.50	67.46	0.03	1.88	161.28	11	9	3
	S3	0.54	480.00	-40.00	0.87	1.55	77.33	5	0	12
	S4	7.21	1865.50	169.59	0.67	1.18	124.98	8	11	0
	S5	0.21	96.00	-8.00	0.99	0.78	7.65	2	0	5
	S6	0.56	248.00	18.83	0.95	1.78	28.85	5	7	3
	S7	0.59	123.00	-6.42	0.98	1.26	16.43	3	4	5
	S8	9.38	898.00	61.50	0.06	2.17	161.20	4	9	2
	Median:	**1.97**	**689.00**	**40.17**	**0.77**	**1.66**	**101.15**	**5.00**	**8.00**	**3.00**

Table B3 cont.

Panel D: GilCarrot

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R^2*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
GilCarrot	GC1	1.34	416.00	13.83	0.90	2.80	103.05	3	7	5
	GC2	1.08	434.50	34.38	0.86	3.00	109.70	5	8	2
	GC3	18.34	1629.50	59.68	0.86	2.30	366.90	5	5	5
	GC4	3.29	481.00	39.25	0.87	3.05	120.65	6	11	1
	GC5	1.37	260.00	11.50	0.94	1.43	37.14	4	10	2
	GC6	5.12	1065.00	98.50	0.08	2.03	217.00	5	5	1
	Median:	**2.33**	**457.75**	**36.81**	**0.87**	**2.55**	**115.18**	**5.00**	**7.50**	**2.00**

Panel E: GilStick

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R^2*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
GilStick	GS1	0.67	370.50	-22.46	0.78	3.90	107.00	7	4	6
	GS2	1.05	522.00	38.83	0.91	1.31	55.26	3	9	3
	GS3	0.18	57.00	-5.18	0.99	1.75	13.40	2	0	3
	GS4	0.78	199.00	12.25	0.95	1.75	29.30	4	8	2
	GS5	2.44	912.00	54.83	0.31	3.48	241.40	4	8	3
	GS7	1.95	667.50	64.55	0.67	3.43	165.80	2	4	1
	Median:	**0.92**	**446.25**	**25.54**	**0.85**	**2.59**	**81.13**	**3.50**	**6.00**	**3.00**

Note: Market B5 in the *Baseline* treatment, C5 in the *Carrot* treatment, and GS6 in the *GilStick* treatment are excluded because they contain subjects who participated in an earlier session of the experiment.

Table B4: Bubble measures for asset Y in Round 2

Panel A: Baseline

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R^2*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
Baseline	B1	2.70	1118.50	93.21	0.58	1.97	150.60	5	12	0
	B2	2.16	1310.00	109.17	0.68	1.50	127.50	3	12	0
	B3	0.80	572.50	-46.46	0.46	1.23	62.53	5	1	7
	B4	2.87	560.00	11.83	0.52	2.63	148.53	8	7	5
	B6	1.54	626.00	24.33	0.39	0.97	47.87	2	7	2
	B7	104.24	3689.00	286.75	0.40	3.09	416.91	11	9	3
	B8	0.51	113.00	0.40	0.97	1.03	11.31	1	3	5
	Median:	**2.16**	**626.00**	**24.33**	**0.52**	**1.50**	**127.50**	**5.00**	**7.00**	**3.00**

Panel B: Carrot

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R^2*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
Carrot	C1	0.43	271.50	-23.77	0.86	1.00	23.38	2	1	7
	C2	0.80	364.50	5.46	0.85	2.25	73.88	3	2	9
	C3	2.10	766.00	-14.42	0.01	2.00	169.54	5	7	4
	C4	8.56	667.00	21.25	0.04	2.03	140.66	5	7	5
	C6	2.26	422.50	-3.29	0.37	3.93	162.88	5	7	5
	C7	3.30	379.00	10.75	0.94	1.77	69.17	6	7	3
	C8	1.35	404.50	-15.05	0.79	1.00	45.40	4	4	5
	C9	1.71	310.00	6.25	0.78	0.50	22.17	1	2	3
	Median:	**1.91**	**391.75**	**1.08**	**0.78**	**1.89**	**71.52**	**4.50**	**5.50**	**5.00**

Panel C: Stick

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R^2*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
Stick	S1	3.42	1565.00	134.73	0.42	1.94	248.77	2	6	2
	S2	4.51	841.00	71.00	0.16	1.40	109.18	4	7	3
	S3	0.59	453.50	-41.23	0.92	1.48	81.65	3	0	7
	S4	1.29	509.50	-41.77	0.76	1.13	38.60	2	2	8
	S5	0.37	34.50	-3.14	1.00	1.10	6.50	2	0	4
	S6	1.16	322.00	25.64	0.93	1.08	28.90	4	6	1
	S7	0.36	353.50	-32.14	0.91	1.09	40.26	4	0	9
	S8	8.23	755.50	48.55	0.01	1.46	109.77	3	4	3
	Median:	**1.23**	**481.50**	**11.25**	**0.84**	**1.26**	**60.95**	**3.00**	**3.00**	**3.50**

Table B4 cont.

Panel D: GilCarrot

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R²*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
GilCarrot	GC1	1.03	277.00	-3.00	0.92	1.53	33.40	2	5	4
	GC2	1.51	448.50	33.38	0.84	2.23	76.55	6	10	2
	GC3	16.01	1654.50	57.96	0.86	1.75	260.08	10	6	5
	GC4	5.21	476.50	31.38	0.84	2.50	94.15	5	9	1
	GC5	2.46	302.00	14.83	0.81	1.43	37.83	6	8	4
	GC6	1.23	588.50	-34.21	0.39	1.73	76.57	9	3	7
	Median:	**1.99**	**462.50**	**23.10**	**0.84**	**1.74**	**76.56**	**6.00**	**7.00**	**4.00**

Panel E: GilStick

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R²*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
GilStick	GS1	1.55	429.00	-31.17	0.49	3.10	145.65	5	3	7
	GS2	1.36	793.50	65.71	0.83	1.14	83.74	2	11	1
	GS3	0.32	60.00	-0.82	0.98	1.23	6.75	1	2	1
	GS4	2.76	595.00	53.40	0.37	1.85	66.93	3	6	2
	GS5	2.56	920.50	51.29	0.22	3.88	288.18	7	8	3
	GS7	1.96	924.50	75.38	0.65	2.66	109.57	2	11	1
	Median:	**1.75**	**694.25**	**52.35**	**0.57**	**2.25**	**96.66**	**2.50**	**7.00**	**1.50**

Note: Market B5 in the *Baseline* treatment, C5 in the *Carrot* treatment, and GS6 in the *GilStick* treatment are excluded because they contain subjects who participated in an earlier session of the experiment.

Tables B5 and B6 below display the respective Round 1 and Round 2 bubble measure values from the individual markets of the *RelInfo* treatment, which is unique to the study in Chapter 4. The bubble measure values of the *Baseline* treatment and the other Chapter 4 treatments *Tournament* (*GilTournament)*, which consists of the markets of the *Carrot* and *Stick* (*GilCarrot* and *GilStick*) treatments pooled together, can be found in Tables B1-B4 above.

Table B5: RelInfo treatment bubble measures for asset X and Y in Round 1

Panel A: Asset X

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R^2*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
RelInfo	R1	0.52	1092.00	-91.00	0.15	4.43	528.93	11	0	12
	R2	0.72	203.00	10.25	0.88	2.58	41.58	4	5	1
	R3	1.34	502.00	-31.50	0.59	4.05	176.68	8	4	7
	R4	2.19	771.50	60.13	0.50	1.80	96.49	5	9	2
	R5	0.68	267.00	11.92	0.84	1.93	42.55	3	7	2
	R6	4.15	774.00	61.17	0.48	1.40	89.40	9	8	1
	R7	0.23	115.50	-5.29	0.97	3.03	31.49	2	2	2
	R8	1.40	412.50	23.86	0.56	1.71	69.00	8	6	3
	Median:	**1.03**	**457.25**	**11.08**	**0.57**	**2.25**	**79.20**	**6.50**	**5.50**	**2.00**

Panel B: Asset Y

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R^2*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
RelInfo	R1	0.39	1104.50	-92.04	0.31	3.70	413.00	11	0	12
	R2	0.59	128.50	-2.63	0.97	1.73	20.28	3	4	2
	R3	1.39	497.50	-32.13	0.66	2.78	133.05	8	4	8
	R4	2.08	770.50	58.29	0.52	1.46	95.60	3	10	2
	R5	1.11	375.00	19.17	0.67	2.08	53.88	3	6	2
	R6	4.34	726.00	46.00	0.16	1.63	101.20	4	4	3
	R7	1.06	225.00	2.92	0.88	4.43	65.43	4	4	2
	R8	2.08	556.50	29.96	0.52	1.57	64.83	9	8	3
	Median:	**1.25**	**527.00**	**11.04**	**0.59**	**1.90**	**80.51**	**4.00**	**4.00**	**2.50**

Table B6: RelInfo treatment bubble measures for asset X and Y in Round 2

Panel A: Asset X

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R^2*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
RelInfo	R1	0.69	691.50	-57.63	0.38	2.35	192.60	8	0	8
	R2	0.45	131.00	-10.27	0.95	1.23	5.40	5	2	3
	R3	3.97	454.00	7.58	0.97	2.18	83.13	11	6	5
	R4	0.56	185.50	1.04	0.94	1.40	22.14	2	6	3
	R5	0.35	161.50	11.38	0.98	1.60	20.33	3	7	2
	R6	1.04	290.00	23.33	0.86	1.03	22.17	2	2	1
	R7	0.53	149.50	2.04	0.93	2.57	23.69	2	3	1
	R8	0.51	304.00	-27.64	0.95	1.03	26.91	4	0	7
	Median:	**0.55**	**237.75**	**1.54**	**0.94**	**1.50**	**22.93**	**3.50**	**2.50**	**3.00**

Panel B: Asset Y

	Market	*Amp.*	*Tot. Disp.*	*Avg. Bias*	*H-R^2*	*Turn.*	*Norm. Dev*	*Dur.*	*Boom Dur*	*Bust Dur*
RelInfo	R1	0.43	874.00	-72.83	0.59	2.05	163.15	11	0	12
	R2	0.58	204.00	-15.67	0.91	1.23	26.90	3	2	6
	R3	3.53	510.00	-5.50	0.86	2.28	106.20	9	5	5
	R4	0.83	175.50	-8.23	0.95	0.94	20.89	1	2	4
	R5	0.88	275.50	21.63	0.92	1.23	17.60	2	10	1
	R6	1.30	346.00	14.45	0.77	0.97	23.40	3	3	3
	R7	0.35	158.50	1.79	0.94	2.71	37.43	3	4	3
	R8	0.40	480.50	-40.04	0.95	1.17	44.34	5	0	12
	Median:	**0.70**	**310.75**	**-6.86**	**0.92**	**1.23**	**32.16**	**3.00**	**2.50**	**4.50**

Appendix B3: Participant instructions for Chapter 3 and 4 experiments

The written instructions provided to participants in the experiment pertaining to <u>Chapters 3 and 4</u> are shown below (next page). These instructions relate to sessions where the market screen for asset X was displayed on the left-hand side of the screen – the instructions for sessions where Y was displayed on the left are qualitatively the same. Treatments vary according to how earnings are calculated, which is addressed in section 6 of the instructions – this section was unique to each treatment. Treatments also vary according the amount of relative performance feedback given, which is covered in Section 5.

1. General Instructions

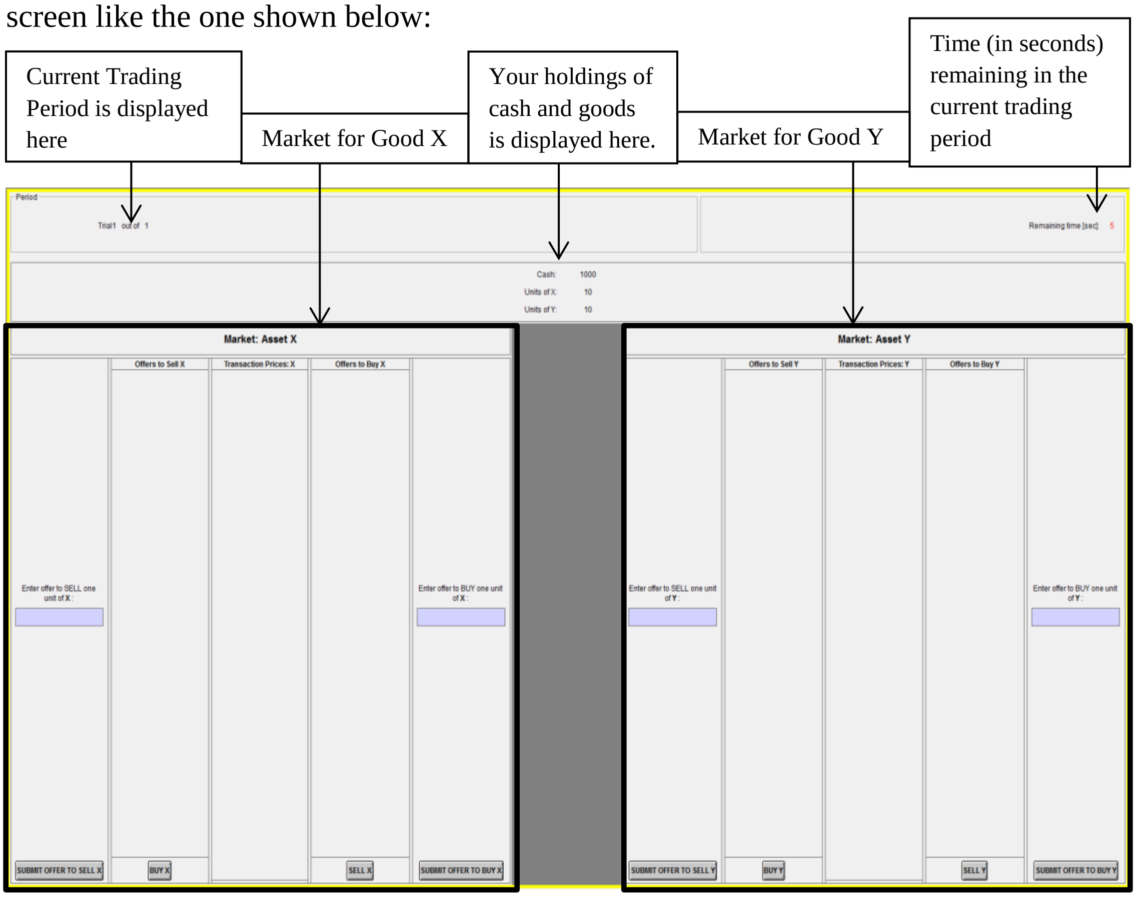
This is an experiment in the economics of market decision-making. The instructions are simple and if you follow them carefully and make good decisions, you may earn a considerable amount of money, which will be paid to you, in cash, at the end of the experiment. The experiment will consist of a sequence of trading periods in which you will have the opportunity to buy and sell in a market. All trading will be in terms of *francs*. The cash payment to you at the end of the experiment will be in Australian dollars, rounded up to the nearest 5 dollars. The conversion rate is ____ francs to 1 dollar.

The experiment will last no more than 2.5 hours, and will include up to 30 minutes of instructions and practice. Please do not speak with any other participants during the experiment. Please also remember to switch off your mobile phone. Failure to comply with these rules will result in your exclusion from the experiment and the forfeiture of all payments.

2. How to Use the Computerised Market

Before proceeding, we introduce the market interface that you will be using for the remainder of the experiment. Please note that any actions you take during this demonstration will **<u>not</u>** count towards your earnings or influence your position later in the experiment.

In this experiment, you will have the opportunity to buy and sell two different goods, called X and Y, in separate markets. In each trading period, you will see a computer screen like the one shown below:



Market: Good X

The market for good X is displayed on the **left-hand** side of your screen. All activity in relation to good X is shown and conducted here.

When you would like to offer to sell a unit of X, use the text area entitled “Enter offer to sell one unit of X” in the first column on the left. In that text area you can enter the price at which you are offering to sell a unit of X, and then select “Submit Offer To Sell X”. Please do so now. Type in a number in the appropriate space, and then click on the button labelled “Submit Offer To Sell X”.

You will notice that 8 numbers, one submitted by each participant in your market, now appear in the second column from the left, entitled “Offers to Sell X”. Your offer is listed in blue. Submitting a new offer will replace your previous offer.

The lowest offer-to-sell price will always be on the top of that list and will, by default, be selected. You can select a different offer by clicking on it. It will then be highlighted. If you select “Buy X”, the button at the bottom of this column, you will buy one unit of X for the currently selected sell price. Please purchase a unit now by selecting an offer and clicking the “Buy” button. Since each of you had offered to sell a unit of X and attempted to buy a unit of X, if all were successful, you all have the same number of units of X you started out with. This is because you bought one unit of X and sold one unit of X.

You may make an offer to buy a unit of X by selecting “Submit Offer to Buy X.” Please do so now. Type a number in the text area “Enter offer to buy one unit of X”, then press the button labelled “Submit Offer To Buy X”. All offers to buy X appear under the column entitled “Offers to Buy X”. The highest offer-to-buy price will always be on top of that list and will, by default, be selected. You can accept any of the offers-to-buy by selecting the offer and then clicking on the “Sell X” button. Please do so now.

The middle column of the market, labelled “Transaction Prices: X”, shows the prices at which X has been bought and sold in this period. The most recent transaction will be listed at the top.

Market: Good Y

The market for good Y is displayed on the **right-hand** side of your screen. All activity in relation to good Y is shown and conducted here. The layout of this market is identical to the market for X. The trading rules and procedures for posting and accepting offers to buy and sell Y are also the same.

To post an offer to sell a unit of Y, use the text area entitled “Enter offer to sell one unit of Y” and then select “Submit Offer To Sell Y”. Please do so now.

You can purchase a unit of Y by clicking the button “Buy Y” at the bottom of the column called “Offers to Sell Y”. Once again, the lowest offer-to-sell price is listed at the top and is selected by default. You can accept any offer by selecting it before clicking “Buy Y”. Please purchase a unit of Y now.

To make an offer to buy a unit of Y, type a number into the text area entitled "Enter offer to buy one unit of Y" and then select "Submit Offer To Buy Y". Please do so now.

These offers are listed in the column "Offers to Buy Y". To accept an offer, click "Sell Y" at the bottom of this column. The highest offer-to-buy price is selected by default. You can accept any of the offers by selecting it before clicking "Sell Y". Please do so now.

The middle column of the market, labelled "Transaction Prices: Y", shows the prices at which Y has been bought and sold in this period. The most recent transaction will be listed at the top.

Other features of both markets:

When you buy a unit of a good (i.e. X or Y), your Cash balance decreases by the price of the purchase. Any other existing offer to buy that good submitted by you is also cancelled. When you sell a unit of a good, your Cash balance increases by the price of the sale, and any other existing offer to sell that good submitted by you is cancelled.

You can participate in both markets at the same time.

If you make offers to buy in both markets at the same time, and say your offer to buy X is accepted first, then your offer to buy Y remains standing as long as you have enough Cash after the purchase of X to honour it, and vice versa. If you do not have enough Cash, then your offer in the second market is cancelled. Similarly, if you have a standing offer to buy in one market, and accept another trader's sell offer in the second market, then your offer to buy in the first market is cancelled if your remaining Cash balance is less than the amount of your offer.

You will now have about 10 minutes to buy and sell in both markets. This is a practice period. **Your actions in the practice period do not count toward your earnings and do not influence your position later in the experiment.** The only goal of the practice period is to master the use of the interface. Please be sure that you have successfully submitted offers to buy and offers to sell in both markets. Also be sure that you have accepted buy and sell offers in both markets. If you have any questions, please raise your hand and the experimenter will come by and assist you.

3. Specific Instructions for this Experiment

This experiment consists of you and 7 other traders. At the beginning of the experiment, all traders will be endowed with a portfolio consisting of 5 units each of two types of goods, called ‘X’ and ‘Y’, and 1950 francs in Cash.

The experiment consists of 12 periods, each lasting 3 minutes. In each period, two separate markets will operate in which you may buy and/or sell units of good X and Y respectively. Both goods can be considered assets with lives of 12 periods, and your inventory of X and Y carries over from one trading period to the next. Note that your cash balance and inventory of assets cannot fall below zero.

At the end of each trading period, each unit of X pays an identical dividend, which is randomly determined by the computer. The possible dividend values and the associated likelihoods are shown below:

Asset: X

Dividend		**Likelihood**
10	→	$^1/_2$
30	→	$^1/_2$

Since each dividend is equally likely, the average dividend per period for X is 20 francs.

Each unit of Y also pays an identical dividend at the end of each period, randomly determined by the computer. The possible dividend values and the associated likelihoods are shown below:

Asset: Y

Dividend		**Likelihood**
0	→	$^4/_5$
100	→	$^1/_5$

The average dividend per period for asset Y is 20 francs ($0 \times {}^4/_5 + 100 \times {}^{15}/\quad = 20$).

The dividend draws for X and Y are independent across trading periods. This means that for both assets, the likelihood of a particular dividend in a period is not affected by the dividends in previous periods. In addition, the dividend draws for X and Y are independent of each other. This means that the occurrence of a particular dividend for X does not affect the likelihood of a particular dividend for Y, and vice versa.

Each unit of X and Y expires worthless after the final dividend is paid at the end of period 12.

4. Average Holding Value Table

You can use the table at the end of this document to help you make decisions. It calculates the average amount of dividends you will receive if you hold a unit of an asset in your inventory for the rest of the market, or equivalently, how much in dividends you give up, on average, when you sell a unit at any time. Each of the 5 columns of the table is described below:

1. *Ending Period*: indicates the last trading period of the market, period 12.
2. *Current Period*: indicates the period during which the average holding value is being calculated.
3. *Number of holding periods*: This is equivalent to the number of times a dividend can be received if a unit of an asset is held in your inventory from the current period to the end of the market.
4. *Average Dividend Per Period*: gives the average amount that the dividend will be in each period for each unit of the asset that is held in your inventory. The number in this column is 20. This is because the average dividend in each period for both X and Y is 20 francs. Since both types of assets have the same average dividend per period, you can use this table to determine the average holding value for both X and Y.
5. *Average Holding Value Per Unit of Inventory*: gives the expected total dividend for the remainder of the market for each unit of an asset that is held in your inventory for the rest of the market. That is, for each unit you hold in your inventory for the remainder of the market, you will receive on average the amount listed in column 5 in dividends. Equivalently, it tells you how much in future dividends you give up on average when you sell a unit in the current period. The number in column 5 is calculated by multiplying the numbers in columns 3 and 4.

Example: Suppose that there are 4 periods remaining. Since the dividend paid on a unit of X has a 50% chance of being 10 and a 50% chance of being 30, the dividend is in expectation 20 per period for each unit of X. Since the dividend paid on a unit of Y has an 80% chance of being 0 and a 20% chance of being 100, the dividend in expectation is also 20 per period for each unit of Y. If you hold a unit of X or Y for 4 periods, the total dividend paid on that unit over the 4 periods is in expectation 4×20 = 80.

5. Summary Screen

At the end of each trading period, a status report will appear on screen for 30 seconds. It displays the following information:

- Your Cash balance before the payment of dividends. This is calculated as:

  CASH BEFORE DIVIDENDS = BEGINNING OF PERIOD CASH

  + (PERIOD SALES REVENUE – PERIOD EXPENDITURE ON PURCHASES)

- The dividends paid by X and Y in this period.
- The number of units of X and Y in your inventory at the end of the period.
- The total amount of dividends you receive this period. This is calculated as:

  PERIOD TOTAL DIVIDEND = (END-OF-PERIOD UNITS OF X × DIVIDEND PER UNIT OF X FOR THE PERIOD)

  + (END-OF-PERIOD UNITS OF Y × DIVIDEND PER UNIT OF Y FOR THE PERIOD)

- Your Cash balance at the end of the period, which is calculated as follows:

  END-OF-PERIOD CASH = CASH BEFORE DIVIDENDS + PERIOD TOTAL DIVIDEND

- Your Account Total. This is equal to your end-of-period Cash plus the value of your holdings of X and Y.

  In periods 1 through to 11, your end-of-period holdings of X and Y are valued at their respective median traded price in that period. So, your Account Total at the end of period 1-11 is calculated as:
  ACCOUNT TOTAL = END-OF-PERIOD CASH

  + (END-OF-PERIOD UNITS OF X × MEDIAN TRADED PRICE OF X DURING PERIOD)

  + (END-OF-PERIOD UNITS OF Y × MEDIAN TRADED PRICE OF Y DURING PERIOD)

  Since all units of X and Y expire worthless after the final dividend payment at the end of period 12 (i.e. at the end of the market), your Account Total at the end of period 12 is equal to your end-of-period Cash balance:

  ACCOUNT TOTAL (end of period 12) = END-OF-PERIOD CASH

- The average Account Total in your market. ****← this point does <u>not</u> appear in the *Baseline* treatment instructions, but does appear for all other treatments****
- Your rank out of the 8 participants in your market, based on your Account Total. A rank of 1 indicates the highest Account Total; a rank of 2 indicates the second-highest Account Total, and so on. ****← this point <u>only</u> appears in the instructions for the *GilCarrot* and *GilStick* treatments****

After seeing the summary screen, press the “Continue” button to go to the next period. The next period will begin once everyone has pressed the “Continue” button, or once the 30 seconds have elapsed, whichever comes first.

6. Your Earnings

**** *Baseline* and *RelInfo* only: ****

Your earnings from this market will equal the balance of your Account Total at the end of the market. Remember that this is equal to your Cash balance at the end of the market.

Note that you do not have to calculate your earnings by yourself. The computer does all the work.

**** *Carrot* only: ****

Your earnings from this market will depend on your performance relative to the other traders in your market. Your performance is measured by comparing the balance of your Account Total at the end of the market (i.e. your final Cash balance) to the average end-of-market Account total/Cash balance in your market. Your payoff is calculated as follows:

$$Earnings_i = \begin{cases} 3000 & if \quad C_i < C^* \\ 3000 + 2(C_i - C^*) & if \quad C_i \geq C^* \end{cases}$$

where C_i is your final Account Total/Cash balance and C* is the average final Account total/Cash balance in your market.

Example: Suppose that the average end-of-market Cash balance in your market is 3500 francs. If your final Cash balance is say 3200 francs, you will earn 3000 francs. On the other hand, if your final Cash balance is say 4500 francs, you will earn 3000 + 2×(4500 – 3500) = 5000 francs.

Note that you do not have to calculate your earnings by yourself. The computer does all the work.

**** *Stick* only: ****

Your earnings from this market will depend on your performance relative to the other traders in your market. Your performance is measured by comparing the balance of your Account Total at the end of the market (i.e. your final Cash balance) to the average end-of-market Account total/Cash balance in your market. Your payoff is calculated as follows:

$$Earnings_i = \begin{cases} 0 & if \quad C_i < \frac{1}{2}C^* \\ 3000 & if \quad \frac{1}{2}C^* \leq C_i \leq C^* \\ 3000 + 2(C_i - C^*) & if \quad C_i > C^* \end{cases}$$

where C_i is your final Account Total/Cash balance and C* is the average final Account total/Cash balance in your market.

Example: Suppose that the average end-of-market Cash balance in your market is 3500 francs. If your final Cash balance is say 1000 francs, you will earn 0 francs from this market. If your final Cash balance is 3200 francs, you will earn 3000 francs. On the other hand, if your final Cash balance is say 4500 francs, you will earn 3000 + 2×(4500 – 3500) = 5000 francs.

Note that you do not have to calculate your earnings by yourself. The computer does all the work.

***** GilCarrot* only: *****

Your earnings from this market will depend on your performance relative to the other traders in your market. The size of your payoff is determined by your rank at the **end** of the market (i.e. period 12), and is calculated as follows:

Rank	Your Earnings (francs)
1 ← largest final Account Total/Cash balance	10,000
2	4,000
3	4,000
4	4,000
5	4,000
6	4,000
7	4,000
8 ←smallest final Account Total/Cash balance	4,000

Note that you do not have to calculate your earnings by yourself. The computer does all the work.

***** GilStick* only: *****

Your earnings from this market will depend on your performance relative to the other traders in your market. The size of your payoff is determined by your rank at the **end** of the market (i.e. period 12), and is calculated as follows:

Rank	Your Earnings (francs)
1 ← largest final Account Total/Cash balance	10,000
2	4,000
3	4,000
…	4,000
'Last' ←smallest final Account Total/Cash balance	0

Note that you do not have to calculate your earnings by yourself. The computer does all the work.

Average Holding Value Table

Ending Period	Current Period	Number of Holding Periods	×	Average Dividend Per Period	=	Average Holding Value Per Unit in Inventory
12	1	12		20		240
12	2	11		20		220
12	3	10		20		200
12	4	9		20		180
12	5	8		20		160
12	6	7		20		140
12	7	6		20		120
12	8	5		20		100
12	9	4		20		80
12	10	3		20		60
12	11	2		20		40
12	12	1		20		20

Appendix B4: End-of-experiment Questionnaire

Participants of the experiment relating to Chapter 3 and 4 completed the following survey after the market stage. The electronic version of the survey is shown, which was generated using the Qualtrics© software (Qualtrics, Provo, UT). A paper version of this survey was administered for roughly half of the total number of sessions (the first half). Section 1 of the survery is a modified version of the questionnaire used by Ackert et al. (2001), Section 2 is the Cognitive Reflection Test developed by Frederick (2005), while Sections 3-5 comprise the 30-item Domain-Specific Risk-Taking (DOSPERT) Scale developed by Blais and Weber (2006).

Qualtrics Survey Software 28/03/2015 11:27 pm

Section 1

What year are you in university?

What department/school are you in at university (e.g. finance, economics)?

What is your gender?

- Male
- Female

What is your age?

How interesting did you find this experiment? (select the appropriate number)

Not very interesting						Very interesting
1	2	3	4	5	6	7

Have you every traded securities for yourself or others?

- Yes
- No

https://asb.qualtrics.com/ControlPanel/Ajax.php?action=GetSurveyPrintPreview&T=6VZ5eZc0g8iE4G2hlszbxY Page 1 of 8

Have you ever participated in the management of an investment portfolio?

- Yes
- No

Compared to the money available to you from alternative sources, how would you characterise the amount of money earned for participating in this experiment? (select the appropriate number)

Nominal Amount						Considerable Amount
1	2	3	4	5	6	7

How would you characterise your attitude towards risk while participating in the market? (select the appropriate number)

Very risk averse						Very risk-taking
1	2	3	4	5	6	7

Describe as best you can the trading/investment strategy you followed, including any changes in strategy between the two rounds of the market.

If you made an error in entering a price, or clicked the wrong button at any stage during the experiment, please tell us exactly what went wrong and in which trading period.

If you wish to leave any feedback for the experimenters regarding this experiment (e.g. the instructions), please do so in the space below.

Section 2

Please answer the following questions

A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?

If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

Section 3

For each of the following statements, please indicate the **likelihood** that you would engage in the described activity or behavior if you were to find yourself in that situation. Provide a rating from *Extremely Unlikely* to *Extremely Likely*, using the following scale:

	Extremely Unlikely	Moderately Unlikely	Somewhat Unlikely	Not Sure	Somewhat Likely	Moderately Likely	Extremely Likely
Admitting that your tastes are different from those of a friend.	○	○	○	○	○	○	○
Going camping in the wilderness.	○	○	○	○	○	○	○
Betting a day's income at the							

horse races.							
Investing 10% of your annual income in a moderate growth mutual fund.							
Drinking heavily at a social function.							
Taking some questionable deductions on your income tax return.							
Disagreeing with an authority figure on a major issue.							
Betting a day's income at a high-stake poker game.							
	Extremely Unlikely	Moderately Unlikely	Somewhat Unlikely	Not Sure	Somewhat Likely	Moderately Likely	Extremely Likely
Having an affair with a married man/woman.							
Passing off somebody else's work as your own.							
Going down a ski run that is beyond your ability.							
Investing 5% of your annual income in a very speculative stock.							
Going whitewater rafting at high water in the spring.							
Betting a day's income on the outcome of a sporting event.							
Engaging in unprotected sex.							
Revealing a friend's secret to someone else.							
	Extremely Unlikely	Moderately Unlikely	Somewhat Unlikely	Not Sure	Somewhat Likely	Moderately Likely	Extremely Likely
Driving a car without wearing a seat belt.							
Investing 10% of your annual income in a new business venture.							
Taking a skydiving class.							
Riding a motorcycle without a helmet.							
Choosing a career that you truly enjoy over a more secure one.							
Speaking your mind about an unpopular issue in a meeting at work.							
Sunbathing without sunscreen.							
Bungee jumping off a tall bridge.							
	Extremely Unlikely	Moderately Unlikely	Somewhat Unlikely	Not Sure	Somewhat Likely	Moderately Likely	Extremely Likely
Piloting a small plane.							

Walking home alone at night in an unsafe area of town.							
Moving to a city far away from your extended family.							
Starting a new career in your mid-thirties.							
Leaving your young children alone at home while running an errand.							
Not returning a wallet you found that contains $200.							

Section 4

People often see some risk in situations that contain uncertainty about what the outcome or consequences will be and for which there is the possibility of negative consequences. However, riskiness is a very personal and intuitive notion, and we are interested in **your gut level assessment of how risky** each situation or behavior is.

For each of the following statements, please indicate **how risky you perceive** each situation. Provide a rating from *Not at all Risky* to *Extremely Risky*, using the following scale:

	Not at all risky	Slightly risky	Somewhat risky	Moderately Risky	Risky	Very Risky	Extremely Risky
Admitting that your tastes are different from those of a friend.							
Going camping in the wilderness.							
Betting a day's income at the horse races.							
Investing 10% of your annual income in a moderate growth mutual fund.							
Drinking heavily at a social function.							
Taking some questionable deductions on your income tax return.							
Disagreeing with an authority figure on a major issue.							
Betting a day's income at a high-stake poker game.							
	Not at all risky	Slightly risky	Somewhat risky	Moderately Risky	Risky	Very Risky	Extremely Risky
Having an affair with a married man/woman.							
Passing off somebody else's work as your own.							
Going down a ski run that is beyond your ability.							
Investing 5% of your annual income in a very speculative stock.							
Going whitewater rafting at high water in the spring.							
Betting a day's income on the							

outcome of a sporting event.							
Engaging in unprotected sex.	○	○	○	○	○	○	○
Revealing a friend's secret to someone else.	○	○	○	○	○	○	○
	Not at all risky	Slightly risky	Somewhat risky	Moderately Risky	Risky	Very Risky	Extremely Risky
Driving a car without wearing a seat belt.	○	○	○	○	○	○	○
Investing 10% of your annual income in a new business venture.	○	○	○	○	○	○	○
Taking a skydiving class.	○	○	○	○	○	○	○
Riding a motorcycle without a helmet.	○	○	○	○	○	○	○
Choosing a career that you truly enjoy over a more secure one.	○	○	○	○	○	○	○
Speaking your mind about an unpopular issue in a meeting at work.	○	○	○	○	○	○	○
Sunbathing without sunscreen.	○	○	○	○	○	○	○
Bungee jumping off a tall bridge.	○	○	○	○	○	○	○
	Not at all risky	Slightly risky	Somewhat risky	Moderately Risky	Risky	Very Risky	Extremely Risky
Piloting a small plane.	○	○	○	○	○	○	○
Walking home alone at night in an unsafe area of town.	○	○	○	○	○	○	○
Moving to a city far away from your extended family.	○	○	○	○	○	○	○
Starting a new career in your mid-thirties.	○	○	○	○	○	○	○
Leaving your young children alone at home while running an errand.	○	○	○	○	○	○	○
Not returning a wallet you found that contains $200.	○	○	○	○	○	○	○

Section 5

For each of the following statements, please indicate **the benefits** you would obtain from each situation. Provide a rating from **1 to 7**, using the following scale

	No benefits at all			Moderate benefits			Great benefits
	1	2	3	4	5	6	7
Admitting that your tastes are different from those of a friend.	○	○	○	○	○	○	○
Going camping in the wilderness.	○	○	○	○	○	○	○
Betting a day's income at the horse races.	○	○	○	○	○	○	○

Investing 10% of your annual income in a moderate growth mutual fund.	○	○	○	○	○	○	○
Drinking heavily at a social function.	○	○	○	○	○	○	○
Taking some questionable deductions on your income tax return.	○	○	○	○	○	○	○
Disagreeing with an authority figure on a major issue.	○	○	○	○	○	○	○
Betting a day's income at a high-stake poker game.	○	○	○	○	○	○	○

	No benefits at all			Moderate benefits			Great benefits
	1	2	3	4	5	6	7
Having an affair with a married man/woman.	○	○	○	○	○	○	○
Passing off somebody else's work as your own.	○	○	○	○	○	○	○
Going down a ski run that is beyond your ability.	○	○	○	○	○	○	○
Investing 5% of your annual income in a very speculative stock.	○	○	○	○	○	○	○
Going whitewater rafting at high water in the spring.	○	○	○	○	○	○	○
Betting a day's income on the outcome of a sporting event.	○	○	○	○	○	○	○
Engaging in unprotected sex.	○	○	○	○	○	○	○
Revealing a friend's secret to someone else.	○	○	○	○	○	○	○

	No benefits at all			Moderate benefits			Great benefits
	1	2	3	4	5	6	7
Driving a car without wearing a seat belt.	○	○	○	○	○	○	○
Investing 10% of your annual income in a new business venture.	○	○	○	○	○	○	○
Taking a skydiving class.	○	○	○	○	○	○	○
Riding a motorcycle without a helmet.	○	○	○	○	○	○	○
Choosing a career that you truly enjoy over a more secure one.	○	○	○	○	○	○	○
Speaking your mind about an unpopular issue in a meeting at work.	○	○	○	○	○	○	○
Sunbathing without sunscreen.	○	○	○	○	○	○	○
Bungee jumping off a tall bridge.	○	○	○	○	○	○	○
	No benefits at all			Moderate benefits			Great benefits
	1	2	3	4	5	6	7

Qualtrics Survey Software 28/03/2015 11:27 pm

Piloting a small plane.	○	○	○	○	○	○	○
Walking home alone at night in an unsafe area of town.	○	○	○	○	○	○	○
Moving to a city far away from your extended family.	○	○	○	○	○	○	○
Starting a new career in your mid-thirties.	○	○	○	○	○	○	○
Leaving your young children alone at home while running an errand.	○	○	○	○	○	○	○
Not returning a wallet you found that contains $200.	○	○	○	○	○	○	○

References

Abel, Andrew B., 1990, Asset prices under habit formation and catching up with the Joneses, *The American Economic Review* 80, 38-42.

Abreu, Dilip, and Markus K. Brunnermeier, 2003, Bubbles and crashes, *Econometrica* 71, 173-204.

Acker, Daniella, and Nigel W. Duck, 2006, A tournament model of fund management, *Journal of Business Finance & Accounting* 33, 1460-1483.

Ackert, Lucy, Narat Charupat, Bryan Church, and Richard Deaves, 2006a, An experimental examination of the house money effect in a multi-period setting, *Experimental Economics* 9, 5-16.

Ackert, Lucy F., Narat Charupat, Bryan K. Church, and Richard Deaves, 2006b, Margin, short selling, and lotteries in experimental asset markets, *Southern Economic Journal* 73, 419-436.

Ackert, Lucy F., Narat Charupat, Richard Deaves, and Brian D. Kluger, 2009, Probability judgment error and speculation in laboratory asset market bubbles, *The Journal of Financial and Quantitative Analysis* 44, 719-744.

Ackert, Lucy F., and Bryan K. Church, 2001, The effects of subject pool and design experience on rationality in experimental asset markets, *Journal of Psychology and Financial Markets* 2, 6 - 28.

Ahern, Kenneth R., Ran Duchin, and Tyler Shumway, 2014, Peer effects in risk aversion and trust, *Review of Financial Studies* 27, 3213-3240.

Allen, Franklin, and Douglas Gale, 2000, Bubbles and crises, *The Economic Journal* 110, 236-255.

Allen, Franklin, and Gary Gorton, 1993, Churning bubbles, *The Review of Economic Studies* 60, 813-836.

Andrade, Eduardo B, Terrance Odean, and Shengle Lin, 2013, Bubbling with excitement: An experiment.

Ang, James S, Dean Diavatopoulos, and Thomas V Schwarz, 2010, The creation and control of speculative bubbles in a laboratory setting, in Cheng-Few Lee, Alice C Lee, and John Lee, eds.: *Handbook of quantitative finance and risk management* (Springer US).

Arkes, Hal R., and Catherine Blumer, 1985, The psychology of sunk cost, *Organizational Behavior and Human Decision Processes* 35, 124-140.

Arkes, Hal R., Cynthia A. Joyner, Mark V. Pezzo, Jane Gradwohl Nash, Karen Siegel-Jacobs, and Eric Stone, 1994, The psychology of windfall gains, *Organizational Behavior and Human Decision Processes* 59, 331-347.

Azmat, Ghazala, and Nagore Iriberri, 2010, The importance of relative performance feedback information: Evidence from a natural experiment using high school students, *Journal of Public Economics* 94, 435-452.

Baghestanian, Sascha, Paul J Gortner, and Joel J Van der Weele, 2015, Peer effects and risk sharing in experimental asset markets.

Bagnoli, Mark, and Susan G. Watts, 2000, Chasing hot funds: The effects of relative performance on portfolio choice, *Financial Management* 29, 31-50.

Bakshi, Gurdip S., and Zhiwu Chen, 1996, The spirit of capitalism and stock-market prices, *The American Economic Review* 86, 133-157.

Bault, Nadège, Giorgio Coricelli, and Aldo Rustichini, 2008, Interdependent utilities: How social ranking affects choice behavior, *PLoS ONE* 3, e3477.

Bault, N., M. Joffily, A. Rustichini, and G. Coricelli, 2011, Medial prefrontal cortex and striatum mediate the influence of social comparison on the decision process, *Proc Natl Acad Sci U S A* 108, 16044-9.

Bebchuk, Lucian A, and Holger Spamann, 2009, Regulating bankers' pay, *geo. Lj* 98, 247.

Becker, Brian E., and Mark A. Huselid, 1992, The incentive effects of tournament compensation systems, *Administrative Science Quarterly* 37, 336-350.

Becker, Gary S., Kevin M. Murphy, and Iván Werning, 2005, The equilibrium distribution of income and the market for status, *Journal of Political Economy* 113, 282-310.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, 1992, A theory of fads, fashion, custom, and cultural change as informational cascades, *Journal of Political Economy* 100, 992-1026.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, 1998, Learning from the behavior of others: Conformity, fads, and informational cascades, *The Journal of Economic Perspectives* 12, 151-170.

Blais, Ann-Renée, and Elke U. Weber, 2006, A domain-specific risk-taking (dospert) scale for adult populations, *Judgment and Decision Making* 1, 33.

Blanes i Vidal, Jordi , and Mareike Nossol, 2011, Tournaments without prizes: Evidence from personnel records, *Management Science* 57, 1721-1736.

Blinder, A.S., 2009. 'Crazy Compensation and the Crisis', *The Wall Street Journal*, 28 May, accessed 16 January 2015, <http://www.wsj.com/articles/SB124346974150760597>

Bothner, Matthew S., Jeong-han Kang, and Toby E. Stuart, 2007, Competitive crowding and risk taking in a tournament: Evidence from nascar racing, *Administrative Science Quarterly* 52, 208-247.

Boyce, Christopher J., Gordon D.A. Brown, and Simon C. Moore, 2010, Money and happiness: Rank of income, not income, affects life satisfaction, *Psychological Science* 21, 471-475.

Brown, Jennifer, and Jin Li, 2010, Going for it: The adoption of risky strategies in tournaments.

Brown, Jeffrey R., Zoran Ivković, Paul A. Smith, and Scott Weisbenner, 2008, Neighbors matter: Causal community effects and stock market participation, *The Journal of Finance* 63, 1509-1531.

Brown, Keith C., W. V. Harlow, and Laura T. Starks, 1996, Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry, *The Journal of Finance* 51, 85-110.

Brunnermeier, Markus K, and Martin Oehmke, 2012, Bubbles, financial crises, and systemic risk, (National Bureau of Economic Research).

Bull, Clive, Andrew Schotter, and Keith Weigelt, 1987, Tournaments and piece rates: An experimental study, *Journal of Political Economy* 95, 1-33.

Bursztyn, Leonardo, Florian Ederer, Bruno Ferman, and Noam Yuchtman, 2014, Understanding mechanisms underlying peer effects: evidence from a field experiment on financial decisions, *Econometrica* 82, 1273-1301.

Busse, Jeffrey A., 2001, Another look at mutual fund tournaments, *The Journal of Financial and Quantitative Analysis* 36, 53-73.

Buunk, Abraham P., and Frederick X. Gibbons, 2007, Social comparison: The end of a theory and the emergence of a field, *Organizational Behavior and Human Decision Processes* 102, 3-21.

Cabral, Luís M. B., 2003, R&d competition when firms choose variance, *Journal of Economics & Management Strategy* 12, 139-150.

Caginalp, Gunduz, David Porter, and Vernon L. Smith, 1998, Initial cash/asset ratio and asset prices: An experimental study, *Proceedings of the National Academy of Sciences of the United States of America* 95, 756-761.

Caginalp, Gunduz, David Porter, and Vernon L. Smith, 2000, Overreactions, momentum, liquidity, and price bubbles in laboratory and field asset markets, *Journal of Psychology & Financial Markets* 1, 24-48.

Caginalp, Gunduz, David Porter, and Vernon L. Smith, 2001, Financial bubbles: Excess cash, momentum, and incomplete information, *The Journal of Psychology and Financial Markets* 2, 80-99.

Camerer, Colin, 1989, Bubbles and fads in asset prices, *Journal of Economic Surveys* 3, 3-41.

Chan, Kenneth S., Vivian Lei, and Filip Vesely, 2013, Differentiated assets: An experimental study on bubbles, *Economic Inquiry* 51, 1731-1749.

Charness, Gary, and Brit Grosskopf, 2001, Relative payoffs and happiness: An experimental study, *Journal of Economic Behavior & Organization* 45, 301-328.

Charness, Gary, and Matthew Rabin, 2002, Understanding social preferences with simple tests, *The Quarterly Journal of Economics* 117, 817-869.

Chen, Hsiu-lang, and George G. Pennacchi, 2009, Does prior performance affect a mutual fund's choice of risk? Theory and further empirical evidence, *The Journal of Financial and Quantitative Analysis* 44, 745-775.

Cherry, Todd L., Peter Frykblom, and Jason F. Shogren, 2002, Hardnose the dictator, *The American Economic Review* 92, 1218-1221.

Cherry, Todd L., Stephan Kroll, and Jason F. Shogren, 2005, The impact of endowment heterogeneity and origin on public good contributions: Evidence from the lab, *Journal of Economic Behavior & Organization* 57, 357-365.

Cheung, Stephen L., and Andrew Coleman, 2014, Relative performance incentives and price bubbles in experimental asset markets, *Southern Economic Journal* 81, 345-363.

Cheung, Stephen L., Morten Hedegaard, and Stefan Palan, 2014, To see is to believe: Common expectations in experimental asset markets, *European Economic Review* 66, 84-96.

Chevalier, Judith , and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105, 1167-1200.

Childs, Jason, and Stuart Mestelman, 2006, Rate-of-return parity in experimental asset markets*, *Review of International Economics* 14, 331-347.

Clark, Andrew E., Paul Frijters, and Michael A. Shields, 2008, Relative income, happiness, and utility: An explanation for the easterlin paradox and other puzzles, *Journal of Economic Literature* 46, 95-144.

Clark, Andrew E., and Andrew J. Oswald, 1996, Satisfaction and comparison income, *Journal of Public Economics* 61, 359-381.

Clark, Jeremy, 2002, House money effects in public good experiments, *Experimental Economics* 5, 223-231.

Corgnet, Brice, Roberto Hernán-González, Praveen Kujal, and David Porter, 2014, The effect of earned versus house money on price bubble formation in experimental asset markets, *Review of Finance*.

Dass, Nishant, Massimo Massa, and Rajdeep Patgiri, 2008, Mutual funds and bubbles: The surprising role of contractual incentives, *Review of Financial Studies* 21, 51-99.

DeMarzo, Peter M., Ron Kaniel, and Ilan Kremer, 2008, Relative wealth concerns and financial bubbles, *The Review of Financial Studies* 21, 19-50.

Dijk, Oege, Martin Holmen, and Michael Kirchler, 2014, Rank matters–the impact of social competition on portfolio choice, *European Economic Review* 66, 97-110.

Duesenberry, James Stemble, 1949. *Income, saving, and the theory of consumer behavior*.

Dufwenberg, Martin, Tobias Lindqvist, and Evan Moore, 2005, Bubbles and experience: An experiment, *The American Economic Review* 95, 1731-1737.

Dupor, Bill, and Wen-Fang Liu, 2003, Jealousy and equilibrium overconsumption, *The American Economic Review* 93, 423-428.

Ehrenberg, Ronald G., and Michael L. Bognanno, 1990, Do tournaments have incentive effects?, *Journal of Political Economy* 98, 1307-1324.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, 2003, Incentive fees and mutual funds, *The Journal of Finance* 58, 779-804.

Elton, Edwin J., Martin J. Gruber, Christopher R. Blake, Yoel Krasny, and Sadi O. Ozelge, 2010, The effect of holdings data frequency on conclusions about mutual fund behavior, *Journal of Banking & Finance* 34, 912-922.

Eriksson, Tor, Anders Poulsen, and Marie Claire Villeval, 2009, Feedback and incentives: Experimental evidence, *Labour Economics* 16, 679-688.

Falk, Armin , and Andrea Ichino, 2006, Clean evidence on peer effects, *Journal of Labor Economics* 24, 39-57.

Fehr, Ernst, and Klaus M. Schmidt, 1999, A theory of fairness, competition, and cooperation, *The Quarterly Journal of Economics* 114, 817-868.

Festinger, Leon, 1954, A theory of social comparison processes, *Human Relations* 7, 117-140.

Fischbacher, Urs, 2007, Z-tree: Zurich toolbox for ready-made economic experiments, *Experimental Economics* 10, 171-178.

Fisher, Eric O'N, and Frank S. Kelly, 2000, Experimental foreign exchange markets, *Pacific Economic Review* 5, 365-387.

Frederick, Shane, 2005, Cognitive reflection and decision making, *The Journal of Economic Perspectives* 19, 25-42.

Frey, Bruno S., and Alois Stutzer, 2002, What can economists learn from happiness research?, *Journal of Economic Literature* 40, 402-435.

Friedman, Milton, and L. J. Savage, 1948, The utility analysis of choices involving risk, *Journal of Political Economy* 56, 279-304.

Frydman, Cary, 2015, What drives peer effects in financial decision-making? Neural and behavioral evidence, (working paper).

Gaba, Anil, and Ajay Kalra, 1999, Risk behavior in response to quotas and contests, *Marketing Science* 18, 417-434.

Gaba, Anil, Ilia Tsetlin, and Robert L. Winkler, 2004, Modifying variability and correlations in winner-take-all contests, *Operations Research* 52, 384-395.

Galí, Jordi, 1994, Keeping up with the Joneses: Consumption externalities, portfolio choice, and asset prices, *Journal of Money, Credit and Banking* 26, 1-8.

Garland, Howard, 1990, Throwing good money after bad: The effect of sunk costs on the decision to escalate commitment to an ongoing project, *Journal of Applied Psychology* 75, 728-731.

Genakos, Christos, and Mario Pagliero, 2012, Interim rank, risk taking, and performance in dynamic tournaments, *Journal of Political Economy* 120, 782-813.

Gilpatric, Scott M., 2009, Risk taking in contests and the role of carrots and sticks, *Economic Inquiry* 47, 266-277.

Gómez, Juan-Pedro, Richard Priestley, and Fernando Zapatero, 2009, Implications of keeping-up-with-the-Joneses behavior for the equilibrium cross section of stock returns: International evidence, *The Journal of Finance* 64, 2703-2737.

Goriaev, Alexei, Theo E. Nijman, and Bas J. M. Werker, 2005, Yet another look at mutual fund tournaments, *Journal of Empirical Finance* 12, 127-137.

Green, Jerry R., and Nancy L. Stokey, 1983, A comparison of tournaments and contracts, *Journal of Political Economy* 91, 349-364.

Greiner, Ben, 2004, The online recruitment system orsee 2.0 - a guide for the organization of experiments in economics, (University of Cologne, Department of Economics).

Gromb, Denis, and Dimitri Vayanos, 2010, Limits of arbitrage, *Annual Review of Financial Economics* 2, 251-275.

Grund, Christian, J. A. N. Höcker, and Stefan Zimmermann, 2013, Incidence and consequences of risk-taking behavior in tournaments — evidence from the nba, *Economic Inquiry* 51, 1489-1501.

Hanke, Michael, Jürgen Huber, Michael Kirchler, and Matthias Sutter, 2010, The economic consequences of a tobin tax—an experimental analysis, *Journal of Economic Behavior & Organization* 74, 58-71.

Hannan, R. Lynn, Ranjani Krishnan, and Andrew H. Newman, 2008, The effects of disseminating relative performance feedback in tournament and individual performance compensation plans, *The Accounting Review* 83, 893-913.

Haruvy, Ernan, Yaron Lahav, and Charles N. Noussair, 2007, Traders' expectations in asset markets: Experimental evidence, *The American Economic Review* 97, 1901-1920.

Haruvy, Ernan, and Charles N. Noussair, 2006, The effect of short selling on bubbles and crashes in experimental spot asset markets, *The Journal of Finance* 61, 1119-1157.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2004, Social interaction and stock-market participation, *The Journal of Finance* 59, 137-163.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2005, Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers, *The Journal of Finance* 60, 2801-2824.

Hong, Harrison, José Scheinkman, and W. E. I. Xiong, 2006, Asset float and speculative bubbles, *The Journal of Finance* 61, 1073-1117.

Hu, Ping, Jayant R. Kale, Marco Pagani, and Ajay Subramanian, 2011, Fund flows, performance, managerial career concerns, and risk taking, *Management Science* 57, 628-646.

Huber, Jürgen, and Michael Kirchler, 2012, The impact of instructions and procedure on reducing confusion and bubbles in experimental asset markets, *Experimental Economics* 15, 89-105.

Hvide, Hans K., 2002, Tournament rewards and risk taking, *Journal of Labor Economics* 20, 877-898.

Isaac, R. Mark, and Duncan James, 2003, Boundaries of the tournament pricing effect in asset markets: Evidence from experimental markets, *Southern Economic Journal* 69, 936-951.

Ivković, Zoran, and Scott Weisbenner, 2007, Information diffusion effects in individual investors' common stock purchases: Covet thy neighbors' investment choices, *Review of Financial Studies* 20, 1327-1357.

James, Duncan, and R. Mark Isaac, 2000, Asset markets: How they are affected by tournament incentives for individuals, *The American Economic Review* 90, 995-1004.

José A. Scheinkman, and Wei Xiong, 2003, Overconfidence and speculative bubbles, *Journal of Political Economy* 111, 1183-1220.

Kahneman, Daniel, and Amos Tversky, 1979, Prospect theory: An analysis of decision under risk, *Econometrica* 47, 263-291.

Kaustia, Markku, and Samuli Knüpfer, 2012, Peer performance and stock market entry, *Journal of Financial Economics* 104, 321-338.

Kempf, Alexander, and Stefan Ruenzi, 2008, Tournaments in mutual-fund families, *Review of Financial Studies* 21, 1013-1036.

Kempf, Alexander, Stefan Ruenzi, and Tanja Thiele, 2009, Employment risk, compensation incentives, and managerial risk taking: Evidence from the mutual fund industry, *Journal of Financial Economics* 92, 92-108.

Kindleberger, Charles P, and Robert Z Aliber, 2011. *Manias, panics and crashes: A history of financial crises* (Palgrave Macmillan).

King, RR, Vernon L Smith, Arlington W Williams, and M Van Boening, 1993, The robustness of bubbles and crashes in experimental stock markets, in R. Day, and P. Chen, eds.: *Non linear dynamics and evolutionary economics* (Oxford Press, Oxford).

Kirchler, Michael, Jürgen Huber, and Thomas Stöckl, 2012, Thar she bursts: Reducing confusion reduces bubbles, *American Economic Review* 102, 865-83.

Kleinlercher, Daniel, Jürgen Huber, and Michael Kirchler, 2014, The impact of different incentive schemes on asset prices, *European Economic Review* 68, 137-150.

Koski, Jennifer Lynch, and Jeffrey Pontiff, 1999, How are derivatives used? Evidence from the mutual fund industry, *The Journal of Finance* 54, 791-816.

Kräkel, Matthias, and Dirk Sliwka, 2004, Risk taking in asymmetric tournaments, *German Economic Review* 5, 103-116.

Kuhnen, Camelia M., and Agnieszka Tymula, 2012, Feedback, self-esteem, and performance in organizations, *Management Science* 58, 94-113.

Lahno, Amreim, and Marta Serra-Garcia, 2015, Peer effects in risk taking: Envy or conformity?, *Journal of Risk and Uncertainty* 50, 73-95.

Lazear, Edward P., and Sherwin Rosen, 1981, Rank-order tournaments as optimum labor contracts, *Journal of Political Economy* 89, 841-864.

Lee, Jungmin, 2004, Prize and risk-taking strategy in tournaments: Evidence from professional poker players.

Lei, Vivian, Charles N. Noussair, and Charles R. Plott, 2001, Nonspeculative bubbles in experimental asset markets: Lack of common knowledge of rationality vs. Actual irrationality, *Econometrica* 69, 831-859.

Levitt, Steven D., and John A. List, 2007, What do laboratory experiments measuring social preferences reveal about the real world?, *Journal of Economic Perspectives* 21, 153-174.

Linde, Jona, and Joep Sonnemans, 2012, Social comparison and risky choices, *Journal of Risk and Uncertainty* 44, 45-72.

Long, J. Bradford de, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Positive feedback investment strategies and destabilizing rational speculation, *The Journal of Finance* 45, 379-395.

Lugovskyy, Volodymyr, Daniela Puzzello, and Steven Tucker, 2009, An experimental study of bubble formation in asset markets using the tâtonnement pricing mechanism, (University of Canterbury, Department of Economics and Finance).

Luttmer, Erzo F. P., 2005, Neighbors as negatives: Relative earnings and well-being, *The Quarterly Journal of Economics* 120, 963-1002.

Malkiel, Burton G., 2003, The efficient market hypothesis and its critics, *Journal of Economic Perspectives* 17, 59-82.

Mas, Alexandre, and Enrico Moretti, 2009, Peers at work, *American Economic Review* 99, 112-45.

Nalebuff, Barry J., and Joseph E. Stiglitz, 1983, Prizes and incentives: Towards a general theory of compensation and competition, *The Bell Journal of Economics* 14, 21-43.

Nieken, Petra, and Dirk Sliwka, 2010, Risk-taking tournaments – theory and experimental evidence, *Journal of Economic Psychology* 31, 254-268.

Noussair, Charles, Stephane Robin, and Bernard Ruffieux, 2001, Price bubbles in laboratory asset markets with constant fundamental values, *Experimental Economics* 4, 87-105.

Noussair, Charles N., and Owen Powell, 2010, Peaks and valleys: Price discovery in experimental asset markets with non-monotonic fundamentals, *Journal of Economic Studies* 37, 152-180.

Oechssler, Jörg, Andreas Roider, and Patrick W. Schmitz, 2009, Cognitive abilities and behavioral biases, *Journal of Economic Behavior & Organization* 72, 147-152.

O'Keeffe, Mary, W. Kip Viscusi, and Richard J. Zeckhauser, 1984, Economic contests: Comparative reward schemes, *Journal of Labor Economics* 2, 27-56.

Oxoby, Robert J., and John Spraggon, 2008, Mine and yours: Property rights in dictator games, *Journal of Economic Behavior & Organization* 65, 703-713.

Palan, Stefan, 2013, A review of bubbles and crashes in experimental asset markets, *Journal of Economic Surveys* 27, 570-588.

Porter, David P., and Vernon L. Smith, 1994, Stock market bubbles in the laboratory, *Applied Mathematical Finance* 1, 111-128.

Porter, David P., and Vernon L. Smith, 1995, Futures contracting and dividend uncertainty in experimental asset markets, *Journal of Business* 68, 509.

Qiu, Jiaping, 2003, Termination risk, multiple managers and mutual fund tournaments, *European Finance Review* 7, 161-190.

Rajan, Raghuram G., 2006, Has finance made the world riskier?, *European Financial Management* 12, 499-533.

Rajan, R.G., 2008. 'Bankers' Pay is Deeply Flawed', *Financial Times*, 9 January, accessed 16 January 2015, < http://www.ft.com/intl/cms/s/0/18895dea-be06-11dc-8bc9-0000779fd2ac.html#axzz3T9uTJuUR>

Robin, Stéphane, Katerina Straznicka, and Marie-Claire Villeval, 2012, Bubbles and incentives: An experiment on asset markets, *GATE-Groupe d'Analyse et de Théorie Économique Lyon-St Étienne Working Paper*.

Robson, Arthur J., 1992, Status, the distribution of wealth, private and social attitudes to risk, *Econometrica* 60, 837-857.

Rosen, Sherwin, 1986, Prizes and incentives in elimination tournaments, *The American Economic Review* 76, 701-715.

Roussanov, Nikolai, 2010, Diversification and its discontents: Idiosyncratic and entrepreneurial risk in the quest for social status, *The Journal of Finance* 65, 1755-1788.

Scherbina, Anna, and Bernd Schlusche, 2014, Asset price bubbles: A survey, *Quantitative Finance* 14, 589-604.

Schoenberg, Eric J., and Ernan Haruvy, 2012, Relative performance information in asset markets: An experimental approach, *Journal of Economic Psychology* 33, 1143-1155.

Schwerter, Frederik, 2013, Social reference points and risk taking, (Bonn Econ Discussion Papers).

Shue, Kelly, 2013, Executive networks and firm policies: Evidence from the random assignment of mba peers, *Review of Financial Studies* 26, 1401-1442.

Sirri, Erik R., and Peter Tufano, 1998, Costly search and mutual fund flows, *The Journal of Finance* 53, 1589-1622.

Smith, Vernon L., 1962, An experimental study of competitive market behavior, *Journal of Political Economy* 70, 111-137.

Smith, Vernon L., Gerry L. Suchanek, and Arlington W. Williams, 1988, Bubbles, crashes, and endogenous expectations in experimental spot asset markets, *Econometrica* 56, 1119-1151.

Smith, Vernon L., Mark van Boening, and Charissa P. Wellford, 2000, Dividend timing and behavior in laboratory asset markets, *Economic Theory* 16, 567-583.

Stöckl, Thomas, Jürgen Huber, and Michael Kirchler, 2010, Bubble measures in experimental asset markets, *Experimental Economics* 13, 284-298.

Sutter, Matthias, Jürgen Huber, and Michael Kirchler, 2012, Bubbles and information: An experiment, *MANAGEMENT SCIENCE* 58, 384-393.

Taylor, Jonathan, 2003, Risk-taking behavior in mutual fund tournaments, *Journal of Economic Behavior & Organization* 50, 373-383.

Thaler, Richard H., and Eric J. Johnson, 1990, Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice, *Management Science* 36, 643-660.

Tomlin, Damon, Andrea Nedic, Deborah A. Prentice, Philip Holmes, and Jonathan D. Cohen, 2013, The neural substrates of social influence on decision making, *PLoS ONE* 8, e52630.

Tran, Anh, and Richard Zeckhauser, 2012, Rank as an inherent incentive: Evidence from a field experiment, *Journal of Public Economics* 96, 645-650.

Tsetlin, Ilia, Anil Gaba, and Robert L Winkler, 2004, Strategic choice of variability in multiround contests and contests with handicaps, *Journal of Risk and Uncertainty* 29, 143-158.

Veblen, Thorstein, 1899. *The theory of the leisure class* (Project Gutenberg, Champaign, Ill).

Vendrik, Maarten C. M., and Geert B. Woltjer, 2007, Happiness and loss aversion: Is utility concave or convex in relative income?, *Journal of Public Economics* 91, 1423-1448.

Wagner, Wolf, 2013, Performance evaluation and financial market runs, *Review of Finance* 17, 597-624.

Weber, Elke U., Ann-Renée Blais, and Nancy E. Betz, 2002, A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors, *Journal of Behavioral Decision Making* 15, 263-290.