THE UNIVERSITY OF MELBOURNE

Department of Finance

FNCE40006

Finance Research Essay

Can Incomplete Markets be Informationally Efficient? An Experimental Study

Anirudh Suthakar

November 2018

Declaration

This essay is the sole work of the author whose name appears on the title page. It contains no material which the author has previously submitted for assessment at the University of Melbourne or elsewhere. To the best of the author's knowledge, the essay contains no material previously written or published by another person except where reference is made in the text of the essay.

…………………………………………….

Signature of Student

Acknowledgements

Firstly, my deepest thanks to Dr. Felix Fattinger for his patient and dedicated help throughout the semester, even when abroad. I benefited greatly from his supervision and instruction.

Secondly, all members of the Brain, Mind and Markets Laboratory – Prof. Peter Bossaerts, Dr. Carsten Murawski, Dr. Nitin Yadav, Dr. Elizabeth Bowman, Harvey Huang, Xiaping Lu, Charlie Wang, Pablo Franco and Jeremy Metha – deserve special acknowledgement for their assistance and friendliness throughout the year. I certainly felt I was part of the team and am happy I chose to work with 'the Lab' on my essay.

Thirdly, a sincere thanks to my close friends Michelle Lee and Katherine Xu, who have contributed the small moments of pleasure which made this year bearable.

Finally, I would like to express my gratitude to my family without whom this year would have been incalculably harder. Thank you for tolerating my moody periods and understanding my stress.

Abstract

We study information aggregation in markets where the correct valuation of securities requires market participants to solve instances of the knapsack problem, a generic, computationally complex problem. In particular, we consider the impact of an incomplete market structure, with only one traded security, on information aggregation and participant performance. We demonstrate experimentally that: (i) Incomplete market prices outperform complete market prices; (ii) Participants better solve the knapsack problem when there is only one traded security, compared to a complete set of Arrow-Debreu securities; (iii) Participants learn from market prices and use this information to improve their knapsack solution; (iv) Contrary to the literature, overconfidence has no impact on participants' trading behaviour.

Table of Contents

Table of Tables

Table of Figures

"The most significant fact about [the price] system is the economy of knowledge with which it operates, or how little individual participants need to know in order to be able to take the right action."

F.A. Hayek, The Use of Knowledge in Society, 1945.

I Introduction

It is generally understood that market prices must fully reveal all information such that market participants are equipped to make correct decisions. Therefore, only complete markets can optimally support information aggregation and revelation in markets (Grossman, 1981). This fundamental principle has underpinned the study of market information aggregation in finance, forming the basis for the Efficient Markets Hypothesis ("EMH") (Malkiel and Fama, 1970). However, Hayek (1945) insists that market prices need reveal only the most crucial information for participants to able to take the right action. In this paper, we will show that incomplete markets can outperform complete markets, a result that confounds the long-held consensus in finance.

The basis of our study is the Knapsack Problem ("KP"), a generic problem from the computer science literature (Kellerer, Pferschy and Pisinger, 2004). The KP is a model of *the* optimization problem posited by Hayek (1945) – the problem that all economic agents in an economy seek to solve. We perform an experimental study which not only overcomes a scarcity of relevant data but also allows for a more rigorous evaluation of the topic through careful design and the implementation of controls.

Our results demonstrate that incomplete markets can aggregate information more effectively than complete markets because they more parsimoniously represent the information that market participants actually need to solve the problem. This result corresponds to predictions from the theory of computation – to solve the KP, algorithms need *just* the most important information, rather than everything (Horowitz and Sahni, 1974).

This work provides a starting point for more detailed introspection within finance about whether complete markets are as necessary for the effective aggregation of information as previously believed. Our results also have real-world relevance – they provide a new perspective about why financial markets around the world (for instance, stock exchanges) gravitate towards structuring themselves as incomplete markets.

This paper will be organized as follows. Section II will review the current state of the literature on market information aggregation and motivate the research questions underlying this paper. Section III will describe the hypotheses that this paper will investigate in detail. Section IV will detail the design of the experiments. Section V will address the methodology of the analysis, while Sections VI and VII will describe and discuss the results. Finally, Section VIII will consider the implications of our results and conclude the paper.

II Literature Review and Motivation

Experimental Markets

There is a rich exploration of markets' information aggregation characteristics in the experimental literature. Foundational work by Plott and Sunder (1982) tests the predictions of Rational Expectations Equilibrium ("REE") models when some market participants are ex-ante fully informed. In such an equilibrium, market prices are fully revealing, and all participants are thus fully informed. They find that prices rapidly converge to their REE levels. Further work by the same authors (Plott and Sunder, 1988) began to shed light on the impact that security design can have on the information quality of market prices. By controlling for market variables other than the type of traded security, they show that previously inefficient markets can become efficient when the 'right' securities are used. Numerous other researchers have since contributed to the field including Forsythe and Lundholm (1990), who study conditions under which markets can reach an REE. The body of work making up the field continues to grow, with important recent research conducted by Page and Siemroth (2018). These authors attempt to quantify the actual amount of public and private information reflected in market prices. Their findings support the idea that markets reflect most public and some private information. However, they find evidence against the concept of 'strong-form' informational efficiency as prices are still far away from reflecting all private information. It is clear that the literature has made a significant contribution towards better understanding the information aggregation capabilities of markets. However, there are questions that currently extant research has yet to fully confront. This paper will contribute to the field by beginning to answer these questions.

The experimental literature generally focuses on investigating how well markets solve a broadly similar valuation 'problem'[1]. The fundamental structure of the problem is such that assuming a prior and randomly sampling the information held by participants[2] would generally reveal the solution to the problem. However, there exist classes of problems where such random sampling is inefficient and ineffective (Bossaerts, Bowman, Huang, Murawski, Tang and Yadav, 2018). The KP is one such problem. It is a constrained combinatorial optimisation problem (Kellerer et al., 2004) that involves a set of items, each with a value and a weight. To solve the problem, one must find the combination of items that has the maximum value whilst meeting an overall weight constraint. Importantly, the KP is noted by Bartholdi (2008) to be a generic problem with applications in fields from economics to engineering. It is similar in form to problems that real-world financiers attempt to solve (Kellerer et al., 2004) – for instance, the selection of an optimal portfolio of assets[3]. There is no reason to expect, ex-ante, that markets respond to such problems in the same way they do to problems in which random sampling is effective in finding the solution. The literature, in general, fails to address what happens to markets' information aggregation capabilities when faced with this type of problem. Therefore, this paper is positioned to make a novel contribution to the experimental investigation of markets' information aggregation characteristics by addressing this gap.

The complexity of the KP is an important feature of the problem. It is described in the computer science literature as a computationally difficult problem (Kellerer et al., 2004). A property of the KP that contributes to this difficulty is that one cannot verify that any solution is optimal without first checking the entire set of possible item combinations

[1] The solution to such a problem is effectively the correct pricing of a security/securities.
[2] This follows the approach of Savage (1972), which has become a fundamental concept in traditional theories of decision-making.
[3] The analogy would be as follows: A manager faces a total funding and/or risk constraint and must optimally select different assets with a return (value) and cost/risk (weight).

(Meloso, Copic and Bossaerts, 2009). This paper will follow Bossaerts et al. (2018) in using the Sahni-*k* metric to describe the difficulty of a given KP instance. This measure is based on the number of steps a particular computer algorithm would have to take to solve a given KP instance. This algorithm begins by selecting a subset of *k* items, where *k* is an integer (positive or zero). It then considers all remaining items, and orders them based on their value-to-weight ratio. Applying the tertiary 'greedy' algorithm, it begins filling the knapsack with the remaining items based on their value-to-weight ratio. The optimal solution to the KP instance is determined through this process as the knapsack with the highest value that does not exceed the weight constraint. The *k* value of an instance can be understood as meaning that the instance can be solved by a Sahni-*k* algorithm with the same *k* value. A higher *k* value indicates a more complex instance. Sahni-*k* is relevant because it explains not only how well computers can solve the KP, but individuals too (Bossaerts and Murawski, 2016). Thus, Sahni-*k* allows us to precisely select instances that vary in difficulty and examine how this difficulty impacts how markets and individuals behave.

The KP has been used as the basis for prior research on the information aggregation characteristics of experimental markets. Bossaerts et al. (2018) utilize a market based on a complete set[4] of Arrow-Debreu securities (Arrow and Debreu, 1954). Participants were asked to solve five separate instances of the KP, each with a different difficulty as measured by Sahni-*k*. In addition, they were asked to simultaneously trade securities in an online marketplace based on their belief of the solution to the KP. Each security corresponded to a single item in the whole set of items for that KP instance. If an item was in the optimal knapsack, the corresponding security would pay a one dollar liquidating dividend at the end of the round; if not, the security would be worthless. Importantly, the price of these securities,

[4] Each security in the market had a unique correspondence with one item in the relevant KP instance.

bounded between zero and one dollar, could be interpreted by participants as a probability – the probability, according to the market, that the item corresponding to that security was in the optimal knapsack. The authors conclude from this experiment that the average participant reached a knapsack solution that was closer to the optimal solution than markets did[5]. They also find that some participants used information from market prices to improve their attempted solutions of the KP. Specifically, less informed participants were able to improve the composition of their knapsack through information revealed by market prices.

The conclusions from Bossaerts et al. (2018) provide a strong basis for further work. They confirm that markets appear to behave according to the Noisy Rational Expectations Equilibrium ("NREE")[6] theory, when noise is defined by Computational Complexity Theory ("CCT")[7]. The authors also show that market prices assist participants, to an extent, in improving their solution to the problem. However, their poor performance relative to the average participant lies in stark contrast to accepted fundaments of economic and financial theory. Proposals to utilize market prices in determining policy, especially with respect to bank regulation, are rife in the literature (Flannery, 1998; Flannery, 2016). These proposals would not be sensible if market prices were not relatively informative. For instance, the literature on contingent capital often applies conversion trigger points based in some way on a bank's stock price (McDonald, 2013). When such a trigger point is reached, the security is

[5] The market 'solution' was derived by using the probability interpretation of market prices to simulate a suggested market knapsack.

[6] A theory of information aggregation that posits some 'noise' exists in a market that prevents prices from being fully revealing (Diamond and Verrecchia, 1981). The nature of this 'noise' is not well-defined (Biais, Bossaerts and Spatt, 2004; De Long, Shleifer, Summers and Waldmann, 1990)

[7] This is motivated by the finding that humans struggle with the same problems that computers find difficult, such that noise in market prices may arise from the computational complexity of the valuation problem underlying the market (Bossaerts and Murawski, 2017; Bossaerts and Murawski, 2016). A growing body of work focuses on connecting markets and computational complexity. Maymin (2011), for instance, posits that if certain computational problems are intractable, markets cannot be even weakly efficient.

converted from debt to bank equity. If prices in the stock market are relatively uninformative, their value in guiding the appropriate conversion of these securities is called into question.

This idea is key to the motivation for this paper. We will seek to improve upon the information quality of market prices when markets are asked to solve a computationally complex problem. We will do this by changing the type of securities traded on the market.

Security Design

The literature documents a vast array of factors that can impact the information aggregation characteristics of markets. These include factors related to market participants – how many there are and how information is distributed amongst them – and factors related to market design – the nature of the traded securities and how participants trade[8]. Of these factors, security design offers a particularly fruitful avenue through which market outcomes could be improved. The experimental literature in the area is sparse but promising. Plott and Sunder (1988) reveal the importance of security design for information aggregation purposes. The lessons from their work are reinforced and expanded upon by Plott and Chen (2002), who use past experimental results to guide the creation of a successful real-world Information Aggregation Mechanism ("IAM"). In collaboration with Hewlett-Packard ("HP"), the authors use markets with Arrow-Debreu securities to forecast future outcomes. They find that predictions from these markets consistently out-performed official HP forecasts, despite low employee participation in the markets.

A common result in this branch of research is that for the purpose of market information aggregation, a set of securities that operate in the 'item' space can be more effective than a single security that operates in the 'value' space (Plott and Sunder, 1988; Plott, 2000; Forsythe

[8] For instance, through an open double-auction process or 'over-the-counter'.

and Lundholm, 1990). In the context of the KP, the security framework from Bossaerts et al. (2018) operates in the 'item' space. It completely spans all possible items in each KP and offers information on what the market 'thinks' about the structure of the optimal knapsack. This type of market is commonly referred to in the literature as 'complete' (Flood, 1991). A single compound security[9] instead operates in the 'value' space – its market price offers information on what the market 'thinks' is the total value of the optimal knapsack. However, the inability of complete markets to accurately reflect information in the work by Bossaerts et al. (2018) casts doubt on whether the 'accepted wisdom' in the literature is meaningful when markets and market participants must solve a computationally complex problem. For instance, the processing of information from complete markets may be difficult for participants, as they are required to consider prices in numerous markets at once. A single compound security does not suffer from this issue. This paper will thus consider whether a security design that involves compound securities that operate in the 'value' space (e.g. those employed in Plott and Sunder, 1988) can achieve better results than those employed in Bossaerts et al. (2018).

Overconfidence

Adopting this type of security design may also have impacts on participants' trading behaviour. In contrast to Arrow-Debreu securities, whose price is rigidly bounded, compound securities can be traded at a wider range of prices[10]. This change could conceivably affect how participants forecast the price of the security and bring into play idiosyncratic personality factors like overconfidence. For instance, there is greater potential for speculation and price bubbles when compound securities are used (Plott and Chen, 2002).

[9] 'Compound' in this case can be considered as meaning that there is no accounting for different items in the construction of the security – its price should reflect information about the whole knapsack.
[10] In the experiment, these prices ranged from a minimum of zero dollars to a maximum of thirty dollars.

The importance of these factors has been thoroughly investigated in the experimental literature. Biais, Hilton, Mazurier and Pouget (2005) show that calibration-based overconfidence ("CBO")[11] has a negative impact on participants' trading performance when there is asymmetric information in the market. The relevance of CBO to participants' trading behaviour is reinforced by Deaves, Luders and Luo (2009) who find that CBO and the better-than-average ("BTA")[12] effect are linked to more market activity. Although these authors note that disentangling BTA and CBO can be a fraught task (Deaves et al., 2009), Regner, Hilton, Cabantous and Vautier (2003) report that these types of overconfidence are uncorrelated with each other. Whilst the nature of this experiment is significantly different to prior work, there is reason to believe that overconfidence effects may be relevant. Speirs-Bridge, Fidler, McBride, Flander, Cumming and Brugman (2010) report that, in general, overconfidence increases with the difficulty of the problem, as well as when performance feedback is limited. The KP is known to be a computationally hard problem (Bossaerts and Murawski, 2016) and performance feedback is minimal, as it is difficult to verify if a given solution is correct. Therefore, there is a strong rationale supporting the examination of these effects with respect to their potential impact on participants' trading behaviour.

The measurement of these two types of overconfidence is a topic of some contention in the literature. Deaves et al. (2009) and Biais et al. (2005) use a measure known as the confidence-interval ("CI") test to measure CBO. This test asks participants to respond to a series of general-knowledge questions with numeric intervals such that they have a specified level of confidence that these intervals contain the true answer. Work by Soll and Klayman (2004) and Speirs-Bridge et al. (2010) disputes the efficacy of this test and claim it does not accurately measure CBO. This paper will thus follow a modified approach developed by

[11] CBO is defined by Odean (1998) as the tendency to overestimate the accuracy of your information.

[12] BTA is defined by Svenson (1980) as overconfidence relative to peers.

Langnickel and Zeisberger (2016), who implement a two-stage technique which improves on the overestimation problems[13] of the CI test. To measure BTA, this paper will follow the approach suggested by Deaves et al. (2009), which is advantageous for its ease-of-use.

[13] The authors showed that participants tended to respond in the same way to CI tests with different levels of confidence. This leads to exaggerated and incoherent measures of overconfidence.

III Hypotheses

It is straightforward to understand that there will likely be differences in how participants derive and use price information to solve the problem in our setting. A single compound security gives market participants *direct* information on the overall value of the KP's optimal solution. This is the solution to the problem that all traders, and hence, the market, are trying to solve. Arrow-Debreu securities of the type used in Bossaerts et al. (2018) offer only *indirect* information about this value, by providing information about the composition of items in the optimal KP solution. On this basis, a single compound security would seem to be advantageous in terms of the simplicity of the information its price should reflect.

However, other converse effects also seem possible. The price of a single security can no longer be easily interpreted as a probability, as they could when markets were complete. Arrow-Debreu securities may thus better facilitate participants' use of price information to modify their KP solution, as the information presented in prices can be directly linked to a particular item in that KP instance. This paper will seek to understand what changes – in terms of how well markets and market agents solve a generic (Bartholdi, 2008) computationally complex problem – when securities are designed such that markets are no longer complete. There are three key questions that this paper will seek to answer, which lead directly to the hypotheses that will be investigated.

The first question considers whether markets and participants in this experiment do better than those in Bossaerts et al. (2018). This is an integral part of understanding whether markets with securities that operate in the value space convey more information through prices than markets with securities that operate in the item space. This question motivates the following pair of 'relative performance' hypotheses.

The first among these two hypotheses is concerned purely with the quality of market prices, and what they tell us about the solution to the underlying KP. It can be understood as seeking to comprehend what a completely uninformed agent would think is the solution to the KP, given that they are only able to observe market prices.

Hypothesis 1

Relative Price Performance – When the securities used in the market operate in the value space, rather than the item space, the information content of market prices will be higher (regardless of the difficulty of the underlying KP[14]).

The second among these two hypotheses considers how participants perform in solving the KP. We seek to understand whether changing the type of security used in the market can assist participants in solving a given KP.

Hypothesis 2

Relative KP Performance – When the securities used in the market operate in the value space, rather than the item space, a higher proportion of participants will be able to solve the underlying KP.

The second question considers the nature of information 'flows' between markets and participants. This question motivates the 'information transfer' hypothesis; which reflects two ideas that both examine how, if at all, participants use information embedded in market prices. The first idea is whether participants can learn from information contained in prices to improve their attempted solution of the KP, as shown in Bossaerts et al. (2018). The

[14] As measured by Sahni-*k*.

second idea considers whether information held by participants is transmitted back to the market via participants' trading behaviour.

Hypothesis 3

Information Transfer – Market agents will be able to learn from market prices when securities in the market operate in the value space and will transmit information back into the market through submitting orders.

The third question considers whether this security design can make participants' overconfidence a factor in how they behave. It motivates the 'overconfidence' hypothesis, which will examine whether relative or individual overconfidence influences how participants trade and attempt to solve the KP.

Hypothesis 4

Overconfidence - More overconfident market participants will expend less effort on solving the KP and be more active in the marketplace.

IV Experimental Design

a. Participants

Participants for the experiments were drawn primarily from the University of Melbourne, with a minority of individuals drawn from The Royal Melbourne Institute of Technology[15]. Five experimental sessions were held in total, with an average of 19 participants each. Each participant had to be between 18 and 35 years of age and have normal or corrected-to-normal vision.

The final sample included a total of 77[16] participants (age range: 18 to 34 years, mean age: 23 years, standard deviation: 3.5 years, gender breakdown: 31 Male, 46 Female).

b. Experimental Task

Participants attempted five instances of the KP[17] while simultaneously trading in an online marketplace[18]. They were also required to complete two questionnaires – one in between trading rounds, and the other after the completion of all trading rounds. The instructions distributed to participants are reported in Appendix IV-a.

Knapsack Problem Instances

For each instance, participants were asked to select a subset from a set of items with given values and weights, to maximise the total value of selected items subject to a total weight constraint. The instances chosen matched exactly[19] the instances studied in Bossaerts et al.

[15] This research was approved by The University of Melbourne Human Research Ethics Committee (Ethics ID: 1852127.2), in accordance with the World Medical Association Declaration of Helsinki. All participants provided written and informed consent before taking part in the research.
[16] Only four out of five sessions made up the final sample, due to a change in the payoff structure.
[17] Accessible at http://bmmlab.org/games.
[18] Accessible at https://adhocmarkets.com/.
[19] One instance was modified to ensure only one combination of items was able to reach the optimal value. This change is highlighted in the instance specifications (See Appendix IV-b).

(2018). This decision was made to ensure that a direct comparison was possible between the information contained in market prices in this experiment and the experiment in Bossaerts et al. (2018). Choosing different instances would unnecessarily complicate such a direct comparison.

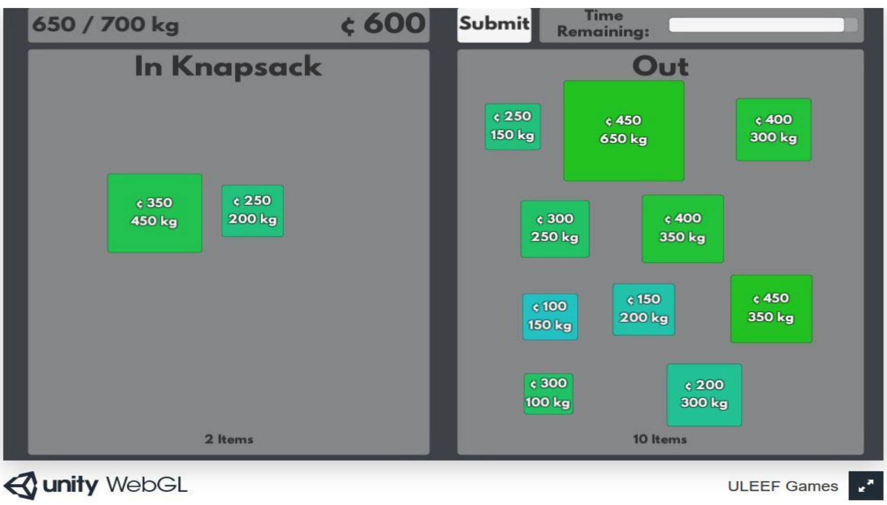
Formally, participants attempted to solve the maximisation problem defined by Equation (1) in every instance, where *'i', 'w', 'v'* and *'C'* denote the item, item weight, item value and total capacity of the knapsack, respectively:

$$\max \sum v_i * w_i \text{ subject to } \sum w_i * x_i \; \leq C \text{ and } x_i \; \epsilon \; \{0,1\} \qquad (1)$$

The user interface of the knapsack solver that participants used is shown in Figure 1. For each participant in every instance, data on each item moved in or out of the knapsack, as well as the time each move occurred was recorded. The solutions submitted by each participant were also recorded.

Figure 1 – Knapsack Problem Interface

Above: The knapsack solver. Participants were shown the instance and could attempt solutions by moving items from the OUT panel (right) to the IN KNAPSACK panel (left). Capacity, capacity used, and knapsack value were displayed on top (left).



Trading in the Online Marketplace

Each round of trading corresponded to one separate instance of the KP. Participants were allowed approximately ten minutes to simultaneously trade and attempt to solve the given instance of the KP. The market was organized as a continuous double-sided open book auction. This structure was chosen to replicate the organization of prominent electronic global stock markets including the pan-European EuroNext exchange and the New York Stock Exchange. Each participant also had access to an online Microsoft Excel[20] spreadsheet. This spreadsheet allowed participants to convert knapsack values into corresponding security payoffs[21]. The user interfaces of the two systems, online marketplace and Excel spreadsheet, are shown in Figure 2.

For each participant in every round, timestamped information on every placed sell and buy order (this includes traded orders) was collected.

Questionnaires

Two questionnaires were distributed to participants during the experiment. The first asked participants, after each round of trading, to record how confidently they could predict their trading performance in the round (on a scale of one to ten). This questionnaire is reported in Appendix IV-c. The second questionnaire contained a series of interval creation tasks that replicated those in Biais et al. (2005). Following this, a further question elicited how well participants felt they had performed in trading, relative to all others who have completed or will complete the same experiment. This questionnaire is reported in Appendix IV-d.

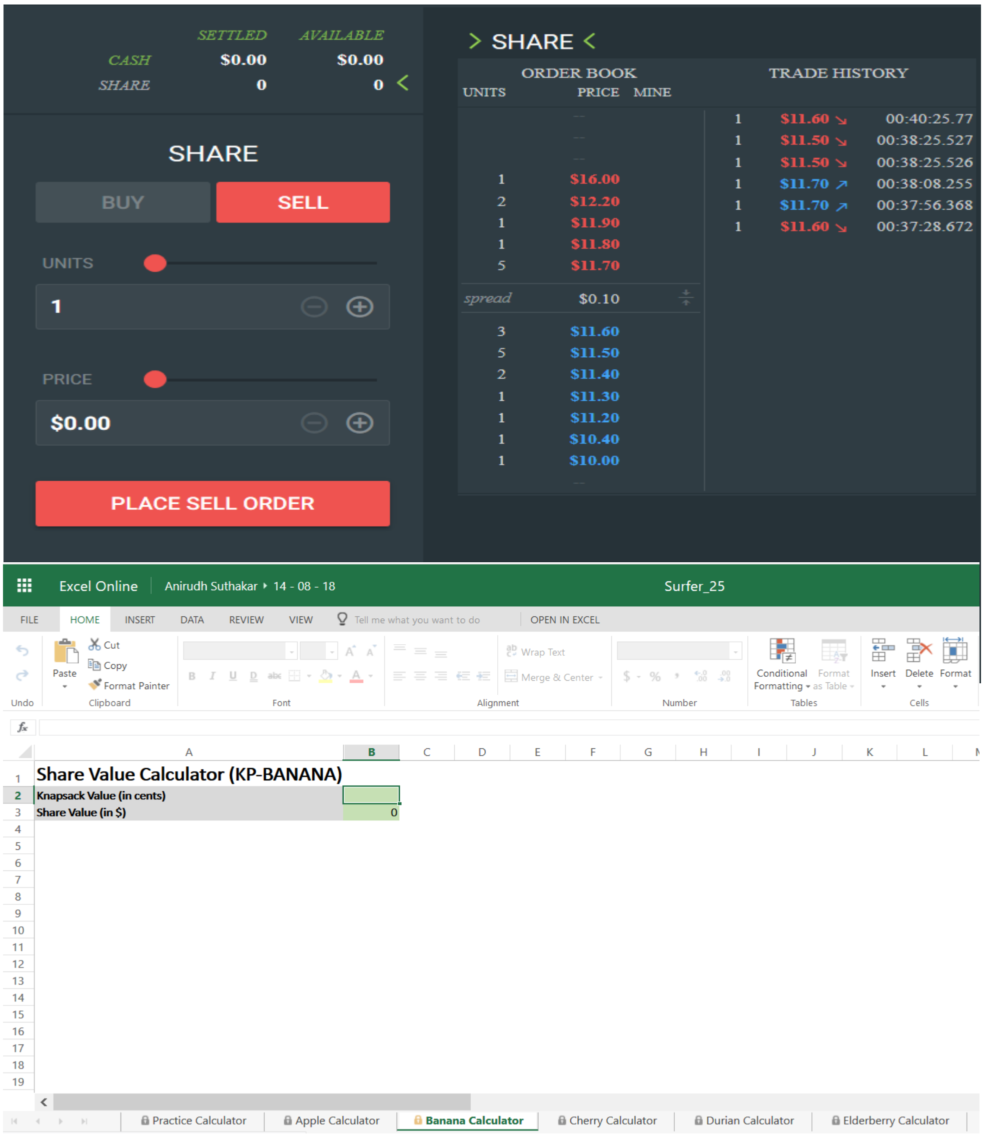
[20] The spreadsheets were hosted online on Microsoft OneDrive.

[21] The payoffs were in terms of experimental dollars, which were converted to Australian dollars ("AUD") at a rate of 2.5:1.

Figure 2 – Online Trading Platform and Online Excel Spreadsheet Interface

Top: Online marketplace. Share denotes the name of the traded security. The panel on the left is the order form, where participants could submit buy and sell limit orders, as well as view their settled and available holdings. The book of limit orders is listed in the 'ORDER BOOK' column, divided by side (Blue: bids, Red: asks) and listed by price level. The rightmost column, labelled 'TRADE HISTORY' contains a listing of past trades, ordered chronologically.

Bottom: Excel spreadsheet. Participants could enter a knapsack value from the knapsack solver (see Fig. 1) into the highlighted cell. The cell below then provides what would be the true payoff of the security if the inputted knapsack value was the optimal value.



Participant Allocations

Participants' initial allocations of shares and experimental cash were designed with two motives in mind. The first was that they should be 'fair'. This meant that, across all five trading rounds, all participants would receive the same ex-ante payout, based on their initial allocation. The second motive was for participants' allocations to encourage trading and allow participants to trade based on personal beliefs. Endowments were split into two 'classes', with each class being approximately evenly[22] represented in each experimental session. One class was bestowed with a larger amount of cash and fewer securities, whereas the other class received more securities and less cash. The specific endowments for each class can be found in Appendix IV-e. Each participant received the same initial allocation for all trading rounds.

Mapping Knapsack Values to Payoffs

The chosen instances of the KP varied in computational complexity, as measured by the Sahni-*k* value of each instance. One way in which this difficulty presented itself was in the number of solutions relatively close to the unique, optimal solution. An easier instance allows for fewer combinations with a total value close to the optimal solution, whereas a more difficult instance allows for more combinations near the optimum.

The first session employed a mapping structure where the achieved knapsack value translated one-to-one into the payoff of the traded security. For instance, a knapsack value of 1500 implied a final payoff equal to 15 experimental dollars per security. However, with this mapping, the incentives for participants to search for higher-value knapsacks decrease in the distance to the optimal value. This lack of incentive was particularly pronounced for more

[22] Due to odd numbers of participants, an exactly even distribution of allocations was not possible in the second, fourth and fifth sessions.

difficult instances. In these instances, there exist many possible combinations with values relatively close to the optimal value.

Therefore, a different mapping structure was required. After careful consideration, a Sigmoid transformation was chosen to provide a generic structure for an incentive-compatible knapsack to payoff mapping. The parameterization was chosen such that each instance's mapping function was centred on the optimal solution of that instance. This feature was important to ensure that participants could not guess the optimal solution of future instances based on their experience from solving prior ones. A mapping function that centred all instances at the same point would be vulnerable to participants trading predominantly on the basis of past knowledge, rather than due to the knapsack at hand and current market prices.

The nature of the Sigmoid function is that it is sharply increasing around the central inflection point, and almost flat near both tails. As the value-distance between feasible knapsacks tends to decrease with instance difficulty, the applied transformation ensured high incentives to search for more valuable knapsacks, even for the most difficult instances. Concerns about participants being able to extract excess information from the shape of the mapping function were alleviated by the rounding of security payoffs to one decimal place. This modification effectively removed any curvature of the Sigmoid transformation around the optimal knapsack value[23]. Thus, participants were prevented from applying trial-and-error techniques to find the optimal knapsack value by simply manipulating the mapping function. To illustrate this, Table 1 provides an overview of each respective instance. From this table, we can see that the instance with a Sahni-k value of 1 would have its optimal knapsack value of 1615 mapped to a security payoff of 16.1 experimental dollars. Figure 3 provides an illustration of the Sigmoid transformation applied to this same instance.

[23] There was no clear slope in each instance for (at least) the four knapsacks closest to the optimal knapsack.

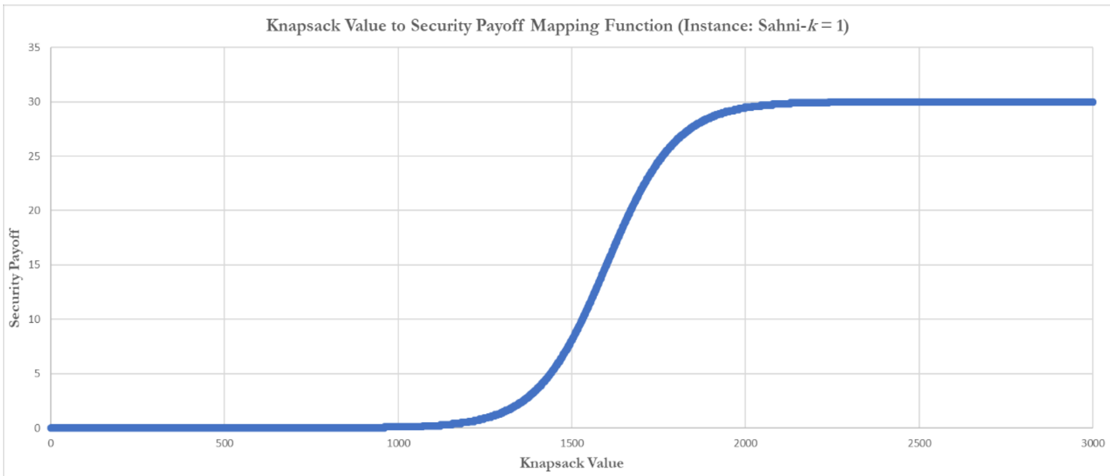
Table 1 – Value of the Optimal Knapsack and Corresponding Security Value in Each Instance

The first column identifies each instance by Sahni-*k*. The second column reports the value of the optimal knapsack for each respective instance. The third column reports the payoff on the security in each instance, derived through each instance's unique mapping function.

Instance (represented by Sahni-k)	**Value of the Optimal Knapsack**	**Mapped Payoff**
0	984	9.8
1	1615	16.1
2	1220	12.2
3	1160	11.6
4	1371	13.7

Figure 3 – Sigmoid Transformation Mapping Knapsack Value to Security Payoff for the instance with Sahni-*k* = 1

This function transforms a knapsack value input (x-axis) into a security payoff output (y-axis) for the instance with Sahni-*k* = 1.



The applied mapping function also relates the computational complexity of knapsack problems to the difficulty of real world problems. For instance, the task of evaluating a company's true stock price is one requiring precision in a particular 'range of uncertainty'. It is a trivial matter to ascertain whether a company's value is closer to zero or several billion dollars. It is considerably more difficult, however, to pinpoint a company's exact stock price – e.g. 48, 49 or 50 dollars. Thus, a mapping function which is flat at the extremities but steep in the region of interest suits the generic nature of the valuation problems tackled by financial markets, and participants in those markets.

Participant Payoffs

Participants' payoffs were designed to reward participants for their attendance and incentivize them to perform well in each trading round. These payoffs consisted of three components. Firstly, each participant was paid a 10 AUD show-up fee. Secondly, in each trading round, a fixed reward of 2 AUD was paid for submitting at least one solution to the knapsack problem at hand, independent of the quality of that submission[24]. Finally, participants were paid based on their trading performance. Individual trading round performance was calculated as the sum of final cash holdings plus the optimal knapsack implied payoff for each unit of the security held post trading. At the end of the experiment, the trading performance of one randomly selected round was converted into AUD at a ratio of 2.5:1. This random selection protocol was applied to ensure that participants remained equally motivated to perform across rounds, and is standard in the literature (Bossaerts et al., 2018). The payoffs to participants ranged from 44 AUD to 70 AUD, with an average of 50.8 AUD.

[24] A fixed reward for submitting at least one solution in each round ensured that while participants were incentivized to attempt a solution to the KP, they could not hedge between the KP and the marketplace.

V Methodology

Our data analysis methodology featured four approaches, corresponding to each of the four hypotheses.

a. Understanding Relative Price Performance

To evaluate whether market prices in this experiment were closer to the true value of the asset than prices in Bossaerts et al. (2018), the following approach was adopted.

First, a measure of price convergence was defined. This measure was constructed to understand what knapsack value was suggested by market prices. Each feasible knapsack was transformed into its payoff equivalent via the Sigmoid mapping function. A 'convergence band' was then created with upper and lower limits defined in Equation (2), below. Here, x_0 is a given knapsack value, x_{+1} is the closest feasible knapsack value above x_0, and x_{-1} is the closest feasible knapsack value below x_0.

$$\left[\frac{x_{+1}}{2} + x_0, \frac{x_{-1}}{2} + x_0\right) \qquad (2)$$

A market was judged to have converged to the x_0 knapsack if prices were eventually 'captured' within this band[25]. This implied market knapsack could then be directly compared to the simulated[26] market knapsacks in Bossaerts et al. (2018).

[25] Meaning that prices at some point entered and did not exit the range of values defined by Equation (2).

[26] Each Arrow-Debreu security had a price between 0 and 1. The final price of each security was interpreted as the probability of that item being chosen by the 'market' in a knapsack. 50,000 simulations were then run to create a distribution of 'market' knapsacks. The average of this distribution thus represented an estimate of the 'market' knapsack for that instance.

b. Understanding Relative KP Performance

To evaluate whether participants were better at solving the KP in this experiment, the following two-step approach was adopted. The first step compared the proportion of participants who solved each instance[27] of the KP in this experiment with the proportions from Bossaerts et al. (2018). The second step consisted of comparing the actual knapsack distributions from Bossaerts et al. (2018) and the distributions from this experiment. This was required because the first step, while important, could not provide holistic information as it only considered participants who reached the optimal solution. By comparing the full distributions of achieved knapsacks between both experiments, a statement could be made on whether participants in general performed better in one experiment or the other. This would complement the results from the first step and provide a richer comparison.

c. Understanding the Information Transfer Mechanism

To evaluate whether participants used information embedded in prices to improve their knapsack and influence their trading behaviour, the following approach was adopted. Three key tests were conducted to understand different aspects of the information transfer mechanism.

Price Information and Knapsack Performance

The first test sought to understand whether market participants used price information to improve their performance in the KP. We first defined a measure of participant KP performance over a time interval[28] (denoted *'KP Performance')* as the highest knapsack value

[27] One instance of the KP (Sahni-k = 4) used in Bossaerts et al. (2018) was not directly comparable with the similar instance used in this paper. This was due to the presence of multiple solutions with different Sahni-*k* in that instance, which was corrected in this paper. Results for the actual knapsack distributions for this instance are noted with this caveat.

[28] An interval of 15 seconds was used. Use of intervals overcame analytical issues with insufficient observations.

an individual had achieved until the point in time defined by the end of that interval, divided by the value of the solution to that instance. To define price information, a variable corresponding to the average of the highest bid prices in the market over a time interval (denoted '*Bidprice*') was used[29]. Given that both variables displayed trending, non-stationary behaviour over time, first differences were taken. The final specification[30] included lags of Δ*Bidprice* along with dummy variables for instance complexity[31] (denoted '*Complexity (1)*' to '*Complexity (4)*') and variables for relative (denoted '*BTA*') and individual (denoted '*CBO*') overconfidence. This was estimated via a generalised linear mixed effects ("GLME") Gaussian model with participant fixed effects to account for unobserved heterogeneity within participants. This specification is reported as Equation (3), below:

$$\begin{aligned}\Delta KP\ Performance &= \alpha_i + \beta_1 * \Delta Bidprice_{i,t} + \beta_2 * \Delta Bidprice_{i,t-1} + \beta_3 * \Delta Bidprice_{i,t-2} \\ &+ \beta_4 * Complexity\ (1) + \beta_5 * Complexity\ (2) + \beta_6 \\ &* Complexity\ (3) + \beta_7 * Complexity\ (4) + \beta_8 * BTA + \beta_9 * CBO \\ &+ \varepsilon_{i,t} \quad (3)\end{aligned}$$

The rationale for this specification was that if participants used market prices to improve their solution of the KP, there would be a relation between price increases and KP performance. If a participant noticed that bid prices were rising, they may take this as a signal to improve their knapsack solution and understand whether that price rise is justified or not.

[29] Given that price information is vaguely defined, the traded price and the ask price were used in alternative specifications.
[30] In this and all following regressions, the final specification was selected primarily using the Akaike Information Criterion ("AIC").
[31] To avoid perfect multicollinearity issues, one dummy (Complexity (0)) is omitted as per standard practice.

Price Information and Effort in the Knapsack

We recognized that there may be a different way that price information is reflected in participant behaviour. As the KP is a difficult problem, participants may not necessarily be able to improve their performance, even if they are trying to do so based on information from market prices. The second test sought to account for this, by understanding whether price information at least spurred participants to put more effort into solving the KP. To do this, we first defined two new variables. The first variable, *'Activity'*, proxied the 'effort' participants put towards solving the KP by counting the number of times items were moved in and out of the knapsack over a time interval[32]. The second variable, *'Relative Performance'*, was a measure of participant performance relative to the market. This was defined as the ratio of the highest knapsack value (in payoff terms) a participant had achieved until the point in time defined by the end of that interval, to the market price[33]. This variable was also differenced to remove non-stationarity.

The data was split into two sub-samples – one which only considered participants performing better than the market[34], and one which only considered those doing worse. The rationale for this was that if participants did use market prices to determine how much effort to devote to solving the KP, those performing worse would work harder to 'catch-up' with the market. Similarly, those performing better would exert relatively less effort, as they know they are in a 'superior' position to the market. The final specification regressed *'Activity'* on *'ΔRelative Performance'* and its lags for each sample, along with complexity dummies, overconfidence

[32] An interval of 10 seconds was used. Alternate intervals of 15 and 20 seconds were also used.
[33] The market price was defined as the midpoint between the average of the prevailing ask and bid prices over the interval.
[34] This would indicate a *Relative Performance* value greater than unity.

variables and a time trend[35]. This was estimated via a GLME Poisson[36] model with participant fixed effects. This specification is reported below as Equation (4):

$$Activity = \alpha_i + \beta_1 * \Delta Relative\ Performance_{i,t} + \beta_2 * \Delta Relative\ Performance_{i,t-1} + \beta_3 * \Delta Relative\ Performance_{i,t-2} + \beta_4 * Complexity\ (1) + \beta_5 * Complexity\ (2) + \beta_6 * Complexity\ (3) + \beta_7 * Complexity\ (4) + \beta_8 * BTA + \beta_9 * CBO + \beta_{10} * Time + \varepsilon_{i,t} \quad (4)$$

Price Information and Trading Behaviour

In contrast to the first and second tests, which investigated the transfer of information from the market to participants, the third test looked at whether information flowed in the opposite direction. This flow was measured by the number and type of orders that participants submitted to the market. We first defined two new variables that counted the number of buy or sell orders (denoted *'Bidcount'* and *'Askcount'* respectively) that a participant submitted in a given time interval[37]. The *'ΔRelative Performance'* measure from the second test was used as a regressor in this specification. The data was again split into two sub-samples, as in the second test. The rationale for this test was that participants whose current KP solution implied that market prices were overvaluing the security would be more likely to place sell orders, and vice-versa. Thus, information that participants held about the KP solution would be transmitted to the market. The final specifications regressed *'Askcount'* and *'Bidcount'* on *'ΔRelative Performance'* and lags, as well as complexity dummies and

[35] This time trend accounted for a roughly exponential decrease in *Activity* over time for both samples. Plots of this trend are available in Appendix V-a.
[36] This model was chosen as it accounted for the non-negative integer nature of the dependent variable.
[37] An interval of 10 seconds was used. An alternate interval of 15 seconds was also used.

overconfidence variables. These were estimated via two GLME Poisson models, with participant fixed effects. The specifications are reported below as Equations (5) and (6):

$$Askcount = \alpha_i + \beta_1 * \Delta Relative\ Performance_{i,t} + \beta_2 * \Delta Relative\ Performance_{i,t-1} + \beta_3 * \Delta Relative\ Performance_{i,t-2} + \beta_4 * \Delta Relative\ Performance_{i,t-3} + \beta_5 * Complexity\ (1) + \beta_6 * Complexity\ (2) + \beta_7 * Complexity\ (3) + \beta_8 * Complexity\ (4) + \beta_9 * BTA + \beta_{10} * CBO + \varepsilon_{i,t} \quad (5)$$

$$Bidcount = \alpha_i + \beta_1 * \Delta Relative\ Performance_{i,t} + \beta_2 * \Delta Relative\ Performance_{i,t-1} + \beta_3 * \Delta Relative\ Performance_{i,t-2} + \beta_4 * \Delta Relative\ Performance_{i,t-3} + \beta_5 * Complexity\ (1) + \beta_6 * Complexity\ (2) + \beta_7 * Complexity\ (3) + \beta_8 * Complexity\ (4) + \beta_9 * BTA + \beta_{10} * CBO + \varepsilon_{i,t} \quad (6)$$

d. Understanding the Impact of Overconfidence

We tested for the impacts of overconfidence primarily through including a participant-specific measure of CBO and BTA in each of the regressions in Section V, Part c. This would tell us whether overconfidence played a role in how participants performed in the KP and how much effort they put in, as well as whether it impacted their trading behaviour.

We also conducted two overconfidence-specific tests. Given that the literature often links trading performance with overconfidence (Biais et al., 2005), we evaluated whether participants' trading performance[38] could be explained by overconfidence, as well as their final holdings[39]. The rationale for this was that participants who were more overconfident

[38] Trading performance was defined as a participant's earnings per instance.
[39] Final holdings were defined as 0: mix of cash and securities and 1: either all cash or all securities.

may tend to trade more and have final holdings that were more concentrated in one asset. These regressions are specified in Equations (7) and (8). Equation (7) was estimated via a Generalized Linear Gaussian model, whereas Equation (8) was estimated via a Generalized Linear model with a Logistic link function.

$$Trading\ Performance_i = \alpha + \beta_1 * BTA_i + \beta_2 * CBO_i + \varepsilon_i \quad (7)$$

$$Final\ Holdings_i = \alpha + \beta_1 * BTA_i + \beta_2 * CBO_i + \varepsilon_i \quad (8)$$

VI Results

a. Descriptive Statistics

Table 2 - Average Participant Earnings per Session

The first column identifies each session by order. The second column reports the average earnings in that session. The third column reports the standard deviation of earnings in that session.

Session	$	SD ($)
2	49.24	1.38
3	47.00	1.08
4	54.61	4.91
5	51.87	1.64
Overall	50.83	4.11

Table 3 – Average Overconfidence Measures per Session and Overall

The first column identifies each session by order. The second column reports the average CBO level. The third column reports average BTA level.

Session	CBO[40]	BTA[41]
2	0.415	33.0715
3	0.393	35.475
4	0.357	36.667
5	0.450	30.8
Overall	0.404	34.004

Table 4 - Average Number of Item Moves per Participant per Instance

The top row identifies each instance by Sahni-*k*. The second row reports the average number of item moves made in the KP ('Activity') by a participant in that instance.

Sahni-k	0	1	2	3	4
Average Number of Item Moves	16	20	10	17	11

[40] A participant with no calibration overconfidence would have a CBO of 0.000.
[41] A participant with no relative overconfidence would have a BTA of 38.5.

b. On Relative Price Performance

To examine the impact of security design on the informativeness of market prices, we compared the knapsack implied by converged prices from this experiment with the knapsack implied by final prices in Bossaerts et al. (2018). Table 5, below, reports the results of this comparison. The entries in Table 5 evaluate the implied knapsacks by how close they are to the optimal solution for that instance. For instance, an entry ranked as 'At the 5th Best Knapsack' means that market prices corresponded to the fifth highest feasible knapsack possible for that instance[42].

Table 5 – Comparison of the Knapsack Implied by Market Prices from this Experiment with those from Bossaerts et al. (2018)

The first column identifies each instance by Sahni-*k*. The second column reports how close to the optimal solution the market knapsack was in this experiment. The third column reports how close to the optimal solution the market knapsack was in Bossaerts et al. (2018).

Instance (labeled by Sahni-k)	**'Distance' to Optimal Solution of the Market Knapsacks**	**'Distance' to Optimal Solution of Bossaerts et al. (2018) Market Knapsacks**
0	At Optimal	Below the 4th Best Knapsack
1	At Optimal	Below the 10th Best Knapsack
2	At Optimal	Below the 4th Best Knapsack
3	At Optimal	Below the 4th Best Knapsack
4	At the 5th Best Knapsack	Below the 8th Best Knapsack

Appendix VI-a shows the bid-ask spreads for each instance. These reinforce the accuracy of our convergence measure – regardless of complexity, the bid-ask spread tends to zero. Thus, the market does collectively accept the converged price to be accurate.

[42] This was done as comparing knapsacks naively by value is not meaningful. A harder problem will have more solutions closer to the optimal than an easier one.

c. On Relative Knapsack Performance

We extended the analysis of relative performance beyond market prices and considered whether using a compound security improved how participants solved the KP. If this was the case, we would expect that a greater proportion of participants would reach the optimal solution to a given instance in this experiment compared to Bossaerts et al. (2018). Table 6, below, reports the results of this comparison. The entries in Table 6 show that for all instances, a greater proportion of participants in this experiment reach the optimal solution. For instance, 60.59% more participants reached the optimal solution in the second hardest instance (Sahni-k = 3).

Table 6 – Comparison of the Proportion of Total Participants who reached the Optimal Solution in this experiment with those from Bossaerts et al. (2018)

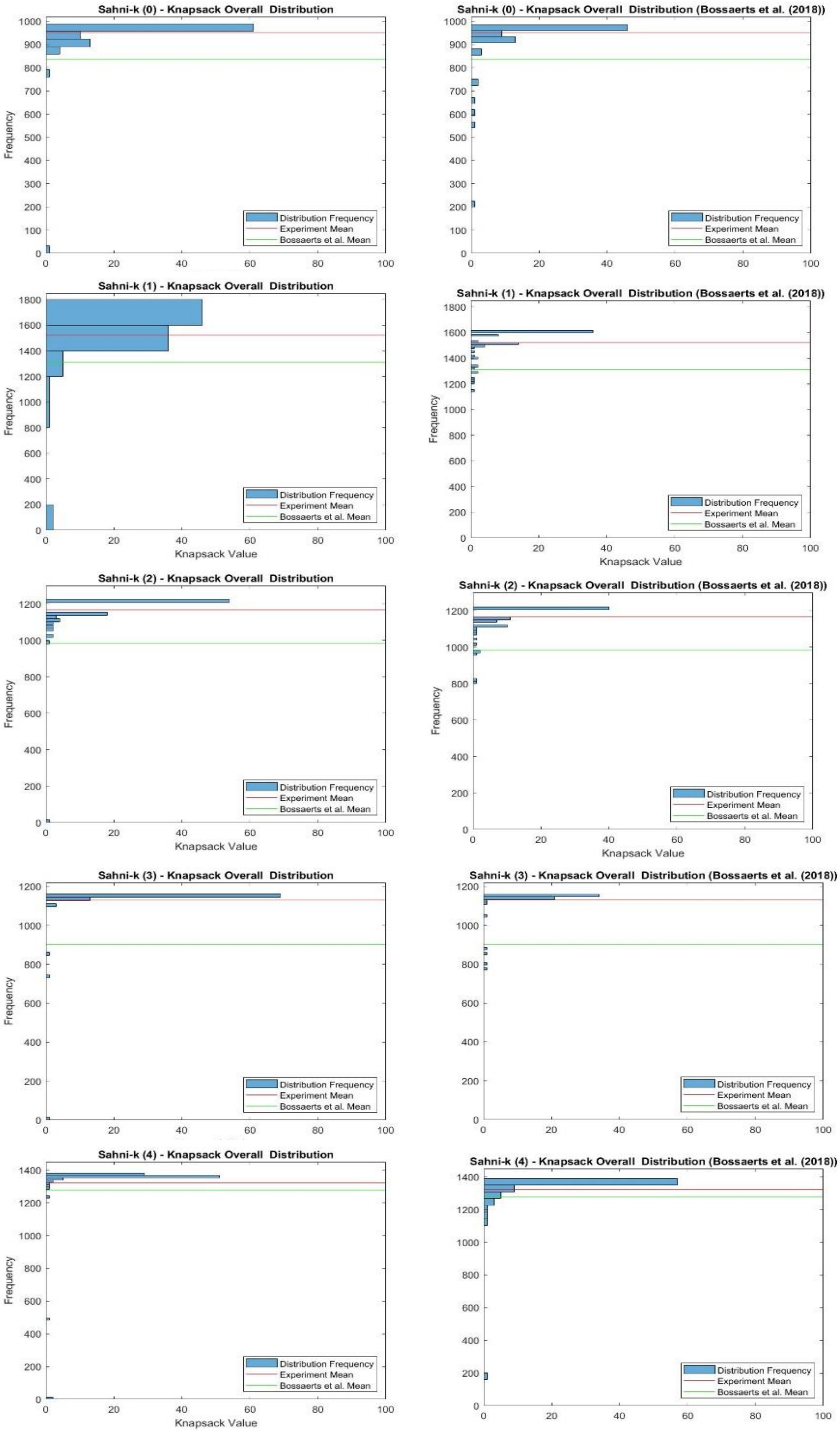
The first column identifies each instance by Sahni-*k*. The second column reports the proportion of total participants who reached the optimal solution in each instance of this experiment. The third column reports the proportion of total participants who reached the optimal solution in each instance of the Bossaerts et al. (2018) experiment. The fourth column reports the percentage change in the proportion of participants who reached the optimal solution in each instance – a positive percentage change here indicates this experiment improved upon Bossaerts et al. (2018).

***Instance (labeled by Sahni*-k)**	**% of Participants who reached the Optimal Solution**	**% of Participants who reached the Optimal Solution in Bossaerts et al. (2018)**	**% Change in the Proportion of Participants who reached the Optimal Solution**
0	67.80%	60.30%	12.44%
1	50.60%	46.20%	9.52%
2	52.80%	47.40%	11.39%
3	43.20%	26.90%	60.59%
4	10.60%	6.40%	65.63%

To understand whether this improvement extended to participants who did not reach the optimal solution, we directly compared the full distribution of submitted knapsacks between both experiments. The results of this comparison are displayed in Figure 3.

Figure 3 – Knapsack Distributions Comparison

Below: A comparison of the values of submitted KP solutions between Bossaerts et al. (2018) (left) and this paper (right). The Green line in each plot corresponds to the mean of submitted KP solutions in Bossaerts et al. (2018), whilst the Red line corresponds to the mean of submitted KP solutions in this paper.



d. On the Information Transfer Mechanism

Price Information and Knapsack Performance

The results of the regression specified by Equation (3) are below, in Table 7. We find that the contemporaneous change in bid price ($\Delta Bidprice_t$), as well as its first and second lags, were significantly ($p < 0.05$) and positively related ($\beta_1 = 0.001$; $\beta_2 = 0.001$; $\beta_3 = 0.001$) with the change in participant knapsack performance ($\Delta KP\ Performance_t$). While the magnitude of the impact of the contemporaneous or lagged terms appears small, it is more sensible to consider these impacts with respect to the reference point of the intercept. In this context, a one dollar change in the bid price from last period would, on average, improve participant knapsack performance by 12.5% relative to their baseline performance[43].

We also find that, relative to the Complexity (0) case, there are significant ($p < 0.1$) and negative relations ($\beta_4 = -0.004$, $\beta_5 = -0.004$, $\beta_6 = -0.006$. $\beta_7 = -0.003$) between complexity and knapsack performance. The overconfidence measures have no effect.

For robustness, the bid price regressors were substituted with ask and trade price variables. Results from the ask price and trade price specifications broadly matched the findings above, and are available in Appendix VI-b.

[43] This is how participants, on average, would perform in the easiest instance with all regressors equal to zero.

Table 7 – Regression Results: Price Information and Knapsack Performance

This table reports the GLME Gaussian regression results for equation (3). For each participant *i,* we estimate equation (3) using data from all trading rounds, divided into 15 second intervals. ΔKP Performance is the first difference of the current running maximum knapsack value divided by the optimal value of the relevant KP instance. ΔBidprice is the first difference of the average of the prevailing market bids in that interval. Complexity (1), (2), (3) and (4) are dummies for the complexity level of an instance – Complexity (0) forms the reference point and is thus not explicitly reported. BTA is the measure of better-than-average overconfidence. CBO is the measure of calibration-based overconfidence. The last five rows report model selection criteria (AIC, BIC and log-likelihood), the number of observations in the dataset and the fixed effects used in the regression. Standard errors are clustered at the participant level and reported in parentheses below the coefficient estimates. Coefficient significance is reported by the following symbols: + $p<0.1$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$.

	Dependent Variable: ΔKP Performance$_t$
ΔBidprice$_t$	0.0010*
	(0.0004)
ΔBidprice$_{t-1}$	0.0010*
	(0.0004)
ΔBidprice$_{t-2}$	0.0010*
	(0.0004)
Complexity (1)	-0.0040**
	(0.0020)
Complexity (2)	-0.0040*
	(0.0020)
Complexity (3)	-0.0060***
	(0.0020)
Complexity (4)	-0.0030^{+}
	(0.002)
BTA	0.00003
	(0.00005)
CBO	-0.0030
	(0.0040)
Constant	0.0080**
	(0.0020)
Participant Fixed Effects	Yes
Observations	7,592
Log Likelihood	13,340
Akaike Inf. Crit.	-26,656
Bayesian Inf. Crit.	-26,573

Price Information and Effort in the Knapsack

The results of the regression specified by Equation (4) are below, in Table 8. For participants who are performing worse than the market, we find that the contemporaneous change in market performance (ΔRelative Performance$_t$) and its first and second lags are significantly ($p < 0.001$, $p < 0.001$, $p < 0.05$) and positively related ($\beta_1 = 0.786$; $\beta_2 = 0.318$; $\beta_2 = 0.252$) with the logarithm of the number of moves[44], on average. The CBO term is also significantly ($p < 0.1$) and negatively related ($\beta_5 = -1.155$) with the logarithm of the number of moves.

For participants performing better than the market, we find that the contemporaneous change in relative performance is significant ($p < 0.05$) and negative ($\beta_1 = -0.006$). The overconfidence measures have no effect for this sample.

In both samples, complexity is generally negatively related to the logarithm of the number of moves made in the knapsack (apart from the coefficient on Complexity (2) in Sub-Sample 2). The Time effect is very significant ($p < 0.001$) and negative in both samples. As expected[45], the effect is stronger in Sub-Sample 2 compared to Sub-Sample 1, with the coefficient almost doubling in magnitude.

For robustness, different time intervals of 15 and 20 seconds were used. Results from these intervals broadly matched the findings above and are available in Appendix VI-c.

[44] Poisson model coefficients are interpreted as 'multiplicative factors'. For instance, a unit increase in the contemporaneous change in relative performance would approximately double the number of moves.
[45] The plots in Appendix V-a show that the exponential trend is much stronger in Sub-Sample 2.

Table 8 – Regression Results: Price Information and Effort in the Knapsack

This table reports the GLME Poisson regression results for equation (4). For each participant *i* in each sample, we estimate equation (4) using data from all trading rounds, divided into 10 second intervals. Activity is the number of items moved in or out of the knapsack in the time interval. ΔRelative Performance is the first difference of the running maximum knapsack value, transformed into its corresponding payoff, divided by the market price. Complexity (1), (2), (3) and (4) are dummies for the complexity level of an instance – Complexity (0) forms the reference point and is thus not explicitly reported. BTA is the measure of better-than-average overconfidence. CBO is the measure of calibration-based overconfidence. Time is the time interval variable, which controls for a time trend in the data. The last five rows report model selection criteria (AIC, BIC and log-likelihood), the number of observations in the dataset and the fixed effects used in the regression. Standard errors are clustered at the participant level and reported in parentheses below the coefficient estimates. Coefficient significance is reported by the following symbols: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001.

	Dependent Variable: Activity_t	
	Sub-Sample 1 (Knapsack worse than market)	*Sub-Sample 2* (Knapsack better than market)
$\Delta\text{Relative Performance}_t$	0.786^{***}	-0.006^{*}
	(0.074)	(0.003)
$\Delta\text{Relative Performance}_{t-1}$	0.318^{***}	-0.0001
	(0.079)	(0.003)
$\Delta\text{Relative Performance}_{t-2}$	0.252^{**}	-0.001
	(0.083)	(0.003)
Complexity (1)	-0.103^{+}	-0.379^{***}
	(0.060)	(0.046)
Complexity (2)	-0.257^{***}	0.021
	(0.049)	(0.051)
Complexity (3)	-0.014	-0.179^{***}
	(0.061)	(0.040)
Complexity (4)	-0.526^{***}	-0.091^{*}
	(0.047)	(0.037)
BTA	0.009	0.0004
	(0.008)	(0.007)
CBO	-1.155^{+}	-0.024
	(0.630)	(0.475)
Time	-0.068^{***}	-0.126^{***}
	(0.001)	(0.001)
Constant	1.943^{***}	2.589^{***}
	(0.368)	(0.284)
Participant Fixed Effects	Yes	Yes
Observations	4,576	6,443
Log Likelihood	-8,310.646	-11,138.500
Akaike Inf. Crit.	16,645.290	22,301.010
Bayesian Inf. Crit.	16,722.920	22,382.580

Price Information and Trading Behaviour

The results of the regressions specified by Equations (5) and (6) are below in Tables 9 and 10, respectively. For participants performing worse than the market, we find that the contemporaneous change in market performance (ΔRelative Performance$_t$) is mildly significant ($p < 0.1$) and negatively related ($\beta_1 = -1.040$) with the logarithm of submitted sell orders. Complexity and overconfidence have no effect for this sample. For participants performing better than the market, we find that relative performance, complexity and overconfidence have no effect on the logarithm of submitted buy orders.

For robustness, a different time interval of 15 seconds was used. Results from this interval displayed some heterogeneity, and are available in Appendix VI-d.

Table 9 – Regression Results: Price Information and Selling Behaviour

This table reports the GLME Poisson regression results for equation (6). For each participant *i* in each sample, we estimate equation (6) using data from all trading rounds, divided into 10 second intervals. Askcount is the number of sell orders submitted in the time interval. ΔRelative Performance is the first difference of the running maximum knapsack value, transformed into its corresponding payoff, divided by the market price. Complexity (1), (2), (3) and (4) are dummies for the complexity level of an instance – Complexity (0) forms the reference point and is thus not explicitly reported. BTA is the measure of better-than-average overconfidence. CBO is the measure of calibration-based overconfidence. The last five rows report model selection criteria (AIC, BIC and log-likelihood), the number of observations in the dataset and the fixed effects used in the regression. Standard errors are clustered at the participant level and reported in parentheses below the coefficient estimates. Coefficient significance is reported by the following symbols: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001.

	Dependent Variable: Askcount_t *Sub-Sample 1* (Knapsack worse than market)
ΔRelative Performance_t	-1.040^{+}
	(0.632)
ΔRelative Performance_{t-1}	-0.471
	(0.609)
ΔRelative Performance_{t-2}	-0.138
	(0.584)
ΔRelative Performance_{t-3}	-0.995
	(0.612)
Complexity (1)	0.168
	(0.226)
Complexity (2)	0.171
	(0.196)
Complexity (3)	0.148
	(0.250)
Complexity (4)	0.267
	(0.208)
BTA	0.002
	(0.005)
CBO	-0.033
	(0.411)
Constant	-2.973^{***}
	(0.282)
Participant Fixed Effects	Yes
Observations	4,576
Log Likelihood	-1,189
Akaike Inf. Crit.	2,402
Bayesian Inf. Crit.	2,479

Table 10 – Regression Results: Price Information and Buying Behaviour

This table reports the GLME Poisson regression results for equation (7). For each participant *i* in each sample, we estimate equation (7) using data from all trading rounds, divided into 10 second intervals. Bidcount is the number of buy orders submitted in the time interval. ΔRelative Performance is the first difference of the running maximum knapsack value, transformed into its corresponding payoff, divided by the market price. Complexity (1), (2), (3) and (4) are dummies for the complexity level of an instance – Complexity (0) forms the reference point and is thus not explicitly reported. BTA is the measure of better-than-average overconfidence. CBO is the measure of calibration-based overconfidence. The last five rows report model selection criteria (AIC, BIC and log-likelihood), the number of observations in the dataset and the fixed effects used in the regression. Standard errors are clustered at the participant level and reported in parentheses below the coefficient estimates. Coefficient significance is reported by the following symbols: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001.

	Dependent Variable: Bidcount_t *Sub-Sample 2* (Knapsack better than market)
$\Delta\text{Relative Performance}_t$	0.006
	(0.037)
$\Delta\text{Relative Performance}_{t-1}$	-0.013
	(0.023)
$\Delta\text{Relative Performance}_{t-2}$	-0.007
	(0.022)
$\Delta\text{Relative Performance}_{t-3}$	-0.016
	(0.021)
Complexity (1)	-0.138
	(0.147)
Complexity (2)	-0.160
	(0.201)
Complexity (3)	0.137
	(0.141)
Complexity (4)	0.356^{*}
	(0.139)
BTA	-0.003
	(0.004)
CBO	-0.182
	(0.299)
Constant	-2.435^{***}
	(0.190)
Participant Fixed Effects	Yes
Observations	6,443
Log Likelihood	-1,900
Akaike Inf. Crit.	3,824
Bayesian Inf. Crit.	3,905

e. On Overconfidence

In addition to the inclusion of overconfidence measures in the information transfer regressions, further tests specific to overconfidence were conducted. The results of these regressions, specified by Equations (7) and (8), are below in Tables 11 and 12 respectively. We find that overconfidence does not have any explanatory power for participants' holdings of assets or their trading performance.

Table 11 – Regression Results: Trading Performance and Overconfidence

This table report the Generalized Linear Gaussian regression results for equation (8). We estimate equation (8) using data from all trading rounds for all instances. Trading Performance measure earnings in dollars. BTA is the measure of better-than-average overconfidence. CBO is the measure of calibration-based overconfidence. The last four rows report goodness-of-fit criteria (R^2 and Adjusted R^2), the number of observations in the dataset and the F-statistic for the key regressors (with degrees of freedom = 2;382). Heteroskedasticity robust standard errors are reported in parentheses below the coefficient estimates. Coefficient significance is reported by the following symbols: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001.

	Dependent Variable: Trading Performance
BTA	0.007
	(0.010)
CBO	0.072
	(0.703)
Constant	49.776^{***}
	(0.433)
Observations	385
R^2	0.001
Adjusted R^2	-0.004
F Statistic	0.269

Table 12 – Regression Results: Final Holdings and Overconfidence

This table report the Generalized Linear with Logistic link function regression results for equation (9). We estimate equation (9) using data from all trading rounds for all instances. Asset Holdings measures the composition of final holdings. BTA is the measure of better-than-average overconfidence. CBO is the measure of calibration-based overconfidence. The last three rows report model selection criteria (AIC and log-likelihood) and the number of observations in the dataset. Coefficient significance is reported by the following symbols: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001.

	Dependent Variable: Asset Holdings
BTA	0.011
	(0.008)
CBO	-0.547
	(0.561)
Constant	-1.680^{***}
	(0.352)
Observations	385
Log Likelihood	-179.755
Akaike Inf. Crit.	365.509

VII Discussion

a. Compound Securities Outperform Arrow-Debreu Securities

We find that markets and participants perform better when the traded securities are compound securities, compared to the Arrow-Debreu securities used in Bossaerts et al. (2018). This contrasts with Plott and Sunder (1988) and the broader consensus in the experimental literature on security design, which is that Arrow-Debreu securities better support information aggregation.

There is some evidence of the impact of complexity on market performance. The only case in which prices did not converge to the true value is when Sahni-k was 4. Thus, only in the most difficult instance was market performance sub-optimal.

Furthermore, there is a marked improvement in the proportion of participants who solve for the optimal solution in each instance relative to Bossaerts et al. (2018). A possible critique is that participants in our experiment were simply better at solving the KP than those in Bossaerts et al. (2018). It is true that we did not control for the quality of individuals' problem-solving skills when selecting participants, as any such control would have been imprecise and unfeasible. However, this experiment followed Bossaerts et al. (2018) in drawing participants from an essentially identical population with the same eligibility requirements – including having no prior experience with the KP in past experiments. Therefore, we can conclude that participant 'quality' is unlikely to be an important driver of our results, and this critique is invalid. Consequently, the increase in the number of participants that solved the KP can be taken as further evidence of the positive impact that using a single compound security has had.

The scalability and relevance of these findings can be questioned by noting whether an average of nineteen participants per session is sufficient to model 'real' market behaviour, where the number of participants is much higher and there would likely be greater liquidity. However, Meloso et al. (2009) report that experiments with just fifteen participants were able to appropriately model real markets. Our experimental results can therefore be considered relevant and applicable with respect to real-world markets.

b. Participants do Learn from Market Prices

Price Information and Knapsack Performance

We find that there is evidence of information transfer from prices to individual problem solving. Changes in market prices improved participants' KP performance. The significance of both the contemporaneous and lagged changes indicates that participants not only 'track' the market in the present but use past information to improve current performance. This makes sense – the KP is difficult and it may be time-consuming to improve upon a solution. Thus, a change in the price last period may only reflect an improvement in later periods.

Furthermore, Sahni-k is found to be related to how participants perform in the KP, as in Bossaerts and Murawski (2016). Participants perform worse on average in problems harder than the easiest (Sahni-k = 0) problem. This can be explained by prior research (Bossaerts and Murawski, 2016), where it was found that participants appear to use the 'greedy' algorithm as a heuristic to solve the KP. This goes some way to explaining why performance is significantly better when Sahni-k is 0[46] compared to higher complexity instances.

[46] When $k = 0$, the 'greedy' algorithm can solve the KP without needing to change any items.

Price Information, Effort and Trading Behaviour

We also find that participants use their position relative to the market to determine how much effort to commit to solving the KP. The second result tells us that participants performing worse than the market tend to put in more effort in the KP, compared to participants performing better than the market. This is expected and sensible – participants read information from market prices and use it to motivate whether their information set can be improved or not. Combined with the first result, this clearly shows that participants use market prices to improve, or at least try to improve, their performance in the knapsack.

The third result provides a rich illustration of how information flows from participants into the market. There is significant heterogeneity in participants' trading and KP activity depending on their position relative to the market. To illustrate this, we can consider the set of participants performing worse than the market. Those performing relatively worse than the market are expected to be less active in the KP and to submit more sell orders. Those performing relatively better (but still worse than the market) are expected to be more active in the KP and to submit fewer sell orders.

These intriguing results can be explained as follows. Participants who are very far away from the market suffer from uncertainty – they either have an opportunity to sell the security for more than it is worth (if the market is overpriced) or they must commit significant effort to 'catch-up' to the market (if the market is accurate). This result indicates that these participants prefer to sell. This can be explained by considering that CBO had a significant effect on effort in the KP only for those performing worse than the market. Thus, it is plausible that these participants were more confident in their private information set, and thus decided to sell rather than chase the market. In contrast, participants closer to the market price may be less vulnerable to such an effect. They see that they are near the market, and are motivated to pursue this price, spending more effort in the KP and submitting fewer sell orders.

The results for the fifteen second time interval (Appendix VI-d) indicate that a time dimension may be relevant in the behaviour of underperforming participants. The relationship between relative performance and submitted sell orders becomes insignificant in this case. This change could be explained by considering the speed at which participants can process and react to market information. Ten seconds may be sufficient time for this, and thus, enlarging the time interval would simply add noise that confounds the relationship.

Interestingly, our results show that participants who are outperforming the market submit the same number of buy orders, regardless of whether they are outperforming by a large or small margin. This seems puzzling – these participants could profit with certainty by buying securities that are currently under-priced. However, this result is easily explained by viewing participants' behaviour through a game-theoretic lens. Regardless of the magnitude of outperformance, participants may be rationally unwilling to reveal their better private information through a standing bid order. The first set of results (Table 7) showed that bid prices can spur participants to improve their KP performance, meaning that submitting such an order could incite other participants to 'catch up'. This would reduce the advantage held by the more informed participant. Instead, these participants would prefer to wait for profitable trades to appear endogenously – for instance, under-priced sell orders from those who are performing worse than the market. The results for the fifteen second interval (Appendix VI-d) also support this strategic explanation. Participants' behaviour does not change – there is still no relationship between relative performance and the number of buy orders submitted. Thus, a time dimension does not appear to be a relevant consideration for this sample, reinforcing the argument that participants who are outperforming are likely to be patiently 'watching' the market.

c. Overconfidence is not Important

The literature suggests that different types of overconfidence play a role in how participants trade and perform. We find no evidence to support this. The overconfidence measures we used were never useful as predictors of the number of bid or ask orders an individual submitted or their performance in the KP. The measures were also unhelpful in understanding participants' final holdings or their trading performance. CBO was found to be relevant in explaining how much effort participants who are performing worse than the market put into the KP, though. Participants with higher CBO tended to make fewer moves than those with lower CBO, on average. This makes sense – participants who overestimate the precision of their private information may not respond as intensely to market signals that this information is inadequate. This result also is also in accord with Regner et al.'s (2003) finding that different measures of overconfidence, like CBO and BTA, are uncorrelated. In this case, the BTA measure is insignificant even when CBO is not.

However, this is a relatively minor finding compared to the irrelevance of both overconfidence measures in all other tests. Thus, this paper provides a counterpoint against the growing consensus that overconfidence is an important part of understanding how participants trade. It also establishes a foundation for future research to investigate why this paper's findings differ from prior studies. For instance, there may be properties of computationally complex problems or our security design that drive this result.

VIII Implications and Conclusion

a. Redefining what it means for a Market to be Efficient

Our results stand in stark contrast to the theoretical framework of market information aggregation set out by Grossman (1981). The contention that a complete set of markets can support a fully revealing equilibrium was rejected in Bossaerts et al. (2018). Here, we find that an incomplete market incorporating one traded security meaningfully outperforms the complete markets tested in Bossaerts et al. (2018).

This result has significant implications for the theoretical and empirical study of market information aggregation. It casts doubt on the fundamental concept of market efficiency as it is understood in finance today. Malkiel and Fama (1970) insist that to be efficient, such that prices can optimally inform individuals, market prices must reveal all information. This is the foundation of the EMH. However, our results should prompt a re-evaluation of whether this definition is sensible. We propose that, along the lines of Hayek (1945), a market can be considered efficient if it provides just enough crucial information – such that an individual can make the right decisions. Thus, this paper provides a basis for future empirical research, which can examine whether real markets are efficient or not under this 'new' definition of efficiency.

This result is also of relevance to the field of security design. It is no longer a given that Arrow-Debreu securities are the ideal choice to induce more informative prices. While we have shown this using a single compound security, future research should look to investigate other security designs that can provide the critical information that individuals need. Such research would not only be beneficial in progressing the literature but could prove useful for better understanding real world markets; for instance, equity markets that use securities similar to the compound securities used in this experiment.

b. Bubbles are not a Factor

Conventional wisdom in the experimental literature suggests that incomplete markets are likely to suffer from price bubbles (Smith, Suchanek and Williams, 1998; Porter and Smith, 2003; Van Boening, Williams and LeMaster, 1993). Much research has focussed on the formation of bubbles in markets that are structured similarly to modern equity markets (Lei, Noussair and Plott, 2001). Plott and Chen (2002) elaborate upon this idea, claiming that markets such as ours – which rely on a single compound security – cannot reliably aggregate information. The authors posit that the inaccuracy of such markets arises in part from their potential to facilitate price bubbles.

We find no evidence to support this consensus in our experiment[47]. We find that prices in our markets can reliably aggregate information, and are, in general, accurate. Indeed, even in markets where prices did not converge to the true value, they never exhibited any sustained overpricing that could be interpreted as evidence for the existence of bubbles.

This finding has implications for real-world security design and future experimental research. Prior research concludes that equity markets are inherently vulnerable to mispricing due to bubbles, but we have shown that this need not be the case. Further research should look to explore what aspects of this experiment ensured that equity-like compound securities were not vulnerable to sustained overpricing.

[47] The design of our experiment is different to Smith et al. (1998). However, as noted, Smith et al. (1998) has been replicated numerous times under various designs and is thus considered to be akin to a generic result in the literature.

c. Conclusion

The consensus view in finance has been that complete markets will necessarily outperform incomplete markets as they provide more information. This theoretical foundation has been made robust by experiments which conclude that markets using Arrow-Debreu securities aggregate information more effectively than those using single compound securities. Our experiments demonstrate that this understanding does not hold when markets are asked to solve a generic computationally complex problem, which mimics the fundamental constrained optimization problem faced by any economic agent. Within this framework, single security markets better aggregate information than their complete counterparts. Thus, this work should persuade researchers to revisit the theoretical and experimental underpinnings of the consensus and reconcile it with our findings.

IX References

Arrow, Kenneth J., and Gerard Debreu. "Existence of an equilibrium for a competitive economy." *Econometrica: Journal of the Econometric Society* (1954): 265-290.

Biais, Bruno, Denis Hilton, Karine Mazurier, and Sébastien Pouget. "Judgemental overconfidence, self-monitoring, and trading performance in an experimental financial market." *The Review of economic studies* 72, no. 2 (2005): 287-312.

Bartholdi, John J. "The knapsack problem." In *Building Intuition*, pp. 19-31. Springer, Boston, MA, 2008.

Biais, Bruno, Peter Bossaerts, and Chester Spatt. "Equilibrium asset pricing and portfolio choice under asymmetric information." *The Review of Financial Studies* 23, no. 4 (2010): 1503-1543.

Bossaerts, Peter, and Carsten Murawski. "How humans solve complex problems: the case of the knapsack problem." *Scientific Reports* 6 (2016): 34851.

Bossaerts, Peter, and Carsten Murawski. "Computational complexity and human decision-making." *Trends in Cognitive Sciences* 21, no. 12 (2017): 917-929.

Bossaerts, Peter, Elizabeth Bowman, Shijie Huang, Carsten Murawski, Shireen Tang and Nitin Yadav. "The Efficient Markets Hypothesis Does Not Hold When Securities Valuation Is Computationally Hard." (2017). Working Paper.

Deaves, Richard, Erik Lüders, and Guo Ying Luo. "An experimental test of the impact of overconfidence and gender on trading activity." *Review of Finance* 13, no. 3 (2008): 555-575.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. "Noise trader risk in financial markets." *Journal of political Economy* 98, no. 4 (1990): 703-738.

Diamond, Douglas W., and Robert E. Verrecchia. "Information aggregation in a noisy rational expectations economy." *Journal of Financial Economics* 9, no. 3 (1981): 221-235.

Flannery, Mark J. "Using market information in prudential bank supervision: A review of the US empirical evidence." *Journal of Money, Credit and Banking* (1998): 273-305.

Flannery, Mark J. "Stabilizing large financial institutions with contingent capital certificates." *Quarterly Journal of Finance* 6, no. 02 (2016): 1650006.

Flood, Mark D. "An introduction to complete markets." *Federal Reserve Bank of St. Louis Review* 73, no. March/April 1991 (1991)

Forsythe, Robert, and Russell Lundholm. "Information aggregation in an experimental market." *Econometrica: Journal of the Econometric Society* (1990): 309-347.

Grossman, Sanford J. "An introduction to the theory of rational expectations under asymmetric information." *The Review of Economic Studies* 48, no. 4 (1981): 541-559.

Hayek, Friedrich August. "The use of knowledge in society." *The American Economic Review* 35, no. 4 (1945): 519-530.

Horowitz, Ellis, and Sartaj Sahni. "Computing partitions with applications to the knapsack problem." *Journal of the ACM (JACM)* 21, no. 2 (1974): 277-292.

Kellerer, Hans, Ulrich Pferschy, and David Pisinger. "Introduction to NP-Completeness of knapsack problems." In *Knapsack Problems*, pp. 483-493. Springer, Berlin, Heidelberg, 2004.

Klayman, Joshua, Jack B. Soll, Claudia Gonzalez-Vallejo, and Sema Barlas. "Overconfidence: It depends on how, what, and whom you ask." *Organizational behavior and human decision processes* 79, no. 3 (1999): 216-247.

Langnickel, Ferdinand, and Stefan Zeisberger. "Do we measure overconfidence? A closer look at the interval production task." *Journal of Economic Behavior & Organization* 128 (2016): 121-133.

Lei, Vivian, Charles N. Noussair, and Charles R. Plott. "Nonspeculative bubbles in experimental asset markets: Lack of common knowledge of rationality vs. actual irrationality." *Econometrica* 69, no. 4 (2001): 831-859.

Malkiel, Burton G., and Eugene F. Fama. "Efficient capital markets: A review of theory and empirical work." *The journal of Finance* 25, no. 2 (1970): 383-417.

Maymin, Philip Z. "Markets are efficient if and only if P= NP." *Algorithmic Finance* 1, no. 1 (2011): 1-11.

McDonald, Robert L. "Contingent capital with a dual price trigger." *Journal of Financial Stability* 9, no. 2 (2013): 230-241.

Meloso, Debrah, Jernej Copic, and Peter Bossaerts. "Promoting intellectual discovery: patents versus markets." *Science* 323, no. 5919 (2009): 1335-1339.

Odean, Terrance. "Volume, volatility, price, and profit when all traders are above average." *The Journal of Finance* 53, no. 6 (1998): 1887-1934.

Plott, Charles R. "Markets as information gathering tools." *Southern Economic Journal* 67, no. 1 (2000): 2-15.

Plott, Charles R., and Kay-Yut Chen. "Information aggregation mechanisms: Concept, design and implementation for a sales forecasting problem." (2002). Working Paper.

Plott, Charles R., and Shyam Sunder. "Rational expectations and the aggregation of diverse information in laboratory security markets." *Econometrica: Journal of the Econometric Society* (1988): 1085-1118.

Porter, David P., and Vernon L. Smith. "Stock market bubbles in the laboratory." *The Journal of Behavioral Finance* 4, no. 1 (2003): 7-20.

Régner, Isabelle, Denis Hilton, Laure Cabantous, and Stéphane Vautier. *Overconfidence, miscalibration and positive illusions*. Working paper, University of Toulouse, 2004.

Savage, Leonard J. *The Foundations of Statistics*. Courier Corporation, 1972.

Smith, Vernon L., Gerry L. Suchanek, and Arlington W. Williams. "Bubbles, crashes, and endogenous expectations in experimental spot asset markets." *Econometrica: Journal of the Econometric Society* (1988): 1119-1151.

Van Boening, Mark V., Arlington W. Williams, and Shawn LaMaster. "Price bubbles and crashes in experimental call markets." *Economics Letters* 41, no. 2 (1993): 179-185.

X Appendix

IV. Experimental Design

IV-a

INSTRUCTIONS

Summary

Over 5 rounds, you will be able to trade shares in an online marketplace. The traded share's payoff equals the solution of a "knapsack problem". This is a problem where one is asked to find the right items from a collection of possible items that fit a given knapsack while maximising its value.

During the experiment, you will have access to two websites and a spreadsheet. One website will be an online trading platform, and the other will be a platform where you can try out solutions to the knapsack problem at hand, to determine the highest value of the knapsack, and thus whether you should buy or sell shares. The spreadsheet will allow you to calculate the payoff per share for different knapsack values. You earn money by buying shares whose price is lower than the optimal value of the knapsack, and by selling shares whose price is higher than the optimal value of the knapsack. You will also earn a fixed $2 reward for each suggested solution you submit to the knapsack problem (independently of whether this solution is correct or not). After each trading round, you will be asked to answer a short question about your confidence in predicting your trading performance.

Following the completion of online trading (i.e., after all five rounds have been completed), you will be asked to answer a questionnaire. This questionnaire contains a series of interval-creation tasks, where you must respond to a general knowledge question with a numerical interval, such that you are 90% sure that the true answer to the question is contained within the interval.

At the end of the experiment, your earnings from trading will be determined by one randomly selected trading round. Your performance in each round will be equal to the sum of remaining cash and the payoffs from your final share holdings in that round, plus the reward for submitting a suggested solution to the Knapsack problem.

Setting: Knapsack Problems

In knapsack problems, one is given a list of items and asked to optimally load them in a knapsack. Each item has a *weight* and a *value*. The knapsack has a *weight limitation* that may prevent one from loading all items in the knapsack. Given this limitation, one is asked to determine the optimal load, i.e., the combination of items to be put in the knapsack that maximises its total value. The total value of a proposed knapsack equals the sum of the values of the individual items.

In each of the five rounds (plus a practice round) you can see the corresponding knapsack problem online, at http://bmmlab.org/games. You should use the Mozilla Firefox web browser to access this website (see Table 3). Log into the website with the ID and password you are given, and navigate to "Play Knapsack Game," where you pick the problem corresponding to the market from the drop-down list (see Table 1 below for a list of problem identifiers). An example of the interface is shown in Figure 1 below. You will be asked to refresh this webpage and login again between each round, in order to navigate to the next problem.

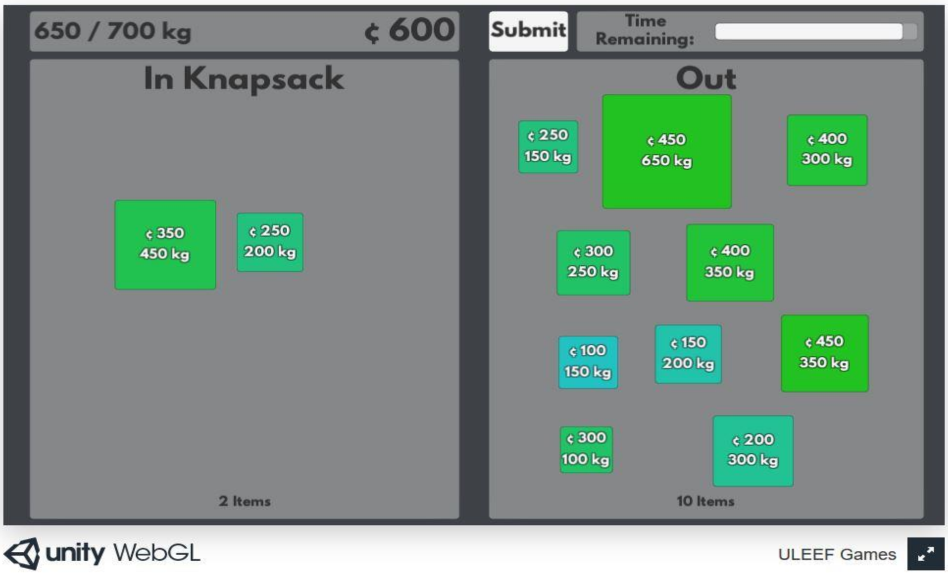


Figure 1: Knapsack Problem Interface

Round	Knapsack Problem Name	Market Name
Practice	KP-Practice	KPM-Practice
1	KP-Apple	KPM-Apple
2	KP-Banana	KPM-Banana
3	KP-Cherry	KPM-Cherry
4	KP-Durian	KPM-Durian
5	KP-Elderberry	KPM-Elderberry

Table 1: Knapsack problem and market names

You can try for yourself which items you think should be in the knapsack, by moving the items you want in the knapsack in the "IN" panel, while keeping the others in the "OUT" panel (which is to the right in the above picture, but at other times will be to the left). Items are identified by their WEIGHT (their size increases with weight) and VALUE (their colour changes from blue to green as value increases). Before the round ends, you should SUBMIT your suggested solution by clicking on the "Submit" button. You only have to submit your solution as long as trading is open (see below). (There is a white bar next to the submit button which indicates "time remaining"; please ignore it.) You earn $2 when you submit your suggested solution, independent of whether it is correct or not.

Important: You are not allowed to access any webpages other than the two you are instructed to use (the knapsack problem and the market pages)! Failure to do so will lead to exclusion from the experiment.

Earnings from Trading in the Online Market

You can earn money by *trading in a market* where shares can be bought or sold. There will be only one type of traded shares. You will be able to determine the payoff per share through an online spreadsheet. You can enter any achieved knapsack value into the spreadsheet, which will then provide you with the corresponding payoff per share. Of course, whether this payoff equals the true payoff per share depends on whether you have entered the optimal knapsack value into the spreadsheet.

If you believe that shares are currently trading *above* the true share value, you can earn money by *selling* shares. In contrast, if you believe that shares are currently priced *below* their true value, you can earn money by *buying* shares. For example, if you think the optimal knapsack value is 1995 cents, and the spreadsheet tells you that the corresponding payoff per share is $12, you would prefer to buy shares priced below $12 and sell shares priced above $12. Note, you should enter knapsack values in cents into the spreadsheet.

Your total earnings from each trading round will consist of (i) the amount of final cash holdings, plus (ii) the sum of payoffs from each share that you hold at the end of trading.

At the end of the experiment, **one** round will be randomly selected as "payment round". Your earnings from cash and shares will be converted into Australian dollars at a ratio of 2.5:1. So if, for example, you end up with 3 shares plus $14 cash, and the optimal knapsack value is $19.95 (1995 cents), then your payment in Australian dollars from that round will be equal to:

$$[(12 \ * \ 3 \, shares) + 14 \, cash] \div 2.5 = 20.00 \text{ AUD.}$$

Trading in the Online Market

Trading takes place through an electronic trading platform called *Flex-E-Markets*. In *Flex-E-Markets* you submit *limit orders*, which are orders to buy or sell at a price you determine, or, if possible, at any better price. Transactions take place from the moment a buy order with a higher price crosses a sell order with a lower price. Orders remain valid until you cancel them or the marketplace closes. You will be given ample opportunity to train yourself in submitting and canceling orders.

You can access *Flex-E-Markets* as follows: use your logon information slip and log onto https://adhocmarkets.com/ using the ID labelled "Roused-Nipper" , as well as the email on your information slip and the same password you have used to access the knapsack problem. You should use the Google Chrome web browser to access the online market (see Table 3). You should then navigate to the market name for the corresponding round (see Table 1 above for a list of market names). Each market will be open for a pre-determined time period (approximately 10 minutes). Instructors will notify you at halftime as well as one minute before markets close.

	Website	Web Browser to use
Knapsack Problem	http://bmmlab.org/games	Firefox
Online Market	https://adhocmarkets.com/	Chrome

Table 3: Websites to use and suggested Web Browsers

Questionnaires

After each round of trading, you will be asked to fill in the question sheet labelled *Questionnaire I*. These questions will ask you about the self-assessment of your trading performance in the previous trading round. After **all** rounds of trading have been completed, you will be asked to fill in the question sheet labelled *Questionnaire II,* which has three sections in total. The first section will ask you to answer ten general knowledge questions, whereas the second and third sections consist of a few follow-up questions.

IV-b

(Items corresponding to the weights in **bold** *are those that have been modified from the original values in Bossaerts et al. (2018))*

Instance	**4**	*Scale Factor*	3	
Total Capacity	4500	*Total Value*	1731	
Sahni-k	0	*Total Weight*	10344	
7	*Item no.*	*Density*	*Value*	*Weight*
*	**1**	0.74	111	150
	2	0.087804878	216	2460
	3	0.151428571	318	2100
*	**4**	0.695652174	96	138
*	**5**	0.204545455	135	660
	6	0.133962264	213	1590
*	**7**	0.214953271	69	321
*	**8**	0.244444444	132	540
*	**9**	0.195402299	255	1305
*	**10**	0.172222222	186	1080
	Sum		1731	10344
Value of Optimal	984			

Instance	**1**			
Total Capacity	1900	*Total Value*	3840	
Sahni-k	1	*Total Weight*	10014	
4	*Item no.*	*Density*	*Value*	*Weight*
	1	0.666666667	500	750
	2	0.862068966	350	406
*	**3**	0.895390071	505	564
*	**4**	0.848739496	505	595
	5	0.797011208	640	803
*	**6**	0.889570552	435	489
	7	0.725429017	465	641
	8	0.282485876	50	177
	9	0.666666667	220	330
*	**10**	0.674603175	170	252
	Sum		3840	5007
Value of Optimal	1615			

Instance	**3**	*Scale Factor*	10	
Total Capacity	8500	*Total Value*	1910	
Sahni-k	2	*Total Weight*	16250	
5	*Item no.*	*Density*	*Value*	*Weight*
	1	0.11627907	150	1290
	2	0.097222222	140	1440
	3	0.038961039	30	770
	4	0.038961039	30	770
*	**5**	0.151515152	100	660
	6	0.15	90	600
*	**7**	0.152173913	280	1840
*	**8**	0.152173913	280	1840
*	**9**	0.135371179	310	2290
*	**10**	0.135869565	250	1840
	11	0.109589041	240	2190
	12	0.013888889	10	720
	Sum		1910	16250
Value of Optimal	1220			

Instance	**2**			
Total Capacity	1044	*Total Value*	1655	
Sahni-k	3	*Total Weight*	1738	
5	*Item no.*	*Density*	*Value*	*Weight*
*	**1**	1.463414634	300	205
	2	1.388888889	350	252
*	**3**	1.136363636	400	352
*	**4**	1.006711409	450	447
	5	0.412280702	47	114
	6	0.4	20	50
	7	0.285714286	8	28
	8	0.278884462	70	251
*	**9**	0.263157895	5	19
*	**10**	0.25	5	20
	Sum		1655	1738
Value of Optimal	1160			

Instance	7			
Total Capacity	1300	Total Value	1979	
Sahni-k	4	Total Weight	1881	
7	Item no.	Density	Value	Weight
*	**1**	1.046875	201	192
	2	1.05	84	80
*	**3**	1.066037736	113	106
*	**4**	1.052083333	303	288
*	**5**	1.070754717	227	212
*	**6**	1.041493776	251	**241**
*	**7**	1.066115702	129	121
*	**8**	1.05	147	140
	9	1.048780488	86	82
	10	1.058333333	127	120
	11	1.051094891	144	137
	12	1.030864198	167	**162**
	Sum		1979	1881
Value of Optimal	1371			

Items with weight in **boldtype** have been modified from Bossaerts et al. (2018). The modifications were:

1. Weight changed from 240 to 241(Item no. 6)
2. Weight changed from 160 to 162 (Item no. 12)

IV-c

Questionnaire II

Please fill in the ID provided to you (the ID that you have used to access the knapsack game)

User ID:

Below are five questions. Each question asks you to rate, on a scale from 1 to 10, how well you think you know how much you earned during the prior round of trading. Following every round of trading, please write down your answer to the corresponding question. For example, following session 1, you should answer question 1 *only*. Following session 2, proceed to answer question 2, and so on. Be as honest as you can in these answers. The answers to these questions will not affect your payoff from trading.

Question 1

In the trading round you have just completed, how well would you say that you could evaluate how much money you made?

Answer 1

I have **no idea** how much I earned through trading	1	2	3	4	5	6	7	8	9	10	I am **completely certain** about how much I earned through trading

Question 2

In the trading round you have just completed, how well would you say that you could evaluate how much money you made?

Answer 2

I have **no idea** how much I earned through trading	1	2	3	4	5	6	7	8	9	10	I am **completely certain** about how much I earned through trading

Question 3

In the trading round you have just completed, how well would you say that you could evaluate how much money you made?

Answer 3

	1	2	3	4	5	6	7	8	9	10	
I have **no idea** how much I earned through trading											I am **completely certain** about how much I earned through trading

Question 4

In the trading round you have just completed, how well would you say that you could evaluate how much money you made?

Answer 4

	1	2	3	4	5	6	7	8	9	10	
I have **no idea** how much I earned through trading											I am **completely certain** about how much I earned through trading

Question 5

In the trading round you have just completed, how well would you say that you could evaluate how much money you made?

Answer 5

	1	2	3	4	5	6	7	8	9	10	
I have **no idea** how much I earned through trading											I am **completely certain** about how much I earned through trading

IV-d

Questionnaire II

Please fill in the ID provided to you (the ID that you have used to access the knapsack game)

User ID:

Section A - *Part I*

On the next page, you will see 10 text descriptions of numerical general knowledge questions. Please provide a lower bound and upper bound estimate for each value. The units you should use when answering a question are provided in square brackets following the question. You should choose your estimates such that you are 90% certain that the true value lies within your range. This means that, on average, 9 out of 10 true values should fall into the respective ranges.

You can find an example question below:

What is the population of the city of Berlin?

If you are 90% sure that the city of Berlin has between 2 million and 5 million inhabitants, you would answer as follows:

Question	Lower Bound	Upper Bound
What is the population of the city of Berlin? [people]	2,000,000	5,000,000

If you have any queries, please feel free to ask one of the experimenters for assistance. Please answer these questions independently and without collaborating with others.

Question	Lower Bound	Upper Bound
Martin Luther King's age at death [years]		
Length of the Nile River (in Africa) [kilometres]		
Number of countries that are members of OPEC (the Organization of Petroleum Exporting Countries) [countries]		
Number of books in the Old Testament (the first part of the Christian bible, followed by the New Testament) [books]		
Weight of an empty aeroplane (Boeing 747) [tonnes]		
Year in which the composer Johann Sebastian Bach was born (a famous German musician) [years]		
Average gestation period (time spent in the womb) of an elephant [months]		
Diameter of the moon [kilometres]		
Air distance from London to Tokyo (direct flight) [kilometres]		
Deepest known point in the oceans [metres]		

Section A - *Part II*

In this part, you will be asked a series of three questions related to your answers to Part I above. Please answer these questions independently and without collaborating with others.

Question 1

How many true values of the 10 questions from the previous page do you expect to lie within your provided intervals?

Answer 1

None of the true values lie within my intervals	0	1	2	3	4	5	6	7	8	9	10	**All** 10 of the true values lie within my intervals

Question 2

How many true values of the 10 questions from the previous page do you expect to lie within the provided ranges of the median participant?

Answer 2

None of the true values lie within the median participant's intervals	0	1	2	3	4	5	6	7	8	9	10	**All** 10 of the true values lie within the median participant's intervals

Question 3

How difficult did you find the 10 questions from the previous page?

Answer 3

Very Difficult	1	2	3	4	5	**Very Easy**

Section B

Please answer these questions independently and without collaborating with others.

<u>Question</u>

Out of the 100 (yourself included) people doing this experiment (not just those in today's session), how many do you think will make more money from trading than you?

Please write your answer as a number from 0 (indicating no-one did better than you) to 99 (indicating everyone did better than you) in the box below.

<u>Answer</u>

IV-e

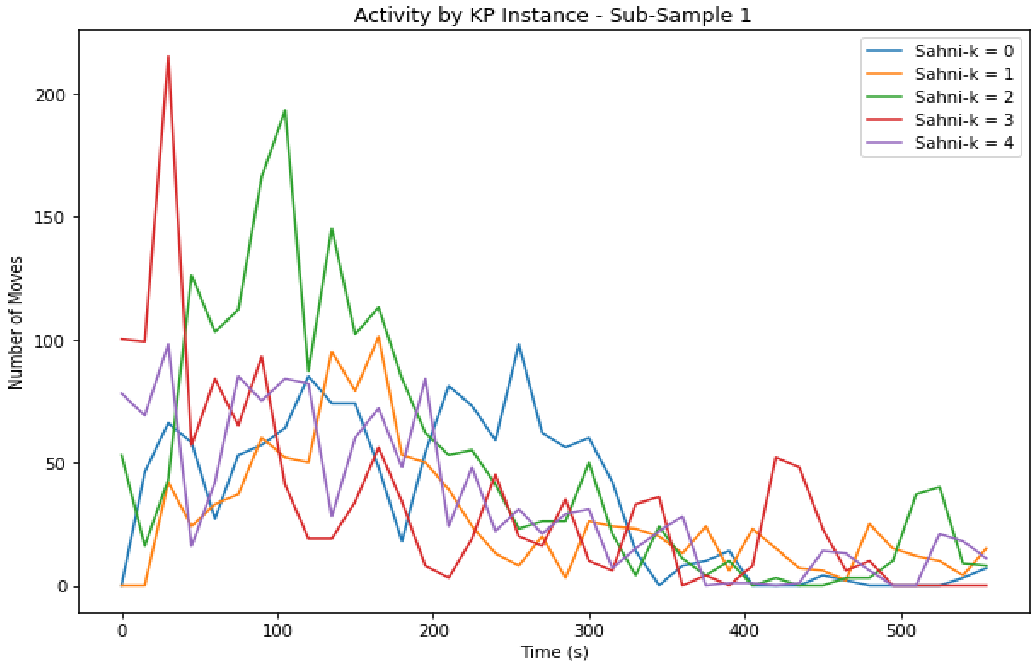
	Cash-Heavy Class	**Security-Heavy Class**
Cash	50	25
Securities	2	4

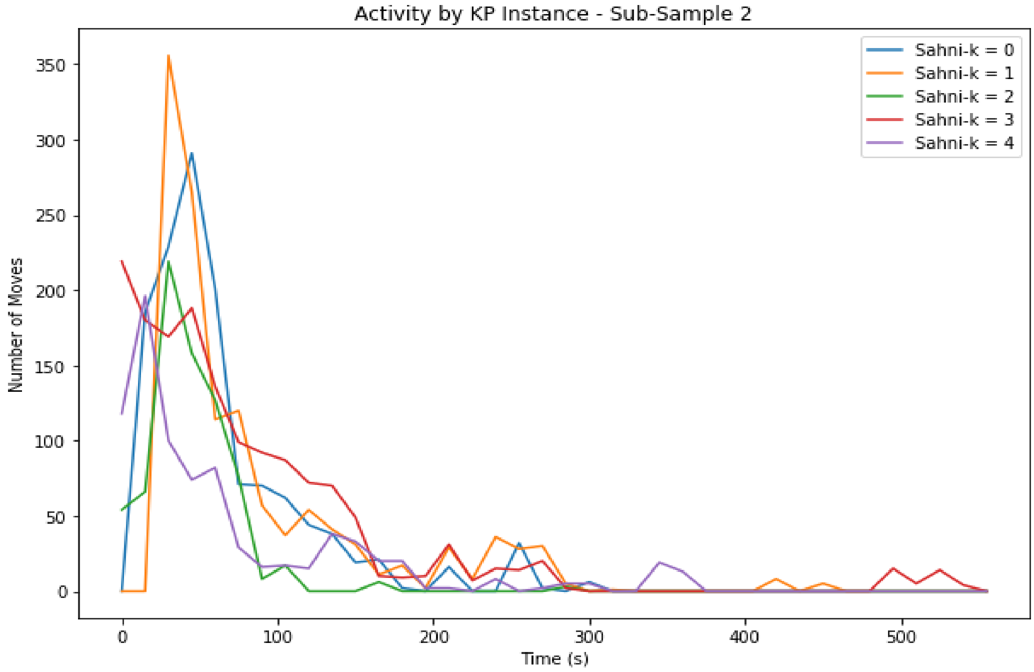
V. Methodology

V-a

Activity over Time, per instance

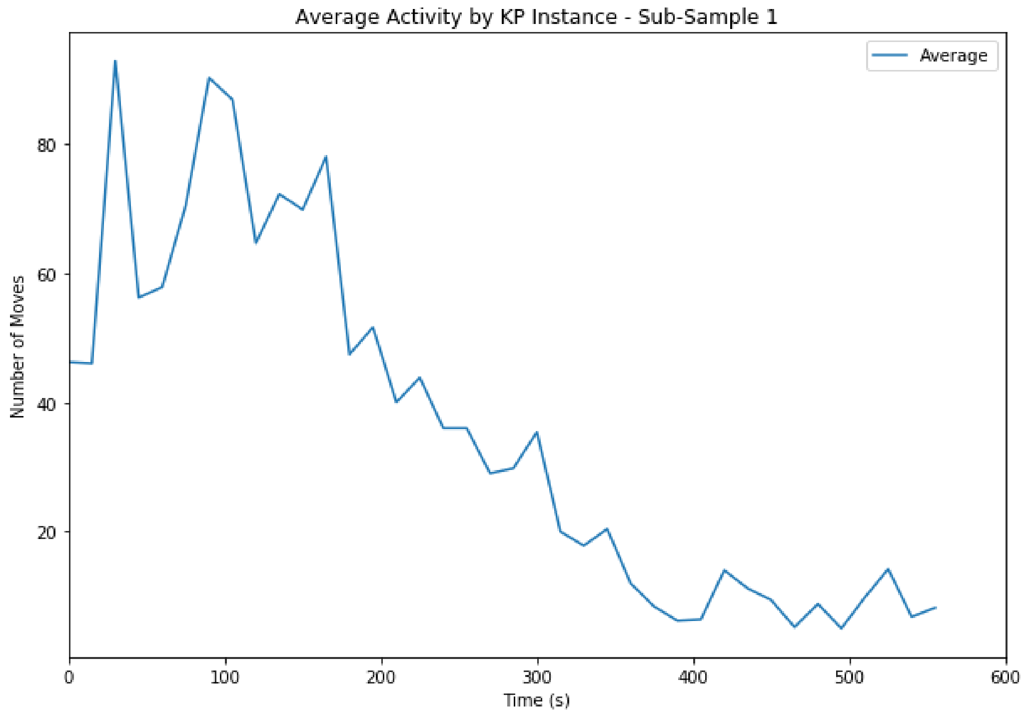
Below: Graphs of the *Activity* over time for sub-sample 1 and 2. Evident is the sharp, somewhat exponential decrease in *Activity* as time passes, in both samples and for all instances.

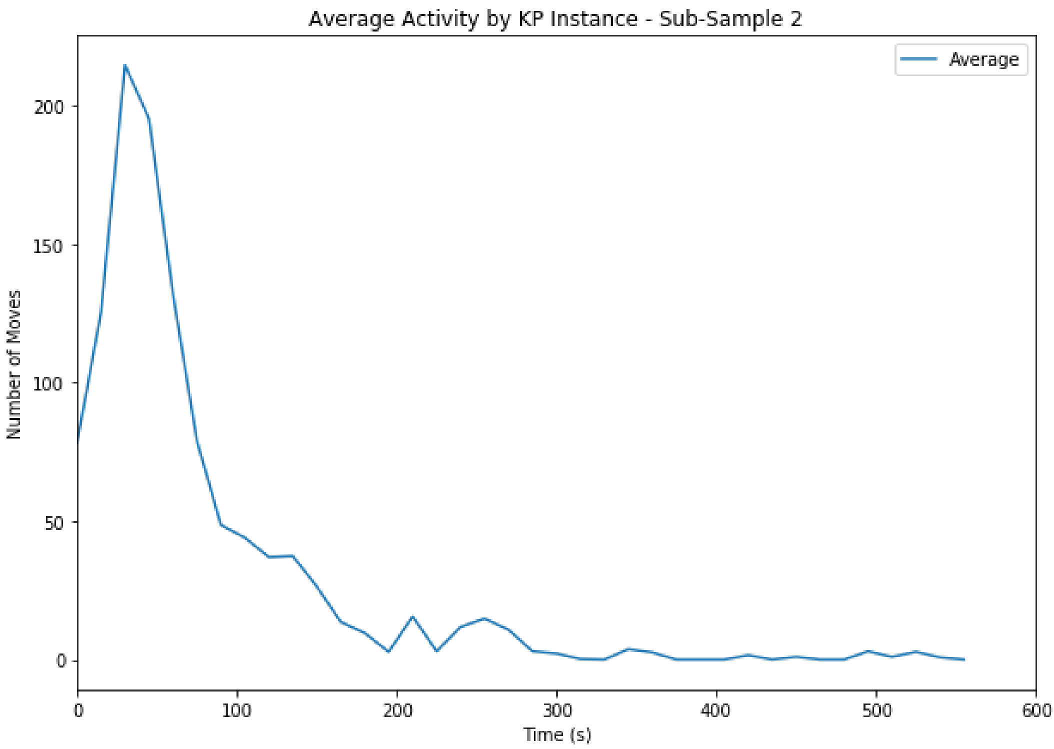




Average Activity over Time

Below: Graphs of the average *Activity* over time for all instances for sub-sample 1 and 2. Evident is the sharp, somewhat exponential decrease in *Activity* as time passes, in both samples and for all instances.



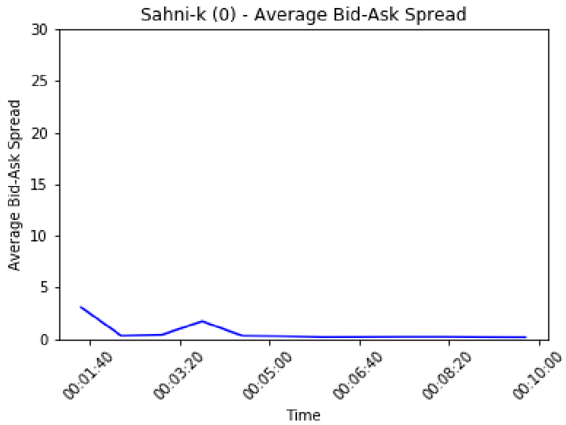


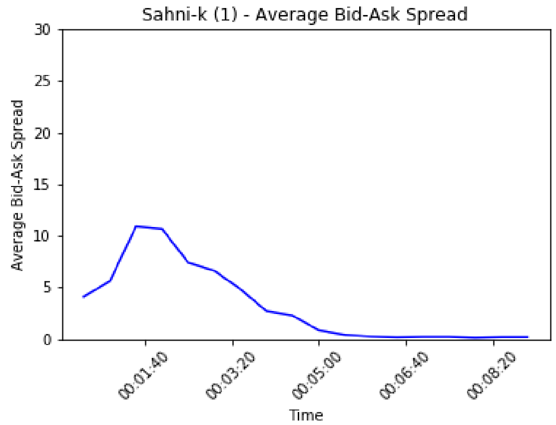
VI. Results

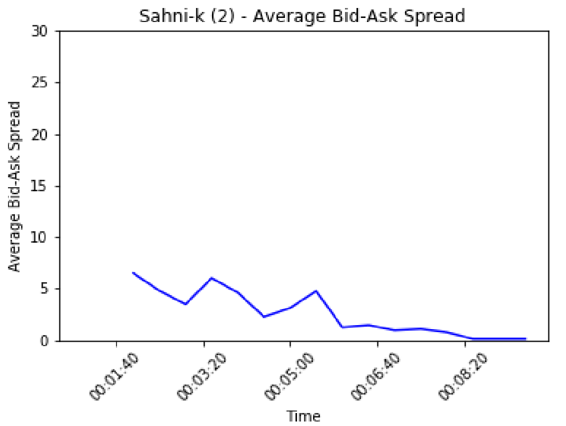
VI-a

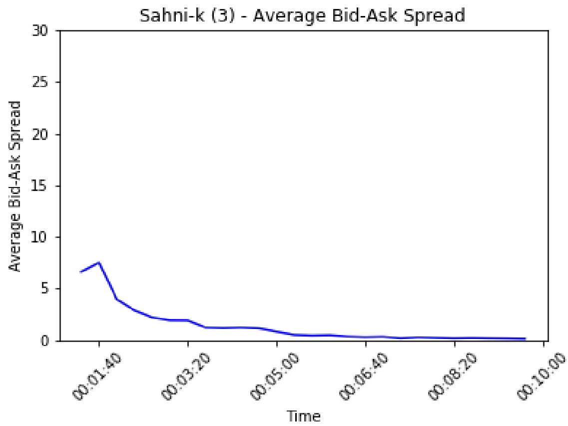
Average Bid-Ask Spread

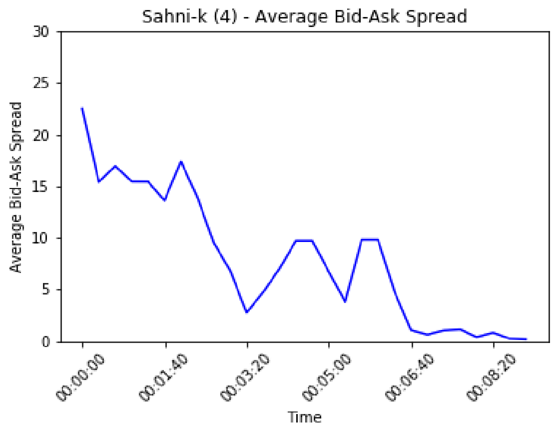
Below: Graphs of the average Bid-Ask spread per instance. Averages were taken across windows of 30 seconds. Evident is the convergence pattern exhibited by all markets – the bid-ask spread tends to zero for all instances, indicating that the market converges to a single price.











VI-b

Supplementary Regression Results: Price Information and Knapsack Performance

This table reports the GLME Gaussian regression results for equation (3). For each participant *i,* we estimate equation (3) using data from all trading rounds, divided into 15 second intervals. ΔKP Performance is the first difference of the current running maximum knapsack value divided by the optimal value of the relevant KP instance. ΔAskprice is the first difference of the average highest market ask across that interval. ΔTradeprice is the first difference of the average traded price across that interval. Complexity (1), (2), (3) and (4) are dummies for the complexity level of an instance – Complexity (0) forms the reference point and is thus not explicitly reported. BTA is the measure of better-than-average overconfidence. CBO is the measure of calibration-based overconfidence. The last five rows report model selection criteria (AIC, BIC and log-likelihood), the number of observations in the dataset and the fixed effects used in the regression. Robust standard errors are reported in parentheses below the coefficient estimates. Coefficient significance is reported by the following symbols: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001.

	Dependent Variable: ΔKP $Performance_t$	
$\Delta Askprice_t$	0.001^{*}	
	(0.0004)	
$\Delta Askprice_{t-1}$	0.001^{*}	
	(0.0004)	
$\Delta Askprice_{t-2}$	0.001^{*}	
	(0.0004)	
$\Delta Tradeprice_t$		0.002^{***}
		(0.001)
$\Delta Tradeprice_{t-1}$		0.001^{**}
		(0.001)
$\Delta Tradeprice_{t-2}$		0.001^{**}
		(0.0005)
Complexity (1)	-0.004^{**}	-0.001
	(0.002)	(0.001)
Complexity (2)	-0.004^{*}	0.0001
	(0.002)	(0.001)
Complexity (3)	-0.006^{***}	-0.001
	(0.002)	(0.001)
Complexity (4)	-0.003^{+}	0.0002
	(0.002)	(0.001)
BTA	0.00003	0.00001
	(0.00005)	(0.00004)
CBO	-0.003	-0.001
	(0.004)	(0.003)
Constant	0.008^{**}	0.003
	(0.002)	(0.002)
Participant Fixed Effects	Yes	Yes
Observations	7,592	6,801
Log Likelihood	13,340	14,318

Akaike Inf. Crit.	-26,656	-28,612
Bayesian Inf. Crit.	-26,573	-28,531

VI-c

Supplementary Regression Results: Price Information and Effort in the Knapsack (15s Interval)

This table reports the GLME Poisson regression results for equation (4). For each participant *i* in each sample, we estimate equation (4) using data from all trading rounds, divided into 15 second intervals. Activity is the number of items moved in or out of the knapsack in the time interval. ΔRelative Performance is the first difference of the running maximum knapsack value, transformed into its corresponding payoff, divided by the market price. Complexity (1), (2), (3) and (4) are dummies for the complexity level of an instance – Complexity (0) forms the reference point and is thus not explicitly reported. BTA is the measure of better-than-average overconfidence. CBO is the measure of calibration-based overconfidence. Time is the time interval variable, which controls for an exponential time trend in the data. The last five rows report model selection criteria (AIC, BIC and log-likelihood), the number of observations in the dataset and the fixed effects used in the regression. Robust standard errors are reported in parentheses below the coefficient estimates. Coefficient significance is reported by the following symbols: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001.

	Dependent Variable: Activity_t	
	Sub-Sample 1 (Knapsack worse than market)	*Sub-Sample 2* (Knapsack better than market)
$\Delta\text{Relative Performance}_t$	0.394^{***}	0.006^{+}
	(0.070)	(0.003)
$\Delta\text{Relative Performance}_{t-1}$	0.322^{***}	0.001
	(0.081)	(0.002)
$\Delta\text{Relative Performance}_{t-2}$	-0.065	-0.005^{+}
	(0.087)	(0.003)
Complexity (1)	0.079	-0.479^{***}
	(0.065)	(0.050)
Complexity (2)	-0.267^{***}	-0.160^{**}
	(0.055)	(0.055)
Complexity (3)	0.016	-0.323^{***}
	(0.069)	(0.044)
Complexity (4)	-0.406^{***}	-0.185^{***}
	(0.054)	(0.036)
BTA	0.010	-0.002
	(0.009)	(0.007)
CBO	-1.678^{*}	-0.046
	(0.697)	(0.534)
Time	-0.096^{***}	-0.197^{***}
	(0.002)	(0.002)
Constant	2.309^{***}	3.305^{***}
	(0.406)	(0.319)
Participant Fixed Effects	Yes	Yes
Observations	3,004	4,283
Log Likelihood	-6,333	-9,485

Akaike Inf. Crit.	12,691	18,993
Bayesian Inf. Crit.	12,763	19,070

Supplementary Regression Results: Price Information and Effort in the Knapsack (20s Interval)

This table reports the regression results for equation (4). For each participant *i* in each sample, we estimate equation (4) using data from all trading rounds, divided into 20 second intervals. Activity is the number of items moved in or out of the knapsack in the time interval. ΔRelative Performance is the first difference of the running maximum knapsack value, transformed into its corresponding payoff, divided by the market price. Complexity (1), (2), (3) and (4) are dummies for the complexity level of an instance – Complexity (0) forms the reference point and is thus not explicitly reported. BTA is the measure of better-than-average overconfidence. CBO is the measure of calibration-based overconfidence. Time is the time interval variable, which controls for a time trend in the data. The last five rows report model selection criteria (AIC, BIC and log-likelihood), the number of observations in the dataset and the fixed effects used in the regression. Robust standard errors are reported in parentheses below the coefficient estimates. Coefficient significance is reported by the following symbols: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001.

	Dependent Variable: Activity_t	
	Sub-Sample 1 (Knapsack worse than market)	*Sub-Sample 2* (Knapsack better than market)
ΔRelative Performance_t	0.425^{***}	0.021^{*}
	(0.071)	(0.009)
ΔRelative Performance_{t-1}	-0.054	0.005
	(0.079)	(0.005)
ΔRelative Performance_{t-2}	0.058	0.011^{*}
	(0.086)	(0.005)
Complexity (1)	0.050	-0.550^{***}
	(0.072)	(0.061)
Complexity (2)	-0.315^{***}	-0.397^{***}
	(0.060)	(0.066)
Complexity (3)	0.090	-0.391^{***}
	(0.076)	(0.053)
Complexity (4)	-0.600^{***}	-0.251^{***}
	(0.060)	(0.039)
BTA	0.011	-0.006
	(0.015)	(0.008)
CBO	-2.923^{*}	-0.173
	(1.151)	(0.613)
Time	-0.131^{***}	-0.263^{***}
	(0.003)	(0.003)
Constant	2.589^{***}	3.788^{***}
	(0.675)	(0.366)
Participant Fixed Effects	Yes	Yes
Observations	2,190	2,996
Log Likelihood	-4,953.521	-7,591.067
Akaike Inf. Crit.	9,931.041	15,206.130
Bayesian Inf. Crit.	9,999.341	15,278.190

VI-d

Supplementary Regression Results: Price Information and Buying Behaviour (15s Interval)

This table reports the GLME Poisson regression results for equation (7). For each participant *i* in each sample, we estimate equation (7) using data from all trading rounds, divided into 15 second intervals. Bidcount is the number of buy orders submitted in the time interval. ΔRelative Performance is the first difference of the running maximum knapsack value, transformed into its corresponding payoff, divided by the market price. Complexity (1), (2), (3) and (4) are dummies for the complexity level of an instance – Complexity (0) forms the reference point and is thus not explicitly reported. BTA is the measure of better-than-average overconfidence. CBO is the measure of calibration-based overconfidence. The last five rows report model selection criteria (AIC, BIC and log-likelihood), the number of observations in the dataset and the fixed effects used in the regression. Robust standard errors are reported in parentheses below the coefficient estimates. Coefficient significance is reported by the following symbols: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001.

	Dependent Variable: Bidcount_t *Sub-Sample 2* (Knapsack better than market)
ΔRelative Performance_t	-0.010
	(0.019)
ΔRelative Performance_{t-1}	-0.003
	(0.017)
ΔRelative Performance_{t-2}	-0.010
	(0.016)
Complexity (1)	-0.082
	(0.147)
Complexity (2)	-0.089
	(0.200)
Complexity (3)	0.211
	(0.141)
Complexity (4)	0.366^{**}
	(0.140)
BTA	-0.003
	(0.004)
CBO	-0.179
	(0.293)
Constant	-2.052^{***}
	(0.186)
Participant Fixed Effects	Yes
Observations	4,283
Log Likelihood	-1,684.992
Akaike Inf. Crit.	3,391.984
Bayesian Inf. Crit.	3,461.963

Supplementary Regression Results: Price Information and Selling Behaviour (15s Interval)

This table reports the GLME Poisson regression results for equation (7). For each participant *i* in each sample, we estimate equation (7) using data from all trading rounds, divided into 15 second intervals. Askcount is the number of sell orders submitted in the time interval. ΔRelative Performance is the first difference of the running maximum knapsack value, transformed into its corresponding payoff, divided by the market price. Complexity (1), (2), (3) and (4) are dummies for the complexity level of an instance – Complexity (0) forms the reference point and is thus not explicitly reported. BTA is the measure of better-than-average overconfidence. CBO is the measure of calibration-based overconfidence. The last five rows report model selection criteria (AIC, BIC and log-likelihood), the number of observations in the dataset and the fixed effects used in the regression. Robust standard errors are reported in parentheses below the coefficient estimates. Coefficient significance is reported by the following symbols: + p<0.1; * p<0.05; ** p<0.01; *** p<0.001.

	Dependent Variable: Askcount_t *Sub-Sample 1* (Knapsack worse than market)
$\Delta\text{Relative Performance}_t$	-0.616
	(0.522)
$\Delta\text{Relative Performance}_{t-1}$	0.395
	(0.453)
$\Delta\text{Relative Performance}_{t-2}$	-0.341
	(0.481)
Complexity (1)	0.216
	(0.226)
Complexity (2)	0.214
	(0.198)
Complexity (3)	0.297
	(0.248)
Complexity (4)	0.313
	(0.210)
BTA	0.001
	(0.005)
CBO	-0.036
	(0.416)
Constant	-0.616
	(0.522)
Participant Fixed Effects	Yes
Observations	3,004
Log Likelihood	-1,073.438
Akaike Inf. Crit.	2,168.876
Bayesian Inf. Crit.	2,234.850