NEW PRODUCT DEVELOPMENT 2.0: Preference Markets
How Scalable Securities Markets Identify Winning Product Concepts & Attributes

Ely Dahan, Arina Soukhoroukova, and Martin Spann

This is a preprint of an Article accepted for publication in the
Journal of Product Innovation Management
© 2009 Product Development & Management Association

Running Title: Preference Markets

Ely Dahan, MIT Sloan School (Visiting), One Amherst Street, E40-173, Cambridge, MA 02142 and
UCLA Anderson School, Los Angeles, 110 Westwood Plaza, B-514, Los Angeles, CA 90095, (310) 985-9703, fax (310) 206-7422, edahan@ucla.edu
Arina Soukhoroukova, University of Passau, Department of Marketing and Innovation, Innstr. 27,
94032 Passau, Germany, as@ideamarkets.com, + 49 851 509 2423, (fax) 49 851 509 2422
Martin Spann, University of Passau, Department of Marketing and Innovation, Innstr. 27, 94032
Passau, Germany, spann@spann.de, + 49 851 509 2421, (fax) 49 851 509 2422

Keywords: Product Testing, New Product Development, Virtual Stock Markets, Product Design,
Preference Markets, Information Markets, Scalable Market Research

Acknowledgements: The authors thank Martin Bernhardt, Barnaby Feder, Craig Fox, Mark
Garmaise, John Hauser, Oliver Hinz, Rich Johnson, Adlar Kim, Andrew Lo, Donald Morrison,
Tomaso Poggio, Ricardo dos Santos, Hyun Shin, Bernd Skiera, Vicky Smith, Sven Theysohn,
Research Conference attendees and judges for their valuable feedback.
New Product Development 2.0: *Preference Markets*

How Scalable Securities Markets Identify Winning Product Concepts & Attributes

**Abstract**

Preference markets address the need for scalable, fast and engaging market research in new product development. The Web 2.0 paradigm, in which users contribute numerous ideas that may lead to new products, requires new methods of screening those ideas for their marketability, and preference markets offer just such a mechanism. For faster new product development decisions, we implement a flexible prioritization methodology for product features and concepts, one that scales up in the number of testable alternatives, limited only by the number of participants. New product preferences for concepts, attributes and attribute levels are measured by trading stocks whose prices are based upon share of choice of new products and features. We develop a conceptual model of *scalable preference markets*, and test it experimentally. We find that benefits of the methodology include speed (less than one hour per trading experiment), scalability (question capacity grows linearly in the number of traders), flexibility (features and concepts can be tested simultaneously), and respondent enthusiasm for the method.
1 Introduction

In an environment of accelerating technology and short product life cycles, one in which a plethora of product concepts and features proliferates, new product development teams need fast and accurate marketing research to filter out the most promising opportunities. Smart phones, video gaming systems, home entertainment, information appliances, and other durable goods require development teams to prioritize literally hundreds of design decisions (Thompson, Hamilton and Rust 2005). There is a need to bridge the front end- and design phases by narrowing many features and concepts down to those key, make-or-break success factors. This requires a fast prioritization methodology, one that scales up in the number of testable product features and concepts.

The quantity of new product concepts and features to be evaluated will steadily increase, driven by the Web 2.0 paradigm, in which users volunteer new product and feature ideas over the internet. This new form of “collaborative creativity” generates thousands of possibilities, and demands new methods of identifying the more marketable ideas, and screening out those with lower potential. In traditional market research, the more features or product concepts to be studied, the greater the number of participants and the cost and time required. Limits on the number of questions for participants derive from bounded rationality (Simon 1955), respondent fatigue (Shugan 1980), and time constraints. Faced with too many questions, respondents may resort to simplifying heuristics, even with tasks involving as few as 10-20 product features (Yee et al. 2007; Gilbride and Allenby 2004).

In this article we propose a flexible new approach to test preferences for large numbers of new product features and concepts through the use of scalable preference markets. Preference markets offer an ideal first-cut screening mechanism, thereby complementing other methods such as conjoint analysis and concept testing which perform better on a
limited number of attributes and product concepts. By relying on the wisdom of crowds (Surowiecki 2004), preference markets identify potential winning and losing ideas. By engaging in stock trading, in which the price of each stock represents the degree of preference for a product attribute level, new feature or fully integrated product concept, participants reveal their own preferences and their expectations of others’ new product preferences, and converge towards an equilibrium which captures the consensus view.

Preference markets address the need for scalable, fast and engaging market research by combining elements of three methods: (1) actual financial markets, in which huge numbers of securities undergo continuous valuation through a fluid network of individuals trading with each other, (2) opinion surveys, which measure individual preferences, and (3) prediction markets, which measure expectations of future events through the market pricing mechanism using virtual stock markets.

While preference markets build upon these three approaches, they differ from them in important ways. Financial markets operate more efficiently when they are “thick,” i.e., when the number of traders exceeds the number of securities being traded (Fama 1970). Even though no single individual has the capacity to follow every security, every security is still traded by a large number of individuals. Unlike in real financial markets, where traders self-select securities, we control our experiments by assigning traders to specific bundles of stocks, thus ensuring that every stock gets traded.

In opinion surveys, respondents answer each question once, do not learn from each other, and typically express self preferences for new product features and concepts. While stock trading outcomes are related to those measured by opinion surveys, they differ substantially. Preference market participants “answer” each question multiple times by buying and selling stocks throughout a trading task. Traders may learn from, and be
influenced by, the behavior of fellow traders. And they may base their trading decisions on self preferences, on expectations of others’ preferences, or on some combination of both. Indeed, Hoch (1987, 1988) shows that aggregating the opinions of heterogeneous individuals produces different results than averaging those individuals’ expectations of others. While not completely eliminating biases, aggregation of diverse opinions frequently outperforms those of individual “experts”, particularly if responses are weighted based on competence or confidence (van Bruggen, Lilien and Kacker 2002). Larrick and Soll (2006) highlight the error-reduction benefits of aggregating individual estimates when the “truth” is bracketed by the heterogeneous estimates provided by individuals. But they also point out that most people have poor intuition about the benefits of aggregating these estimates. The present research suggests that the market pricing mechanism may do a better job aggregating multiple perspectives than would individuals using intuition alone.

Scalable preference markets also continue the trend towards Internet-based market research, yielding benefits such as speed, adaptive interactivity, and task engagement (Dahan and Hauser 2002; Sawhney, Verona and Prandelli 2005). Other research has recognized the challenge of respondent fatigue, and addressed it through adaptive questioning (Sawtooth Software 1999; Touibia et al. 2003), more engaging tasks such as user design (Park, Jun and MacInnis 2000; Randall, Terwiesch and Ulrich 2007; Liechty, Ramaswamy and Cohen 2001), and task simplification as in self-explicated questioning (Kivetz, Netzer and Srinivasan 2004). Preference markets build upon these Internet benefits, but add competition and interactivity to enhance the respondent experience and align incentives for truth-telling.

Finally, previous research on prediction markets has used stock trading to forecast actual outcomes such as election results, movie box office receipts, or sporting event outcomes (see Table 1 for a summary of prior research on prediction and preference markets). In addition
to this published research, firms such as Microsoft (Proebsting 2005) employ internal prediction markets.

Table 1: A Sampling of Prior Research on Prediction- and Preference Markets

<table>
<thead>
<tr>
<th>Few Stocks</th>
<th>Prediction Markets (actual outcomes)</th>
<th>Preference Markets (no actual outcomes)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spann and Skiera 2003: 5 web-enabled cell services</td>
<td>Dahan et al. 2008: Study 1: 11 bicycle pumps; Study 2: 8 Laptop bags Study 3: 8 crossover vehicles</td>
</tr>
<tr>
<td></td>
<td>Chen and Plott 2002: 12 HP printers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wolfers and Zitzewitz 2004b: 3 war-related securities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spann et al. 2009: 10 to 15 movie related stocks per market</td>
<td></td>
</tr>
<tr>
<td>Many Stocks</td>
<td>Forsythe et al. 1992; Forsythe, Rietz and Ross 1999: multiple political races</td>
<td>The present research: 56/64 smart phone features and integrated products in two studies</td>
</tr>
<tr>
<td></td>
<td>Pennock et al. 2001: 161 events on the Foresight Exchange</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spann and Skiera 2003: 152 HSX.com movies</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Elberse and Eliashberg 2003; Elberse 2007: HSX.com movies</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Servan-Schreiber et al. 2004: 208 games on TradeSports and NewsFutures</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cowgill, Wolfers and Zitzewitz 2008: 270 markets in 2.5 years</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spann and Skiera 2009: 837 soccer games in multiple markets</td>
<td></td>
</tr>
</tbody>
</table>

Preference markets, on the other hand, do not predict actual outcomes, nor are they based upon external information. Rather, they measure expectations of others’ new product preferences, based upon individual self preferences combined with insights about others. While prediction markets typically run for weeks or longer, preference markets require only minutes, as there is no outside “news” to affect the market. For example, Dahan et al. (2008) evaluate product concepts in stock trading tasks that run less than an hour.
Participants are presented with new product concepts and then trade securities representing the competing designs. In effect, traders place bets on those concepts which they expect to curry favor with their fellow traders. Their results exhibit high consistency and reliability across trading experiments and against independent surveys. Shin and Dahan (2009) show that preference markets are more like market research than they are like finance, and also validate the use of volume-weighted average prices (VWAP).

To the best of our knowledge, the present article is the first to test the scalability of such markets for measuring new product preferences over a larger number of concepts and features than any individual trader can process. In effect, we are building upon the “Wisdom of Crowds” effect, and adding to it a twist on Simon (1955), the unbounded rationality of crowds.

In the context of NPD, this article is also the first to analyze not only customer preferences for integrated product concepts, but for product attributes and attribute levels at varying prices, including product and service brands. Such a decompositional approach frequently proves useful to NPD managers when making design tradeoffs and positioning new products in the marketplace.

The aim of this research is to propose a conceptual model and methodology of scalable preference markets to handle large numbers of product concepts and features, position the methodology in the context of other methods of new product research, validate the approach through empirical tests, and derive insights about the application of scalable preference markets to marketing problems. The article is structured as follows. We match potential applications of preference markets to the different phases of new product development in Section 2. In Section 3, we discuss previous research on prediction markets and connect
these to our conceptual model of preference markets. Section 4 details the preference market methodology. Section 5 analyzes empirical data collected from repeated preference markets for smart phones and their attributes. And section 6 concludes with a general discussion, managerial insights, limitations, and areas of future research.

2. Application of Preference Markets to Phases of New Product Development (NPD)

Preference markets may apply during four phases of new product development. In the early idea generation and concept selection phases, preference markets can narrow potential concepts and product attributes to a manageable number, focusing resources where they will yield the greatest marginal benefit. In the later detailed design and testing & launch phases, preference markets can help assess price sensitivity, detailed new product feature preferences, and optimal advertising and promotion. However, at these later stages (especially during testing & launch), preference markets are only a special case of prediction markets that forecast the market potential of a product prior to introduction. A primary distinction of preference markets in the latter NPD stages is that the concepts tested need not ultimately be launched, and actual outcome is not required as is the case for prediction markets. For example, an NPD team might use preference markets to test potential advertising campaigns, price points or distribution strategies prior to product launch. Only one option will be realized based on the new product preferences of the traders, but the lack of actual outcomes does not prevent preferences from being measured. Table 2 summarizes our conclusions about who should participate in preference markets, which stock types can be tested, how these markets could be implemented and why running preference markets at four key stages of new product development makes sense.
From Table 2, we see that preference markets appear to be particularly beneficial in the early stages of NPD as a way of prioritizing design decisions and allocating resources. They complement other market research methods such as conjoint analysis and virtual concept testing which perform better with a limited number of attributes and concepts, and which are geared to individual preference measurement. In contrast to prediction markets, which have been shown to be effective at predicting future outcomes, preference markets may be vulnerable to unreliable measurement if respondents randomly construct product preferences on the fly. Were this to be the case, traders’ own preferences and expectations of each other’s preferences would not only be quite heterogeneous, but also quite noisy, and market prices would not capture consensus preferences accurately. This would manifest in the form of low correlations between self preferences and expectations of other’s preferences. It would also result in low correlations between repeated tests and with external validation data. One objective of the present research is to test whether preference market security
prices are based on traders’ underlying new product preferences and their expectations of others’ preference prior to trading or are rather more random outcomes based on constructed preferences and the “gaming” aspect of the market exercise.

It is worth noting that the traders in both trading experiments we conducted had highly heterogeneous new product preferences and came from very different personal backgrounds. Thus, we would expect poor correlations between securities prices and other preference measures unless preference markets are truly effective at measuring underlying preferences that exist prior to the start of trading.

In order to test the ability of markets to measure actual new product preferences, we develop a conceptual model in Section 3 and test the model empirically in Section 5.

3 Conceptual Model of Preference Markets

A conceptual model of scalable preference markets builds upon prior work on financial markets and experimental economics, including information and prediction markets, as well as traditional market research. Four hypotheses are linked: (1) individual new product preferences lead to expectations about others, (2) rational traders use their expectations to decide when to buy and sell securities, (3) market prices aggregate information and beliefs held by individuals, and (4) individuals learn from markets. These four effects combine to explain how market prices measure people’s beliefs about others’ new product preferences.

3.1 The Wisdom (and Biases) of Crowds: Connecting the SELF to OTHERS

The task of estimating the market success of new products requires experts to distinguish between their own self-preferences and those of others, which may or may not be similar. Prior research has demonstrated that individuals’ self preferences can bias their expectations of others’ preferences (Hoch 1987). Yet, aggregating individual opinions, even biased ones, produces surprisingly accurate and objective estimates of the consensus of
opinion (c.f., Surowiecki 2004, Lorge et al. 1958, Larrick and Soll 2006). In prediction markets with actual outcomes, such as the Iowa Electronic Market for the 1988 presidential race (Forsythe et al. 1992), “62% of the Bush supporters bought more Bush stock than they sold, while 68% of Dukakis supporters bought more Dukakis stock than they sold.” In marketing surveys, respondents may overweight their own opinions when estimating the new product preferences of others, a simple result of “false consensus”. So we expect aggregate self preferences to be strongly correlated to aggregate expectations of others, while being subject to the effect of biases at the individual level. To the extent that individuals may lean in opposite directions - i.e., that some individuals prefer an option while others reject that option - this will result in preference heterogeneity, which the market pricing mechanism will have to correct through the process of trading.

In the context of new product development, respondents’ expectations of others’ preferences may provide insight, albeit indirect and at the aggregate level, into individual preferences.

H1a. Wisdom: Self-preferences for new product features and concepts provide insight about others’ preferences, therefore mean self-preferences and mean expectations of others’ preferences will be highly correlated.

H1b. Bias: Individuals who prefer a new product concept or attribute level have higher expectations of others’ preferences for that option than do individuals who reject that option.

3.2 Rational Expectations: How expectations of OTHERS affect ORDERS to buy and sell

Rational, profit-maximizing investors utilize personal knowledge in determining the value of a stock, and make trading decisions based upon this knowledge (Lucas 1972). This principle, well-established in financial markets, is also evident in experimental markets, where traders have an incentive to reveal their version of the truth (Smith 1982; Plott and Sunder 1982). Experimental markets have been shown to be quite robust to manipulation by some traders because other traders take the possibility of manipulation into account when setting their own expectations (Hanson, Oprea and Porter 2005). Even in opinion surveys where participants are rewarded for making insightful observations, reward-
maximizing contributors factor in their expectations of others’ reactions (Toubia 2006). Similarly, we expect portfolio-maximizing traders in preference markets to buy and sell based upon their expectations of others, rather than on their own self-preferences for product features and concepts. Therefore, we expect stock prices to more closely correlate with expectations of others than with self preferences.

H2. Focus on Others: *Expectations of others’ preferences for new product features and concepts affect individual buy/sell decisions for stocks more than do self preferences, so stock prices will correlate more highly with expectations of others.*

In new product development, design teams must consider preference heterogeneity with regards to market segmentation and targeting. After all, individuals, not crowds, make product purchase decisions. But NPD teams also benefit from learning aggregate new product preferences, especially when there are many possible product concepts, attributes and levels. Identifying the standout ideas and key features early on, and filtering out weaker ones, enables NPD teams to focus resources on higher-leverage opportunities.

### 3.3 PRICES: How markets achieve consensus based upon ORDERS

In a competitive economy, the price of a good reveals its value (Hayek 1945). Similarly, in financial markets, the efficient market hypothesis posits that stock prices aggregate all information known to traders (Fama 1970; Fama 1991). Experimental markets confirm the theory and converge towards “truth” within a few iterations, even when information is dispersed (Plott and Sunder 1988; Forsythe, Palfrey and Plott 1982). Importantly, not all market participants are equal in their influence on prices. In financial and prediction markets, traders with greater knowledge or certainty exert greater influence on prices, thus weighting their opinions more heavily. Elberse and Eliashberg (2003) and Elberse (2007) use the HSX prices to incorporate ex-ante expectations of the revenues of new motion pictures. Cowgill, Wolfers and Zitzewitz (2008) demonstrate that participants are able to
learn from their initial mistakes. “Informed” traders with private information effectively set market prices for the “less informed” (Oliven and Rietz 2004).

For prices to be truly informative, profit-maximizing participants should be rewarded based upon real outcomes (Smith, Suchanek and Williams 1988). But real outcomes may be hard to come by at the front end of new product development. For example, only a few of many product concepts or features may actually be launched. Product development teams may not be able to afford to wait for actual market outcomes because speed-to-market, and the first-mover-advantage that results from speed, are frequently key success factors.

For such research questions, the prerequisite that securities be linked to actual outcomes may have to be relaxed, and may either be replaced with an alternative such as actual survey results, or at least the belief on the part of traders that their portfolio valuation depends on such external survey results. As long as traders behave consistently with the belief that their portfolio performance will be measured based on actual, observable results, then market prices should reflect traders’ expectations of those results. Even lacking observable outcomes, preference markets can utilize the market pricing mechanism to efficiently aggregate consensus expectations.

H3. Consensus: The preference market pricing mechanism summarizes the consensus of individual offers to buy and sell stocks representing product features and concepts, so preference market prices will correlate highly to mean expectations of others’ preferences for new product features and concepts.

3.4 LEARNING from market PRICES

In financial markets, investors continuously observe prices, and rapidly update valuations in response to news and exogenous events (MacKinlay 1997). In this context, stock price shocks can become newsworthy in their own right. Upon receiving news or observing stock price changes, traders may respond by adding further volatility (Blanchard and Watson 1982; Timmermann 1993). Of course, one risk of this price-based communication is the
possibility that traders will learn the wrong thing from each other, and jump on a “misguided bandwagon.” Smith, Suchanek and Williams (1988) show that inexperienced traders may produce market bubbles and crashes. A few marginal traders with strong (but inaccurate) beliefs can influence a majority of traders who are weaker in their beliefs, leading to herding behavior. Further, the underlying individual new product preferences themselves can be influenced by such communication. Salganik, Dodds and Watts (2006) show that the individual preferences change dramatically when people are exposed to the preferences of others’. Additionally, publicly posted opinion surveys, such as critical reviews, polls, or even web chat, may affect individual opinions (c.f., Chevalier and Mayzlin 2006). In the case of preference markets, the potential for herding behavior suggests that repeated, independent market simulations are called for to verify inter-test reliability. On the other hand, for fashion goods or those with network externalities, where individual new product preferences are heavily influenced by others, the inter-trader communication inherent in preference markets offers potential insight about how preferences can evolve.

Regarding learning, preference markets are simpler than financial and prediction markets. Rather than outside news, the only new information revealed to traders is the stock price itself. Therefore, we expect traders in preference markets to learn from stock prices, update their own beliefs, and converge in their expectations of others.

H4. Updating: The process of trading will cause traders to update their expectations of others’ preferences for new product features and concepts, and will reduce the variability in these expectations across traders.
3.5 A Conceptual Model of Preference Markets

Figure 1 integrates sections 2.1 through 2.4 and our four hypotheses into a conceptual model of preference markets, and highlights the five types of data collected in our studies.

Figure 1: A Conceptual Model of Preference Markets for New Products & Attributes

The first circle represents SELF preferences, that is each trader's new product preferences for him- or herself. These preferences influence expectations of OTHERS' preferences [H1a and H1b] as depicted in the second, “outward-looking” circle. Traders’ rational expectations of OTHERS inform buy/sell decisions [H2], as shown in the ORDERS square in the middle of Figure 1. For example, the value of a specific stock traded in a preference market can be linked to the percentage of people who choose to buy the product feature or concept linked to this stock. If a trader believes that more than 20% of people prefer the Motorola brand, for example, he or she should buy the stock when the market price is $20. Conversely, if a trader believes that less than 20% prefer Motorola, he or she should sell at that point. Market PRICES are determined through executed trades. Traders with stronger convictions hold greater sway in determining stock PRICES [H3]. Traders may engage in LEARNING, due to updated expectations of others based upon newly posted prices [H4].
A Methodology for Preference Markets

Two objectives of running a preference market are: (1) eliciting truthful measurement of trader expectations, and (2) negotiating consensus about new product preferences.

In order to achieve these objectives, it is imperative that: (a) traders understand the connection between each stock and the underlying product attribute-level or integrated concept it represents, (b) the number of stocks being traded by each trader is manageable, (c) stock trading evokes judgments similar to actual product evaluation, i.e., at least some traders must have insight about preferences for the underlying new products, and (d) traders have the incentive to reveal the “truth” when trading (c.f., Spann and Skiera 2003; Wolfers and Zitzewitz 2004a). With these four objectives in mind, below we identify four design decisions relevant to preference markets, discuss our choices for each design decision, and highlight some methodological contributions.

4.1 Defining the Securities

We define individual securities so that stock prices measure strength-of-preference for a particular product, product concept, brand, feature, attribute level, or bundle of attributes. Traditionally, market research studies impose parallelism by focusing on one question type for a given task. For example, conjoint analysis asks respondents to evaluate product attribute levels, while concept testing has them compare product concepts or real products (Huber et al. 1993). In section 4.2, we show how to easily “mix-and-match” actual products, product concepts, and all manner of features and attributes in a single preference market.

Stocks represent either binary product choices (e.g., “FM Tuner Included” or “FM Tuner Not Included”), or mutually exclusive ones (e.g., “Brick”, “Slide Open”, or “Flip-Phone” form).

To accurately capture new product preferences, traders must understand how a stock’s price reflects strength-of-preference. Stock prices can be defined as “average rating on a 1-100
scale,” “number of units that will be sold,” or “percentage of people choosing an option.” To simplify trading, the scale for stock prices should be common across all stocks. Product features and concepts must be vividly and clearly communicated to the traders.

4.2 Experimental Design for Scalability: The Unbounded Rationality ofCrowds

After connecting stocks to product characteristics, we must connect individual traders with stocks. In other stock markets, traders self-select stocks, typically trading only a tiny percentage of the universe of securities. We seek both the scalability of financial markets and the control of experimental research, and achieve this by assigning small groups of traders to small groups of stocks. A contribution of the present research is the unique and highly flexible experimental design in which participants trade certain stocks within their subgroup and other stocks across multiple subgroups. Traders can be assigned to the subgroups randomly, or based on interest, product expertise, or market segment. A sample experimental design appears in Figure 2.

![Figure 2: Organizing Traders and Stocks into Groups, each Trading a Subset](image)

Trading stocks across subgroups ensures communication through the price mechanism, in effect putting all preference measures on a common scale. Multiple product categories can be studied simultaneously by assigning some traders to one product category, other traders
to a second category, and a third group of traders to both categories. Similar experimental designs may assess price elasticity by measuring preferences over identical products and attributes at varying price points. Such “price-elasticity” stocks may be traded independently by separate groups of traders, to avoid context and anchoring effects. An additional use would be to measure inter-test reliability by having multiple groups trading identical sets of stocks, just not with each other (while still having some stocks common across groups in order to maintain the common scale alluded to earlier).

Regardless of how traders are matched to stocks, it is important to recruit at least some traders who have strong preferences in the product category, or who can provide insight about others’ preferences. A representative sample of consumers as market participants is not required, as long as traders possess insight about customers in that product market (Forsythe et al. 1999; Spann and Skiera 2003). Further, a sufficient number of traders is needed, determined by the need for liquidity in each stock and the number of stocks each trader can handle. For example, if an individual trader can manage twenty stocks, and each stock needs to be traded by twenty people, then twenty stocks could be traded by twenty people in a trader subgroup.

The population from which participants are drawn depends on the product expertise and degree of secrecy required. Access might be restricted to invited consumers or employees in order to guard intellectual property embedded in the stock definitions. Or a narrow sample may be recruited out of a population with particular interest in the product category. On the other hand, an open access market, such as HSX or TradeSports allows users to self-select based on their degree of interest in each stock, potentially improving motivation.
4.3 Market pricing mechanism

The duration of trading can be extremely short because no exogenous information enters a preference market. An efficient trading mechanism for such brief markets is an Internet-based double-auction mechanism (c.f., Guarnaschelli, Kwasnica & Plott 2003, Sunder 1992). An equilibrium regarding stock values can form within minutes, and prices stabilize in under an hour. To avoid boundary effects at the end of trading, traders are given a rough timeframe for trading, and the market is stopped randomly, after say 30-45 minutes.

4.4 Incentives

Incentives should induce truth telling and active involvement by traders. In financial and prediction markets, traders try to maximize portfolio value because their payoffs are proportional to the liquidation value of all stocks and cash. Preference markets need not provide payoffs to each trader, nor to reveal actual outcomes. Rather than basing incentives on realized outcomes, they can be based on closing or average stock prices, which are endogenous to the market, and act as a surrogate for actual outcomes. The lack of actual outcomes may leave preference markets vulnerable to pricing bubbles and gaming. So one could generate exogenous “truth” by conducting an independent preference survey, and using its results as the actual outcome for each stock. In addition, beyond rewarding trader performance and accuracy, one might reward effort (e.g., number of trades).

To reduce the cost of compensating every trader based on final portfolio values, we might randomly select prize winners based on the ranking of each portfolio within a trading subgroup. Maximizing one’s expected reward would still be consistent with maximizing one’s portfolio value, even if one is not the top trader within the subgroup. The short duration of preference markets add to the intrinsic reward of competing, since within minutes of completing the market, one discovers one’s ranking among all traders. The incentives need to be high enough to attract traders. Details are provided in Appendix 1.
5 Empirical Studies

In order to test the feasibility and accuracy of scalable preference markets and to test our conceptual model from Section 3, we ran two studies within the smart phone product category, including a laboratory test involving MBA students as well as a field test at a multinational corporation involving managers and engineers. The results from the two studies support the hypotheses and conceptual model, but must be viewed as the preliminary empirical studies that they are. We hope that others will replicate our studies to further validate the preference markets method.

5.1 Study 1: Laboratory Test of Scalable Preference Markets

Our first study tests the key aspects of preference markets: scalability, flexibility and learning in a laboratory test with a student sample.

Study Design and Procedure

We modeled 56 different design and concept stocks with different types of scales: 31 binary feature levels, 19 mutually exclusive feature levels, and 6 full phone concepts (Appendix 2). 116 MBA student respondents (a 38% response rate) were recruited to complete two surveys in advance of stock trading. In advance of trading, each participant completed (1) a SELF survey, as shown in Figure 3(a) for all 56 options, to be compared with (2) a second survey of expectations of OTHERS for 20 of the 56 options.

Binary feature levels such as “Bluetooth ($49)” were surveyed using radio buttons with only two options, “yes” or “no.” “Yes,” the response given by 40% of the 116 people surveyed, meant that the respondent would add Bluetooth to his or her smart phone were that feature priced at $49 retail. Mutually exclusive feature levels also used radio buttons, with between two to six alternatives, exactly one of which had to be selected. We note that SELF preferences were highly heterogeneous, evidenced by a median coefficient of variation of
76% across all 56 stocks (a coefficient of variation of 100% being the highest possible in this context, which would occur if 50% of respondents chose an option and the rest rejected it).

After the SELF survey, those respondents who would be trading a particular stock also answered a question about OTHERS’ preferences, “What percentage of participants would buy this feature?” For example, among the 41 respondents who were about to trade the “Bluetooth” stock, the average answer to the question about OTHERS was 26% (s.d. 19%).

After trading, a POST survey (56% response rate) asked traders to provide updated estimates of others’ preferences. In this experiment surveys of SELF, OTHERS, and POST can be compared against stock trading, enabling us to test the four hypotheses that comprise the conceptual model in Figure 1.

Figure 3: Updated Multi-Screen User Interface for Survey and Trading

113 of the 116 survey respondents opted to participate in the stock trading experiment on a university holiday, 93 of them in person in two classrooms, and 20 off-site logged into the market over the Internet. To test scalability, we develop an experimental design consisting of six groups of traders and six overlapping groups of stocks, as shown in Figure 4.
Figure 4: Scalable Preference Markets Design

A continuous double auction market mechanism was implemented, as in Forsythe (1992). In such a system, a trade is executed only when a seller’s asking price is at or below a buyer’s bid price. No initial prices or orders were set up in advance, and all endowments within a trader group were identical: 100 shares of each of twenty securities, and $15,000 in virtual cash. The $15,000 amount was chosen to provide sufficient liquidity, while not encouraging excessive speculative behavior, and represented approximately 25% of the expected portfolio value.

The user interface, depicted in Figure 3(b), provided traders with short descriptions and images, real-time trading information, and supply and demand data in the open order book. During the fifty minute duration of the experiment, traders attempted to maximize their respective portfolios, including the market value of all stocks and cash, by executing a total of 1,680 trades. Each stock was traded between 5 and 150 times, confirming the effectiveness of the experimental design in scaling up in the number of traders.

As an incentive, a $50 reward was offered to the “winners,” who were randomly drawn from amongst all traders within each of the six groups in Figure 4. The probability of winning depended on the trader’s rank within his or her trading group, based upon total portfolio value at ending market prices, at a randomly chosen ending time, so that those who had
performed well had a higher probability of winning. Further, traders with the highest total portfolio values were announced publicly, a form of recognition within this competitive peer group. These incentives were designed to induce traders to reveal their true beliefs, even if they were not performing particularly well. Additionally, an award was offered for the highest portfolio based on actual survey results using the SELF survey data.

**Results**

Prediction markets use closing prices since the efficient market hypothesis suggests that the most recent stock price summarizes all known information (Fama 1970), but in the context of preference markets, given their lack of external information, every offer to buy and sell expresses the belief of an individual trader, and each executed trade represents an agreement between at least two traders. So, analyzing the entire data set with volume-weighted average prices (VWAP) captures a broader range of opinions. The first four rows and columns of Table 3 summarize the relationships between stock trading and the SELF-, OTHERS-, and POST surveys in Study 1 and for the next Study 2. All correlations are significant at the $p < 0.001$ level.

Given the highly heterogeneous new product preferences of the 116 survey respondents, we conclude from the Pearson correlation of 0.880 between the SELF and OTHERS surveys, that respondents are accurate in estimating each other’s preferences. Thus, H1a is supported, as is Figure 1’s link between SELF and OTHERS. As for H1b, the hypothesis that expectations of others are biased by self preferences, the surveys support this hypothesis as well, since individuals who chose an option for themselves had higher expectations of the percentage of others who would choose that option (37% higher on average across all 56 stocks) than those who rejected an option, and those who rejected an option for themselves had 18% lower expectations, on average. These differences were significant at the $p < 0.001$ level for 46 of
the 56 stocks. These biases work against overall accuracy, and yet are overcome by the “wisdom of crowds” effect (H1a) and by the market pricing mechanism [H3].

Table 3: Correlations based on Volume-Weighted Average Prices (VWAP) for Studies 1 & 2

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>2</th>
<th>2</th>
<th>2</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>SELF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHERS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRICES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POST</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 SELF          1 OTHERS  1 PRICES  1 POST  2 SELF  2 OTHERS  2 PRICES  2 POST

1 OTHERS  .880***
1 PRICES  .622***  .750***
1 POST    .717***  .832***  .924***
2 SELF    .653***  .677***  .723***  .651***
2 OTHERS  .685***  .752***  .769***  .733***  .863***
2 PRICES  .525***  .620***  .661***  .626***  .767***  .829***
2 POST    .560***  .647***  .741***  .696***  .714***  .837***  .910***

SELF: Mean survey results for individuals’ choices
OTHERS: Mean pre-trading expectations of others’ choices
PRICES: Preference market results: volume-weighted average stock prices (VWAP)
POST: Mean post-trading expectations of others’ choices
***: $p < .001$

A subset of 39 stocks were common to both Studies 1 and 2.

As seen in row three of Table 3, preference market PRICES correlate better with expectations of OTHERS ($\rho = 0.750$) than with SELF preferences ($\rho = 0.622$), supporting H3.

Study 1 also supports H2, the hypothesis that traders base their stock buying and selling decisions upon expectations of OTHERS more so than on SELF preferences. By aggregating all of the orders to buy and sell each stock, and comparing the volume-weighted mean value of the order prices against the self and others survey, stock by stock, we observe that the correlation between ORDERS and OTHERS ($\rho = 0.76$ at the $p < .0001$ level) is higher than the correlation between ORDERS and SELF ($\rho = 0.65$ at the $p < .01$ level).

Finally, Study 1 also reveals that trading stocks results in a significant amount of learning among traders, and that H4 is supported. Specifically, traders update their beliefs about
others based on the stock prices they observe, as seen in the increased \( \rho \) of 0.924 \((p < .001)\) between stock \textsc{prices} and the \textsc{post} survey, as compared to the lower correlation of 0.832 between \textsc{prices} and the pre-trading version (\textsc{others}) of the same survey. Further, the coefficient of variation in estimates of others was reduced from an average of 65\% in the \textsc{others} survey to 55\% in the \textsc{post} survey across all 56 stocks (a statistically significant reduction for 40 of the 56 stocks at the \( p < 0.001 \) level). So, it appears that the process of trading causes participants to converge towards a consensus of opinion. The learning aspects of scalable preference markets could be particularly useful for product categories in which individual new product preferences are shaped by others, such as fashion goods or those with network externalities.

**Discussion**

Study 1 showed that preference markets can be *scalable* by virtue of an experimental design that matches traders with a convenient number of stocks, and creating trading links between the groups. Further, we found that we can incorporate various types of stock types for single attributes as well as full concepts into a preference market and test these simultaneously. Importantly, study 1 supports hypotheses H1a through H4, consistent with Figure 1’s conceptual model of preference markets.

**5.2 Study 2: Field Test of Scalable Preference Markets**

Our second study tests how well scalable preference markets perform outside of a purely academic environment, under real world conditions. We conduct the field test with managers, designers, and engineers within a large firm that is directly engaged in the smart phone industry. Further, we test the limits of scalability by reducing the number of traders, most of whom trade remotely, while increasing the number of stocks. As in study 1, we test the four hypotheses of our conceptual model.
Study Design and Procedure

With the help of an internal innovation team at the firm, 63 people participated in the on-site experiment, of whom 15% were in marketing & sales, 65% in technical positions, 5% in finance and the remaining 15% in other functional areas. Participants reported an average of 4.5 years of industry experience and 53% claimed a management position. The experiment was conducted at the firm’s corporate headquarters, with over 60% of participants accessing the market remotely from their offices, after having completed SELF and OTHERS surveys in advance. The remote participants learned how the experiment worked through a live, 15-minute video web cast with audio questions and answers. The experiment employed the same user-interface and experimental design as Study 1. Six groups were formed, ranging in size from 5 to 15 traders, with 21 to 24 stocks each (Figure 2). While Study 1 had approximately two traders per stock (113 traders, 56 stocks), Study 2 was twice as intensive, with an average of only one trader per stock (63 traders, 64 stocks). In discussion with the firm's executives we defined 64 stocks encompassing of mutually exclusive features, 39 of which could be compared against those in Study 1, and 25 of which were new, and included recent advances and features of interest to the firm.

Results

Referring to the fifth through eighth rows and columns of Table 3, the correlation of 0.863 between the SELF-and OTHERS surveys confirms that respondents are accurate in estimating each other’s new product preferences, supporting H1 as in Study 1. Again, preference market PRICES capture expectations of OTHERS ($\rho = .829$) better than SELF preferences ($\rho = .707$), supporting H2 and H3. For the 29 respondents (46% response rate\textsuperscript{5}) who completed the POST survey, the hypothesis that they learned from trading stocks, H4, is supported by the higher $\rho$ of 0.910 between the POST survey and STOCK prices (as compared with $\rho = .829$), and by the reduction in the average coefficient of variation (c.v.) from 68% for
the OTHERS survey to 41% in the POST survey. In fact, learning manifests itself in the form of statistically significantly reduced coefficients of variation for 61 of the 64, or 95%, of the smart phone features.

For the 39 stocks common to Studies 1 and 2, the students and firm participants diverged somewhat in their SELF preferences ($\rho =$ only .661 between Study 1 and Study 2), so it is not surprising that Study 2’s STOCK prices were weaker predictors of Study 1’s SELF preferences ($\rho = .525$) than they were of study 2’s SELF preferences ($\rho = .707$). Differences between the groups’ results may be due to time-varying new product preferences (Studies 1 and 2 took place 20 months apart), differences in how stocks were defined, and the distinction between students and professionals. Considering all of these differences, we are encouraged by the degree of convergent validity between Studies 1 and 2.

In effect, Study 2 was a real-world replication of Study 1. First, it demonstrated that scalable preference markets perform well in the field, with managers and employees trading in an efficient manner. Specifically, we learned that the majority of traders mastered the user interface and were able to trade remotely from their offices. Further, the high ratio of one stock-per-trader was still sufficient to achieve accurate results. Mitigating the high stock-per-trader ratio and remote participation rate were the high level of participants’ market expertise and the use of easier, mutually exclusive questions. Study 2 produced remarkable results in a very short time, with fewer people, over a very larger number of questions. The wisdom-of-crowds-, expertise-aggregation-, and learn-from-trading effects were all evident.
5.3 Respondent enthusiasm for surveys and preference markets

In addition to the POST survey questions about updated expectations of OTHERS, we asked respondents in Studies 1 and 2 (69% and 87% response rates, respectively) about their relative preference between surveys and stock trading. The results are shown in Figure 5.

Figure 5: Which Method Did Respondents Prefer: Survey or Stock Trading?

There was near-unanimity in preference for stock trading over surveys. Scalable preference markets differ from surveys in that they include elements of competition, interaction, gaming, learning, and the opportunity to gain recognition and win prizes, which might explain the strong result. In addition, 75% of the industry experts in Study 2 expressed a willingness to participate in a preference market again.

6 Discussion

We developed a conceptual model of preference markets and tested a scalable version of the method that worked well in practice. Through two studies, we validated the model that SELF preferences influence expectations of OTHERS, which in turn are reflected in stock PRICES. Of course, reverse causality, in which one’s expectations of others’ tastes and preferences may help form self preferences, may also explain some of our results. Were that the case, measuring expectations of others would be all the more important. But given that
we observed high variation between traders in their expectations of others it seems likely that individuals have more confidence in their self preferences than in their expectations of others, so that causality is more likely to be from SELF to OTHERS. This is further evidenced by the fact that two independent groups of traders had higher correlations between their average expectations of others (0.752) than in their average individual preferences (0.653). In other words, preference markets address the challenges of heterogeneous new product preferences quite well.

Our results suggest that scalable preference markets offer an effective tool for product development teams, especially when large numbers of design decisions need to be prioritized. For example, the top 5-10 stocks may merit further study via conjoint analysis. The number of features and concepts that can be tested scales in the number of traders, with one trader per stock representing a minimum. Respondents express a strong preference for trading stocks over answering surveys. And they learn from each other while trading, updating their expectations in a way that converges towards a clearer consensus.

Despite these promising results, some issues remain: external validity, comparison with conventional methods and directions for future research.

6.1 External Validity

Validating methods with actual, external data poses a challenge in new product development research, as many of the ideas tested may not exist. And even in the case of existing features and concepts, access to accurate data may be limited. Instead, we look to new product releases and comparisons to prior market research studies for at least some degree of validation of the accuracy of our results.

Looking across both experiments, several clear trends emerge in the data.
Table 4 shows that five smart phone traits were preferred by the majority, even at a price premium, in virtually every survey and preference market. These five features can be interpreted as “must haves,” while ten others were consistently rejected by over two thirds of respondents. The rejected smart phone aspects may represent low-priority, or niche, design considerations. From a marketing perspective, the features in the middle represent differentiation opportunities that merit further study. Scalable preference markets facilitate “triage” of customer preferences; design teams may prioritize opportunities and focus their product development efforts.

<table>
<thead>
<tr>
<th>Preferred by a Majority</th>
<th>Heterogeneous Preference</th>
<th>Rejected by a Majority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Size &amp; Weight (3-4”)</td>
<td>Oper. System (Microsoft rising)</td>
<td>Hands Free Operation</td>
</tr>
<tr>
<td>Color Display (320x240+)</td>
<td>Memory Capacity &amp; Battery Life</td>
<td>Bluetooth, Infrared, USB</td>
</tr>
<tr>
<td>Camera (quality rising)</td>
<td>Mini- Keybd. vs. 12-key vs. Stylus</td>
<td>GPS (but rising)</td>
</tr>
<tr>
<td>Verizon Cell Network</td>
<td>WiFi Capability and Push Email</td>
<td>FM radio, Video Camera</td>
</tr>
<tr>
<td>Black or Silver Phone</td>
<td>Slot types (SD rising)</td>
<td>Changeable Faceplates</td>
</tr>
<tr>
<td></td>
<td>MP3 vs. TV</td>
<td>European Compatibility</td>
</tr>
<tr>
<td></td>
<td>Phone Brands and Models</td>
<td>e-Wallet</td>
</tr>
</tbody>
</table>

Table 4 also presents an interesting example of external validity, in that one would expect leading smart phone manufacturers to launch new products conforming to these results. And, as shown in Figure 6, Nokia, Motorola, and BlackBerry launched smart phones in 2006 that largely fit the table and appeared to be converging towards a dominant design. On January 27, 2007, Apple shook up the smart phone market by humanizing the dilemma of keypad vs. mini-keyboard vs. stylus user interface with its innovative touch screen interface, which has the added benefit of greater screen real estate in many applications. The iPhone included all of the “preferred by a majority” features on the left side of Table 4 except for the cell network, for which Apple opted to strategically partner with AT&T, and with the exception of Bluetooth, left out all of the features “rejected by the majority.”
Further, we compare stock prices from Studies 1 and 2 against self-stated preferences for 14 and 11 features, respectively, from two individual surveys conducted independently of the present research. The 2004 study, with 518 MBA student respondents, was published recently in a leading journal, and the more recent 2005 study, also with MBA respondents, is part of a working paper. Both studies focused on new methods of conjoint analysis.

Table 5: Indication of External Validity (Pearson correlations between stock results and two independent studies, 14 and 11 attributes compared)

<table>
<thead>
<tr>
<th></th>
<th>1 PRICES</th>
<th>2 PRICES</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004 Study (n=518)</td>
<td>.802**</td>
<td>.714*</td>
</tr>
<tr>
<td>2005 Study (n=206)</td>
<td>.885**</td>
<td>.769*</td>
</tr>
</tbody>
</table>

**: p < .05, *: p < .10

The correlations, which range from 0.714 to 0.885, relate the two sets of stock market results from our experiments to the two external studies, and provide further evidence of external validity. We note that in Table 5, Study 1’s stock trading, also conducted with MBA students, correlates slightly better with the external data than does Study 2, which was conducted with industry experts.

The data support the hypotheses, and offer a reasonable degree of external validity, leading to the conclusion that preference markets can be quite useful to new product development teams in measuring product concept and attribute preferences as part of NPD.
Comparison of Preference Markets with Conventional Methods, Limitations and Directions for Future Research

Table 6 compares preference markets with other methods, and highlights their scalability. Preference markets complement other methods by narrowing a large number of potential product features and concepts to a manageable set that can be further analyzed at the individual level using the other approaches. Further, distinct benefits of preference markets over survey-type methods are interaction, competition, and learning among participants. More importantly, preference markets scale up in the number of respondents much more easily than surveys.

An important limitation of scalable preference markets is that they do not measure individual preferences. Our results demonstrate that markets achieve a consensus about expectations of average preferences, and do not provide insight about distinct individuals. To measure heterogeneity, methods such as conjoint analysis are better suited to the task. Of course, one could devise markets that attempt to estimate segment-level preferences by cleverly defining stocks and trader groups using the approach taken in Figure 4, but this would still be a blunt instrument for measuring individual differences.

Implementation of preference markets in firms requires the firm or outside consultants it may engage to develop trading software and infrastructure. Respondents need to be taught the mechanics of trading and the underlying meaning of each stock. The key outcomes, the stock prices themselves, become known to all traders immediately, so data security may pose a problem. And the market mechanism itself pulls no punches; the consensus view, whether positive or negative, becomes instantly transparent. Champions of specific product ideas may not readily accept negative outcomes, a challenge with any market research, but one which might be exacerbated by the immediacy of preference markets.
The present research is limited by the fact that we only studied a single product category. It remains to be seen how well the method will translate in contexts in which the innovation type, product type, or customer characteristics vary. We leave this to future research, but expect that preference markets may perform in a surprisingly robust way much as do heterogeneous investors in financial markets in evaluating numerous industries and firms.

Table 6: Comparison with conventional methods

<table>
<thead>
<tr>
<th>Description</th>
<th>User Design</th>
<th>Conjoint Analysis</th>
<th>Self Explicated</th>
<th>Preference Markets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantages</td>
<td>Individuals customize optimal products</td>
<td>Individuals rate-, rank- or choose feature bundles</td>
<td>Individuals rate importances of unbundled features</td>
<td>Trader groups achieve consensus through trading</td>
</tr>
<tr>
<td></td>
<td>Identifies optimal feature bundles from many combinations; engaging task</td>
<td>Quantifies tradeoffs over a finite number of features; measures individual utility</td>
<td>Quantifies individual trade offs over more features; easier task</td>
<td>Measures consensus preferences over many features and concepts; scalable; engaging, fun</td>
</tr>
<tr>
<td>Disadvantages</td>
<td>Does not measure trade-offs; setup costs can be high</td>
<td>Task difficulty, response error, complex analysis</td>
<td>Potential problem of “everything is important”</td>
<td>Group preferences only; simultaneous participation needed</td>
</tr>
<tr>
<td>Best Fit Applications</td>
<td>Customized goods; optimal bundles; Key feature go/no go’s</td>
<td>Optimal design/price and positioning for a few key decisions</td>
<td>When conjoint is too difficult or costly, or too many features</td>
<td>Narrow many options, group consensus; when speed is key</td>
</tr>
</tbody>
</table>

Optimizing the application of preference markets will require further research. We hope that fellow researchers will experiment with the design options listed in Appendix 1, so that the accuracy and applicability of scalable preference markets can be improved. Specifically, future research might study: (a) new stock types such as “no buy” or “price-elasticity” stocks (i.e., product and feature take-up rates at varying retail prices), (b) new incentives and rewards, (c) better tests of internal and external validity, (d) the impact of product-related “news” revealed during trading, and (e) the connection between individual trading behavior and individual preference.

Considering the scalability, flexibility, speed, and attractiveness to respondents of preference markets, we anticipate that the methodology will gain adherents over time, enabling firms and their product development teams to prioritize the features and concepts that address the consensus opinions of the market.
References


## Appendix 1: Design of Preference Markets

### Design of Virtual Securities

| Stock definition (Stimuli) |  
|---------------------------|---|
| - Product features        |   |
| - Product concepts or real products |   |
| Stock structure |  
| - Binary                |   |
| - Mutually exclusive     |   |
| - Bundles                |   |
| Include price for the option or not |   |

### Experimental Design

| Respondents |  
|-------------|---|
| - How many respondents? |   |
| - Which population: Convenient sample, consumers, insiders, outsiders, users, experts, managers |   |
| - Open or closed access |   |
| - How many repetitions? |   |
| - Remote access or central location? |   |
| Assignments of stocks |  
| - Random |   |
| - Self selection / Rule-based |   |
| - Sorting / Filtering |   |
| Information display |  
| - Allow outside information |   |
| - Show last price, quantities, volume, order book, past data, rankings |   |

### Market Mechanism Design

| Trading actions |  
|-----------------|---|
| - Limit orders, market orders, restrictions |   |
| - Short selling, options |   |
| - Position limits and price limits |   |
| - Trading fees or no trading fees |   |
| Trading mechanism |  
| - Call auction |   |
| - Market maker (i.e., dealer) |   |
| - Pari-mutuel engine |   |
| - Double auction |   |
| Duration, Trading hours |  
| - Initial conditions |   |
| - Initial prices |   |
| - Endowment |   |
| - “Bank” |   |

### Incentives

| Reward type |  
|-------------|---|
| - Non-monetary (Recognition, enjoyment) |   |
| - Monetary rewards |   |
| Rewarding rules |  
| - Not based on performance: everyone, random sample |   |
| - Performance based: tournament, sample, everyone, proportional to the portfolio value |   |
| Rewarding |  
| - Accuracy: best predictors |   |
| - Effort: trading behavior |   |
| Winner determination |  
| - Actual prices |   |
| - Other replications |   |
| - “truth” |   |
| - Exogenous: parallel market research |   |
Appendix 2: Three Stock-types for Study 1

(19) Mutually Exclusive Smartphone Feature Levels and (6) Mutually Exclusive Smart Phones
(each of the 6 categories sums to 100%)

(31) Binary Smartphone Feature Levels
(each garners between 0% and 100% “share” at the feature price shown)
1 Traders are informed prior to trading that stock prices are defined as the percentage of people who chose to purchase each product or feature, a specific brand in this case, at a given retail price.

2 Securities markets involve many more than just four design decisions, as delineated in Appendix 1.

3 No significant differences were observed between the respondents and non-respondents to the POST survey in terms of trading activity, offers to buy and sell, and trader performance.

4 All of our key results hold when using closing prices, as they are highly correlated with VWAP, but we note that VWAP slightly improves the accuracy of these results as shown in Dahan et al. (2008).

5 There are no significant differences between the 29 individuals who responded to the POST survey and the 34 who did not in the amount of trading activity, offers to buy and sell, and performance.