Under Review

Asset Pricing under Computational Complexity

Peter Bossaerts, Elizabeth Bowman, Shijie Huang, Carsten Murawski, Shireen Tang and Nitin Yadav

August 28, 2018

ASSET PRICING UNDER COMPUTATIONAL COMPLEXITY

Peter Bossaerts, Elizabeth Bowman, Shijie Huang, Carsten Murawski, Shireen Tang and Nitin Yadav

We study asset pricing in a setting where correct valuation of securities requires market participants to solve instances of the 0-1 knapsack problem, a computationally "hard" problem. We identify a tension between the need for absence of arbitrage at the market level and individuals' tendency to use methodical approaches to solve the knapsack problem that are effective but easily violate the sure-thing principle. We formulate hypotheses that resolve this tension. We demonstrate experimentally that: (i) Market prices only reveal inferior solutions; (ii) Valuations of the average trader are better than those reflected in market prices; (iii) Those with better valuations earn more through trading; (iv) Price quality is unaffected by the size of the search space, but decreases with instance complexity as measured in the theory of computation; (v) Participants learn from market prices, in interaction with trading volume.

Keywords: Efficient Markets Hypothesis, Computational Complexity, Financial Markets, Noisy Rational Expectations Equilibrium, Absence of Arbitrage.

All authors are of the Brain, Mind & Markets Laboratory, Department of Finance, Faculty of Business and Economics, The University of Melbourne, Victoria 3010, Australia. Corresponding author: Peter Bossaerts; email: peter.bossaerts@unimelb.edu.au; phone: +61.3.3095.3257.

date: August 28, 2018

"An intellect which at a certain moment
would know all forces that set nature in motion,
and all positions of all items of which nature is composed,
if this intellect were also vast enough to submit these data to analysis [...];
for such an intellect nothing would be uncertain."

Pierre-Simon de Laplace, A Philosophical Essay on Probabilities, 1814

1. INTRODUCTION

The Efficient Markets Hypothesis (EMH) is one of the foundations of modern asset pricing theory. Markets are called "efficient" when "security prices fully reflect all available information" (Fama, 1991, p. 1575). Theoretical analyses of this statement usually envisage a setting where the value of securities is a function of information that is distributed in the economy, and correct valuation requires aggregation of the different bits of information (Grossman and Stiglitz, 1976).[1] As such, uncertainty can be resolved efficiently through *sampling*, because a law of large numbers holds: one could start with a randomly chosen agent, collect her information, then visit additional agents, and add their information; as more agents are sampled, one obtains a gradually improving estimate of the true value of the security at hand, eventually reaching perfect precision (see, e.g., Grossman (1976, 1978); Hellwig (1980); Diamond and Verrecchia (1981)).

Here, we consider a fundamentally different setting. Uncertainty emerges because of computational complexity, instead of a lack of information. Specifically, we envisage a situation where correct valuation of securities require market participants to combine publicly available information *in the right way*. As such, valuation is a combinatorial problem that is *computationally complex*.[2]

[1]The market microstructure literature sometimes takes a different approach, where better-informed compete on the basis of the *same* information (see, e.g., Holden and Subrahmanyam (1992)). There, the issue is whether prices eventually reveal their information, and how (see Bossaerts et al. (2014) for an experimental investigation.)

[2]We follow computer science convention and define computational complexity as the amount of computational resources required to solve a given problem (Arora et al., 2010).

To put this differently: uncertainty about the value of a security is the consequence of not knowing what information is pertinent, rather than not having the information in the first place. We would claim that this characterization of uncertainty is germane to the "fundamentalist" approach to finance: the value of equity, for instance, requires one to think about what projects belong to a firm, how much and what types of debt it should issue, where it should operate, etc. But even in standard quantitative analysis, our characterization of uncertainty has become more relevant recently: many market participants have access to large data sets, so the issue is not unavailability of data; instead, the problem is that of not knowing which data are relevant.

We formalize this as follows. We envisage a situation where the values of securities depend on the solution of instances of the 0-1 knapsack problem (KP), a combinatorial optimization problem (Kellerer et al., 2004). In the KP, the decision-maker is asked to find the sub-set of items of different values and weights that maximizes the total value of the knapsack (KS), subject to a weight constraint. The KP is computationally hard. There is no known algorithm that both finds the solution and is efficient, that is, can compute the solution in polynomial time.[3] Intuitively, it is "easy" to verify that a particular set of items achieves a given total value, but it is "hard"

Note that recent analyses of "complexity" in finance and economics either take a less formal approach (e.g., Skreta and Veldkamp, 2009), or model complexity in probabilistic terms (e.g., Spiegler, 2016). Problems that require resources (usually in terms of time) beyond those available are referred to as "computationally intractable" (Cook, 1983).

[3] In computational complexity theory, an algorithm is called "efficient" if the rate at which the amount of time to compute the solution grows as the size of the computational problem increases, is upper-bounded by a polynomial. The KP can be solved with dynamic programming. This algorithm is polynomial in the number of items; however, the memory required to implement dynamic programming grows exponentially. Therefore, dynamic programming substitutes exponentially growing time to address memory for exponentially growing time to compute. Experiments demonstrate that humans do not appear to use dynamic programming when solving the KP, which is not surprising: human working memory is very small (Murawski and Bossaerts, 2016).

to find the set of items with the highest total value.[4]

One could argue that there is really no uncertainty here, since all the information (knapsack capacity, items, item values and weights) is publicly available. There is a long tradition in decision theory, however, to insist that situations like this represent uncertainty. De Finetti, for instance, defines an event as a decision problem, i.e., a question that can be answered in only two ways: yes (true) or no (false). Uncertainty then emerges when the answer cannot be determined purely using logical inference (De Finetti, 1995). The question is then reduced to: is it appropriate to use the tools of probability theory in order to find the right answer? Can we pretend that a law of large numbers holds, sample, update one's prior belief (about the answer) using Bayes' law, and thereby reduce uncertainty? Some, like Kolmogorov, would claim that the answer is no (Kolmogorov, 1983).

For Savage (Savage, 1972), the answer is yes. As long as the decision-maker exhibits choices that satisfy a minimal number of rationality restrictions (such as the sure-thing principle), her actions can be modeled *as if* she has a utility profile over outcome quantities, and beliefs that she updates using Bayes' law. Whether she learns effectively is not an issue; she chooses coherently, and remains coherent as she samples. See Bossaerts et al. (2018).

To be sure, in the context of the KP, randomly trying out knapsacks and learning using Bayes' law is extremely ineffective.[5] We shall illustrate this with an example later on, where an agent who adheres to Savage's principles would find the optimum with 95% chance only after 400 trials, in a problem

[4]Technically, the version of the KP used here is the optimization version of the problem. This problem is NP-hard. The corresponding decision version of the problem is NP-complete. The optimisation version is at least as hard as the corresponding decision version.

[5]Sampling does not have to be random, but the probabilistic law generating outcomes has to be known, as in e.g., importance sampling (Hastings, 1970), for otherwise Bayes' law cannot be applied. But one cannot know the probabilistic law that generates knapsack values given the optima without having solved the KP in the first place, but then learning is pointless.

where there are merely 82 possible capacity-filled knapsacks... It would be better to just list all the possible knapsacks and pick the best one.

Recent research has demonstrated that humans follow effective, methodical approaches when solving the KP. Murawski and Bossaerts (2016) discovered that humans do not sample knapsacks, and instead follow strategies that are reminiscent of the algorithms that computer scientists use to solve such problems (Kellerer et al., 2004). While not always guaranteeing a solution, these algorithms are far more efficient than random sampling. One such algorithm is the *greedy algorithm*, whereby items are put into the knapsack in decreasing order of the value-over-weight ratio, until reaching capacity. As a result, humans perform in the KP as do computers: effort required increases, and resulting performance decreases, as the objective computational complexity of the instance increases.[6]

But these strategies easily violate Savage's axioms. The greedy algorithm, for instance, is inconsistent with the sure-thing principle, and it is simple to show why. Consider two KPs with the same capacity, namely, 2. One KP has three items with value and weight pairs (2,1), (3.5,2) and (1,1), while the second one has only the last two items. Which would the decision-maker choose if she is to maximize value? The greedy algorithm would have her choose the second one, since it reaches a value of 3.5, while the first one only generates a value of 3. Yet everything she can obtain with the second one she can also get with the first one since the items in the first problem are in a superset of the second one.

When convenience is called for, finance scholars have no issue with violations of the Savage axioms, and hence, expected utility. Mean-variance optimization, for instance, violates Savage's axioms; like the greedy agorithm,

[6]Evidence is emerging that in simpler computational problems, such as that of determining the least costly fund investment under a variety of fee structures, humans do not follow methodical approaches, but instead learn through trial-and-error. See Anufriev et al. (2018).

it too violates the sure-thing principle. Yet mean-variance optimization is effective when deriving optimal portfolios of many securities. It would be cumbersome to do the same with preference profiles that satisfy Savage's axioms, such as expected log utility. So, finance scholars use mean-variance optimization, opting for efficacy over rationality.

At the market level, however, violations of the Savage axioms would be objectionable. Indeed, to violate the sure-thing principle is tantamount to allowing for arbitrage opportunities. The evolution of market prices should be free of arbitrage opportunities; price updates should be as if using Bayes' law. This is the Fundamental Theorem of Asset Pricing, of course (Dybvig and Ross, 2003). As such, we must not expect market prices to reflect the effective solution algorithms that individuals appear to follow.

This fundamental tension between how markets should behave and how individuals in those markets do behave was already pointed out in Dybvig and Ross (2003) in the context of violations of expected utility related to loss and ambiguity aversion. Experiments have illuminated how individuals can violate rationality assumptions yet market prices do not allow for arbitrage opportunities; see, e.g., Bossaerts et al. (2010); Asparouhova et al. (2015). The conclusion is simple: markets must not be modeled after the behavior of their participants.

What could be the consequences of a tension between a market that does not allow for arbitrage opportunities at the cost of learning effectiveness, and individuals whose actions may violate the sure-thing principle, but makes them learn faster? Here, we hypothesized that, in a market where valuation depends on the solution of a 0-1 KS problem: (i) Market prices can only reveal inferior solutions, (ii) Individual traders will, on average, do better than the market, (iii) Individuals who find the solution earn more through trading.

We test these predictions with a laboratory experiment, organized as follows. We endowed participants with shares of several securities (10 or 12), each of which corresponded to an item in a given instance of the KP. All securities lived for a single period, after which they paid a liquidating dividend. The dividend equalled one dollar if the corresponding item was in the solution of the instance of the KP; and zero otherwise. This makes the securities akin to Arrow securities, were it not that multiple securities could pay off at once, since in general more than one item could be in the optimal knapsack. After markets opened, participants traded the securities in a computerised continuous open-book system (a version of the continuous double auction where infra-marginal orders are kept in the system until cancelled).[7] All participants were provided with the same information about the instance to be solved, that is, each item's value and weight, and the total capacity of the knapsack. Participants had access to a computer program where they could try out candidate solutions. This also meant that we could track their solution attempts.[8]

There exists an asset pricing paradigm that is consistent with Savage's axioms yet does make the the predictions that prices will not fully reveal the right information, and where the average trader performs better than the market. This is the Noisy Rational Expectations Equilibrium (NREE) pioneered in Grossman (1977). There, noise is usually modeled as uncertainty over aggregate supplies (Grossman and Stiglitz, 1976), or uncertainty about the demand from "irrational" traders (Long et al., 1990), or randomness in non-traded risk (Biais et al., 2003). In our setting, one can think of noise as uncertainty generated by computational complexity. As such, we hypothesized that price quality would decrease in the level of noise, measured in terms of instance complexity.

We measure instance complexity using the *Sahni-k* metric. This metric

[7]We used the software developed by Adhoc Markets (http://www.adhocmarkets.com).

[8]The program is part of the ULEEF GAMES suite and and can be accessed at http://uleef.business.utah.edu/games.

increases in the number of items that have to be put in the knapsack before the greedy algorithm can be used to complete the knapsack and obtain the optimal solution. Murawski and Bossaerts (2016) showed that Sahni-k explains individual performance. In choosing knapsack instances, we made sure that Sahni-k was unrelated to the number of possible filled knapsacks, so that reduction of price quality would not merely the result of an increase in the size of the search space. As a matter of fact, it would be most interesting to discover that price quality does not decrease with the number of possible full knapsacks.[9]

In NREE, individuals manage to read prices, despite the inferior quality of the information revealed in those prices (compared to the average trader's). We wondered whether in our setting too, traders would improve their solutions because of the availability of prices, and if so, how. Because we tracked individuals' attempts at solving the KS instances at hand, we were in the position to shed light on the channels through which prices make individuals revise their solutions. In earlier work, we had shown that more participants find the optimal solution in the presence of markets than if they were asked to solve the same instances individually (Meloso et al., 2009a).

We find broad support for all our hypotheses: (i) Market prices reveal little information about the optimal solution; in fact, they behave as if a representative agent is merely randomly sampling knapsacks; (ii) The average trader performs better than the market; (iii) Individuals who are close to the solution earn more through trading; (iv) Market (and individual) performance decreases in instance complexity, or Sahni-k, and not in the size of the search space; (v) Individuals manage to improve their solutions by reacting to high volume in low-priced securities.

[9]In the classical paradigm of valuation through aggregation of disparate information, asset pricing theory predicts that increase in the number of possible values does not impede full revelation, as long as this number is finite (Radner, 1979).

Our findings are in sharp contrast with those from experiments that build on the classical uncertainty paradigm. There, information bits are dispersed among participants, and correct valuation merely requires one to add the bits together. When security payoffs are common-knowledge (i.e., everyone knows how everyone else values the payoffs given aggregate information), controlled experiments with centralized, double auctions, have confirmed that markets manage to aggregate available information (Plott and Sunder, 1982, 1988; Plott, 2000).[10] The aggregation is so good that participants will refuse to pay for information (Sunder, 1992; Copeland and Friedman, 1992). That is, the Grossman-Stiglitz paradox emerges (Grossman and Stiglitz, 1976).

The remainder of the paper is organized as follows. In the next section, using a simple example inspired by our experimental setup, we illustrate how random sampling is far less effective in the 0-1 knapsack problem than when information merely needs to be aggregated. Section 3 presents the experimental paradigm. Section 4 reports the results. We then discuss the further implications in a Conclusion.

2. SAMPLING IN THE PRESENCE OF COMPUTATIONAL COMPLEXITY: AN ILLUSTRATION

Here, we illustrate how resolution of uncertainty is fundamentally different when uncertainty emerges because of lack of information *vs.* when it emerges because of computational complexity. We assume an agent who adheres to Savage's principles (Savage, 1972), and hence, assigns beliefs to the possible values, and gradually updates these based on random sampling and Bayes' law. We refer to this agent as the "Savage agent."

We first consider a situation of lack of information. There, one needs to

[10]The ability of *decentralized* markets to aggregate dispersed information is hotly disputed, with Wolinsky (1990) and Duffie and Manso (2007) taking opposite views. See Asparouhova and Bossaerts (2017) and Asparouhova et al. (2017) for experimental evidence.

gather information, and we assume, as in the classical paradigm, that a law of large numbers holds, so that information collection and subsequent averaging gradually lead one close to the truth. Imagine there are ten securities that pay a liquidating dividend of 1 or 0 dollars, with equal chance. Each period ("trial"), our agent receives a signal, drawn from a Bernoulli distribution with $p = 0.70$ if the final dividend is 1, and with $p = 0.30$ otherwise. The agent uses Bayes' law to update beliefs about the final payoff. Assuming a quadratic loss function, the agent's valuation is the posterior mean, starting from an unconditional estimate of 0.5. The top panel of Fig. 1 shows how valuations of the securities quickly separate between those that will end up paying 1 and those that expire worthless.

As such, repeated sampling gets one closer to the fundamental value: the chance that one deviates too far from knowing the truth decreases with sample size. As time progresses, the chance of one valuation to veer off (say, from close to 1 down to 0) is drastically reduced. That is, even for finite samples (finite number of signals), the valuation estimate is "probably approximately correct." Computer scientists would refer to this situation as one of *finite sample complexity* (Valiant, 1984), that is, a close approximation of the true value can be obtained with a finite number of samples.

Now consider a different problem. Here, all the information is available from the beginning, but uncertainty emerges because of computational complexity. Consider securities whose terminal values depend on the solution of an instance of the 0-1 KP (instance is #8 in Murawski and Bossaerts (2016)). There is a security for each available item. If the item is in the optimal solution, the security's terminal value equals 1; otherwise it expires worthless. As before, we assume that the agent assigns a prior value to each security, say 0.5. She then gradually adjusts beliefs by sampling, as follows. Each trial, the agent randomly tries a subset of the items that fills the KS to capacity. The agent then computes the value of the KS and compares to

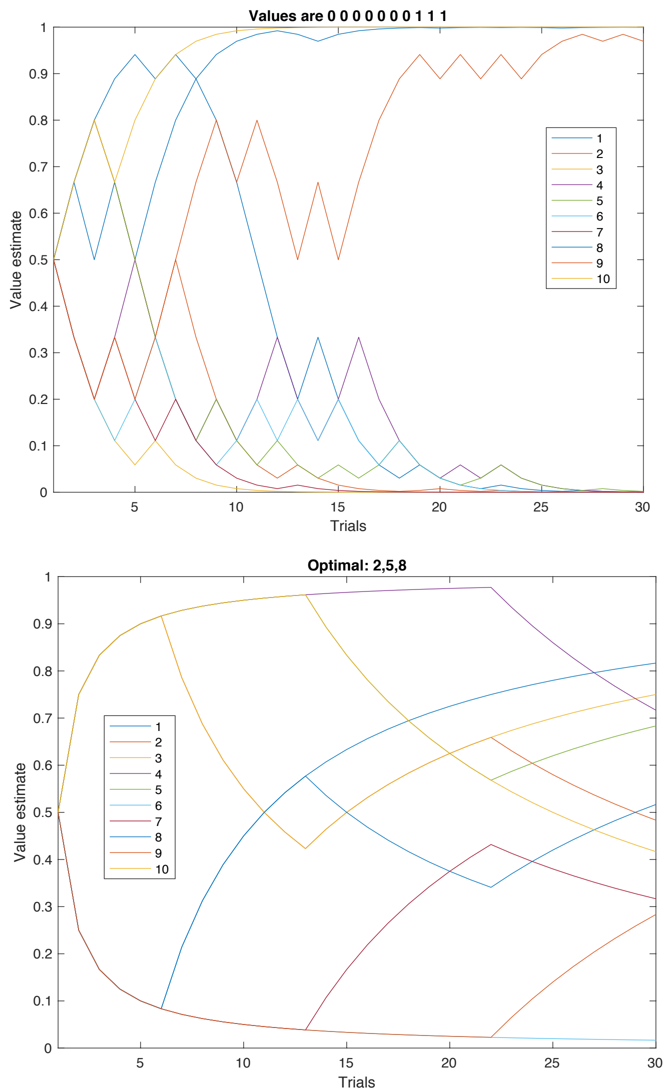


FIGURE 1.— **Simulations When Valuations Require Averageing (Top)** *vs.* **Combinatorics (Bottom).** Top: Evolution of valuations of ten securities where correct valuations require information aggregation; as sampling increases over time, the probability of being correct increases gradually. This is the situation in traditional theoretical analyses of EMH. Terminal values of the 10 securities are listed on top, in ascending order of security number. Bottom: Evolution of prices of ten securities in a situation where correct valuations require combining information in the right way; as sampling increases over time, the probability of being correct – even approximately – does not increase gradually. Only the securities with numbers listed on top ("Optimal") pay a liquidating dividend of 1; all others expire worthless. See text for details.

the maximum value obtained in previous trials. If the new KS value is less than the previous maximal value, then she constructs signals as follows. The signal for a security equals 1 when the corresponding item was in the earlier (better) solution, and 0 otherwise. If the new KS value is higher than the trailing maximum, then the signal for a security equals 1 when it is in the current solution; otherwise the signal is 0. The agent then updates security valuations in trial t by weighing the valuation in the previous trial $t-1$ by $(t-1)/t$ and the signal in the present trial by $1/t$. This may not sound like Bayesian because it does not use the true likelihood (of observing full knapsack values given the optimal solution). However, this is without consequence, because a Savage agent does not need to use the true likelihood: subjective beliefs are allowed. We merely posit that there exists a belief that is justified by the reinforcement learning rule.

The bottom panel of Fig. 1 illustrates the evolution of securities values under the above updating rule. Even after 30 trials, it is still unclear which item is in the optimal solution (which security will pay one dollar). Worse, valuations can erratically move from high to low and *vice versa* even in later trials. Since there are only a finite, albeit large, number of possible capacity-filled knapsacks, the algorithm will eventually find the optimal solution. This will happen when, just by chance, the best knapsack is drawn. But this takes time. In the present situation, there are (a mere) 82 possible capacity-filled knapsacks, so the chance of drawing the best knapsack in any trial equals 1/82. This implies, among others, that the chance that the algorithm finds the optimal knapsack within 30 trials is about 5%; it takes approximately four hundred trials to bring this chance up to 95%.

As such, assigning beliefs to the outcomes and learning by sampling, while very effective when the cost of processing information is low, is highly ineffective when the cost of processing information is high. In our example, it would have been better to just list all 82 possible capacity-filled knapsacks and pick the optimum. This simple, deterministic procedure, finds the optimum for sure in 82 steps, while the sampling approach may not find it in 5% of the cases even after a much as *four hundred* trials. Of course, the deterministic approach quickly becomes ineffective too. In some of the KP instances we will consider next, there are up to 400 capacity-filled knapsacks.

3. EXPERIMENTAL DESIGN

Participants

Participants were recruited from the University of Melbourne community, in four experimental sessions with 18 (one session) or 20 (three sessions) participants. To be eligible, participants had to be current students of the University of Melbourne aged between 18 to 30 years old with normal or corrected-to-normal vision. The final sample included a total of 78 participants (age range: 18 to 26 years, mean age = 22, standard deviation = 4, gender: 44 male, 34 female). The study was approved by the University of Melbourne Human Research Ethics Committee (Ethics ID: 1647059.1) and was conducted in accordance with the World Medical Association Declaration of Helsinki. All participants provided written informed consent.

Task

Participants attempted five instances of the 0-1 knapsack problem, while simultaneously trading in an online marketplace. In each instance, participants selected items with given values and weights from a set, in order to maximize the total combined value within the weight constraint for the selection. Formally, participants were asked to solve the following maximization problem:

$$\max_x \sum_{i=1}^{I} x_i v_i \;\text{ s.t. } \sum_{i=1}^{I} x_i w_i \le C \text{ and } x_i \in \{0,1\}\, \forall i,$$

where i, w, v and C denote the item number, item weight, item value and knapsack capacity, respectively. The number of items in an instance varied between 10 and 12. Instances were taken from two prior studies (Meloso et al., 2009b; Murawski and Bossaerts, 2016). The sequence of instances in an experimental session was counterbalanced across sessions. Parameters for the five instances are provided in Table I.

The KP instances were made available electronically on a computer interface where participants could try out different solutions.[11] The software recorded every move of an item into and out of the knapsack. Importantly, the software did not indicate if a candidate solution was optimal.

Each item in an instance mapped into a security in the online marketplace. Therefore, between 10 to 12 markets were available in each instance. Exchange was organized using the continuous double-sided open book system, like most purely electronic stock markets globally. Trading was done on the online experimental markets platform Flex-E-Markets.[12] Participants traded for 15 minutes, or in later rounds, less. The user interfaces of the two systems, knapsack solver and online marketplace, are shown in Fig. 2.

Sessions started with a brief motivation and a reading of the instructions, after which participants were given ample time to familiarise themselves with knapsack solver, online marketplace, and how to exploit, through trading, knowledge acquired when attempting to solve a practice instance. Motivation, instructions and practice took a total of one hour. After a break, we ran the (five) rounds that counted for final earnings. In total, a session

[11] The application is part of a game suite called ULEEF GAMES; it can be accessed at http://uleef.business.utah.edu/games.

[12] See http://www.flexemarkets.com.

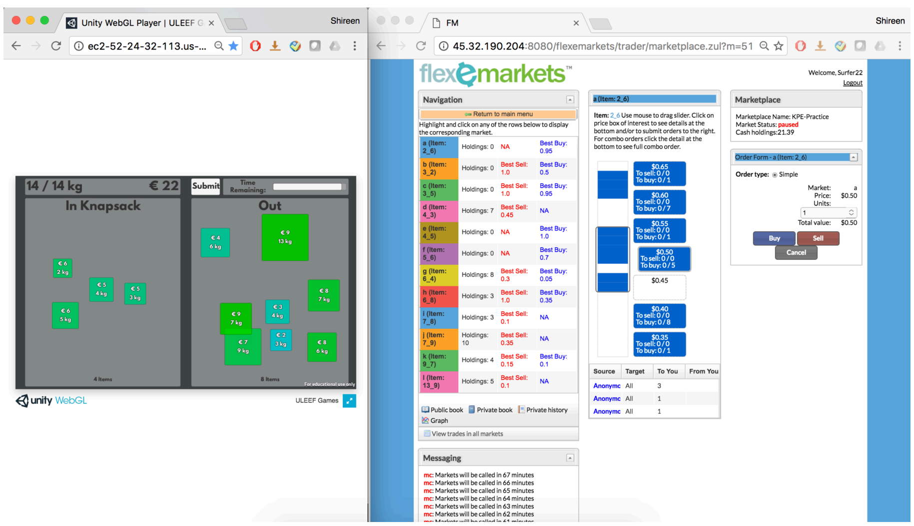


FIGURE 2.— Knapsack Problem Interface and Online Trading Platform

Left: Knapsack solver, where 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 KS value were displayed on top (left). Right: Online marketplace. Each of the 12 colour-coded markets corresponded to an item in the KS, and promised to pay one dollar if the item was in the optimal solution. The book of limit orders was shown in the middle, listed by price level, and colour-coded (blue: bids; red: asks). To the right is the order form, where participants could submit buy and sell limit orders, or cancel previous orders.

took between two and two-and-a-half hours.

Participant instructions including the timeline of a typical experimental session can be found in the Supplementary Online Material (SOM).

We recorded every order and trade in the marketplace, as well as every move into or out of the knapsack from each participant. Timestamps were synchronized between the knapsack solver and the online marketplace.[13]

Participant incentives

Participants took positions in the items they believed to be in the optimal knapsack by buying shares of the corresponding security. They could also sell shares corresponding to items they believed not to be in the solution; short sales were not permitted, however. In every instance, participants were endowed with $25 in cash holdings (Australian dollars, approximately eighteen U.S. dollars), and 12 shares randomly allocated to securities. The price range of a share was bounded between $0 and $1.

Final earnings consisted of: (i) liquidating dividends for the shares held at market close; (ii) any change in cash holdings between the beginning and end of trading. Earnings were cumulative across instances. Additionally, participants received a fixed reward ($2) for submitting a proposed solution through the knapsack solver, as well as a show-up fee of $5.

Initial Allocations

We designed initial allocations of securities to induce trade, by concentrating individual endowments in particular markets. While initial allocations were randomized, they were "fair" in the sense that all participants received the same number of shares in correct items across the five instances. Ini-

[13]The marketplace server became unreachable during the second round in the final session, so we have no trade data for that round. This denial of service originated with the server service provider, and hence was beyond our control.

tial allocations were such that \$31.20 in liquidating dividends were paid per participant on average. Participants were not told that they had "fair" initial allocations. We imposed fairness in the belief that earnings would suffer from the Hirshleifer Effect if prices were to fully reveal all available information (see Discussion for more details).

Although the concepts of "risk" and "risk aversion" have yet to be defined precisely in the context of computational complexity, we intuited that uncertainty about the solution to a KP instance would induce participants to diversify holdings across multiple securities. As a result, participants would trade not only because of perception of superior information. We eliminated aggregate risk by ensuring that there were an equal numbers of shares across securities. This was meant to avoid price distortions that could arise from differences in relative supplies.

Additionally, payment for submission of a solution through the KP interface was fixed and independent of whether the submission was correct. This made it impossible for participants to hedge between trading in the marketplace and submission of solutions through the KP interface.

Expert Traders

All participants were given the same information about the KP instance in a session at the beginning of each session. Thus, there was no information heterogeneity and no information asymmetry. However, we did not provide participants with the solutions, and since our instances were hard while performance was variable (the proportion of participants who submitted the correct solution varied between 6.4% and 60.3%; see SOM for details), heterogeneity (and asymmetry) arose spontaneously as participants started to search for the correct solution. Because all traders were given the same information, we cannot really talk about "informed" and "uninformed" traders when referring to those who found the optimal solution and those who failed to find it. So, for the purposes of our study, we define the former as an "Expert Trader." However, we re-emphasize that even expert traders may not have been aware that they knew the solution. Such is the nature of the KP...

4. RESULTS

Descriptive Statistics

Each of the 78 participants solved five instances of the KP (390 attempts in total). We first looked at computational performance of participants, that is, participants' ability to find optimal solutions. To this end, we examined the proportion of participants that were able to solve an instance. Overall, 37.2% of attempts were correct. Performance varied both by participant (min = 0, max = 1, $SD = 0.26$) and experimental session (min = 0.33, max = 0.47, $SD = 0.06$).

Next, we investigated whether computational performance in an instance was related to the instance's complexity. We measured instance complexity with Sahni-k. This metric increases with both the number of computational steps and the amount of memory required to solve an instance. Intuitively, Sahni-k is equal to the number of items that have to be selected into the knapsack before the knapsack can be filled up using the greedy algorithm to find the solution. The greedy algorithm fills the knapsack by selecting items in decreasing order of the ratio of value over weight until none of the remaining items fits into the knapsack. If Sahni-k equals 0, the greedy algorithm generates the solution of the instance. If k is greater than 0, the Sahni algorithm generates all feasible k-element subsets, fills up the knapsack using the greedy algorithm and finds the set of items with the highest total value. The Sahni-k of the instances in this study varied from 0 to 4; they are listed in Table I.[14]

[14]Note that the number of sets the algorithm considers is given by the binomial coefficient with n equal to the number of items in the instance and k equal to Sahni-k.

The proportion of participants who solved the instance correctly decreased from 60.3% when Sahni-k was equal to 0, to 6.4% when Sahni-k was equal to 4. To test the negative relation between computational performance and Sahni-k, we estimated a mixed-effects model with a binary variable set to 1 if an attempt was correct as dependent variable (0 otherwise), a fixed effect for Sahni-k and random effects (varying intercepts) for participant and experimental session. We found a significant main effect of Sahni-k ($\beta = -0.578$, $p < 0.001$). The pattern confirms the validity of Sahni-k as a measure of instance difficulty for humans, first documented in Meloso et al. (2009b) and Murawski and Bossaerts (2016).[15] This means that the negative relation between Sahni-k and computational performance previously documented at individual level Murawski and Bossaerts (2016) also exists at the level of markets (see SOM for further details).

We used the number of items participants moved into and out of the knapsack as a proxy for effort. The mean total number of moves in an instance was 24.7 (min = 3, max = 123, $SD = 19.5$; descriptive statistics can be found in SOM). To test whether the number of moves depended on instance complexity, we related the number of moves a participant made, to Sahni-k of the instance (mixed-effects model with a fixed effect for Sahni-k and random effects for participant and session). We found a positive effect of Sahni-k on the number of item moves ($\beta = 1.649$, $p < 0.05$). This means that participants expended more effort on harder instances. This finding is consistent with Murawski and Bossaerts (2016), who also found a positive correlation between proxies of instance difficulty and effort, using a larger number of instances than in the present study.

The number of subsets being considered increases with k and quickly grows beyond the capacity of human working memory. For an instance with 12 items, the number of sets considered if k is equal to 1, is 1 whereas if k is equal to 4, the number of sets considered equals 495.

[15]For a comparison of Sahni-k as a measure of instance complexity with other measures, see Murawski and Bossaerts (2016).

The mean number of trades per session was 135.0 (min = 91, max = 202, SD = 28.0). The mean number of trades varied by both session (min = 117.4, max = 161.6, SD = 18.9) and instance (min = 127.2, max = 165.7, SD = 17.0). The number of trades did not vary with instance difficulty as measured with Sahni-k (one-way ANOVA, $F(4, 14) = 1.27$, $p > 0.1$). The number of trades in items in the optimal solution did not differ from the number of trades in items that were not in the optimal solution (two-sample t-test, $t(52) = -0.811$, $p > 0.1$).

Fig. 3 plots the evolution of trade prices in the third round of the first session. Each security is indicated by the weight and value of the corresponding item (weight_value) and whether the item is in the optimal knapsack. Notice how prices do not monotonously decrease towards zero (indicating that the corresponding item is "OUT") or one (the corresponding item is "IN"). Shares in item 129_15, for instance, bounce back and forth between a minimum of 25 cents and a maximum of 95 cents. By the end, they were trading at 60 cents. The shares expired worthless because the corresponding item was not in the optimal knapsack (more descriptive statistics on prices, including range of prices per item, stratified by instance difficulty, are available in the SOM).

Finally, participants earned \$6.32 on average per instance (min = −1.66, max = 11.41, SD = 1.92) and 31.58 on average in total (min = −8.30, max = 49.25, SD = 9.59) from trading in the market.[16] To test whether the earnings from trading depended on instance complexity, we related earnings of a participant in an instance, to Sahni-k of the instance (mixed-effects model with a fixed effect for Sahni-k and random effects for participant and

[16]Instance earnings could be negative because participants were allowed to buy, and hence spend cash, on shares that eventually did not pay a dividend; the change in cash was subtracted from total dividends earned on final share holdings. Earnings cumulated across instances. At the end of the experimental session, participants were paid the minimum of \$25 and cumulative earnings plus sign-up reward of \$5 plus per-instance submission reward of \$2.

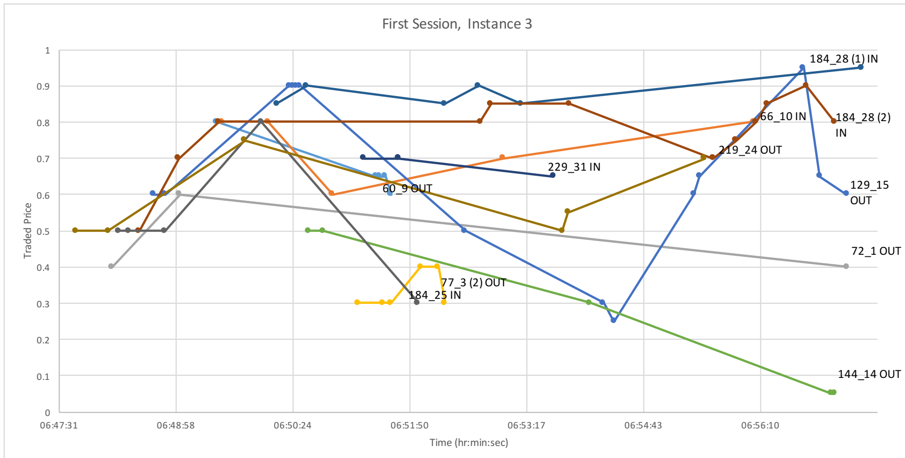


FIGURE 3.— Evolution of Traded Prices in One Instance.

Shown are time series of transaction prices for the 3rd instance in the first session. Series are identified by the weight and value of the corresponding item (weight_value) and whether the item was in the optimal knapsack ("IN," in which case the shares paid one dollar; when "OUT," they expired worthless).

session on the intercept). We did not find a systematic relation between Sahni-k and earnings from trading ($\beta = -0.177$, $p > 0.1$).[17]

Testing our predictions

In the following, we test the five predictions.

Prediction 1: Securities Prices Reveal Only Inferior Solutions

The first prediction is that markets do not reveal optimal solutions. In our sessions, there was always at least one participant who solved to instance at hand, and often multiple participants (see SOM). Consequently,

[17]Mean earnings were highest in the instance with Sahni-k equal to 0 and lowest in the instance with Sahni-$k = 2$. Mean earnings amounted to \$8.41, \$5.13, \$5.03, \$5.93 and \$7.08 for Sahni-k 0 to 4, respectively. Mean maximum payoff was \$21.05 in the instance with Sahni-k equal to 0, compared to \$12.50, \$13.70, \$16.25, and \$11.95 for instances with Sahni-k equal to 1 to 4, respectively. Notice the non-monotonicity of earnings. This is related to the way we allocated shares, which, as mentioned before, attempted to balance incentives to trade and fairness.

if prices were to perfectly reflect available knowledge, they should allow us to formulate an algorithm that always finds the optimum. We hypothesized this not to be the case.

To evaluate the prediction, we constructed a "market performance" metric, defined as the distance from the optimal solution, in item space, of "market solutions" implied by trade prices. Distance in item space is measured as a score which is incremented by one point if a correct item was in the submitted knapsack *or* an incorrect item was left out of the knapsack. The score is subsequently scaled by dividing by the total number of items in the knapsack.

To construct "market solutions," we interpreted the last traded price of an item as the market's "belief" that the item belonged in the optimal solution. We then bootstrapped a "market knapsack" by drawing without replacement items based on these beliefs and filling the knapsack until capacity was reached. We computed the performance score of this "market knapsack." We repeated this procedure 10,000 times. We then averaged the resulting market performance scores across the bootstraps. If prices correctly valued securities, and hence, correctly revealed an instance's solution, we would draw only from "IN" items, and hence, obtain a perfect performance metric.

Information revealed in market prices allowed us to reach a performance score of 78% in case of the instance with Sahni-k equal to 0. This score decreased monotonically to 68%, 64%, 57% and 52% as Sahni-k increased from 1 to 4. We conclude that market prices always revealed inferior solutions, and that the solutions became worse as instance complexity increased.

Prediction 2: The Average Trader Outperforms The Market

In the next step, we examined whether the market did better in solving the instances than the average participant. Specifically, we compared the performance of the "market knapsack" to the performance of the knapsacks submitted by individual participants. To do so, we constructed performance scores for individuals the way we did for the market. We found that the market knapsack performed worse than the knapsack submitted by the average participant for every level of computational complexity (Fig. 5). In three out of five instances, the market's score was significantly worse (two-sample t test with unequal variances; Bonferroni-Holm family-wise error correction at the $p = 0.05$ level). Performance of the market decreased with instance difficulty (Sahni-k; slope $= -0.06$, $p < 0.001$), at the same rate as for individual participants.

Prediction 3: Expert Traders Make Money

To determine whether Expert Traders earned more money from trading, we correlated individual computational performance in an instance with earnings (in Australian dollars) from trading in the marketplace. We computed the former as distance in item space of submitted knapsack from the optimal solution, using the score computed as described above.

The mean payoff in an instance for Expert Traders (those who found the solution of the instance) was \$8.21 (min $= -0.85$, max $= 21.05$, $SD = 3.51$), compared to \$5.11 (min $= -10.55$, max $= 13.7$, $SD = 3.93$) of the remaining participants. Moreover, in *every* instance, the highest payoff among all participants was earned by an Expert Trader.

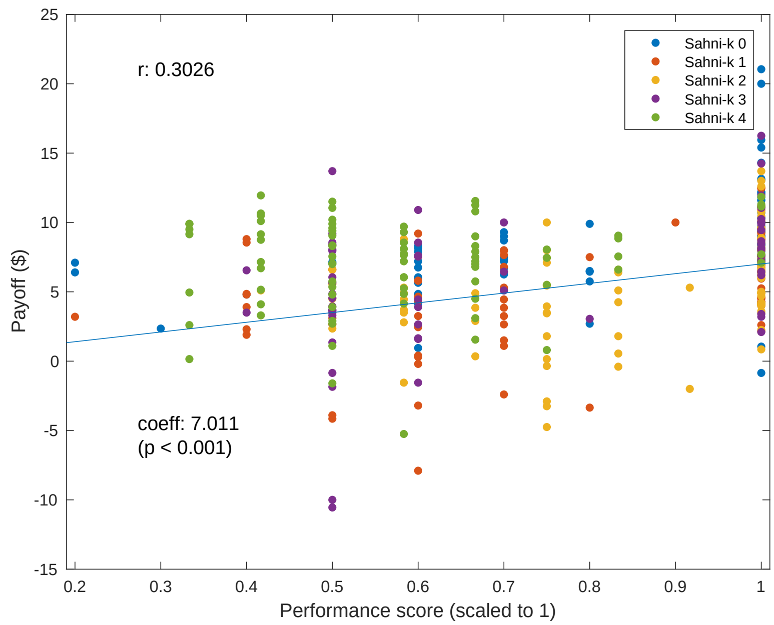
To test whether the ability to earn money increased in expertise, we related computational performance, as measured by the score described above, to earnings from trading. A higher score implies that the participant had more correct items in the submitted knapsack and more incorrect items out of the submitted knapsack, and as such, the participant knew the correct value of more securities than someone with a lower score. We estimated a generalized linear mixed-effects model with earnings in a session as dependent variable, a fixed effect for score (computational performance) and random intercepts for participant and instance. We found a significant effect of score ($\beta = 7.011$, $p < 0.001$). A ten percentage point increase in computational performance score was associated with additional earnings of 70 cents (see Fig. 4, top panel).

Considering overall session earnings per participant and computational performance, we found that they were highly correlated as well: an increase in performance score of 10 percentage points increased earnings in the marketplace by almost $4 on average (Fig. 4, bottom panel). Note that only two participants received a perfect score, implying that only two participants solved all five KP instances correctly.

Prediction 4: Market Performance Decreases with Instance Complexity, Not with Size of Search Space

We proposed that a measure of computational complexity unrelated to the number of possible filled knapsacks determines price quality. One way to measure computational complexity of the instance at hand is to use the Sahni-k metric introduced before. As Sahni-k increases, both the number of computational steps and the amount of memory required to solve the instance increases. Errors in computations or memory retrieval accumulate, and hence, noise in the proposed solution increases, or the computer runs out of computational resources. As mentioned before, we propose Sahni-k since this metric was found to predict individual performance in solving KP instances (Murawski and Bossaerts, 2016). Sahni-k is inspired by algorithms that approximate KP solutions. The higher Sahni-k, the more likely a computer using those algorithms will not find the optimum.

We found that market performance (performance of knapsack solutions one can construct from trade prices) gradually decreased as Sahni-k increases (Fig. 5). The decrease in performance is highly significant (univari-



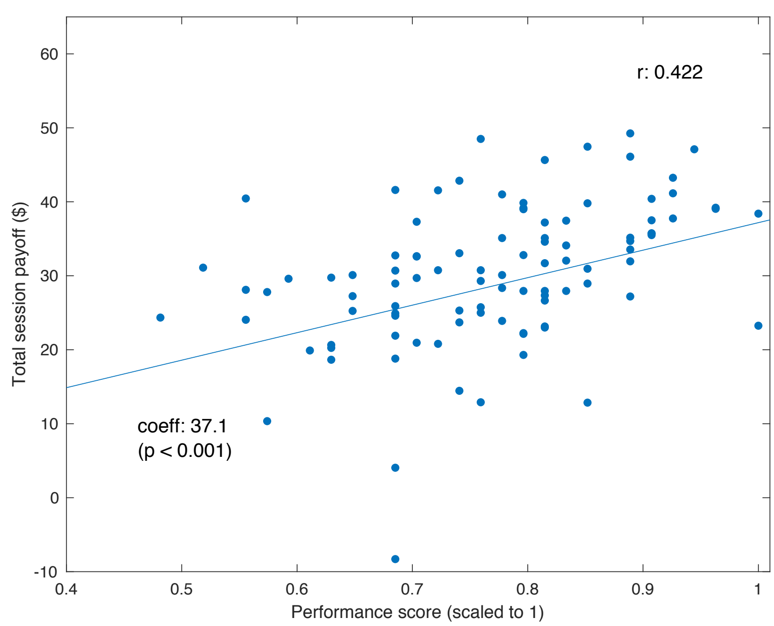


FIGURE 4.— **Individual Trading Payoff Against Performance Score Of Submitted Knapsack.** Individual payoff from trading, per instance (top; colour-coded by difficulty) and per session (bottom), against average performance score for the submitted solutions; performance score is measured as distance in item space of the submitted knapsack; see text for details; per-instance observations are stratified by Sahni-k. Slope coefficient per instance (mixed-effects regression; see text) equals 7.01 ($p < 0.001$); Pearson correlation $r = 30.3\%$. Slope coefficient per participants equals 37.1 ($p < 0.001$); Pearson correlation $r = 42.2\%$.

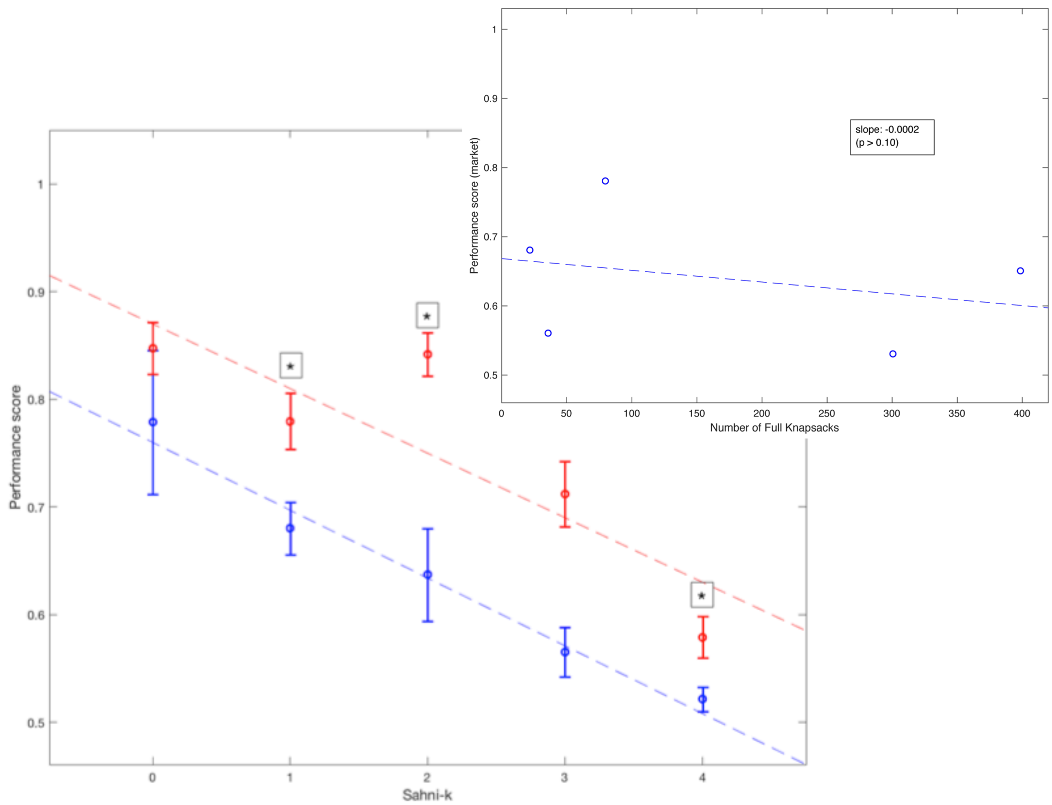


FIGURE 5.— **Performance Of Instance Solutions Generated From Last Trade Prices Compared To Average Individual Submissions.** Market performance score, measured as distance in item space of "market knapsack" obtained from simulations based on last trade prices (blue), against corresponding average score of individual participants (red), across all sessions, stratified by instance difficulty (Sahni-k). Error bars cover +/- 1 standard error (Market: across sessions; Individuals: across individuals in all sessions). Dashed lines are corresponding regression lines (slope of market line: -0.006; p value < 0.001)). Asterisks in boxes indicate that the average individual score is significantly higher than market score at $p = 0.05$ significance level, corrected for multiple comparisons. INSERT: Market performance score against number of possible knapsacks filled at capacity. Dashed lines: regression line.

ate LS of mean performance on Sahni-k: slope $= -0.06$, $p < 0.001$).

In contrast, market performance did not change with the number of possible knapsacks filled to capacity; see Insert of Fig. 5 ($|\text{slope}| < 0.001$, $p > 0.10$). That is, merely increasing the size of the search space does not lower price quality.

Prediction 5: Traders Improve Their Solutions by Reacting to High Volume in Low-Priced Securities

In traditional "rational expectations" asset pricing models (Radner, 1979), lesser informed traders are assumed to know the mapping from states to prices and use this mapping to infer states from observed prices. In this setting, the term "state" refers to an expectation based on all the information available in the economy. Here, we can interpret "state" as the correct solution to the KP instance.

We investigated whether traders indeed "read" prices. Specifically, we studied to what extent trade induced traders to re-visit and improve their knapsack, moving it closer to optimum in item space. Overall, descriptive statistics of moves of correct and incorrect items in and out of knapsacks can be found in SOM.

Prior research has shown that an important reason why humans may not find the optimum is because they tend not to re-consider incorrect items that they put into the knapsack early on (Murawski and Bossaerts, 2016). Poor episodic memory may explain this reluctance to re-visit early moves. Here, we ask whether trade in such items made it more likely that uninformed traders took them out.

To test whether participants improved their knapsack based on information available in the market, we regressed the probability of removing an incorrect item that was included early on, onto the price of, and the trading volume in, the corresponding security, as well as their interaction using a generalized linear model with a logit link function.[18] We consider an item to be included "early on" if the participant moved it into the knapsack within the first two minutes of trading. We found that there was a significant main effect of price ($\beta = -3.491$, $p < 0.001$). This means a 10% decrease in price of a security was associated with a 11% increase in the probability of removing the corresponding item from the knapsack.

The effect was stronger when the security corresponding to the item was traded more heavily (interaction term of price and trading volume, $\beta = 13.725$, $p < 0.05$). However, there was no main effect of trading volume ($\beta = -7.125$, $p = 0.068$). The negative relation between security price and probability of removal is displayed in Fig. 6.

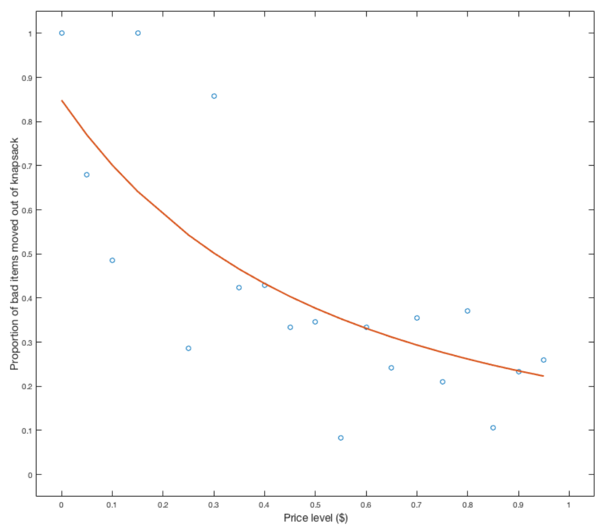


FIGURE 6.— **Impact Of Trade Prices On Removal Of Incorrect Items** Incorrect items chosen into knapsack within two minutes of trading: Fraction of those items that are eventually removed, as a function of average trade price. Red line: least squares fit of $\beta/(1+x)^2$ where x = price level, $\beta = 0.85$ ($p < 0.01$).

These findings suggest that trade, combined with low prices, induced participants to re-consider incorrect inclusions of items, thus improving overall

[18]Model selection analysis based on the Bayesian Information Criterion suggested that a model with random effects for participant, instance difficulty (Sahni-k) and session fitted worse than a fixed-parameter regression.

computational performance, and hence, securities valuation.

We did not find an effect on item selection from trade in high-priced securities. We conjecture that short-sale restrictions contributed to the asymmetry between high-priced and low-priced securities: if an item was deemed to be overpriced, participants could only sell shares that they already owned, and hence could not put more pressure on prices.

5. CONCLUSION

In this study, we investigated to what extent computational complexity affected market outcomes. Like with loss or ambiguity aversion (Dybvig and Ross, 2003), there is a fundamental tension between individual behaviour and the need for market prices to be free of arbitrage opportunities. Individuals solve computationally "hard" problems in a methodical way, which, although highly effective, easily leads to violations of the sure-thing principle. These violations cannot be exposed at the market level, because they would amount to arbitrage opportunities. We formulated hypotheses aimed at resolving the tension. These hypotheses are in line with predictions of the Noisy Rational Expectations Equilibrium, provided one interprets "noise" as (instance) complexity.

We reported results from a markets experiment aimed at testing our predictions. We found that prices generally revealed incorrect solutions, and that price quality deteriorated as instance complexity increased, though price quality was no worse when the search space increased (number of possible capacity-filled knapsacks). We also documented that Expert Traders (those who submitted the optimal solution) earned significantly more from trading on their information; in general, the closer one was to the correct solution, the more was earned through trading. Market prices converted into candidate solutions that were worse than those submitted by the average participant. Information revealed in prices did feed back into individual problem solving, but only in conjunction with volume.

A number of remarks are in order, which should put our findings in perspective.

We found that traders learned from prices, which is one of the core tenets of "rational expectations" asset pricing theory. To our knowledge, we are the first to show this directly, by recording participants' thinking (their attempts to solve the valuation problems) and how it changed as a function of prices. Here, we discovered that markets actually mitigated a strong bias that keeps individuals from discovering the correct solution of KP instances, namely, hesitation to re-consider items that were added to the knapsack early on (Murawski and Bossaerts, 2016). Prices, in conjunction with trade, made participants re-visit parts of the solution that they had constructed within the first two minutes of trading.

Our finding suggests that markets may provide a powerful mechanism to help individuals solve complex problems. This raises the issue whether a market mechanism is preferable to mechanisms that provide clearer incentives for problem solving, such as a prize system, where only the first to submit the correct solution wins a prize. Meloso et al. (2009b) reports that more participants manage to find the correct solution of KP instances with markets than with a prize mechanism.

This last finding is relevant for the debate on the desirability of patents as a way to promote intellectual discovery, for two reasons. First, the prize system is analogous to the current patent system. Second, intellectual discovery can be thought of as the solution of a combinatorial optimization problem such as the KP (Boldrin and Levine, 2002).[19] Evidently, markets may facilitate intellectual discovery better than patents.[20]

[19] A recent empirical study corroborates this claim, by showing that patents filed with the U.S. Patent Office between 1790 and 2010 were mostly for inventions that combined existing technologies in novel ways rather than opening up fundamentally new avenues of exploration (Youn et al., 2015).

[20] To illustrate how, consider markets in Li, Na, Mg and other chemicals (we are

The conclusion that markets may trump patents is important for economic history as well. It is generally accepted that the patent system provided the main impetus for technological advances over the last century and a half (Khan and Sokoloff, 2001). However, at the same time markets penetrated all parts of life, and it may very well have been that markets were the major facilitator of innovation rather than patents. This conjecture is consistent with historical evidence that technological advances can be far bigger during epochs with share trading but without patents than in epochs with only patents (Nuvolari, 2004).

Our findings allow one to put into perspective the evidence on EMH (the Efficient Markets Hypothesis) from historical analyses of field data. Fama (1991, 1998) surveys a vast body of studies that appears to suggest that EMH holds, because anomalies could be attributed to chance (sometimes the null will be rejected) or to methodological errors. Behavioral finance scholars reject this conclusion, starting with Bondt and Thaler (1985), who claimed that long-run price reversals are caused by overreaction. Among others, Hirshleifer (2001) summarizes the evidence against EMH, and argues against chance or methodological mistakes.

Our experiment could shed new light on the EMH controversy. When heterogeneous information emerges because agents hold dispersed information from which security prices can easily be computed, markets may be expected to satisfy EMH. But it is obvious that not all real-world situations can be described as such. Instead, correct valuation may be computationally hard, at which point the conclusions from our experiment become important, and

---

using standard chemical abbreviations). These are potential components of future battery technology. To determine which chemicals markets to invest in requires one to assess which component, or maybe combination of components, will be required for the best battery technology. This casts intellectual discovery squarely in terms of the KP, and to profit from it, one needs markets as we organized them. Inventors are induced to participate in the marketplace, and earn money by buying those components that they believe are in the best battery technology, while selling others.

violations of EMH will ensue. Examples of high computational complexity includes corporate valuation when valuation depends on how the corporation is structured (Which departments should be spun off? Which companies should be acquired in order to create "synergistic" valuation effects? Etc.).

To measure computational complexity, we used standard notions from computer science. These notions are based on a theoretical computational model (Turing machine), and it was not *a priori* obvious that they would extend to human computing. Recent evidence suggests that computational complexity theory does extend to decision-making by humans (Murawski and Bossaerts, 2016), and the findings reported here corroborate this notion. Therefore, evidence is mounting that the theory of computation is universal, in support of the Church-Turing thesis (Church, 1934; Turing, 1936).

In the present study, we showed that computational complexity theory extends even to markets: we found that markets performed worse when valuation had higher computational complexity, where complexity was measured in terms of Sahni-k. This metric, one possible measure of computational resources required to solve an instance, tracks human performance in solving KP instances (Murawski and Bossaerts, 2016).

Finally, we chose to have participant trade Arrow securities whose payoffs depended on whether the corresponding item was in the optimal solution. One could envisage alternative security designs. For instance, we could have let participant trade a single security that paid the value of the optimal solution to the KP at hand. According to traditional asset pricing logic, the latter should generate worse outcomes: a given knapsack value may translate into multiple possible knapsack compositions, many of which could be far off the optimal knapsack. That is, one security may not be enough to "invert" from knapsack values to choice of items. Intuitively, directly trading one Arrow security per item seemed to be the most effective way for participants to share information about potential solutions to the KP at hand. The market with Arrow securities seems to be more "complete." Whether this is true remains to be proven empirically. We leave this for future work. If pricing is more accurate, and more participants find the optimal solution when only one security is traded (whose payoff depends on the value of the optimal knapsack), then we would definitely know that asset pricing under computational complexity is fundamentally different from classical asset pricing theory. Our intuition that "more securities is better" would fail.

Contributions: Bossaerts designed the experiment, Tang run experiment and conducted the statistical analysis; Bossaerts, Murawski and Tang wrote the manuscript; Huang, Bowman and Yadav contributed to software design and data acquisition.

Acknowledgment: Financial support from a R@map chair at The University of Melbourne (Bossaerts) is gratefully acknowledged. The experiment reported here was approved by the University of Melbourne Human Research Ethics Committee (Ethics ID: 1647059.1) and was conducted in accordance with the World Medical Association Declaration of Helsinki. All participants provided written informed consent. We are grateful for commentary received from many colleagues, especially from Sébastien Pouget, whose discussion at the Financial Management Association 2017 meetings led us to re-evaluate the results in the context of the Noisy Rational Expectations Equilibrium. A previous version of this paper circulated under the title "Information Aggregation under Varying Levels of Computational Complexity: Experiments."

REFERENCES

Anufriev, M., T. Bao, A. Sutan, and J. Tuinstra (2018): "Fee Structure and Mutual Fund Choice: An Experiment," Working Paper Series 45, Economics Discipline

Group, UTS Business School, University of Technology, Sydney.

Arora, S., B. Barak, M. Brunnermeier, and R. Ge (2010): "Computational complexity and information asymmetry in financial products," in *ICS*, 49–65.

Asparouhova, E. and P. Bossaerts (2017): "Experiments on Percolation of Information in Dark Markets," *The Economic Journal*, 127, F518–F544.

Asparouhova, E., P. Bossaerts, J. Eguia, and W. Zame (2015): "Asset pricing and asymmetric reasoning," *Journal of Political Economy*, 123, 66–122.

Asparouhova, E., P. Bossaerts, and W. Yang (2017): "Costly Information Acquisition in Decentralized Markets: An Experiment," *SSRN Working Paper*.

Biais, B., P. Bossaerts, and C. Spatt (2003): "Equilibrium Asset Pricing under Heterogeneous Information," Available from author's webpage.

Boldrin, M. and D. Levine (2002): "The Case against Intellectual Property," *The American Economic Review*, 92, 209–212.

Bondt, W. F. M. D. and R. Thaler (1985): "Does the Stock Market Overreact?" *The Journal of Finance*, 40, pp. 793–805.

Bossaerts, P., C. Frydman, and J. Ledyard (2014): "The Speed of Information Revelation and Eventual Price Quality in Markets with Insiders: Comparing Two Theories*," *Review of Finance*, 18, 1–22.

Bossaerts, P., P. Ghirardato, S. Guarnaschelli, and W. R. Zame (2010): "Ambiguity in asset markets: Theory and experiment," *The Review of Financial Studies*, 23, 1325–1359.

Bossaerts, P., N. Yadav, and C. Murawski (2018): "Uncertainty and computational complexity," *Philosophical Transactions of the Royal Society B*, under review.

Church, A. (1934): "The Richard Paradox," *The American Mathematical Monthly*, 41, 356–361.

Cook, S. A. (1983): "An overview of computational complexity," *Communications of the ACM*, 26, 400–408.

Copeland, T. E. and D. Friedman (1992): "The market value of information: Some experimental results," *Journal of Business*, 241–266.

De Finetti, B. (1995): "The logic of probability," *Philosophical Studies*, 77, 181–190.

Diamond, D. W. and R. E. Verrecchia (1981): "Information aggregation in a noisy rational expectations economy," *Journal of Financial Economics*, 9, 221–235.

Duffie, D. and G. Manso (2007): "Information percolation in large markets," *The American economic review*, 97, 203–209.

Dybvig, P. H. and S. A. Ross (2003): "Arbitrage, state prices and portfolio theory," *Handbook of the Economics of Finance*, 1, 605–637.

Fama, E. (1998): "Market Efficiency, long-term returns, and behavioral finance," *Journal of Financial Economics*, 49, 283–306.

Fama, E. F. (1991): "Efficient Capital Markets: II," *The Journal of Finance*, 46, pp. 1575–1617.

Grossman, S. (1976): "On the efficiency of competitive stock markets where trades have diverse information," *The Journal of finance*, 31, 573–585.

——— (1978): "Further results on the informational efficiency of competitive stock markets," *Journal of Economic Theory*, 18, 81–101.

Grossman, S. J. (1977): "The existence of futures markets, noisy rational expectations and informational externalities," *The Review of Economic Studies*, 44, 431–449.

Grossman, S. J. and J. E. Stiglitz (1976): "Information and Competitive Price Systems," *The American Economic Review*, 66, 246–253.

Hastings, W. K. (1970): "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, 57, 97–109.

Hellwig, M. F. (1980): "On the aggregation of information in competitive markets," *Journal of economic theory*, 22, 477–498.

Hirshleifer, D. (2001): "Investor Psychology and Asset Pricing," *The Journal of Finance*, 56, 1533–1597.

Holden, C. W. and A. Subrahmanyam (1992): "Long-lived private information and imperfect competition," *The Journal of Finance*, 47, 247–270.

Kellerer, H., U. Pferschy, and D. Pisinger (2004): *Knapsack Problems*, Springer Science & Business Media.

Khan, B. Z. and K. L. Sokoloff (2001): "History lessons: the early development of intellectual property institutions in the United States," *The Journal of Economic Perspectives*, 15, 233–246.

Kolmogorov, A. N. (1983): "Combinatorial foundations of information theory and the calculus of probabilities," *Russian mathematical surveys*, 38, 29.

Long, J. B. D., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990): "Noise Trader Risk in Financial Markets," *The Journal of Political Economy*, 98, 703–738.

Meloso, D., J. Copic, and P. Bossaerts (2009a): "Promoting intellectual discovery: patents versus markets," *Science*, 323, 1335–1339.

——— (2009b): "Promoting Intellectual Discovery: Patents vs. Markets," *Science*, 323, 1335.

Murawski, C. and P. Bossaerts (2016): "How Humans Solve Complex Problems:

The Case of the Knapsack Problem," *Scientific Reports*, 6.

Nuvolari, A. (2004): "Collective Invention During the British Industrial Revolution:The Case of the Cornish Pumping Engine," *Cambridge Journal of Economics*, 28, 347–363.

Plott, C. and S. Sunder (1988): "Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets," *Econometrica*, 56, 1085–1118.

Plott, C. R. (2000): "Markets as Information Gathering Tools," *Southern Economic Journal*, 67, 2–15.

Plott, C. R. and S. Sunder (1982): "Efficiency of experimental security markets with insider information: An application of rational-expectations models," *The Journal of Political Economy*, 663–698.

Radner, R. (1979): "Rational Expectations Equilibrium: Generic Existence and the Information Revealed by Prices," *Econometrica*, 47, 655–678.

Savage, L. J. (1972): *The foundations of statistics*, Courier Corporation.

Skreta, V. and L. Veldkamp (2009): "Ratings shopping and asset complexity: A theory of ratings inflation," *Journal of Monetary Economics*, 56, 678–695.

Spiegler, R. (2016): "Competition and Obfuscation," *Annual Review of Economics*, 8.

Sunder, S. (1992): "Market for information: Experimental evidence," *Econometrica: Journal of the Econometric Society*, 667–695.

Turing, A. M. (1936): "On computable numbers, with an application to the Entscheidungsproblem," *J of Math*, 230–265.

Valiant, L. G. (1984): "A theory of the learnable," *Communications of the ACM*, 27, 1134–1142.

Wolinsky, A. (1990): "Information revelation in a market with pairwise meetings," *Econometrica: Journal of the Econometric Society*, 1–23.

Youn, H., D. Strumsky, L. M. A. Bettencourt, and J. Lobo (2015): "Invention as a combinatorial process: evidence from US patents," *Journal of the Royal Society Interface*, 12, 20150272.

TABLE I

Five KS Instances Used In The Experiment. Available items, capacity C, and Sahni-k k, for the instances in the experiment. Density is defined as the ratio of value v to weight w. IN means the item is in the solution; OUT means it is not. NFKS denotes the number of possible knapsacks filled to capacity. Problem 5 has a second solution (not indicated), with a higher weight and higher Sahni-k however.

1 ($k = 1$) — $C = 1,900$, NFKS = 80

Items	1	2	3	4	5	6	7	8	9	10
v	505	435	350	505	640	465	170	220	500	50
w	564	489	406	595	803	641	252	330	750	177
Density	0.90	0.89	0.86	0.85	0.80	0.73	0.67	0.67	0.67	0.28
Solution	IN	IN	OUT	IN	OUT	OUT	IN	OUT	OUT	OUT

2 ($k = 3$) — $C = 1,044$, NFKS = 22

Items	1	2	3	4	5	6	7	8	9	10
v	300	350	400	450	47	20	8	70	5	5
w	205	252	352	447	114	50	28	251	19	20
Density	1.46	1.39	1.14	1.01	0.41	0.40	0.29	0.28	0.26	0.25
Solution	IN	OUT	IN	IN	OUT	OUT	OUT	OUT	IN	IN

3 ($k = 2$) — $C = 850$, NFKS = 399

Items	1	2	3	4	5	6	7	8	9	10
v	28	28	10	9	25	31	15	24	14	3
w	184	184	66	60	184	229	129	219	144	77
Density	0.15	0.15	0.15	0.15	0.14	0.14	0.12	0.11	0.10	0.04
Solution	IN	IN	IN	OUT	IN	IN	OUT	OUT	OUT	OUT
Items	11	12								
v	3	1								
w	77	72								
Density	0.04	0.01								
Solution	OUT	OUT								

4 ($k = 0$) — $C = 1500$, NFKS = 36

Items	1	2	3	4	5	6	7	8	9	10
v	37	32	44	23	45	85	62	106	71	72
w	50	46	180	107	220	435	360	700	530	820
Density	0.74	0.70	0.24	0.21	0.20	0.20	0.17	0.15	0.13	0.09
Solution	IN	IN	IN	IN	IN	IN	IN	OUT	OUT	OUT

5 ($k = 4$) — $C = 1300$, NFKS = 301

Items	1	2	3	4	5	6	7	8	9	10
v	227	129	113	127	303	144	84	147	86	201
w	212	121	106	120	288	137	80	140	82	192
Density	1.07	1.07	1.07	1.06	1.05	1.05	1.05 1.05			
Solution	IN	IN	OUT	OUT	IN	IN	IN	IN	IN	OUT
Items	11	12								
v	251	167								
w	240	160								
Density	1.05	1.04								
Solution	IN	OUT								