MIT Libraries | DSpace@MIT

MIT Open Access Articles

Securities trading of concepts (STOC)

Massachusetts Institute of Technology

Securities Trading of Concepts (STOC)

Ely Dahan[*], Adlar J. Kim[**], Andrew W. Lo[***], Tomaso Poggio[•] and Nicholas Chan[••]

August 5, 2010

[*]   (*corresponding author*) UCLA Medical School, Claremont Graduate University, Princeton University, 2542 Cardigan Court, Los Angeles, CA 90077-1337, elydahan@gmail.com

[**]   J.P. Morgan, New York, NY 10179

[***]   Harris & Harris Group Professor of Finance; Director, MIT Laboratory for Financial Engineering, MIT Sloan School of Management, Cambridge, MA 02142

[•]   Eugene McDermott Professor in the Brain Sciences and Human Behavior, McGovern Institute, Computer Science and Artificial Intelligence Lab, MIT, Cambridge, MA 02139

[••]   Two Sigma Investments, LLC, New York, NY 10012

Securities Trading of Concepts (STOC)

Abstract

Identifying winning new product concepts can be a challenging process that requires insight into private consumer preferences. In order to measure consumer preferences for new product concepts, we apply a securities-trading approach where new product concepts are traded as financial securities: Securities Trading of Concepts (STOC). We apply this method because market prices are well known to efficiently collect and aggregate private information regarding the economic value of goods, services, and firms, particularly when trading financial securities. Our research includes the first application of securities markets to test potential new product concepts, and is the first to compare such an approach against stated-choice, conjoint, constant-sum and longitudinal revealed preference data. In our research, we place STOC in the context of existing methodologies, as well as prior research on prediction markets and experimental economics. In the process, we put forth four hypotheses to test empirically through a series of experiments in multiple product categories, that STOC: 1) is more cost-efficient than other methods; 2) passes validity tests; 3) measures expectations of others; and 4) reveals individual preferences, not just those of the crowd. All hypotheses are confirmed, with the notable exception that STOC does not accurately predict actual product market shares and price sensitivity. Our results also show that traders exhibit bias based on self-preferences when trading. Ultimately, STOC offers two key advantages to traditional market research methods—cost efficiency and scalability. For new product development (NPD) teams deciding where to invest product-development resources, this scalability may be especially important in the Web 2.0 world where customers are constantly interacting with firms and with each other in suggesting numerous product design possibilities that need to be screened.

1 Introduction

New product innovation depends on a firm's ability to both generate many potentially winning concepts and accurately distinguish between winners and losers. Preferences for some new product categories are primarily attribute-based, and are therefore most amenable to decompositional methods such as conjoint analysis. However, the Web 2.0 world, where customers interact with firms and each other in suggesting numerous and complex product design possibilities, demands a more accurate and scalable filtering method for identifying most promising integrated product concepts.

We propose that Securities Trading of Concepts (STOC), wherein new product concepts are traded as financial securities, can potentially serve as this filter, helping to identify winning new product concepts, such as those with market appeal, during the early and lower investment phases of new product development. The reason STOC may work well is that securities markets are well-known to efficiently collect and aggregate diverse information regarding value using the simple summary statistic of price (Hayek 1945). Rather than summarizing economic value, the security prices in STOC measure intensity of preference among the traders for competing new product concepts because traders receive multiple votes and they can show the strength of their votes with the number of their shares that they allocate.

When compared to existing preference measurement research methods such as surveys, voice-of-the-customer methods, conjoint analysis, concept tests, and focus groups, STOC has potential cost and scale advantages due to its attractiveness to respondents and the way in which it collects multiple answers from each respondent. Despite similarities in outward appearance and implementation to prediction markets, STOC differs significantly in that it measures

preferences for concepts that may never be realized rather than *observable* outcomes and controls for the effects of outside news, both of which we'll discuss in more detail shortly.

The present research hypothesizes that: 1) STOC compares favorably to extant methods; 2) STOC passes validity tests; 3) STOC measures expectations of others; and 4) STOC may also reveal individual preferences if traders make buying and selling decisions partially based on their *personal* preferences, and use them to form expectations of *others'* preferences. We test these hypotheses empirically through a series of experiments in multiple product categories. In each experiment, STOC is tested against traditional market research methods applicable to new product concepts.

The present research is aimed at illuminating the relative strengths and weaknesses of STOC, with the hope of better informing the market for preference measurement techniques. The paper proceeds by providing a brief overview of the STOC concept and methodology in section 2. Section 3 compares STOC to existing methods. Section 4 presents four STOC hypotheses, while section 5, the results. Section 6 discusses the results, managerial implications, deficiencies and conclusions.

2   The STOC Concept and Methodology

The STOC methodology centers around hypothetical consumers trading hypothetical securities (each associated with a product or service concept), in virtual stock markets. The five key steps in designing a STOC game are summarized in Table 3. Each participant receives an initial portfolio of cash (virtual or real) and virtual stocks. Participants are also provided with detailed information on the product concepts that include specifications, images, and multimedia illustrations. A typical objective of the STOC game might be for each participant to maximize the value of his or her portfolio, evaluated at the closing stock prices or the volume-weighted-

4

average prices (VWAP). Markets are typically open for 20 to 30 minutes and end at random times. STOC's high speed derives from the lack of outside news which allows prices to converge quickly. If participants play with real money, they will have the opportunity to profit from trading and will conversely bear the risk of losing money. The financial stakes in the game provide incentives for participants to reveal true preferences, process information, and conduct research. If fictitious money is used, prizes can be awarded according to individuals' performances. One can also reward all participants simply for their service.

As in real financial markets, stock prices are determined by supply and demand, which depend on participants' evaluation of their own and others' preferences for the underlying products. Thus, at the market equilibrium, prices should fully reflect all participants' aggregate preference of the products being tested. Traders make trading decisions just as they would in a financial stock market: they assess the values of the stocks, sell overvalued ones and buy undervalued ones, essentially voting on the worth of the underlying products. In this way, a stock's price becomes a convenient index of a product's consumer value. In our STOC tests, all trades require a specific buyer to purchase shares from a specific seller, both of whom have placed a limit order (i.e., we employ a double auction with no market makers); one could also employ market making software that enables one-sided trades in "thin" markets with few traders as Hanson (2003, 2005, 2009) proposes. Technical details of the web-based STOC trading system can be found in Web Appendix 1 at the following URL (*JMR* web appendix URL here).

3 STOC vs. Existing Consumer Preference Measurement Methods

3.1 Potential Benefits vs. Other Market Research Methods

STOC may be compared against existing, well-established methods for estimating consumer preferences. For example, 1) surveys (cf. Burchill and Brodie 1997), 2) voice-of-the-customer methods (cf. Griffin and Hauser 1993), 3) conjoint analysis (cf. Srinivasan and Shocker 1973, Green and Wind 1981, Green and Srinivasan 1990), 4) concept tests (cf. Urban, Hauser and Roberts 1990, Dahan and Srinivasan 2001, Dahan and Hauser 2002), and 5) focus groups (cf. Mahajan and Wind 1992, Calder 1977, Fern 1982).

STOC offers five potential benefits over most preference measurement methods:

1. **Accuracy Due to Incentive Compatibility:** STOC traders have the incentive to trade based on their most accurate perceptions of others' preferences since they are more likely to perform well in the game by doing so. Moreover, STOC overcomes certain social biases because the method doesn't directly reveal potentially sensitive private information or preferences, as might happen with traditional methods of individual data collection. STOC also continuously captures the intensity and degree of preference in ways that surveys typically do not, e.g., by weighting multiple trades of varying volume based on the total amount of capital committed. STOC's validity and accuracy can be empirically tested by comparing its preference measurement results against those of the traditional methods.

2. **Interactive Learning:** STOC traders learn from each other through the price mechanism. Since participants observe each other's valuations of the product concepts, they can update and adjust their own valuations dynamically in the market environment. Learning may reduce response error, especially as STOC prices converge. Learning may be especially useful for fashion goods for which individual preference may depend at least partially on the

opinions of others. A downside of learning is that STOC's revelation of prices (at least to the group of traders) hurts the firm's ability to keep the results private. STOC's learning effects can be measured by asking traders the same preference questions immediately before and after trading, and calculating which responses better capture actual preferences, and which have lower variance.

3. **Scalability:** Unlike surveys, in which the number of questions asked is limited by the capacity of each respondent to answer, securities markets are intrinsically scalable due to the fact that each trader need only evaluate a small subset of the universe of securities. Dahan, et. al. (2010) specifically demonstrate scalability limited only by the total number of traders, and outline experimental designs that group any number of traders into sub groups. The efficiency of the market, and therefore the quality of data collected, improves with the number of participants. This extends to the number of product concepts that may be evaluated—since there is no requirement that each respondent trade every security, the bounded rationality of the traders does not limit the number of concepts that can be evaluated in a STOC market. Limiting the number of securities seen by each trader to a small subset of the full gamut of ideas being tested also mitigates the aforementioned risk of public revelation.

4. **Integrated Product Concepts:** Like Dahan and Srinivasan's (2000) Virtual Concept Testing (VCT), the STOC method is particularly useful relative to conjoint methods when a product cannot be easily quantified, delineated, or represented by a set of attributes (for example, a movie script, fashion item, car body style, or piece of art). Market participants evaluate the integrated product concepts directly, and market prices effectively reflect the overall viability of the concepts, including the ability of a concept to fulfill unarticulated needs. All that is

required is a thorough physical depiction of each concept. While STOC could also be used to measure preferences for product attributes and their levels, as demonstrated by Dahan et. al. (2010), such an approach might make more sense as a complement to conjoint analysis rather than a substitute, for example as a way of filtering from a larger set of possible attributes to a smaller one more manageable in conjoint. STOC may also outperform conjoint in predicting choice share for highly integrated, aesthetic product concepts such as Wii video games since their attributes are harder to define.

5. **Recruiting Cost Economics**

STOC is potentially more cost efficient when compared to other consumer preference testing methods if respondents prefer trading to answering surveys, and if more data is collected per respondent. The economics of any concept testing method depend largely on the costs of recruiting and then compensating a sample of N respondents willing to accurately share their preferences about M concepts. The firm must recruit more than N subjects as response rates are likely to be below 100%, and the costs per person for recruiting and compensation may vary based on the "attractiveness" of the preference measurement method being employed. The total respondent-related costs can be summarized as follows (derived in Appendix 1):

$$(1) \qquad TC = \left(M \cdot \frac{N_{sample}}{q_{respondent}} \right) \times \left(\frac{c_{recruit}}{r\%} + c_{respondent} \right).$$

The five key factors influencing the total respondent cost of measuring preferences for M given concepts, and the challenge underlying each of these factors is summarized in Table 1.

Insert Table 1 here

Hypothesis 1

In fact, we hypothesize that STOC improves all five cost measures simultaneously by being more attractive to respondents (lowering $c_{recruit}$ and improving $r\%$), and, by virtue of being a competition rather than a straight survey, more motivating (improving $q_{respondent}$) and engaging (lowering $c_{respondent}$). The sample size required should be smaller since respondents learn from each other and respond with expectations about the population mean, thus reducing error and variance, and each provide multiple data points (reducing N_{sample}) in the form of trades. These effects are testable, as in the following two hypotheses:

> H1a. *STOC is more attractive to respondents:* Respondents will prefer trading in STOC games over answering market research surveys covering an equal number of product concepts.

> H1b. *STOC requires smaller sample sizes for an equal amount of information.*

Both hypotheses are confirmed, although the results are primarily based on simple observations and surveys. Hypothesis H1a was empirically tested with surveys of 63 real world executives and engineers at a telecommunications firm and 113 MBA students, each of whom completed market research surveys and played STOC games for the same product concepts on the same day. Both groups expressed strong preferences for STOC over the surveys (90% and 85%, respectively).

Hypothesis H1b is demonstrated by virtue of the small sample sizes in each of eleven STOC experiments, with trader groups ranging in size from 18 to 56, with a median size of 38 traders. The traditional market research methods against which the STOC results are compared involved much larger respondent samples, typically 100 to 300 respondents, yet the results from both approaches are comparable in quality, as we shall see shortly.

That STOC would require smaller sample sizes makes intuitive sense. Traditional market research typically requires respondent sample sizes of 50–300, or more, to allow for heterogeneity in preference, and most concept tests measure preferences for 5–12 concepts at a time due to cognitive constraints, so 4–50 respondents per concept tested is typical. Our STOC experiments, on the other hand, reveal that 1–2 traders per concept stock is sufficient to produce stable results. The amount required to recruit potential respondents and compensate those completing the exercise can typically be reduced by half or more due to their preference for playing the game. Thus, by cutting sample size in half and recruiting costs per respondent by a similar amount, STOC could potentially reduce the total respondent costs of a concept preference study by 75% or more. Since the infrastructure costs of running the STOC trading system are essentially fixed, we anticipate long term variable cost savings from STOC testing.

3.2 Prediction Markets

Beyond the other well-established methods for estimating consumer preferences discussed above, there exists non-financial "prediction markets" used to forecast political elections, movie box office data, and other real world outcomes. The Iowa Electronic Markets (IEM)[1] pioneered prediction markets for the purpose of forecasting election results (Forsythe, Nelson, Neumann & Wright 1993). The IEM is operated by the University of Iowa College of Business for research and teaching. These markets are small-scale, real-money futures markets where contract payoffs depend on economic and political events such as elections, meaning that traders receive incentives based on how they predict future events. The IEM features real-money futures markets in which contract payoffs depend on the outcome of political and economic

---

[1] The Iowa Electronic Markets, http://www.biz.uiowa.edu/iem/

events. IEM's predictions have outperformed most national polls.[2] Similarly, the Hollywood Stock Exchange, HSX.com, has provided accurate predictions of movie box office results (Spann and Skiera 2003). The Foresight Exchange (FX)[3] predicts the probability of future events occurring such as changes in the environment, scientific breakthroughs, the collapse of companies, or political and news outcomes. Companies such as Hewlett Packard, Microsoft, Best Buy and Google have employed internal prediction markets to forecast printer sales, software release dates, consumer electronics sales, and software take-up rates.

Such prediction markets share with STOC the benefits of information aggregation, the joy of competitive play, the ability to learn from others, and the incentive to be accurate. Prediction markets focus on actual outcomes, operate for weeks, months, and sometimes years, and incorporate private information and news as it happens. STOC markets, in contrast, focus on concepts that may never come into existence, and therefore may never have actual outcomes, run for a much shorter time period—10–60 minutes typically, and are not influenced or complicated by outside news. In fact, the only information available to STOC traders is the personal preferences they hold, their expectations of others' preferences, and whatever they learn by observing price movements during trading.

3.3 Rational Expectations (RE) Models and Experimental Markets

STOC is also closely related to the literatures in rational expectations (RE) models with asymmetric information and experimental markets. RE is an economic hypothesis that states that agents' predictions of the future value of economically relevant variables is not systematically incorrect because all the errors are random. RE is used in many contemporary macroeconomic models, game theory, and other applications of rational choice theory. In a standard asymmetric

---

[2] *BusinessWeek*, 11/11/96
[3] The Foresight Exchange, http://www.ideosphere.com/fx

information RE model (Grossman, 1981), heterogeneous agents with diverse information trade with each other and, under certain conditions, the market will converge to an equilibrium in which prices fully reveal all relevant information. The most important criterion for convergence is that agents condition their beliefs on market information. In particular, agents make inferences from market prices and quantities about other agents' private information.

The RE model has received considerable attention in the study of experimental markets (Plott and Sunder, 1982, 1988; Forsythe and Lundholm, 1990; Davis and Holt, 1993). Studies of the informational efficiency of a market relative to the RE benchmark fall into two categories: markets with fully informed agents ("insiders") and uninformed agents, and markets with many partially informed agents. In various experimental markets with human subjects, the results for both market structures are the same: markets eventually converge to the RE equilibrium, i.e., information aggregation and dissemination occur successfully.

STOC trading shares some characteristics with such experimental economics markets, and information aggregation and dissemination provide compelling explanation for the success of our experiments. For example, traders who possess superior information about the products, or have high confidence in their beliefs, can be considered "insiders." On the other hand, traders who have little knowledge or opinion of the products can be regarded as the "uninformed." The interaction between the insider and uninformed constitutes information dissemination. What is intriguing about this scenario is that even when a subset of traders ignores the underlying product information and only focuses on market information, the market still converges to efficient prices that aggregate all the relevant information and beliefs. A striking example of just how informationally efficient financial markets can be is provided by Maloney and Mulherin (2003), who document the fact that in the wake of the Space Shuttle Challenger's explosion in 1986, the

stock price of Morton Thiokol—the vendor who was ultimately held responsible for the infamous failed O-ring in the booster rocket—dropped precipitously within minutes after the explosion, and much more so than any other vendors' stock prices during the same period. In our own experiments, post-STOC-trading surveys of the traders (Figure 1) reveal that the majority base their trades more on what is happening during the game than on their underlying preferences for the product concepts, and that their expectations of others changed a lot while trading.

Alternatively, individual traders may form their own beliefs about the products, acknowledging that market prices will depend on aggregate beliefs. This is similar to the information aggregation scenario in which there are no "insiders", but where all traders are partially informed. Even in this case, where no single trader has full information, an RE equilibrium will be reached under very general conditions (Grossman, 1981; Davis and Holt, 1993, Chapter 7).

However, there is one important difference between STOC markets and the other exchanges in the experimental markets literature. In a typical experimental market, subjects' preferences and their information set are fixed and assigned by the researchers. Therefore, even before trading begins, theoretical equilibrium prices can be calculated. In contrast, in a STOC market, neither the subjects' preferences nor their information sets are known—in fact, these are what STOC market trading experiments are meant to discover. This suggests an important practical consideration in implementing STOC markets: the composition of traders should match the population of target consumers as closely as possible, or at least include traders with insight into the preferences of these consumers. For example, if the target population for a particular product is teenage female consumers, a STOC market consisting of middle-age males may not

yield particularly useful preference rankings for that product. However, if the cross section of traders in a STOC market is representative of the target population, the force of market rationality will ensure that the price-discovery mechanism will provide an accurate measure of aggregate preferences. We demonstrate the importance of selecting representative traders in our experiments on Wii video games by having two groups of traders with distinct preferences trade the identical (8) eight video game concepts, with the outcomes differing dramatically.

3.4 Some Possible Challenges Related to STOC and Markets in General

It is important to note some of the possible downsides related to STOC and to markets in general, however. STOC requires simultaneous participation by a group of subjects, the size of which grows linearly in the number of concepts being tested. It also requires specialized trading system software and communications infrastructure. The methodology is laid out in detail in the web appendix. Despite the proven accuracy of prediction markets forecasting election outcomes, movie box office receipts, and sporting events, just to name a few, healthy skepticism about securities markets remains. This skepticism has only been reinforced by the perceived market failures of the last few years, all of which suggests an uphill battle for the diffusion of market-based methods in general, and STOC in particular.

Of course, in general, market-based methods for eliciting information also have certain limitations. Unlike typical marketing research techniques in which information is collected from individuals and aggregated in subsequent analysis, the market method focuses on aggregate beliefs and preferences. While individual heterogeneity is not easily observed from market prices, it does enter the trading process in the form of differences in security valuation. Later we propose a metric of trader bias and show that individual preferences may be recovered using the fact that STOC security prices are stationary. Moreover, virtual concepts markets may be

vulnerable to price manipulations and speculative bubbles because the values of virtual securities hinge on the aggregate beliefs, which are endogenously determined within the same market. Traders may form false beliefs that could cause prices to deviate from their fundamentals. And all of the behavioral critiques that have been leveled against the Efficient Markets Hypothesis in the financial economics literature (see, for example, Shefrin, 2005) apply to concepts markets as well. For these reasons, the market method must be applied with caution, and the consistency of the results must be checked through repeated STOC market tests of the same concepts or other means of validation. The greatest level of vulnerability may occur when traders have a poor sense of their own preferences or of those of other people. This might occur, for example, when the product category is too new for traders to grasp, or when the stimuli shown prior to trading are unclear or confusing (as we demonstrate in one experiment shortly). Another drawback, as already touched on briefly above, is that STOC publicly exposes the tested concepts to a group of respondents, which may be problematic in the case of sensitive, proprietary intellectual property. While other concept testing methods have similar problems, the interactive nature of trading may make it more difficult to contain the problem of information leaks.

4 Testing the STOC Methodology

To validate the STOC method, we compare it against alternative methods of measuring preferences for new product concepts. We present results for eleven STOC experiments, summarized in Table 2, in four product categories, the product attributes of which appear in: Figure 1: Crossover Vehicle Trader Attitudes Before and After Trading ($n = 77$)
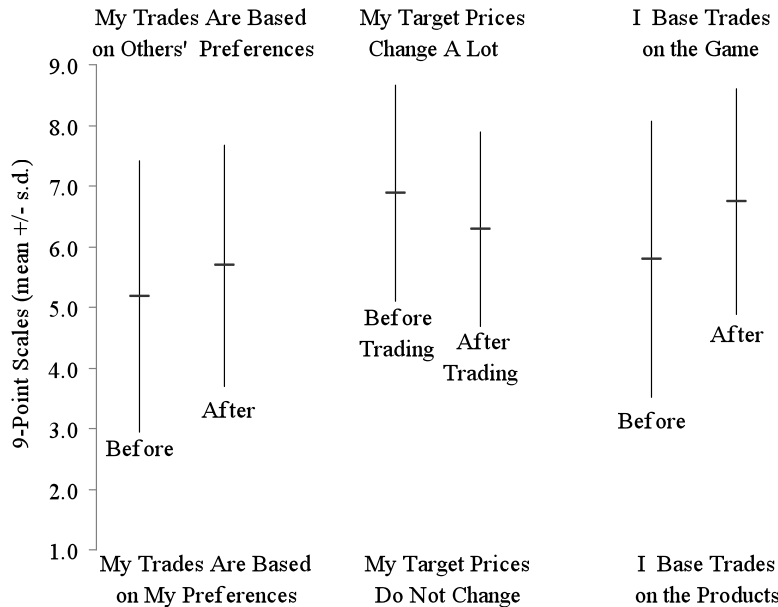
Figure 21) bicycle pumps (Dahan and Srinivasan, 2000), 2) laptop computer messenger bags (Toubia, et. al., 2003), 3) crossover vehicles (Dahan and Hauser, 2002), and 4) Wii video games.

Insert Table 2 here

Traders in seven of the tests were MBA and Executive MBA students, but four additional STOC tests included management level attendees from the MIT Center for Electronic Business conference (crossover vehicle test 3), senior global executives attending executive education classes (crossover vehicle test 4) and UCLA Executive MBA's (Wii video games in tests 2 and 3). All eleven tests were run under controlled conditions in a business school trading laboratory or corporate classroom.

The resulting securities prices, normalized to sum to 100 in each experiment, are compared against six stated-choice and revealed preference approaches: rank-ordered choice, conjoint analysis, virtual concept testing, actual purchase decisions in a simulated store and in actual automotive markets, and constant-sum voting. We find that results across different market experiments in the four product categories are reliable across repeated tests and predictive of stated-choice and revealed preference results, but less so of longitudinal sales data.

More specifically, we begin our experiments in 2000 by conducting two STOC tests of Dahan and Srinivasan (2000)'s bicycle pump concepts, depicted as static images resembling the physical prototypes examined by their respondents.

Insert Figure 3

We employ their physical prototype, conjoint analysis, and static web virtual concept test results for bicycle pumps in hopes of validating the STOC method. Both bicycle pump STOC experiments were run on the opposite coast several years after the original data were collected, and were conducted with the *same* group of traders as a method of measuring test-to-test repeatability.

Toubia, et. al's (2003) laptop PC messenger bags offer an excellent data set for validating STOC, as their data set consists of customizable laptop PC messenger bags sold for real money through a simulated store to 330 MBA students. We focus on eight randomly chosen bags, representing a range of popularity (choice share), that were actually "sold" to 43% of the respondents in their research. To test whether STOC could predict the simulated store choice shares, two STOC tests were run to measure preferences for the same eight bags, but employing two different forms of stimuli: the table shown in Figure 6 and the individual images shown in Figure 7.

17

Beginning in 2000, we also tested eight crossover vehicles, as in Dahan and Hauser (2001), in four STOC experiments. Three crossover vehicles had already been released (Lexus, Mercedes, and BMW) and five were yet-to-be-released (Pontiac, Acura, Buick, Audi and Toyota) at the time of the STOC experiments. In the six following years, we collected unit sales data for each of the eight vehicles from *Ward's Automotive News*. These data are used as a test of external validity and the predictive power of the STOC method.

Finally, in 2009, we conducted three STOC experiments using 19 Wii video game concepts developed by student teams competing with each other in two UCLA MBA courses. Each video game concept included game play software as well as at least one piece of original hardware, i.e. a peripheral device that would enhance game play and work with the Wii console's motion-detecting, force-sensing and/or WiFi features.

Each Wii video game concept was also defined by its feature levels for six conjoint attributes in a study completed by each of the 90 students as well as 160 outside respondents as seen in Figure 1: Crossover Vehicle Trader Attitudes Before and After Trading ($n = 77$)
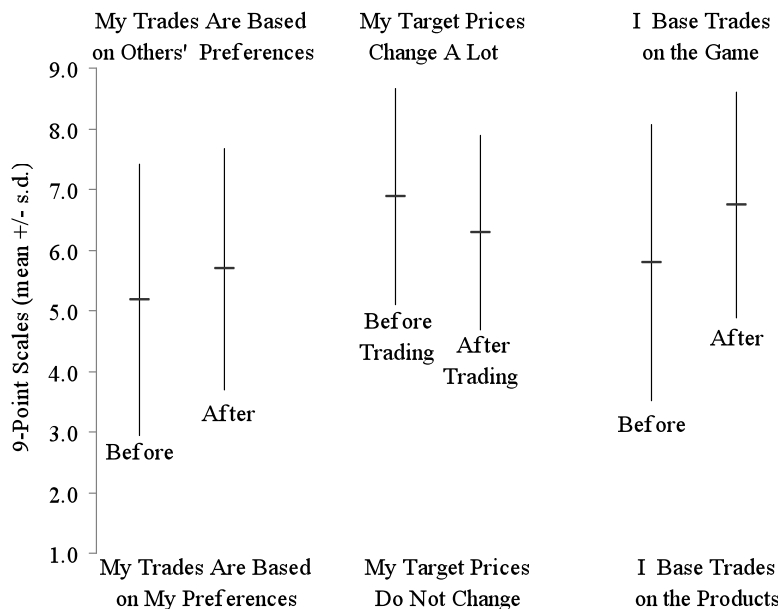
Figure 2.

The concepts proposed by the 35 students competing on 8 teams in an elective course on *New Product Development* are depicted in Figure 8, while those proposed by 55 Executive MBA students comprising 11 teams in a core Marketing course are shown in Figure 9.

Insert Figure 8 and Figure 9

After presenting their concepts to each other at the final session of their respective courses, the students used constant-sum voting to complete three surveys: (1) SELF Preferences for all concepts except their own (which determined project grades); (2) Expectations of Others' Preferences for all eight concepts (with recognition for the "best guessers"); and (3) expectations of the eight average STOC Prices just prior to trading. After trading was completed, the students completed a fourth, constant-sum "Post-STOC" survey utilizing what they had learned from trading to re-estimate the mean preferences for all eight Wii Video Game concepts.

Wii video game Test 2 was identical to Test 1, except that the 55 Executive Education students completing the four surveys and playing the STOC game judged the eight concepts from the *other* class. That is, they had no involvement in the development of those eight concepts. In Test 3, the Executive MBA students evaluated the 11 game concepts that their own teams had developed.

We are grateful for the cooperation of the aforementioned researchers who enabled us to utilize the identical product concept illustrations in our STOC tests. Thus, we are able to compare results for identical market research problems using STOC versus prior methods.

For the first eight STOC experiments, traders were told that their final portfolio valuations would be based on the closing prices of each stock, plus the cash they had left on hand. For the final three experiments testing Wii video game concepts, traders were told that

final portfolio valuations would be based *not* on closing prices, but rather on *mean* prices throughout the trading period (i.e. using volume-weighted-average-prices). A common user interface, shown below, was employed for each of the eleven STOC tests, each of which ran in under one hour including instructions and wrap up.

Insert Figure 11

The STOC tests allow us to confirm hypotheses about: (1) the validity of STOC's preference measurements, (2) whether STOC prices measures individual preferences or traders' expectations about others' preferences, and (3) whether STOC reveals individual preferences even when traders focus on other people. Six hypotheses testing STOC's validity include:

H2a. *Convergent Validity with VCT:* Web-based virtual concept testing (VCT) and constant-sum voting are highly correlated with STOC prices.

H2b. *Convergent Validity with Conjoint Analysis:* Conjoint Analysis and STOC results are highly correlated.

H2c. *Test-to-test Repeatability:* Repeated STOC test results are highly correlated when run with the same traders or two groups of traders with similar preferences, but uncorrelated when run with groups of traders with distinct preferences.

H2d. *Simulated Store Predictive Validity:* STOC results accurately predict simulated store sales.

H2e. *Actual Market Share Predictive Validity:* STOC results accurately predict actual market shares.

H2f. *Product Price Insensitivity:* STOC results correlate more highly with VCT results when product prices and price sensitivity are not included.

The final three Wii video game experiments also allow us to test whether STOC prices more completely reflect individual self-preferences or rather individuals' expectations of others, as summarized in the following two hypotheses.

H3a. *STOC prices reflect E[Others]:* STOC prices *correlate more closely with traders' expectations of others (and with their prior expectations of STOC prices) than with traders' self-preferences.*

H3b. *Learning:* Traders' estimates improve and converge after STOC trading.

Finally, despite the fact that STOC prices capture traders' consensus preferences, we can still test whether trading reveals individual preferences, since individual traders are subject to false consensus effects, and may bias their beliefs about other people based on self-preferences. In this instance, biases are a good thing since they may reveal preferences, which is STOC's primary goal.

H4. *Individual Trading Bias:* Traders exhibit bias *based* on self-preferences when trading, therefore individual trading reveals individual self-preference.

5   Results of STOC Tests

Before reporting the STOC test results, we need to identify the metric that best summarizes trading. Volume-Weighted-Average-Price (VWAP) which capture all of the trades from start to finish during a trading game, weighting each trade price by the number of shares traded, correlates to the validation data better and more consistently than the five alternative metrics shown in Figure 4 (median, low price, mean, closing price, and high price). Therefore, in subsequent reporting we utilize VWAP, normalized to sum to 100, when analyzing STOC results.

Insert Figure 4 here

VWAP's observed superiority is consistent with Shin and Dahan (2010) in which they model whether STOC market prices conform to a stationary or non-stationary process. They analyze the same trading data as in the present research, and employ unit-root tests to verify the

21

stationarity or lack thereof of the mean prices for each security, and find that STOC security prices are, in fact, stationary. Their model is:

$$(2) \qquad P_{i,t} = \alpha_i + \gamma_i \cdot DIFF_{i,t-1} + e_{i,t}, \; e_{i,t} = \rho \cdot e_{i,t-1} + \upsilon_{i,t}, \; \upsilon_{i,t} \sim w.n.$$

where α is an intercept term for each stock, *DIFF* is the difference between the current stock price, *P,* and that stock's volume-weighted average price (VWAP) *up to* time *t,* and the errors, the *e*'s, are autoregressive and follow a white noise process (*w.n.*), i.e. a sequence with zero mean, constant variance, and no serial correlation (Enders 2004).

In Shin and Dahan's empirical analysis, the α_i's for each stock are highly correlated to our VWAP values, and their results support the conclusion that an ideal STOC metric should include all trades. Importantly, applying Shin[4] and Dahan's test of stationarity to all of our STOC data, we find that all of our trading experiments pass the stationarity test at the 95% confidence level. That is, they do not follow a random walk in which the next trade is equally likely to be above the current stock price or below it. In fact, their model correctly predicts whether the price of each stock will rise or fall in the next trade 72.1% of the time vs. a 50-50 guess. This stationarity occurs because there is no outside information entering the market during a STOC game, only the individual preferences and beliefs about others that exist at the time trading begins, and what is revealed through the price mechanism as the game unfolds. One way of interpreting this result is that STOC works like traditional market research in which trades represent a random sample drawn from a stationary distribution. We will utilize this important stationarity result when analyzing the data at the individual trader level for Hypothesis 4 by comparing each trade made against these stationary VWAP's.

---

[4] We thank Professor Hyun Shin for his assistance with these analyses.

Table 4 summarizes the hypothesis testing based on results from all eleven STOC experiments. Ten of the eleven hypotheses are confirmed with remarkably strong results.

Insert Table 4 here

Hypothesis 2 Confirmed

In the majority of cases, STOC correlates well with virtual concept testing and constant-sum voting (H2a confirmed), self-stated preferences and conjoint analysis (H2b confirmed). Most of the experiments showed good test-to-test reliability (H2c confirmed), especially with similar or identical groups of traders, but there were two notable exceptions. The fourth crossover test, with only 16 global executives trading, did not correlate well with the other three STOC tests. Similarly, the second Wii video game STOC test had virtually no correlation with the first Wii video game test, even though both tested the same eight products. The most likely explanation is that the two groups of traders had wildly different preferences (as evidence of the difference between the two groups, we note that they had a correlation of only 0.02 between their average self-preferences). So, while H2c is confirmed and we believe that STOC is reliable and repeatable, we caution that STOC measures consensus preferences of that *particular* group of traders, so as with other market research methods, a representative sample is required.

The results for Hypothesis H2d, that STOC correlates with simulated store sales, seem equivocal because the first test depicting laptop bags in a tabular format failed while the second test depicting the same bags in simpler images succeeded remarkably well. The difference between the two tests was only in the stimuli employed, not the underlying product concepts. Simply put, the tabular format for the eight Laptop PC bags failed miserably, while the simple images of each bag and its featured were much better understood. So H2d is confirmed, given

23

effective stimuli. The lesson is that stimuli matter in STOC, as in other market research methods, and pre-testing of stimuli is advised.

Insert Table 5 here ##

While the success of STOC in predicting market shares in a simulated store suggests that the method could be used in forecasting potential real world sales of product concepts, we urge caution. Simulated stores control for many marketing mix variables, such as advertising, brand equity, and product availability, that are not controlled in actual product markets. The next result highlights the risks of using STOC to predict actual sales.

The only rejected hypothesis is H2e, i.e. that STOC predicts actual market shares, and it is convincingly rejected based on all four crossover vehicle experiments. The failure of STOC to predict actual vehicle market share makes intuitive sense when one considers that STOC measures vehicle *preferences* without respect to vehicle price, not what respondents would *actually buy* given budget and product availability constraints. This result highlights the earlier distinction between prediction markets, which measure expectations about future observable *outcomes*, versus STOC, which measures *preferences* for product concepts.

Confirming Hypotheses H2a, H2b and H2c the results from the two bike pump experiments show a remarkable agreement with those from the Dahan and Srinivasan study despite fundamental differences between the two methods, and wide separation in time and location. The detailed results in Table 5 reveal a high degree of correlation between both STOC tests and Dahan and Srinivasan's original bicycle pump validation data. Table 7 calculates the mean absolute errors between any two sets of choice share predictions, and shows that STOC fared well in measuring preferences based on the Web static images.

Insert Table 6 and Table 7 here ##

Differences include those in the data collection mechanism (a virtual security market in the STOC case, versus a virtual shopping experience in the Dahan and Srinivasan study), the modeling of the predicted market share (the use of relative security prices with STOC versus individual level conjoint analysis in the other study), the questions asked (*the group* prefers in STOC, versus what *you* prefer in the others study), and lastly the subject population (Stanford students of all types with Dahan and Srinivasan study and MIT MBA students with STOC).

Table 8's crossover test results also confirm H2a and H2c. "Self-stated" survey data represent the normalized market shares for vehicles ranked by each individual in the top three out of eight choices. "VCT w/Prices" represent vehicle market shares based on scoring in the top three of eight vehicles using Dahan and Srinivasan's 2000 methodology and accounting for the price of each vehicle when calculating utility. "VCT NO Prices" is the same calculation, but based only upon vehicle preferences *without accounting for vehicle prices* in the utility calculations. And, as before, "STOC" represents the normalized market shares based on the volume-weighted average prices of each of the eight securities.

Insert Table 8 here

Two notable results follow from this analysis:

1. **H2e is rejected: STOC failed to predict actual vehicle sales**: The first row of Table 8, in which correlations to actual 2001–2006 unit sales of the eight vehicles were calculated for each method, reveals that all four STOC tests did *not* predict actual sales, so Hypothesis H2e is rejected. This result confirms our point that STOC markets are not prediction markets, but rather measure underlying preferences among the traders, not intent to purchase. On the other hand, self-stated choices and virtual

concept testing with pricing *did* predict actual unit sales (correlations in the 0.42 to 0.63 range), though the results are not statistically significant. The superiority of VCT and self-stated choice over STOC derives from the distinction between what people *prefer*, and what they are *willing to pay*. STOC zeroes in on preference rather than willingness-to-pay. Also, vehicle prices were not highlighted in STOC tests 1, 2 and 4, and were only emphasized to traders during test 3, which was the only STOC test with at least some predictive value (0.52, but again not significant).

2. **STOC captures preferences well, but not price sensitivity:** All four crossover tests reveal that STOC accurately measured vehicle preferences without considering vehicle pricing. Consider the crossover vehicle STOC results in Table 8. When vehicle prices are left out of the analysis, STOC correlates remarkably well with the virtual concept test results (correlations of 0.80, 0.97, 0.72 and 0.83, respectively, for Tests 1 through 4). But when vehicle prices and respondent price sensitivity *are* factored in, the correlations between STOC and VCT disappears (correlations of - 0.10 to 0.31). STOC traders focus on vehicle preferences, not on willingness-to-pay. So Hypothesis 2f is confirmed: STOC traders reveal their product concept preferences, but not necessarily their intent to buy nor their price sensitivity.

Hypothesis 3 Confirmed

Results for the first two Wii Video Game tests are shown in Table 9. The high correlations between aggregate individual preferences and E[Others] (0.94, 0.94, and 0.88), and between E[STOC Prices] and STOC VWAP (0.93, 0.94, and 0.63) confirm strong wisdom-of-crowds effects. While many individual respondents and traders do not accurately estimate

others' preferences, aggregating their individual beliefs produces an accurate estimate of aggregate preferences because most of the individuals' errors cancel out.

Insert Table 9 here

While the results for the individual tests are remarkably good, the correlations between test 1 and test 2 for the same eight Wii video game concepts are virtually nonexistent, highlighting the earlier point that differences in trader populations will reflect in STOC results.

The tests also confirm H3a, that STOC measures expectations of others more so than self-preferences. In all three tests, the correlations between E[Others] and volume-weighted-average-STOC-prices (VWAP) are significant and higher than those between VWAP and SELF preferences (e.g. 0.91 vs.0.89 in Test 1, 0.79 vs. 0.69 in Test 2, and 0.66 vs. 0.32 in Test 3). Traders rely more heavily on their beliefs about other traders than on their personal preferences when trading. This is especially evident in Test 3 as seen in Table 10.

Insert Table 10 here

Traders learn from each other, confirming Hypothesis H3b as evidenced by Post-STOC trading results outperforming E[STOC Prices] measured just prior to trading in all three tests.

Comparisons to conjoint analysis for the three Wii Video Game tests are shown in Table 11, and indicate that the competition between STOC vs. Conjoint Analysis is essentially a draw.

Insert Table 11 here

STOC outperforms conjoint analysis in estimating SELF preferences for two of the three Wii Video Game tests, as the values in the lower row of Table 11 exceed those in the upper row for tests 1 and 2. Conjoint analysis outperforms in test 3, although neither method does

particularly well in that test. Overall, having the information from *both* methods would be superior to depending on either one alone, and STOC trading of attributes and levels could act as an input into conjoint analysis when trying to narrow from many possible attributes to a manageable few. Also, STOC may prove more useful in measuring preferences for attributes such as product aesthetics and ease-of-use that are difficult to measure using conjoint analysis.

Hypothesis 4 Confirmed

Consider Table 12's four possible outcomes for any given stock trade between two individuals.

Insert Table 12 here

While traders do not know the VWAP in advance of trading, the researcher can precisely measure those numbers in hindsight. We learned earlier that STOC VWAP's are stationary. If each trader's portfolio is ultimately valued at these VWAP's, then each trade can be viewed retrospectively as having been profitable or not. In fact, we recommend that traders be informed that winners will be determined based on final portfolios priced at VWAP, not closing price, to align incentives and reduce gaming in the closing minutes of STOC competitions. Selling shares above VWAP or buying below VWAP, as shown in the black boxes, would be profitable. But *selling below* VWAP might indicate an "abnormal" dislike for a particular stock. And *buying a stock above* VWAP might reveal a strong preference for the concept associated with that stock. Below, we test these hypotheses empirically by looking at the 50% of transactions in the off-diagonal, white cells. The goal is to detect individual preferences in the trading data.

Appendix 2 develops a metric measuring trader n's bias for stock m:

$$(3) \qquad \frac{VWAP_m + \left(\widetilde{z}_{m,n} \times \widetilde{\sigma}_m\right)}{\sum\limits_{m=1}^{M} VWAP_m + \left(\widetilde{z}_{m,n} \times \widetilde{\sigma}_m\right)},$$

where the $\widetilde{z}_{m,n}$ measures trader n's relative bias for stock m using the volume-weighted trades

made by that person for that stock in the off-diagonal cells of Table 11, and $\widetilde{\sigma}_m$ measures the

volume-weighted variance of stock m's price relative to its volume-weighted mean (VWAP).

We note that each trader's biases for all m stocks sum to 100%, and can therefore be compared to

the constant sum survey results. We can compare the individual trader's bias against known

preferences from the individual-level survey data. The key result, depicted in Figure 13, is that

trading reveals preferences at the individual level.

Insert Figure 13 here

Two primary findings regarding individual level preferences confirm Hypothesis H4:

- **Individuals are subject to false consensus biases**: There are strong positive correlations between the survey of self-preference and each respondent's expectation of others. This implies that most individuals' expectations of others is strongly biased by self-preference, consistent with the false consensus bias. Note that we saw a similar effect at the aggregate level for the Wii game concepts in Table 9 and Table 10, where the correlations between mean self-preferences and mean expectations of others were extremely high and significant (0.94, 0.94 and 0.88, respectively).

- **Individuals trade based on their biases**: More interestingly, individual trader preferences measured with the trading bias metric are positively correlated with the survey of self-preference data (mean $\rho = 0.25$), and even more so with the individuals' E[Others] (mean $\rho = 0.46$), which in turn are highly correlated with self-

preferences (mean $\rho = 0.61$). While STOC was not primarily designed to measure preferences at the individual level, the method reveals traders' personal biases for and against the product concepts being tested. As in conjoint analysis, STOC measures implicit preferences rather than requiring explicitly stated preferences.

Insert Figure 14 here

Figure 14 compares the ability of individuals to estimate others' preferences against the extent to which they rely on those estimates when trading, and reveals that individuals who estimated others better relied more on their own preferences when trading, while traders who poorly predicted others' preferences followed the market rather than trading based on their own beliefs. One explanation for this remarkable result is that market prices tend to confirm the beliefs of those who start trading with accurate estimates, while those whose estimates are far off the mark adjust their trading after learning market prices. That is, they learn from the STOC market, providing further confirmation of Hypothesis 3b.

6   Discussion and Conclusions

In this paper we study a novel application of the market mechanism: the use of securities markets to aggregate and infer diverse consumer preferences. We implement this idea in the specific context of eleven product-concept testing studies that aim to predict potential market share for between eight to eleven product prototypes. During our research, we confirm four hypotheses about STOC: 1) games are more popular and cost efficient than surveys; 2) STOC passes validity tests, but is *not* a prediction market; 3) the method measures consensus expectations of others; and 4) uncovers individual preference too, not just those of the crowd.

We observe that, unlike financial and prediction markets, STOC prices regress towards stationary means due to the absence of outside information. Ten of the eleven STOC experiments are remarkably consistent with several common preference measurement techniques. But stimuli must be clear and salient, and traders need to be trained and properly primed prior to trading. STOC reveals individual heterogeneity through trading biases.

We caution that the STOC methodology was not designed to measure individual preferences. Conjoint analysis is likely more effective in that context, especially in measuring price sensitivity and attribute tradeoffs. However, STOC does reveal individual preferences indirectly. But STOC requires simultaneity and technical infrastructure that other methods would not need, thus posing challenges to its widespread adoption. To STOC market reliability in some marketing applications, one may need to anchor the values of the securities to some objective fundamental variables of the corresponding product concepts. To test market share predictions, for example, one could compare security values with the realized market shares of the subset of products that already exist, or, barring the existence of real market share data, with the outcomes of external customer choice surveys. We hope to refine STOC market methods along these and other lines in future research. Finally, confidentiality needs to be considered while using STOC, as product information is revealed in the process. Experimental design solves this issue using STOC's scalability by making each trader aware of a subset of concepts.

Despite the challenges, STOC offers two primary advantages relative to traditional market research methods:

- **Cost efficiency**: The fact that respondents prefer competing in the STOC game to taking surveys, combined with the efficiency of STOC's market pricing

mechanism in aggregating preferences, reduces recruiting and respondent compensation costs by as much as 75% for any given number of concepts.

- **Scalability**: The ability to conduct a large number of concept tests quickly and simultaneously, limited only by the number of respondents, is particularly beneficial. STOC serves as an early phase screening mechanism for concepts flowing from crowdsourcing, open innovation, or particularly prolific product innovation within the firm during new product development.

The efficacy of STOC markets at identifying winning concepts may not surprise economists. After all, Keynes (1936) commented on the similarities between stock selection and a beauty contest over seventy-four years ago:

...professional investment may be likened to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole ..."

The analogy is perhaps more accurate for describing what happens in financial markets and STOC games in the *short run*. After all, over the long run financial stock prices depend not only on investors' subjective beliefs and expectations of others, but also on other objective information such as companies' earning potentials and valuations of assets. On the other hand, the trading experiments presented in this paper are precisely "beauty contests," since values of the virtual securities are derived endogenously from the preferences of the market participants, and their expectations of others' preferences, both of which are largely subjective.

References

"Box Office Muddle: How Will Films Fare From Week to Week," (1996) *Wall Street Journal*, July 21, 2000. *Business Week*, November 11.

Burchill, G. W. and C. H. Brodie (1997). *Voices into Choices*. (Cambridge, MA: Center for Quality Management.

Calder, B. J. (1977). "Focus Groups and the Nature of Qualitative Marketing Research." *Journal of Marketing Research*, 14, 353-64.

Chan, N. (2001) Ph.D. Thesis: Artificial Markets and Intelligent Agents, February.

Chan, N., E. Dahan, A. Lo and T. Poggio (2001). "Experimental Markets for Product Concepts," CBCL Working Paper #200/AI Memo #2001-013, Massachusetts Institute of Technology, Cambridge, MA, July.

Chen, K. and C. R. Plott (2002), "Information Aggregation Mechanisms: Concept, Design and Implementation for a Sales Forecasting Problem," California Institute of Technology.

Dahan, E., A. Soukhoroukova, and M. Spann (2010), "NEW PRODUCT DEVELOPMENT 2.0: Preference Markets" *Journal of Product Innovation Management*, forthcoming, 40 pp.

____ and J. R. Hauser (2002), "The Virtual Customer," *Journal of Product Innovation Management*, 19 (5), 332-53.

____ and V. Srinivasan (2000), "The predictive power of internet-based product concept testing using depiction and animation," *Journal of Product Innovation Management* 17(2), 99-109.

Davis, D. and C. Holt (1993). *Experimental Economics*. Princeton, NJ: Princeton University Press.

Fama, E. F. (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, 25, 383-417.

____ (1991), "Efficient Capital Markets: II," *Journal of Finance*, 46 (5), 1575-617.

Farmer, D. (2002), "Market force, ecology and evolution", *Industrial and Corporate Change* 11, 895–953.

Farmer, D. and A. Lo (1999), "Frontiers of finance: Evolution and efficient markets", *Proceedings of the National Academy of Sciences* 96, 9991–9992.

Fern, E. F. (1982). "The Use of Focus Groups for Idea Generation: The Effects of Group Size, Acquaintanceship, and Moderator on Response Quantity and Quality." *Journal of Marketing Research*, 9, 1-13.

Forsythe, R. and R. Lundholm (1990), "Information aggregation in an experimental market", *Econometrica*, 58, 309-47.

____, F. Nelson, G. Neumann, and J. Wright (1993), "The Iowa Presidential Stock Market: A Field Experiment," *Research in Experimental Economics*, 1-43.

____, T. Palfrey, and C. Plott (1982), "Asset Valuation in an Experimental Market," *Econometrica*, 50 (3), 537-68.

____, T. Rietz, and T. Ross (1999), "Wishes, Expectations and Actions: A Survey on Price Formation in Election Stock Markets," *Journal of Economic Behavior & Organization*, 39, 83-110.

Green, P. E. and Y. Wind (1981), "New way to measure consumers' judgment," in Y. Wind, V. Mahajan & R. N. Cardozo eds., *New Product Forecasting*, D.C. Heath and Company, Lexington, MA, 89-108.

____ and V. Srinivasan (1990), "Conjoint Analysis in Marketing: New Developments With Implications for Research and Practice," *Journal of Marketing*, 3-19.

Griffin, A. J. and J. R. Hauser (1993), "The Voice of the Customer," *Marketing Science*, Winter, 1-27.

Grossman, S. J. (1981), "An introduction to the theory of rational expectations under asymmetric information," *Review of Economic Studies*, 48, 541-559.

Hanson, R. (2003), "Combinatorial Information Market Design," *Information Systems Frontiers* 5(1) January, 105-119.

Hanson, R. (2007), "Logarithmic Market Scoring Rules for Modular Combinatorial Information Aggregation," *Journal of Prediction Markets*, 1(1) February, 3-15.

Hanson, R. (2009), "On Market Maker Functions," *Journal of Prediction Markets*, 3(1), 61-63

Hanson, R., R. Oprea, and D. Porter (2005), "Information Aggregation and Manipulation in an Experimental Market," *Journal of Economic Behavior and Organization*, 60 (4), 449-59

Hayek, F. (1945), "The Use of Knowledge in Society", *American Economic Review, XXXV,* No. 4; September, 519-30.

Keynes, J.M. (1936), Chapter 12 of *General Theory of Employment, Interest and Money*, London: Macmillan, 401 pp.

Maloney, M. and H. Mulherin (2003), "The Complexity of Price Discovery in an Efficient Market: The Stock Market Reaction to the Challenger Crash", *Journal of Corporate Finance* 9, 453–479.

O'Brien, J. and S. Srivastava (1991), "Dynamic stock markets with multiple assets," *Journal of Finance*, 46, 1811-38.

Pennock, D. M. and M. P. Wellman (1997), "Representing Aggregate Belief through the Competitive Equilibrium of a Security Market," in *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*, 392-400.

Pennock, D. M., S. Lawrence, C. Giles, and F. Nielsen (2001), "The real power of artificial markets," *Science*, 291, 987-988.

Plott, C. R. and S. Sunder (1982), "Efficiency of experimental security markets with insider information: An application of rational-expectations models," *Journal of Political Economy*, 90, 663-98.

____ and ____ (1988), "Rational Expectations and the aggregation of diverse information in laboratory security markets," *Econometrica*, 56, 1085-1118.

Servan-Schreiber, E., J. Wolfers, D. M. Pennock, and B. Galebach (2004), "Prediction Markets: Does Money Matter?," *Electronic Markets*, 14 (3), 243-51.

Shin, H. and E. Dahan (2008), "A time-varying model of securities trading of concepts", UCLA working paper.

Smith, V. L. (1982), "Microeconomic Systems as an Experimental Science," *American Economic Review*, 72 (5), 923-55.

Spann, M. and B. Skiera (2003), "Internet-Based Virtual Stock Markets for Business Forecasting," *Management Science*, 49 (10), 1310-26.

Srinivasan, V. and A. D. Shocker (1973), "Estimating the Weights for Multiple Attributes in a Composite Criterion Using Pairwise Judgments," *Psychometrika*, 38, 4, (December), 473-493.

Sunder, S. (1992), "Market for Information: Experimental Evidence," *Econometrica*, 60 (3), 667-95.

Surowiecki, J. (2004), *The Wisdom of Crowds*. New York: Doubleday.

Toubia, O., D. I. Simester, J. R. Hauser, and E. Dahan (2003), "Fast Polyhedral Adaptive Conjoint Estimation," *Marketing Science*, 22 (3), 273-303.

Urban, Glen L., J. R. Hauser and J. H. Roberts (1990), "Prelaunch Forecasting of New Automobiles: Models and Implementation," *Management Science*, Vol. 36, No. 4, (April), 401-421.

Ward's Automotive News, monthly unit sales data by vehicle, 2001-2006.

Wolfers, J. and E. Zitzewitz (2004a), "Prediction Markets," *Journal of Economic Perspectives*, 18 (2), 107-26.

Appendix 1: Derivation of Equation (1)

Equation (1) can be derived by considering the two primary, people-related costs of running a market research study: (a) those costs associated with *recruiting* potential respondents and (b) those for *compensating* people who actually participate.

$$\text{Total Cost} = \overbrace{\left(\frac{\text{Number of}}{\text{respondents needed}} \times \frac{\text{recruits}}{\text{respondent}} \times \frac{\text{Cost per}}{\text{recruit}}\right)}^{\text{Recruiting Costs}} + \overbrace{\left(\frac{\text{Number of}}{\text{respondents needed}} \times \frac{\text{Compensation}}{\text{per respondent}}\right)}^{\text{Respondent Compensation}}$$

$$= \frac{\text{Number of}}{\text{respondents needed}} \times \left(\frac{\text{Cost per recruit}}{\text{response rate}} + \frac{\text{Compensation}}{\text{per respondent}}\right)$$

$$= \overbrace{\left(\frac{M\ Concepts}{being\ tested} \times \frac{Sample\ size}{required} \div \frac{question\ capacity}{respondent}\right)}^{\text{Number of respondents needed}} \times \overbrace{\left(\frac{Cost\ per\ recruit}{response\ rate} + \frac{Compensation}{per\ respondent}\right)}^{\text{Average cost per respondent}},$$

which we summarize with the following notation:

$$(1)\quad TC = \left(M \cdot \frac{N_{sample}}{q_{respondent}}\right) \times \left(\frac{c_{recruit}}{r\%} + c_{respondent}\right).$$

where,

Symbol	Description
N_{sample}	Number of required Respondents
$q_{respondent}$	question capacity per respondent
$c_{recruit}$	cost to recruit each person
$c_{respondent}$	compensation for each respondent
$r\%$	response rate

Appendix 2: The Trading Bias Metric

Below, we develop a trading-based metric to measure individual biases and the preferences those biases imply.

- Denoting stocks using $m = 1,..., M$ and individual traders using $n = 1,..., N$

- Let $P^i_{m,n}$ be the price of the ith trade of stock m for trader n

- $VWAP_m$: volume-weighted average price for stock m.

- $V^i_{m,n}$: volume (# of shares) of the ith of $K_{m,n}$ trades of stock m for trader n

- $s^i_{m,n}$: side of ith trade (buy or sell) stock m for trader n.

- $I(VWAP_m, P^i_{m,n}, s^i_{m,n})$: indicator function which returns 1 if trader n *bought* shares *above* the $VWAP_m$ or *sold* shares *below* the $VWAP_m$ and 0 otherwise.

$$I(VWAP_m, P^i_{m,n}, s^i_{m,n}) = \begin{cases} 1 : s^i_{m,n} = buy, P^i_{m,n} > VWAP_m \quad or \quad s^i_{m,n} = sell, P^i_{m,n} < VWAP_m \\ \\ 0 : otherwise \end{cases}$$

- $\tilde{\sigma}_m$: volume weighted standard deviation of stock m, which equals

$$= \sqrt{\sum_{i=1}^{K_{m,n}} \sum_{n=1}^{N} \left[\left(P^i_{m,n} - VWAP_m \right) \times I\left(VWAP_m, P^i_{m,n}, s^i_{m,n} \right) \right]^2 \times V^i_{m,n} \Bigg/ \sum_{i=1}^{K_{m,n}} \sum_{n=1}^{N} V^i_{m,n} \times I\left(VWAP_m, P^i_{m,n}, s^i_{m,n} \right)}$$

- $\tilde{z}_{m,n}$: volume weighted z-score of stock m for trader n, which equals

$$= \sum_{i=1}^{K_{m,n}} \frac{P^i_{m,n} - VWAP_m}{\tilde{\sigma}_m} \times V^i_{m,n} \times I\left(VWAP_m, P^i_{m,n}, s^i_{m,n} \right) \Bigg/ \sum_{i=1}^{K_{m,n}} V^i_{m,n} \times I\left(VWAP_m, P^i_{m,n}, s^i_{m,n} \right)$$

We now define trader n's bias for stock m as the following standardized metric:

$$(2) \qquad \frac{VWAP_m + \left(\tilde{z}_{m,n} \times \tilde{\sigma}_m \right)}{\sum_{m=1}^{M} VWAP_m + \left(\tilde{z}_{m,n} \times \tilde{\sigma}_m \right)}.$$

Note that each trader's biases for all m stocks sum to 100%, and can therefore be compared to the constant sum survey results.

Table 1: Five Routes to Lower Respondent Costs in Concept Testing

Symbol	Description	Challenge	To Lower Cost	STOC's potential Benefit
N_{sample}	Number of required Respondents	Statistical power	Reduce N_{sample}	Lower due to *interactions* with others and *multiple answers* per respondent
$q_{respondent}$	question capacity per respondent	Bounded rationality	Increase $q_{respondent}$	Higher due to *motivation* and ability to self-select questions to be traded
$c_{recruit}$	cost to recruit people	People avoid surveys	Reduce $c_{recruit}$	Lower because recruits are *attracted* to playing the game, even multiple times
$c_{respondent}$	compensation for respondents	People value their time	Reduce $c_{respondent}$	Lower due to the intrinsic pleasure of playing the *game* itself
$r\%$	response rate	Many people opt out	Increase $r\%$	Higher due to *intrinsic* pleasure of game, desire to play again, *competitiveness*

Table 2: Data Collected for each of Four Product Categories

Method / *Product type*	Experiment	STOC Method	Conjoint Analysis	Virtual Concept Test	Self-Stated Choices	Simulated Store	Longitudinal Sales Data
Bike Pump Concepts	Tests 1 & 2 $n = 28$	9 Pumps; Same traders tested twice	Rank 18 full profiles, est. 10 parameters w/LINMAP $n = 141$	Dahan and Srinivasan '00 VCT Physical, VCT Web; $n = 102, 87$			
Actual Laptop Bags	Test 1 $n = 50$	*Table* of 8 Laptop Bags				Toubia, et. al. 2003 unit shares for 8 bags sold in the simulated store; $n = 143$	
	Test 2 $n = 62$	*Images* of 8 Laptop Bags					
Actual Crossover Vehicles	Test 1 $n = 49$	8 vehicles *No* Prices		VCT with and without Prices	Top 3 of 8 with prices		Cumulative units sold for each of 8 vehicles from 2001-2006 per *Ward's Automotive News*
	Test 2 $n = 43$	8 vehicles *No* Prices		VCT with and without Prices	Top 3 of 8 with prices		
	Test 3 $n = 42$	8 vehicles *With* Prices		VCT with and without Prices	Top 3 of 8 with prices		
	Test 4 $n = 16$	8 vehicles *No* Prices		VCT with and without Prices			
Wii Video Game Concepts	Test 1 $n = 35$	8 Own Wii Video Games	Rank 16 full profiles, est. 10 parameters w/ LINMAP $n = 35$ & 65	Constant Sum Allocation of 100 Points across 8 or 11 Wii Games in (4) Surveys: • SELF Preferences • E[Others' Preferences] • E[STOC prices] • E[Actual Share] after STOC game			
	Test 2 $n = 55$	8 Others' Wii Video Games					
	Test 3 $n = 58$	11 Own Wii Video Games					

Table 3: Five Key Steps to Designing and Executing a STOC Experiment

Step	Description	Challenges	Key Considerations
1	Choose STOC Concepts	Narrowing from many options	Stocks should clearly and concisely depict multiple product concepts that differ from each other. Not every trader has to see every stock.
2	Define STOC Prices	Open-ended vs. precise definition	Traders need to understand the definition of each stock, e.g. "the % of people who prefer this concept" or "market share of this product"
3	Define & Teach Trading Method	Programming & User Interface	The user interface should be easy-to-use, informative about the trading activity for each security and trader performance
4	Trading & Data Collection	Need Simultaneous trading; Trader Errors	Transaction details between any two traders needs to be recorded: security name, volume, price, timing. Traders should be able to review, edit and cancel open orders that have not cleared.
5	Data Analysis	Choosing a metric; What is measured?	The metric should include all information such as the number of shares traded and at which price (not just closing prices)

Table 4: Summary of Eleven Experiments and Hypotheses

Hypotheses	H1. STOC Benefits		H2. STOC Validity Testing — VCT, Conjoint, Test-test Reliability, Simulated Store, Actual Sales, Price Sensitivity						H3. What does STOC Measure?		H4. Measuring Individuals
Eleven Experiments	**H1a** STOC requires fewer respondents	**H1b** STOC preferred over surveys	**H2a** STOC correlates with VCT or Const. Sum	**H2b** STOC correlates with Conjoint	**H2c** STOC tests are repeatable	**H2d** STOC correlates w/simulated store	**H2e** STOC predicts actual car sales	**H2f** STOC measures preferences, not pricing	**H3a** STOC measures Expectations of Others	**H3b** STOC traders learn from each other	**H4** STOC measures individual preferences
Bike Pump 1	✔ (n = 18)		✔ $\rho = 0.81$	✔ $\rho = 0.75$	✔ $\rho = 0.9$						
Bike Pump 2			✔ $\rho = 0.89$	✔ $\rho = 0.83$							
Crossover 1	✔ (n = 49, 43, 42, 18)		✔ $\rho = 0.80$		✔ $\rho \in (0.7, 0.9)$		✘ $\rho = 0.22$	✔ $\rho = 0.8 > -0.1$			
Crossover 2			✔ $\rho = 0.97$		✔ $\rho \in (0.7, 0.9)$		✘ $\rho = 0.03$	✔ $\rho = 0.9 > -0.1$			
Crossover 3			✔ $\rho = 0.72$		✔ $\rho \in (.69, .73)$		✘ $\rho = 0.52$	✔ $\rho = 0.6 > 0.5$			
Crossover 4			✔ $\rho = 0.83$		✘ $\rho \in (0.2, -0.6)$		✘ $\rho = -0.40$	✔ $\rho = 0.7 > -0.2$			
Laptop Bag 1 (Table format)	✔ (n = 16, 26)					✘ $\rho = -0.05$					
Laptop Bag 2 (Image format)						✔ $\rho = 0.80$					
Wii Game 1	✔ (n = 33, 51, 56)		✔ $\rho = 0.91$	✔ $\rho = 0.77$	✘ $\rho = 0.55$				✔ $\rho =.91 >.89$	✔ $\rho =.98 >.93$	✔ $\rho_{mean} = 0.41$
Wii Game 2			✔ $\rho = 0.79$	✔ $\rho = 0.75$					✔ $\rho =.79 >.69$	✔ $\rho =.99 >.94$	✔ $\rho_{mean} = 0.30$
Wii Game 3			✔ $\rho = 0.66$	✘ $\rho = 0.44$					✔ $\rho =.66 >.32$	✔ $\rho =.96 >.63$	✔ $\rho_{mean} = 0.11$
Executive (n=78) **and MBA** (n=78) **Survey**		✔ 90%, 85% prefer STOC									

40

Table 5: Correlations Between (2) Laptop Bag STOC Tests
and Simulated Store Unit Sales

	Test 2 STOC Image Format	Simulated Store Sales
Test 1 STOC Table Format	-0.14	-0.05
Test 2 STOC Image Format		**0.80**

Significance Level: **p < 0.05

Table 6: Correlations Between (2) Bicycle Pump STOC Tests
and Validation Data From Dahan and Srinivasan (2000)

	Conjoint Analysis	Web Static Images	Test 1 STOC	Test 2 STOC
Physical Prototypes	0.99****	0.99****	0.75**	0.82***
Conjoint Analysis		0.98****	0.75**	0.83***
Web Static Images			0.81***	0.89***
Test 1 STOC				0.86***

Significance level: **p<0.05, ***p<0.01, ****p<0.001

Table 7: Mean Absolute Error (MAE) Between (2) Bicycle Pump STOC Tests
and Validation Data From Dahan and Srinivasan (2000)

	Conjoint Analysis	Web Static Images	Test 1 STOC	Test 2 STOC
Physical Prototypes	5.6%	1.7%	7.7%	7.2%
Conjoint Analysis		5.3%	5.7%	5.4%
Web Static Images			8.0%	7.3%
Test 1 STOC				1.4%

Table 8: Correlations Between (4) Crossover Vehicle STOC Tests and Validation Data from Actual Unit Sales, Self-Stated Choices, and Virtual Concept Tests

Hypothesis H2e is rejected based on STOC;s failure to predict actual sales in the top row.

	Test 1 Self-Stated	Test 1 VCT w/Prices	Test 1 VCT NO Prices	Test 1 STOC	Test 2 Self-Stated	Test 2 VCT w/Prices	Test 2 VCT NO Prices	Test 2 STOC	Test 3 Self-Stated	Test 3 VCT w/Prices	Test 3 VCT NO Prices	Test 3 STOC	Test 4 VCT w/Prices	Test 4 VCT NO Prices	Test 4 STOC
Actual Units Sold 2001-2006	0.44	0.58	-0.2	0.22	0.42	0.63*	-0.1	0.03	0.62*	0.48	-0.0	0.52	0.52	-0.4	-0.4
Test 1 Self-Stated		0.54	0.54	0.62*	0.89***	0.79**	0.55	0.64*	0.91***	0.37	0.63*	0.90***	0.74**	0.19	0.29
Test 1 VCT w/Prices			-0.1	-0.1	0.32	0.91***	-0.2	-0.1	0.55	0.97****	-0.0	0.50	0.94****	-0.3	-0.2
Test 1 VCT NO Prices				0.80**	0.44	0.14	0.96***	0.91***	0.51	-0.2	0.95****	0.62*	0.10	0.81**	0.72**
Test 1 STOC					0.58	0.20	0.85***	0.92****	0.61	-0.2	0.90***	0.69*	0.07	0.65*	0.62*
Test 2 Self-Stated						0.59	0.54	0.66*	0.84***	0.12	0.48	0.80**	0.49	0.00	0.07
Test 2 VCT w/Prices							0.08	0.20	0.81**	0.81**	0.27	0.79**	0.95****	-0.0	-0.0
Test 2 VCT NO Prices								0.97****	0.52	-0.4	0.93****	0.62*	0.00	0.76**	0.66*
Test 2 STOC									0.65*	-0.3	0.93****	0.73**	0.07	0.68*	0.57
Test 3 Self-Stated										0.35	0.61	0.97****	0.72**	0.12	0.07
Test 3 VCT w/Prices											-0.2	0.31	0.85***	-0.3	-0.2
Test 3 VCT NO Prices												0.72**	0.19	0.83***	0.73**
Test 3 STOC													0.67*	0.31	0.19
Test 4 VCT w/Prices														-0.1	-0.0
Test 4 VCT NO Prices															0.83**

All four STOC tests correlate highly with VCT when vehicle pricing is ignored, confirming Hypothesis H2f.

Significance level: *p<0.10, **p<0.05, ***p<0.01, ****p<0.001

Table 9: STOC Prices for (8) Wii Game Concepts vs. Constant Sum Surveys of SELF Preferences, E[Others' Preferences], E[STOC Prices], and Post-STOC Estimates in Two Experiments

	Test 1 E[Others]	Test 1 E[STOC]	Test 1 Post	Test 1 STOC VWAP	Test 2 SELF Pref	Test 2 E[Others]	Test 2 E[STOC]	Test 2 Post	Test 2 STOC VWAP
Test 1 (n=35) SELF Preferences	0.94****	0.92***	0.90***	0.89***	0.02	0.09	0.24	0.37	0.39
Test 1 E[Mean(Others)]		0.98****	0.90***	0.91***	0.28	0.35	0.47	0.57	0.59
Test 1 E[STOC Prices]			0.92****	0.93****	0.29	0.40	0.55	0.64*	0.67*
Test 1 Post-STOC				0.98****	-0.0	0.14	0.35	0.51	0.56
Test 1 STOC VWAP					0.00	0.15	0.33	0.50	0.55
Test 2 (n=55) SELF Preferences						0.94****	0.84***	0.72**	0.69*
Test 2 E[Mean(Others)]							0.93****	0.81**	0.79**
Test 2 E[STOC Prices]								0.95****	0.94****
Test 2 Post-STOC									0.99****

Significance level: *$p<0.10$, **$p<0.05$, ***$p<0.01$, ****$p<0.001$

Table 10: STOC Prices for (11) Wii Video Game Concepts vs. Constant Sum Surveys of SELF Preferences, E[Others' Preferences], E[STOC Prices], and Post-STOC Estimates

	Test 3 E[Others]	Test 3 E[STOC]	Test 3 Post	Test 3 STOC VWAP
Test 3 (n=58) SELF Preferences	0.88****	0.92****	0.47	0.32
Test 3 E[Mean(Others)]		0.97****	0.75***	0.66**
Test 3 E[STOC Prices]			0.74***	0.63**
Test 3 Post-STOC				0.96****

Significance level: **$p<0.05$, ***$p<0.01$, ****$p<0.001$

Table 11: Predicting Aggregate SELF preferences: STOC vs. Conjoint Analysis
r-Squared, and Rank Correlation for 3 Tests (**winners** in bold)

	Test 1 $(M = 8, N = 35)$		Test 2 $(M = 8, N = 55)$		Test 3 $(M = 11, N = 58)$	
	r^2	Rank	r^2	Rank	r^2	Rank
Conjoint Analysis	0.60	0.69	0.06	0.14	**0.30**	**0.61**
STOC VWAP	**0.80**	**0.93**	**0.48**	**0.88**	0.11	0.43

Table 12: Trades between individuals are profitable for one person and unprofitable for the other; unprofitable trades reveal individual preferences and trading biases

	Buy a Stock	Sell a Stock
Below VWAP	Profitable Trade	Reveals Dislike
Above VWAP	Reveals Preference	Profitable Trade

Figure 1: Crossover Vehicle Trader Attitudes Before and After Trading ($n = 77$)

My Trades Are Based on Others' Preferences

My Target Prices Change A Lot

I Base Trades on the Game

9-Point Scales (mean +/- s.d.)

Before / After

Before Trading / After Trading

Before / After

My Trades Are Based on My Preferences

My Target Prices Do Not Change

I Base Trades on the Products
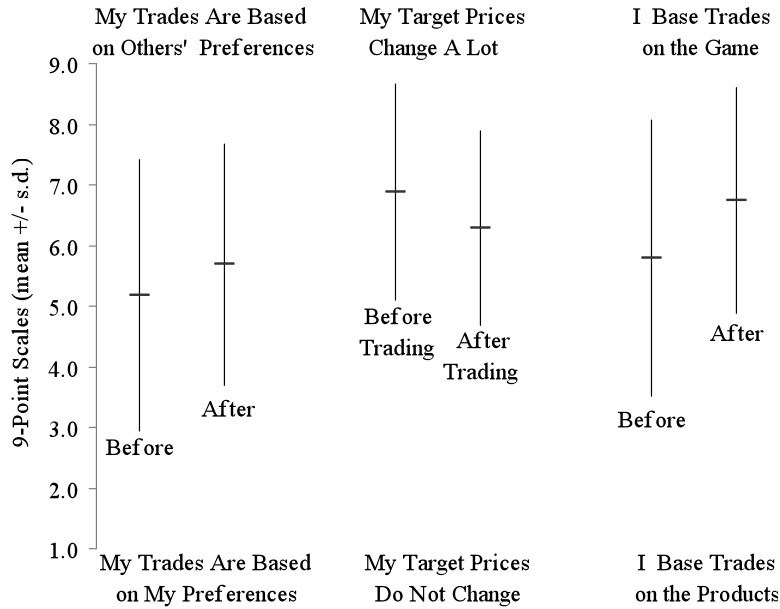
Figure 2: Attributes for Bike Pumps, Crossover Vehicles, Laptop Bags, and Wii Video Games

Our Rating System

Black   Red

Worst   Best

Fast
How long to inflate a tire?
10 minutes → 0.5 minutes

Compact
Does it fit into a small space?
bookbag → pocket

Easy
Who can use this pump?
athletic adults only → even 6-yr-olds

Durable
How durable is it?
barely adequate → indestructible

Our Rating System

Worst   Best

Seating Flexibility How easy is it to reconfigure seating and storage?
Difficult / Fixed → Easy / Very Flexible

Cargo Volume How much storage capacity in cubic feet (min/max)?
15 / 50 → 50 / 100

Fuel Economy What is the EPA gas mileage (city / highway)?
13 City / 17 Hwy → 20 City / 25 Hwy

Horsepower What is the engine's horsepower rating?
150 hp → 300 hp

0-60 Seconds How long to accelerate from 0 to 60 mph?
10 sec. → 6 sec.

Towing Capacity How many pounds can you tow?
2,000 lbs. → 6,000 lbs.

1. Larger size — Adds $9
2. Red color — No price change
3. Logo — Adds $4
4. Handle — Adds $6
5. PDA holder — Adds $6
6. Cell holder — Adds $7
7. Mesh pocket — Adds $4
8. Flap closure — Adds $7
9. Bottom boot — Adds $7

Game Genre: Learn Experience Perform Work Out

# of Players: 1   2-4

Reviews: 52 Unfavorable 73 Mixed 94 Favorable

Effort: Relaxed   High Energy

Online: No   Yes

Price: $69 $149 $299

Figure 3: (11) Bike Pump Product Concepts Underlying the STOC Securities

	Cyclone	AirStik	Solibloc	Gearhead	Silver Bullet	Tire Repair System	Gecko	Epic	Skitzo	Rim-Gripper	2wister

Worst ⬅➡ Best
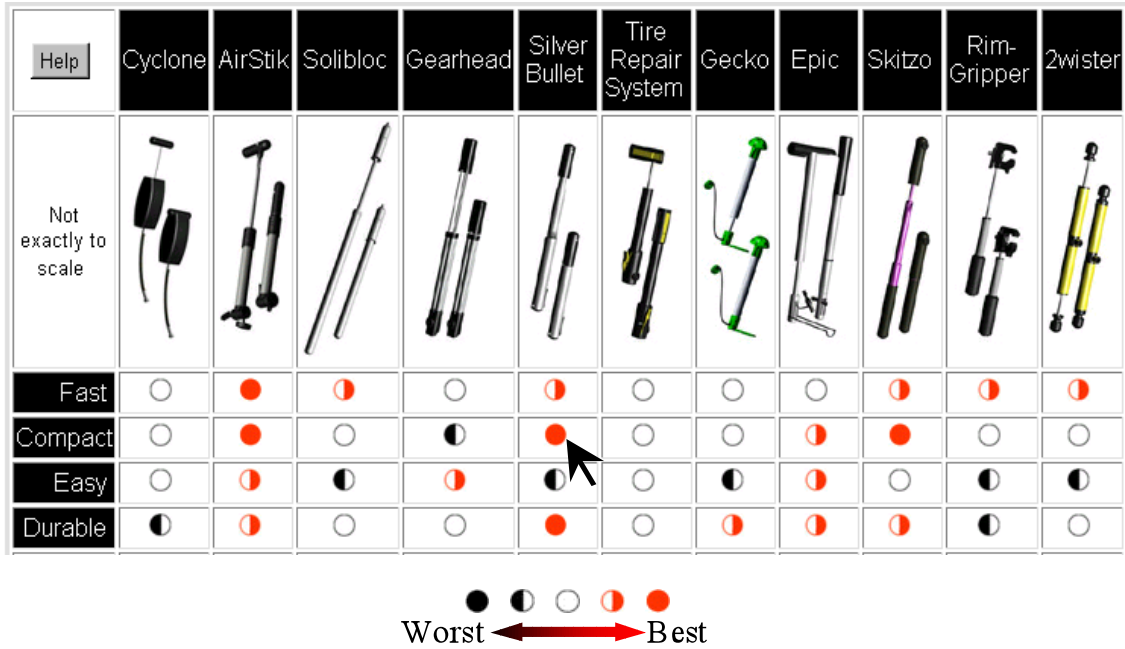
Figure 4: Which STOC Metric Correlates Best to Validation Data?
Mean Correlation to Validation Data for (6) Six Metrics in Five STOC Tests

Volume-Weighted Average Price (VWAP) — 0.85
Median Price — 0.84
Low Price — 0.77
Mean Price — 0.61
Closing Price — 0.52
High Price — 0.29
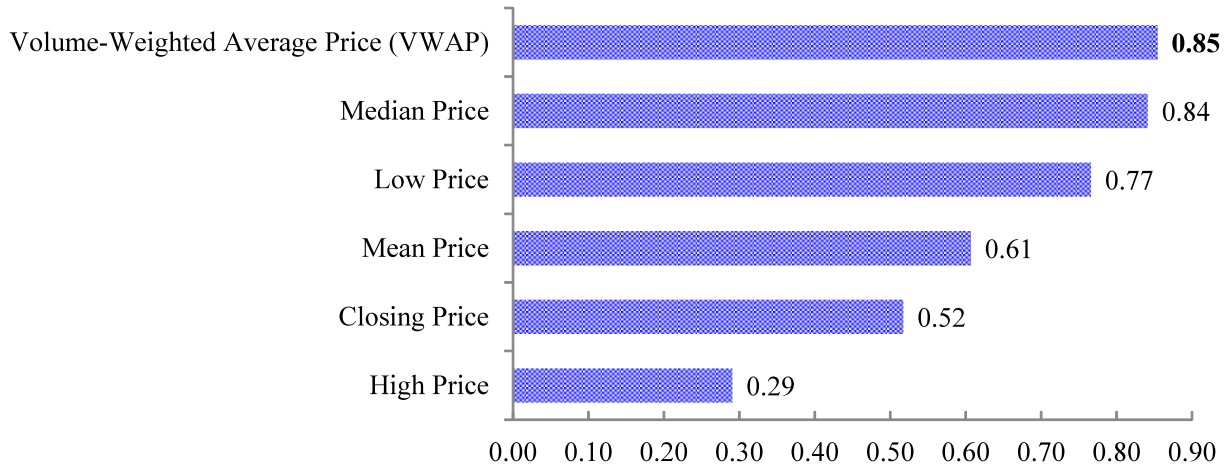
Figure 5: (8) Eight Crossover Vehicles

	Pontiac Aztek	Mercedes-Benz ML320	Acura MD-X	Buick Rendezvous	Lexus RX-300	BMW X-5	Audi All-Road	Toyota Highlander
Seats	5	5 (7 opt.)	7	7	5	5	5 (7 opt.)	5
Seating Flexibility	◐	◐	●	◐	○	○	◐	◐
Cargo Volume	◐	○	●	◐	○	◑	◑	◐
Fuel Economy	●	○	◐	◐	◐	◑	◐	◐
Horsepower	◑	○	◐	◑	○	●	◐	○
0-60 acceleration	◑	◑	◐	◑	○	●	○	○
Towing Capacity	○	◐	○	○	○	●	●	○

Figure 6: (8) Eight Laptop PC Messenger Bags in a Tabular Format

	Bag 3	Bag 4	Bag 8	Bag 9	Bag 10	Bag 13	Bag 15	Bag 16
Price	$89	$88	$99	$80	$95	$79	$78	$87
Size	Medium	Large	Large	Medium	Large	Medium	Medium	Large
Appearance	Black	Red & Black	Black	Black	Red & Black	Red & Black	Red & Black	Black
Logo	No	Yes	No	Yes	No	No	No	Yes
Handle	Yes	Yes	Yes	No	No	No	Yes	Yes
PDA Holder	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Cell Phone Holder	No	No	Yes	Yes	Yes	No	No	No
Mesh Pocket	No	Yes	Yes	Yes	No	Yes	Yes	No
Closure for Sleeve	Full Flap	Velcro Tab	Velcro Tab	Velcro Tab	Full Flap	Velcro Tab	Full Flap	Velcro Tab
Boot	Yes	No	Yes	No	Yes	Yes	No	Yes

Figure 7: Laptop PC Messenger Bags Depicted Visually without a Tabular Format

Figure 8: Wii Video Game Concepts Developed by 35 students on (8) competing teams

(8) Eight
Student-Developed
Wii Video Games

Figure 9: Wii Video Game Concepts Developed by 55 students on (11) competing teams

(11) Eleven
Student-Developed
Wii Video Games

Figure 10: Typical Product Concept Information for
Bike Pumps, Crossover Vehicles, Laptop PC Bags
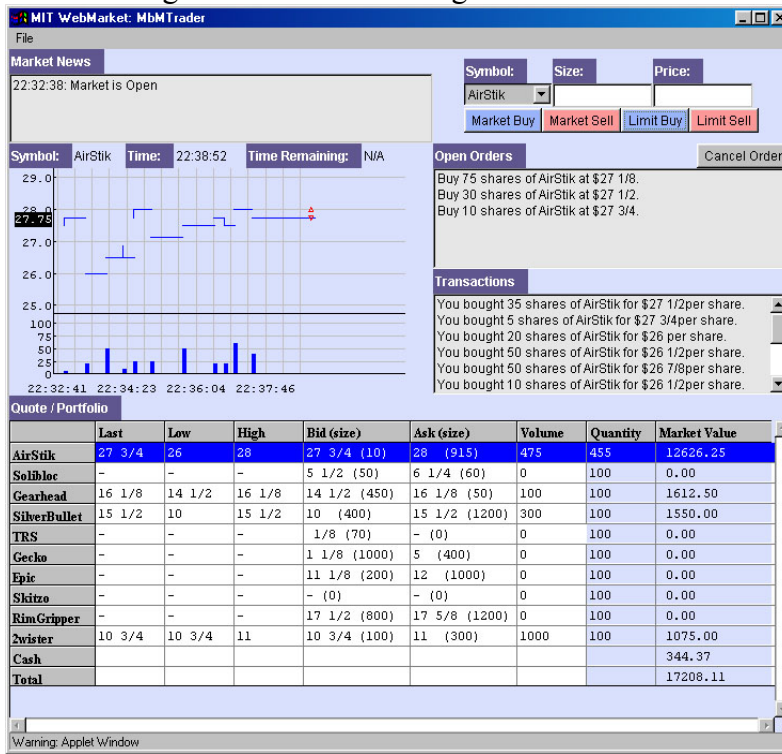
Figure 11: STOC Trading User Interface

49

Figure 12: Price and Volume History of AirStik

Test 1

Test 2
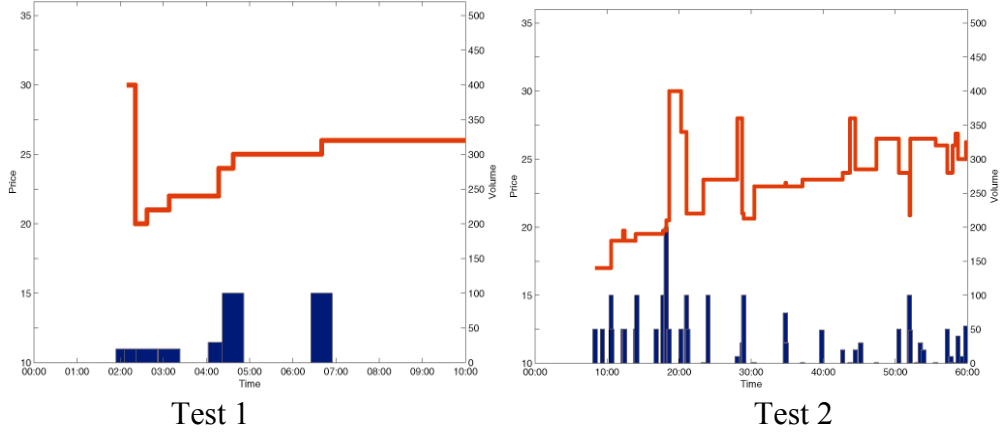
Figure 13: Individuals' trade based on their own expectations of others,
which depend on self preferences due to false consensus

Figure 14: Individuals who better estimate others' preferences tend to
rely more on their own biases when trading

Individual Insight into Others vs.
Extent of Trading Bias
(*n* = 140, r² = 0.5)

○ Test 1, r-sqr = 0.71
□ Test 2, r-sqr = 0.54
+ Test 3, r-sqr = 0.45

Correlation: E[Others] *vs.* Trading Bias

Individual's Correlation: E[Others] vs. Mean E[Others]

Web Appendix 3: STOC Web Market Simulator

In order to conduct market experiments, we developed software, which replicates the mechanisms of real-world market. The STOC Web Market is software implem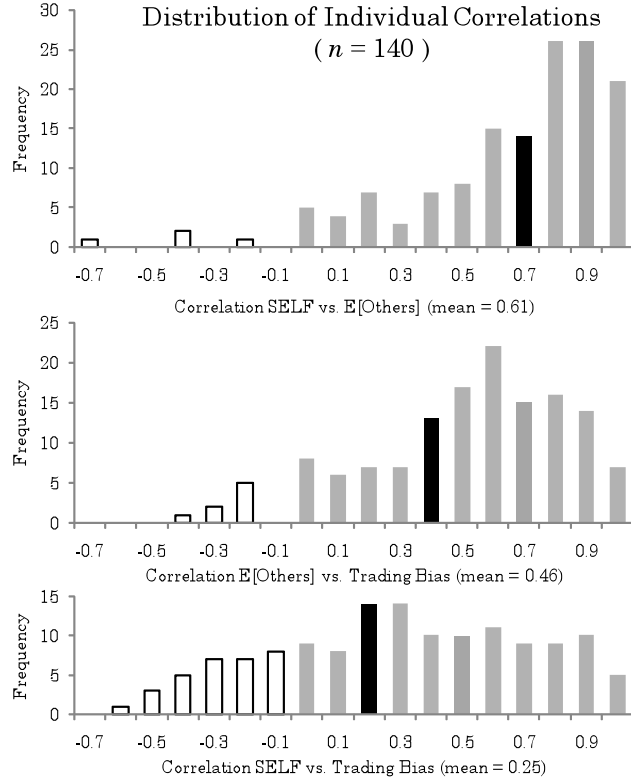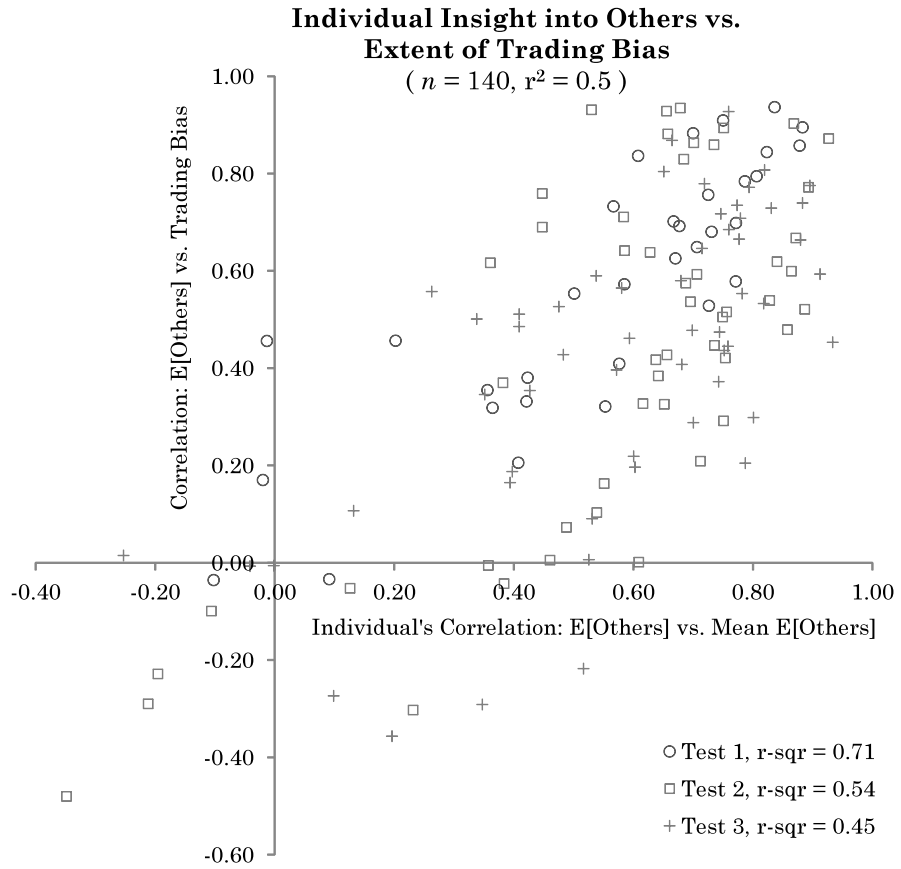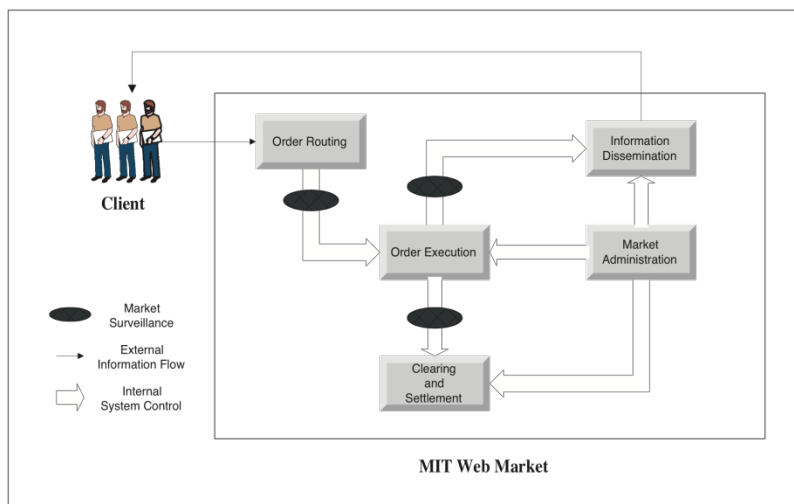ented in Java, which provides a platform for different types of electronic markets on the World Wide Web. We designed the software to use it as control laboratory for software and/or human traders based simulations and experiments.

1.1 Technical Description

Figure 15: Abstract Level Design of STOC Web Market

As shown in Figure 1, an abstract level design of the system defines roles of electronic market. The main roles are broken down into pieces as the following: order routing, match/execution, clearance/settlement, information dissemination, market administration, and market surveillance.

When orders are received from traders, the market is responsible for routing the order to an appropriate internal order book, where market makers (or auctioneers) can match them with other orders. When multiple securities are traded, each order needs to be routed to an appropriate market maker who handles the same security. After orders are matched and executed by market

makers, they are cleared and settled. The clearing/settling refers to updating the market information such as last traded price, and traders' portfolio information. Once orders are settled, traders are notified of updated market information and their portfolio holdings, which is the role defined as information dissemination. Throughout this process, appropriate market surveillances are taken place. The market surveillance is to monitor the orders to make sure that they are not violating the predefined market rules. Lastly, the system needs to be capable of administrating the rules of the market in terms of order matching, executing, clearing, and settling.

The STOC Web Market is designed into several modules to fulfill such roles. The software is implemented in Java with database connectivity to store all trade related information in a database for further analysis. The market participants access the market via Java applet, which gets downloaded into their Web browser. The Java language provides several advantages such as portability, dynamic loading, multi-threading, and object serialization, which make it a convenient platform for implementing the inherent complexity of the market simulator. These factors are especially advantageous to our system since it uses Web technology. They made it possible for all users to interact with the system from any platforms and trade securities concurrently in the market. Also they have made possible to build a robust system with relatively limited time. The Figure 2 represents the Java modules we built for the simulator. In the rest of this section, we describe each module in more detail.

Figure 16: Components of STOC Web Market
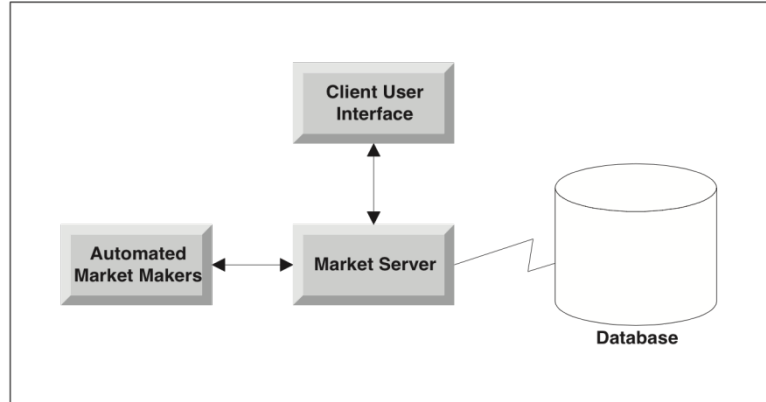
Market Server Module

The market server module serves six major functions: order match/execution, information dissemination, order routing, clearance/settlement, market administration, and brokerage service.

- **Order Match/Execution**: The order match/execution plays a central role in handling and executing the orders. One of the principal advantages of object-oriented programming is that we can easily incorporate different types of order match/execution schemes into the system.

- **Information Dissemination**: The information dissemination is to provide market information to traders. The market server disseminates information in two ways: one is to disseminate information to all traders participating in the market, and the other is to disseminate information to a specific trader. In order to disseminate public information to all traders, the market server uses an object called a 'Ticker Tape', which all traders have direct access to. For example, security information is placed on the 'Ticker Tape' when it is traded, or when the bid/ask prices are altered by newly placed orders. It also provides other general market information – whether market is open or closed, opening time of market,

duration of sessions, and other market news. In order to disseminate information to a specific trader, market server directly sends information to the trader using private channel. Each client maintains the channel for direct communication with the market server. The private information such as trade confirmation and portfolio holdings are sent through this channel.

• **Order Routing**: When orders are received, the market server routes them to an appropriate order book, which consolidates all orders of the same security.

• **Clearance/Settlement**: After the trade has taken place, the market server clears the transaction between two counter parties by transferring cash and securities between buyer and seller. The role of market server as a clearinghouse is to update trader's portfolio in the database. After the update, the market server provides updated portfolio information to the owner of that portfolio. The clearance/settlement can be implemented with different settlement rules and clearing procedures that may apply to different securities. For example, futures market has different settlement rules from those of the options market.

• **Market Administration**: In order to administer the market, the software provides a console window for market administrator as depicted in Figure 3. The console window is to serve five major roles: opening and closing of the market, monitoring traders who are participating in the market, monitoring the market makers, configuring duration of sessions, and disseminating general market information to market participants.

• **Brokerage Service**: The essence of the brokerage service is the maintenance of accounts. Traders open accounts through simple registration process. When a new trader opens up an account, the market server accesses the information stored in the database on how the

initial portfolio shall be created. Based on this information, the market server assigns predefined amount of fictitious money and shares of securities to the newly created account.

Figure 17: A Console for Market Server
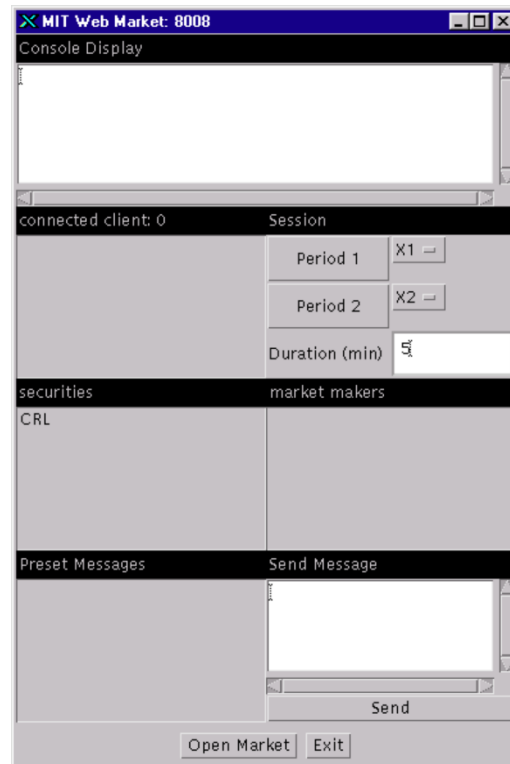
Market Maker Module

The STOC Web Market is a highly configurable market in two aspects, the market institutional structures (trading mechanisms) and the types of securities to be traded. The market structure is abstracted in the model of market makers. For each security traded in the market, different types of market makers can be integrated into the market in a "plug-and-play" manner.

In the current STOC Web Market system, the market maker is not responsible for checking general market rules, but may wish to impose additional rules on its own clients for transactions

placed through the market maker itself. If so, the market maker should give appropriate error message of the rules that are violated to the trader. If no rules are violated, the market maker must note the order to market server module. In addition, if this order allows some sales to proceed, the market maker notes to market server module to consummate the sales. Each market maker should strive to fulfill orders eventually, of course; however, there is no requirement that it be done immediately. Also, multiple market makers can handle the same security.

Figure 5: Client User Interface
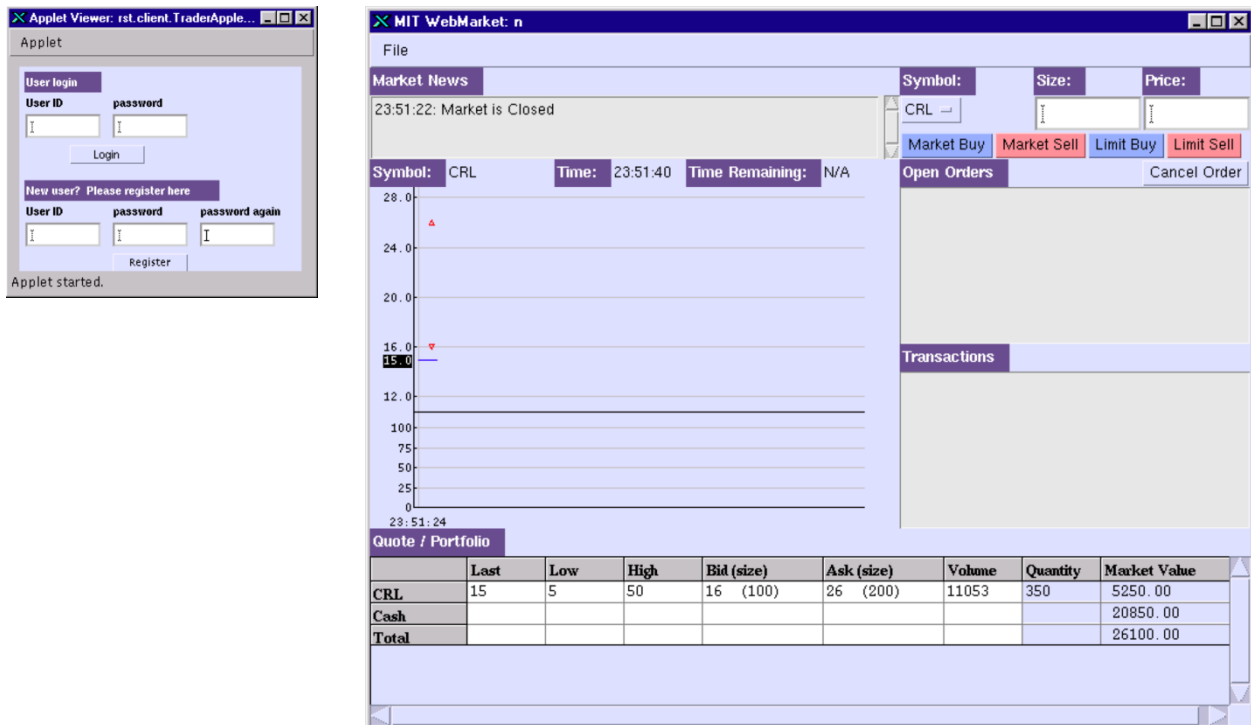
Client User Interface

The traders can access the market using GUI interface developed in Java applet. The applet consists of two parts: signup/login and trading interface (Figure 5). In signup/login screen, trader can either register herself as a new user or login to the market with pre-existing user ID and password. After logging in, the client side protocol allows trader to do the followings using

trading interface: place and cancel orders, obtain updates on ongoing market activity, and obtain information on its own portfolio.

Database Module

The database is divided into two groups of tables; tables for defining the market, and tables for storing trading information.

- **Defining the Market**: There are four entities that define the market; security information, users, initial value of portfolios, and market session. In order to setup a market, all of these information need to be provided. When trader registers as a new user, the trader's information is stored in the 'users' table. At the same time, the trader's initial portfolio is created in the 'portfolios' table based on the initial value, which are predefined in the database.

- **Storing Trading Information**: There are mainly two types of trading information. One is order information and the other is sales information. The order information consists of type of security, side (buy or sell), price, and quantity. After the execution, the sales information is stored in 'sale log' table. It is important to separate the two in the database since the first one is needed to show quote information, and the second one is used to determine the price of securities.

1.2 Summary

In summary, the main features of the STOC Web Market Simulator are: (1) automating trading, (2) modular design for market structure, and (3) an environment that allows effective interactions between traders.