



2010

Improving the Crystal Ball: Consumer Consensus and Retail Prediction Markets

Carol Kaufman-Scarborough
Rutgers University

Maureen Morrin
Rutgers University

Greg Petro
First Insight LLC

Eric T. Bradlow
University of Pennsylvania

Follow this and additional works at: http://repository.upenn.edu/statistics_papers



Part of the [Business Commons](#), and the [Statistics and Probability Commons](#)

Recommended Citation

Kaufman-Scarborough, C., Morrin, M., Petro, G., & Bradlow, E. T. (2010). Improving the Crystal Ball: Consumer Consensus and Retail Prediction Markets. *Journal of Research in Interactive Marketing*, 4 (1), 30-45. <http://dx.doi.org/10.1108/17505931011033542>

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/statistics_papers/11
For more information, please contact repository@pobox.upenn.edu.

Improving the Crystal Ball: Consumer Consensus and Retail Prediction Markets

Abstract

Retail buyers' forecasts, decisions, and subsequent purchases result in billions of dollars of merchandise being purchased and offered for sale by retailers around the world. However, academic research examining this decision process has been limited, and recommendations for improvement almost nonexistent. In the present study, we begin to address this issue by introducing a new approach that compares retail buyers' consensus forecasts with those from a sample of "ordinary" consumers. The potential for incorporating forecasts from ordinary consumers suggests an opportunity to create what are termed retail prediction markets, which offer significant potential to improve the accuracy of buyers' forecasts. We conclude with limitations and areas for future research.

Keywords

Forecasting, Retailing, Buyers, Consumer behavior

Disciplines

Business | Statistics and Probability

Improving the Crystal Ball: Consumer Consensus and Retail Prediction Markets

Carol Kaufman-Scarborough, Maureen Morrin, Greg Petro, & Eric T. Bradlow

August 1, 2006

Carol Kaufman-Scarborough is Professor of Marketing and Undergraduate Program Director at Rutgers University School of Business, Camden, NJ 08102-1656, telephone: 856-225-6592, email: ckaufman@camden.rutgers.edu. Maureen Morrin is Associate Professor at Rutgers University School of Business, Camden, NJ 08102-1656, telephone: 856-225-6713, email: mmorrin@camden.rutgers.edu. Greg Petro is President of First Insight, LLC, Pittsburgh, PA. Eric T. Bradlow is the K.P. Chao Professor, Professor of Marketing, Statistics, and Education, and Academic Director, Wharton Small Business Development Center, The Wharton School of the University of Pennsylvania, 3730 Walnut Street, 761 JMHH, Philadelphia, PA 19104-6340, telephone: 215-898-8255, email: ebradlow@wharton.upenn.edu.

Improving the Crystal Ball: Consumer Consensus and Retail Prediction Markets

Retail buyers' forecasts, decisions, and subsequent purchases result in billions of dollars of merchandise being purchased and offered for sale by retailers around the world. However, academic research examining this decision process has been limited, and recommendations for improvement almost nonexistent. In the present study, we begin to address this issue by introducing a new approach that compares retail buyers' consensus forecasts with those from a sample of "ordinary" consumers. The potential for incorporating forecasts from ordinary consumers suggests an opportunity to create what are termed retail prediction markets, which offer significant potential to improve the accuracy of buyers' forecasts. We conclude with limitations and areas for future research.

Improving the Crystal Ball: Consumer Consensus and Retail Prediction Markets

Large retailers' profits are critically dependent upon the ability of their merchandise buyers to accurately anticipate consumer demand. Retail buyers are responsible for choosing the products comprising the retail assortment that will be resold to the ultimate consumer (Kumar 2001; Ettensen and Wagner 1986). Their forecasts and subsequent purchases result in billions of dollars of merchandise being purchased and offered for sale by retailers around the world. Retail buyers determine the amounts of each item to be purchased, develop and implement pricing and promotional plans, and manage the interaction of these variables over the course of the selling season (McIntyre, Achabal, and Miller 1993).

In recent years, department store customers have voiced dissatisfaction over poor style choices, inadequate sizes, and stockouts (Lazorchak and O'Neal 2001). As a result, major retailers such as the Gap, Dillard's, Federated, and Saks have had to resort to significant discounting of their merchandise to unload products that could not be sold at or near full retail price due to a mismatch between forecasted and actual consumer demand for the items in the store (Derby and Zaczekiewicz 2006; Young 2006). Moreover, some major strategic buying decisions have been reversed in response to negative consumer reactions, such as Saks' recent restoration of their petite collections after scores of complaint letters were received following the collection's elimination (Barbaro 2006). These problems indicate that there is significant room for improvement in the ability of retail buyers to understand consumer preferences and accurately forecast product demand. While prior academic research has examined the congruency between consumers' preferences and retailers' assortment strategies (Morales *et al*,

2005; Van Herpen and Pieters, 2002), such studies do not provide insights regarding consumer reactions to specific products within those planned assortments.

Despite the importance of retail buyers' decisions, academic research examining their decision process has been rather limited and recommendations for improvement almost nonexistent (Davies 1994; Fairhurst and Fiorito 1990). The current research explores the role that input from ordinary consumers might play in improving retailers' forecasts. The potential for incorporating consumer consensus forecasts suggests an opportunity to create what Ray (2006) has termed "prediction markets" in the field of retailing, based on decisions buttressed by the forecasts of both experts and novices, which may offer significant opportunities for enhancing the accuracy of retail buyer forecasts.

The retail buyer decision-making process is a complex one, as buyers are required to plan and manage all the steps within the merchandising cycle for the products under their control (McIntyre, Achabal, and Miller 1993). They must be able to understand consumer preferences, interpret future product trends, procure the best mix of products for their customers (Choi and Gaskill 2000), plan for both short and long product life cycles, set initial prices, and manage markdowns throughout selling periods. Despite this critical role, little is known about the factors that affect a buyer's success, or what types of informational input or decision aids might improve their decision accuracy. Hansen and Skytte (1998) note the lack of analytical development in this field due to scattered and unrelated studies dominated by a concern for analyzing lists of decision criteria used by buyers. Moreover, some of the past studies have relied on an organizational buying model (Sheth 1973, 1981; Shim and Kotsiopoulos (1991); Choi and Gaskill 2000; Fairhurst and Fiorito 1990; Wagner, Ettenson, and Parrish 1989), the applicability of which to the retail buyer decision-making process has been questioned (Kline and Wagner 1994).

Prior research involving actual retail buyers has often been retrospective in nature, with buyers interviewed or asked to complete self-administered questionnaires that involving recalling their decision process from memory (McIntyre, Achabal, and Mille, 1993; Wall, Sommers, and Wilcock, 994; Choi and Gaskill 2000). Recall-based studies are limited in that they depend on respondents' willingness and ability to accurately self-report their decision processes on a post-hoc basis. Another approach has involved exposing retail buyers to hypothetical scenarios and examining the decision criteria used (e.g., Kline and Wagner 1994). This research stream also has its limitations due to the hypothetical nature of the methodology employed.

Despite its limitations, this prior stream of research has importantly shown that buyers rely heavily on their intuition in making buying decisions (e.g., McIntyre, Achabal, and Miller 1993). Although retail buyers sometimes rely on suggestions from the manufacturers' sales representatives in making their decisions (Shim and Kotsiopulos 1991; Fairhurst, Lennon and Yu 1996; Kincaide, Woodard and Park 2002), the most frequently cited information source for buying decisions is typically the retail buyer's own knowledge, (e.g., Hirschman and Mazursky 1982; Klein and Wagner 1994). Interestingly, some research indicates that the most important external source of information used by retail buyers is customer requests (Klein and Wagner 1994), suggesting that buyers might indeed welcome direct consumer input into their buying decisions.

Prior research from the forecasting literature suggests that forecasting accuracy in other domains of expertise can be significantly improved by supplementing managerial intuition with database models (Blattberg and Hoch 1990) or by aggregating estimates from individuals, especially those of experts (see Clemen 1989 for a review). This would suggest that simply aggregating retail buyers' forecasts should enhance their forecast accuracy. Even simple

averaging techniques have been found to significantly enhance accuracy, with combination forecasting having been fruitfully applied in several areas such as meteorology, macroeconomics, etc. The purpose of the present research is to examine a slightly more nuanced notion: namely, whether the consensus forecasts of expert retail buyers might be significantly enhanced with consensus input from ordinary consumers, the ultimate goal of which would be to improve retail buyers' decision accuracy and thus retail profitability.

The present research is predicated on the notion that the "wisdom of crowds" (Surowiecki 2004) can enhance retail buyers' forecasting accuracy. In the current context, the "crowd" refers to ordinary consumers without any particular expertise or experience in the field of retailing. Surowiecki (2004) argues that groups of ordinary (i.e., untrained, non-expert) people can often outpredict even the most knowledgeable of experts because they bring a diversity of opinions and level of independent thinking that is often absent among experts. Examples are cited about how the collective intelligence of a firm's employees often results in more accurate new product forecasts than those of the firm's product managers (Nocera 2006).

Recent internet developments that allow for direct input from non-experts (e.g., Google's page rank algorithm, wikipedia, online social networks, blogs) demonstrate some anecdotal support for the counterintuitive notion that ordinary groups of people "get it right" more often than supposed experts in the field (Nocera 2006). Ray (2006) coined the term "prediction markets" to capture the idea that forecasts which aggregate the input of non-experts with that of experts are "uncannily accurate" in predicting financial trends such as interest rates, inflation, stock prices, etc. We explore the potential for creating prediction markets in the field of retailing in the current research.

What has been lacking, until now, is a format for obtaining non-expert input in a time-efficient manner that can then be utilized by an individual or group of expert decision makers, to harness the consensus input (Fitzgerald 2006) of consumers. In this paper we describe one way to harness the input of ordinary consumers into retailers' product forecasts. We describe a study we conducted that involved obtaining real-time product evaluations from both retail buyers and ordinary consumers of items that would shortly arrive on the store shelves of a major department store. We then monitored each item's sales and profitability over time and compared the in-store performance of each of the items with the respondents' prior evaluations of those items. The overall forecasting accuracy of the retail buyers' evaluations is then compared to that of the consumers as a whole and to a subset of more predictive consumers, and several illustrative product forecasts are discussed. The results strongly indicate that consumer consensus input could improve retailer forecasts.

Below, we first present details of the data collection methods that were used; a novel approach that involves obtaining real-time forecasts of retail buyers and “ordinary” consumers like “you-and-me.” In such a manner, we demonstrate that our approach is both economically and implementably practical. This is followed by a description of the results, which suggests significant improvement would likely be obtained by incorporating consumer consensus forecasts into retail buyers' forecasting decisions, which in turn suggests the potential for the creation of retail prediction markets. We conclude with limitations and areas for future research.

Method

Although there are numerous ways in which one could supplement the judgment of retail buyers (e.g. with historical data, analytical forecasting tools, etc...), the approach taken here is consistent with that of other disciplines where forecasts are improved by utilizing the knowledge of others (Armstrong 2006, <http://www.forecastingprinciples.com/>). While CATI (computer-assisted telephone interviewing) methods were one possible option, we instead chose to implement data collection in a way that may be more quickly amenable to firms and one that allowed full-color product representations, i.e., a real-time online survey. The survey design and subject recruitment process are described next.

Participants

For the purposes of this study, an online survey was completed by a total of 235 respondents, consisting of 19 retail buyers who worked at a major U.S. department store corporation and 216 ordinary (i.e., non-expert) consumers.

Retail Buyer Sample. Access to a sample of retail buyers was critical to this research, as was the choice of store items to which we would apply our method. Nineteen retail buyers participated in the study. They were made available to the researchers by a major U.S. department store corporation, which was promised early access to the study's results. These were buyers who currently purchased apparel/footwear/accessories/jewelry for a major division of the department store chain. The 19 retail buyers were each given \$50 gift certificates to local restaurants for their participation. All of the retail buyers were female, with an average age of 34, and an average of 14.5 years of work experience in the retail industry, of which an average of 9 of those years were spent as a retail buyer.

Each retail buyer evaluated a total of 19 products included in the survey (described in more detail below), although each buyer had personally been responsible for having recently ordered just one or two of these products for their division. To maintain confidentiality, the researchers did not attempt to identify which item had been personally ordered by which buyer. Instead, we analyzed the consensus evaluations of the retail buyers.

Consumer Sample. The consumer sample, which consisted of 216 respondents, was recruited via online announcements to a university community and via word-of-mouth communication. Consumer respondents were paid \$5 for their participation or were entered into a lottery to win one of several \$100 prizes. The consumer sample was about two-thirds female (64.8%), with a mean age of 31 years (range: 16 to 77), and a mean household income of \$72,000. Seventy percent of the consumers reported they normally shopped in department stores. Although this population is slightly wealthier, more highly female, and younger than the population in general, we have no reason to believe that the gains observed here are abnormal or unexpected. Future studies could certainly explore this issue and maybe more importantly explore whether there are other more accurate forecasting populations that could be obtained via (more costly) non-random/convenience sampling that could provide greater benefits.

Product Stimuli

The department store corporation provided us with access to approximately 40 products that would be hitting their stores' shelves a few days before the data collection period began. These 40 items were arrived at by asking each of the 19 buyers who would be taking part in the survey to recommend about two products they needed to make decisions on in the upcoming

season for use in our study. We requested that these items be "new" in the sense that the buyers did not have previous experience with them in the store. Thus, we did not want items for which there was extensive historical sales data or experience that could be used to predict in-store performance (and thus, for which the forecast would have minimal uncertainty).

We chose 19 of these 40 items to be included in the final study, on the basis of providing adequate breadth in terms of product type (e.g., clothing, accessories, shoes, jewelry) and price point range (full retail prices ranging from \$32 to \$360). Just a few days before these 19 items were to arrive at the stores (i.e., before any sales results had been obtained), they were photographed by one of the researchers and uploaded into the online surveys for the retail buyers and consumers. These items are listed in table 1 and an example of one of the product's digital images is provided in Appendix 1. We have masked the brand names and other product details to maintain store confidentiality. The department store also provided the researchers with confidential information on each of the 19 items over time that included in-store sell-through, price discounting, costs, profit margins, etc. which would be used, as described below, in profit calculations.

Survey Instrument

The data collection instrument consisted of an online survey that included full-color, high-quality digital photographs of the 19 products that would soon be sold in a major division of the department store. Respondents filled in an online survey that included digital photographs of the 19 items before the items arrived at the stores for sale to the public. Respondents were told the survey concerned a retail fashion study designed to help us better understand how people make buying decisions in department stores. They were told they would be shown 19 items and

asked about how they felt the typical shopper would respond to each. Because prior research has shown that the correspondence between consumers' own purchase intentions and actual sales is not high (Armstrong and Brodie 1999; Morwitz 2001), we chose to ask respondents how the "typical department store customer" would respond to the products. Respondents to the online survey examined each item (photograph and brief description) and evaluated each product with several ten-point closed-ended questions.

The first question that appeared below the photograph of an item asked: "How likely would the typical department store customer be to purchase this item" on a scale of 1 = Never to 10 = Definitely. Another question asked about the highest price a typical department store customer would be willing to pay for the item. For each item, the manufacturer's suggested retail price (MSRP), which was provided by the department store corporation, served as the midpoint of the price estimate scale; and plus or minus 25% from the MSRP were the scale endpoints. Another question asked whether the respondent would classify the item more as a wardrobe basic or a high fashion item. Respondents were also asked to rate their confidence in their purchase likelihood and price estimate responses. Both the consumers and the buyers each took almost identical surveys, with minor differences in demographic responses (e.g., we asked consumers about household income and buyers about work experience in the industry).

Results

Mean Responses

We first collapsed across the nineteen products and obtained a mean score for each individual for each of the five main dependent measures: purchase likelihood, purchase likelihood confidence, pricing (i.e., highest price willing to pay), pricing confidence, and

fashionableness. We conducted a MANOVA on these five dependent measures as a function of whether the respondent was a consumer or retail buyer. The multivariate tests were significant ($F(5, 229) = 9.06, p < .0001, \eta^2 = .17$). The retail buyers exhibited a higher mean score on three of the five measures: purchase likelihood ($M_C = 6.24$ vs. $M_B = 7.60; F(1, 233) = 17.55, p < .0001, \eta^2 = .08$), purchase likelihood confidence ($M_C = 7.22$ vs. $M_B = 7.97; F(1, 233) = 4.35, p = .038, \eta^2 = .02$), and fashionableness ($M_C = 5.99$ vs. $M_B = 7.42; F(1, 233) = 27.08, p < .0001, \eta^2 = .10$). There were no differences between consumers and retail buyers in their mean estimates for pricing ($p > .35$) or pricing confidence ($p > .50$).

These results suggest that retail buyers tend to be more optimistic about purchase likelihood and are more confident about their purchase likelihood predictions, compared to consumers, as well as more positive in the degree to which they perceive items to be fashionable (vs. more basic in nature). This result, in and of itself, may suggest that part of the reason why widespread price discounting occurs in the major department stores may be due to overly optimistic estimates of purchase likelihood and fashionableness of the products by retail buyers compared to consumers, assuming retail buyers take this into account when forecasting and ordering. Interestingly, there were no overall differences in estimates of how high of a price buyers and consumers thought products would sell for, or their levels of confidence in pricing, although their pricing estimates for individual items often varied significantly (see figure 1). Taken together, these results suggest that retail buyers may make their intuitively-based forecasting decisions based more on their (overly optimistic) forecasts of purchase likelihood and fashionableness, rather than on their forecasts of pricing. If encouraged to base their forecasts more on pricing estimates rather than purchase likelihood or fashionableness, this alone might significantly enhance retail buyers' forecasts.

Insert figure 1 about here

Key Metric: Profitability Gap

We next wanted to measure how well the retail buyers' and consumers' evaluations of the 19 products could predict each product's in-store performance in terms of profitability, a key metric for most retail operations. This allows us to examine the financial impact of the marketing efforts of the retail buyers and their contributions to the firm's marketing productivity (Lehmann 2004; Rust, Ambler, Carpenter, Kumar, and Srivastava 2004; Rust, Lemon, and Zeithaml 2004). In the present context, we focus on the financial accountability of the retail buyers' merchandise selections, pricing policies, and promotional strategies. To do this, we obtained sell-through, cost, and profitability measures from the department store corporation for 10 weeks of sales for each of the 19 items in the store's division that employed the 19 retail buyers. We used this information to assess item profitability by calculating Gross Margin Return on Investment (GMROI).

GMROI is a commonly used measure of buyer performance in the retail industry (McGinnis, Gable, and Madden 1984). GMROI is arrived at by dividing total gross margin dollars by average inventory at retail. It indicates whether a sufficient gross margin is being earned relative to the investment in inventory that is used to generate the gross margin dollars. While other measures can be used to assess buyer performance (such as sales revenues, inventory turnover, etc.), GMROI is often preferred in retailing because it is easy to compute, its components are readily available to all retailers, and more importantly, it measures how

profitably business assets have been deployed, which is consistent with corporate ROI goals (Fairhurst and Fiorito 1990; Bates 1979; McGinnis, Gable and Madden 1984). Moreover, the measure is sensitive to buyer forecasting errors, which may lead to products being liquidated at clearance prices because too much product was bought or it was offered at too high a price, or lost sales lost due to stockouts because not enough product was bought or it was offered at too low a price. Many analysts define retail success as “achieving high gross margins and customer service levels with as little inventory as possible” (Mattila, King, and Ojala 2002).

Since GMROI can vary by department and by product (e.g., some products are bought at a lower price point compared to MSRP which will boost GMROI, but not on the basis of decision effectiveness), we normalized the measure for the nineteen products by dividing each product's actual GMROI by its maximum potential GMROI (i.e., the GMROI that would be obtained if every unit purchased was sold at full price), to arrive at a percentage 'gap' from profit potential. We refer to this metric as the GMROI profitability gap. Each of the 19 items' GMROI and GMROI gaps (in absolute and percentage terms) are provided in table 1. Table 1 indicates that item #16 (the designer suit) was the least profitable of the 19 items and item #5 (the designer brown sandal) was the most profitable, in terms of profitability gap size.

Insert table 1 about here.

Predictiveness of Product Evaluations

To assess the forecasting accuracy of respondents' product evaluations, we calculated correlation coefficients for each respondent between the nineteen products' profitability gaps and the individual's responses for purchase likelihood, pricing, purchase likelihood confidence, pricing confidence, and fashionableness. We report in table 2 mean correlations across all respondents, as well as broken out for the consumer and retail buyer groups. Negative correlations indicate better predictive abilities (e.g., if a respondent thinks a product will sell at a higher price, and the product exhibits a smaller profitability gap, the respondent is a better forecaster).

Insert table 2 about here.

Table 2 indicates that pricing-related estimates of respondents tend to be the best predictors of the items' profit potential. This result holds for both consumers and retail buyers. As a result, we utilized respondents' pricing estimates as the best indicator of an item's profitability (i.e., as a respondent's best forecasting measure).

Note that there are no statistically significant differences between the total consumer and buyer samples on any of the mean correlations. This result may be somewhat disconcerting to retailers, as it suggests their expert buyers on average are no more predictive than a convenience sample of consumers, in terms of predicting in-store performance of the products. However, this

result is in accord with prior research that suggests experts may oftentimes be no more accurate in their predictions than novices (Armstrong 1991).

The next issue we were interested in exploring was whether it is possible to find a more accurate or predictive consumer consensus based on a subset of consumers from the entire convenience sample who could possibly "outpredict" the retail buyer expert consensus group. Note, of course, that one could try and self-select a set of expert retail buyers that could outperform the norm as well. This would be equally valid; however, as we describe below there was not a retail buyer subset of significant magnitude in our sample that would allow for this.

Consumer Subset

In arriving at a group of better consumer predictors, we considered how large such a group should be. A group of 5 to 20 experts has been considered optimal in prior research (Armstrong and Brodie 1999), with an equal weighting of each individual's opinion. To arrive at our more predictive consumer subset, we selected those consumers whose correlation between pricing and profitability gap was less than -0.40 . This approach resulted in a group of 26 consumers, whose mean correlation coefficient between pricing estimate and profitability gap was -0.505 (range: -0.41 to -0.72). We henceforth refer to this group of individuals as the consumer subset. The subset's mean age was 31.7 years, 57.7% were female, and they reported a mean household income of \$98,756. Thus, compared to the entire consumer sample, the consumer subset exhibited a higher income level, and a larger proportion was male.

We utilized a somewhat arbitrary correlation coefficient cut-off of -0.40 also because it resulted in subset size ($n = 26$) about equivalent to that of the retail buyer group ($n = 19$). Note that if we had made the cutoff more stringent, such as at -0.50 , to result in a smaller consumer

subset, to more closely match the size of the retail buyer group, this would have made the consumer subset even more accurate than the retail buyer group. Specific circumstances in the real world could determine just where the accuracy cut-off should be. In addition, just due to pure “randomness and large samples,” as the size of the total consumer panel grows, the ability to find a consumer set with better predictive ability of a given size will increase. As we discuss in our concluding section, while in one sense this is a statistical artifact, in another sense this does suggest an opportunity; however, one would need to validate this group for another decision/context to confirm that pure chance was not driving the result.

The consumer subset ($n = 26$) clearly possessed greater predictive accuracy compared to the retail buyer group ($n = 19$) on the basis of the ability of their pricing estimates to predict item profitability ($r_C = -.51$ versus $r_B = -.16$; $t(43) = 7.54$, $p < .0001$). Alternatively, we can say that for this particular subset of consumers, their pricing estimates explained $(-.51)^2 = .26$ or 26% of the variation in the 19 items' profitability levels. The 19 buyers' pricing estimates, by comparison, explained just $(-.16)^2 = .03$ or 3% of the variation in the items' profitability levels. To obtain an equally predictive group of retail buyers (i.e., a group consisting of individuals meeting the correlation coefficient cutoff of $-.40$ or better) would result in a group of just two retail buyers from the sample of 19.

Table 3 contains some additional information comparing the consumer subset with the retail buyer group. Specifically, it shows that the consumer subset evaluated the 19 items lower than did the retail buyers on purchase likelihood ($p < .0001$) and degree of fashion ($p < .0001$) but not on mean price or self-reported ability to forecast customer demand. Thus, it does not appear that the highly predictive consumer subset believes that items will sell at lower prices overall than the retail buyers, but rather that they can determine which of the items within the set

of products will not command higher prices. That is, the predictive consumer subset is better able to discern price sensitivity within the product assortment. Nevertheless, the consumer subset does not believe they possess greater predictive forecasting skills than do the retail buyers.

Insert table 3 about here.

Illustrative Examples

Next, we provide illustrative examples of how the retail buyers might utilize consumer consensus input in real-time to improve their forecasting accuracy. To do this, we refer to three item-specific examples from our study that were specifically chosen to show differing aspects of how consumer input could be of use. Bar charts in figures 2 to 7 display the distribution of responses to the pricing question for each of these three items (highest price willing to pay) for the consumer subset and retail buyer groups. The height of each bar represents the proportion of the sample choosing that price point for an item.

We first take a look at the least profitable of the 19 items in the assortment of products that was evaluated, item #16, the designer suit (MSRP = \$360). In figure 2 we see that few consumers thought the suit would sell at \$360 (which was the MSRP and midpoint of the pricing scale). Instead, we see most consumers expecting it to sell at the lowest price on the scale = \$252, with decreasing numbers of consumers believing it would sell at successively higher price points. Visually, the heights of the bars for the consumer subset indicate a downward sloping demand curve, indicating a high degree of price elasticity (i.e., price discounting would be

expected to significantly increase unit sales for this item.) We find a similar pattern of expectations for this product from the retail buyers, in figure 3. If the buyer making the decision regarding purchasing this item were to obtain consumer feedback that matched her own fairly pessimistic expectations regarding pricing for this item, the buyer could more confidently reduce (or eliminate) units ordered, or plan for a certain amount of price discounting necessary at an early stage in the selling season to ensure sell-through. In this instance, the consumer input would increase the retail buyer's certainty regarding her intuition that the item is overpriced, which would then result in an adjusted forecast.

Insert figures 2 to 7 about here.

The next illustration is item #1, an item of jewelry consisting of a pendant. In this case, consumers were considerably more bullish on the item's pricing power than were the retail buyers. Figure 4 shows that a substantial proportion of the consumer subset believed the pendant would sell at \$65 (the MSRP and midpoint of the scale). Yet in figure 5, we see the suggestion of a more elastic demand curve (i.e., steeper downward slope) from the retail buyers, with the majority choosing the lowest price point on the scale (\$47). In this case, had the buyers had consumer input, they likely would have resisted discounting this item early on in its life cycle. What occurred in the stores was an almost immediate discounting effort, likely due to weaker price point expectations by the buyer. Margins (as compared to certainty in the previous

example) likely would have been enhanced had the buyer known in advance that consumers would support a higher price point for this item.

Finally, we illustrate what occurred with item #6, a designer sneaker. The pattern of responses from the consumer subset and the retail buyers in this case is representative of what occurred for many of the other items in the product assortment. In this case, consumers were considerably less certain about the product's ability to sell at a higher price point, compared to the set of retail buyers. In figure 6 we see the suggestion of a highly elastic demand curve, indicating that few consumers thought the product would sell at \$98 (MSRP and midpoint of the scale) or higher. This pattern contrasts sharply with the retail buyers' expectations as depicted in figure 7, which suggests an upward-sloping demand curve. In fact, the majority of retail buyers believed the item would sell at a price point considerably higher than the MSRP. If the buyer for this item had access to consumer input prior to placing the order, she could reduce the order quantity or plan for an early and heavy price discounting effort after the product's arrival in store.

Discussion

Sales forecasting is a critical component of marketers' success (Armstrong and Brodie 1999). We demonstrated here that ordinary consumers clearly could provide valuable real-time input to the retail buyer decision-making process. We found not only that a group of ordinary consumers are about as good as a group of expert retail buyers at evaluating and forecasting the in-store performance of products, but that subsets of consumers can easily and inexpensively be found whose predictive capacity considerably exceeds that of an expert retail buying group. In this study the consumer subset's price estimates explained 26% of the variation in the products' profit performance, compared to just 3% explained by the retail buyers' price estimates. The

study presented here demonstrates that consumers' collective input has considerable potential to improve retailers' product forecasts, which could result in more preferred in-store product assortments, more adequate quantities, realistic initial selling prices, less discounting, higher retailer margins, and more satisfied customers. The results suggest, in essence, that retail buyers who incorporate consumers' expectations could create retail prediction markets. A conservative estimate of the value of a one percent improvement in GMROI to a major retailer that generates \$25 billion in annual revenues is \$400 million in incremental margin.

The current approach is made possible by the rapid and widespread growth of the internet, which has effectively removed many of the barriers that previously existed in harnessing ordinary consumers' predictive forecasts with the potential for improving the buying decisions of expert retail buyers. The online survey, for example, allowed respondents to view color photographs of the actual products as they made their predictions, adding realism and timeliness unavailable in paper and pencil questionnaires. As technology enhances the potential for connectedness among consumers, retailers have the opportunity to benefit.

Limitations

The current set of results is of course limited to a single department store corporation's products, a single group of buyers and consumers, and one specific selling season. Future field experiments using choices from larger assortments of items over different purchase cycles would increase the generalizability of results. The ultimate test of such an approach would require repeated testing of the effectiveness of the consumer subset group, to see whether their predictive abilities hold up over time and across decision contexts (given seasonality effects, etc.).

Modifications of the survey may also be a future option that would provide additional insights. Since we often found that respondents chose a price point at the very end of the offered continuum (e.g., consumers often chose the lowest price point on a scale), this would suggest that a wider range of prices would better capture respondents' expectations. If so, this would suggest that an even larger proportion of variance in item profitability could be explained with respondents' expectations. In addition, simpler product descriptions may be needed, as some respondents were not familiar with specific items' apparel types (e.g. "shrug") or their designs ("crochet" tank). It is also possible that buyers indicated higher confidence in their responses due to some level of self-consciousness, particularly when evaluating a product that they themselves had been responsible for buying; albeit this was somewhat rare in our data set.

Some potential threats to the type of approach tested here consist of the possibility for "trolls" or "spin doctors" (Fitzgerald 2005) to deliberately provide inaccurate product evaluations to mislead retailers. However, such efforts would quickly be detected using a system such as the one suggested here involving correlation analysis (inaccurate evaluations would result in that person not being included in the predictive consumer consensus subgroup).

Recommendations for Future Research

Future studies could explore issues related to the optimal size and profile of the consumer subset and the conditions under which such subset choices result in better forecasts. Surowiecki (2004) suggests that collective intelligence depends on the diversity of the group's membership as well as independence of opinion and decentralization of the process. These parameters could be explicitly tested in future research. Are there consumers who are more knowledgeable than others on variables of interest to retailers, such as pricing, which would make them better

forecasters (Magi and Julander 2005)? Is it true that the more that forecasts within a group differ (among either consumers or buyers), the larger should be the number of forecasters within the group (Hogarth 1978)? Are consumers better predictors of items for members of the same gender or age? Does prior retail sales experience increase predictive ability? Larger samples would allow for comparisons involving gender, age, and retail buying experience among members of the consumer sample.

Other areas for research concern the learning that would take place over time with such a system. To what extent are the better consumer forecasters consistent in their accuracy over time? Would the consumers, if provided feedback on the relative accuracy of their forecasts, or financial incentives for forecast accuracy reflect improvement in the quality of their forecasts over time? Similarly, would retail buyers reflect a more effective incorporation of consumer input after greater experience over time? Would the most overly confident retail buyers refuse to incorporate consumer input and thus doom their own predictive accuracy? These are empirical questions that will likely prove of great interest to retailers in the future.

Table 1**Item Profitability**

Item	GMROI	Max Possible GMROI	GMROI Gap	GMROI Gap @ % of Potential
1-pendant	1.40	5.67	4.27	.75
2-designer purse #1	.37	1.08	.71	.66
3-designer purse #2	.36	1.32	.96	.72
4-designer silver sandal	.63	1.19	.57	.47
5-designer brown sandal	.83	1.19	.36	.30
6-designer sneaker	.49	1.23	.74	.60
7-designer black sandal	.24	1.93	1.69	.88
8-designer gold sandal	.75	2.44	1.68	.69
9-silk cami	.37	2.28	1.91	.84
10-gauche jeans	.32	2.03	1.71	.84
11-sequin shrug	.50	1.75	1.25	.71
12-crochet tank	.35	1.47	1.12	.76
13-girl's dress	.39	1.91	1.52	.80
14-brown shrug	1.14	2.10	0.95	.45
15-blue halter	.76	2.14	1.37	.64
16-designer suit	.10	3.00	2.90	.97
17-designer shrug	.29	1.56	1.27	.81
18-designer belt	.94	1.40	.46	.33
19-designer halter	.36	1.46	1.10	.75

(Note: Brand names have been disguised to assure confidentiality.)

Table 2
Mean Correlations Between Profitability Gap and Survey Response

	Purchase Likelihood	Purchase Likelihood Confidence	Pricing	Pricing Confidence	Fashionable
All Respondents (n=235)	-.039	-.036	-.136	-.078	-.101
Consumers (n=216)	-.037	-.034	-.134	-.072	-.104
Range:	(-.60 to .59)	(-.61 to .59)	(-.72 to .39)	(-.61 to .45)	(-.67 to .50)
Retail buyers (n=19)	-.059	-.053	-.156	-.146	-.066
Range:	(-.42 to .38)	(-.43 to .38)	(-.50 to .21)	(-.51 to .27)	(-.35 to -.06)
Test for Differences: Consumers vs. Buyers	p = .681	p = .725	p = .673	p = .152	p = .448

Table 3
Consumer Subset and Retail Buyers

	Consumer Subset (n = 26)	Retail Buyers (n = 19)	Test for Difference (p-value)
Purchase Likelihood	6.12	7.60	.0001
Price	2.90	3.33	.152
Self-reported ability to forecast customer demand	6.81	7.68	.114
Fashionable	6.15	7.42	.0001

Figure 1.



Figure 2.

Item #16

Consumer Subset (n = 26)



Figure 3.

Item #16

Retail Buyers (n = 19)



Figure 4.

Item #1

Consumer Subset (n = 26)

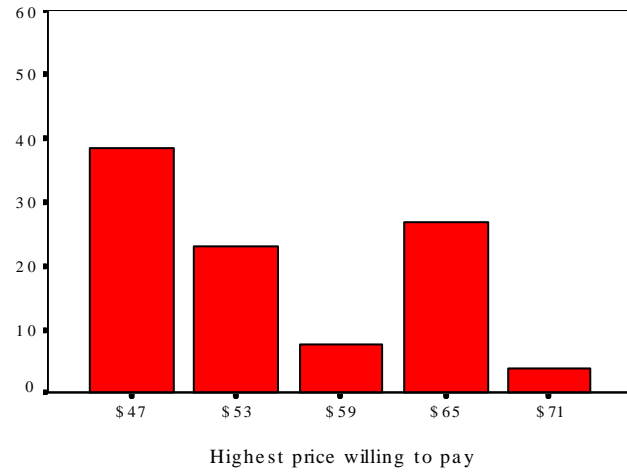


Figure 5.

Item #1

Retail Buyers (n = 19)

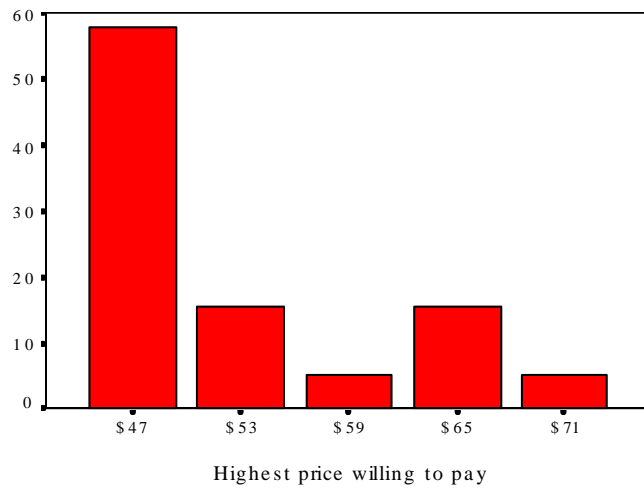


Figure 6.

Item #6

Consumer Subset (n = 26)

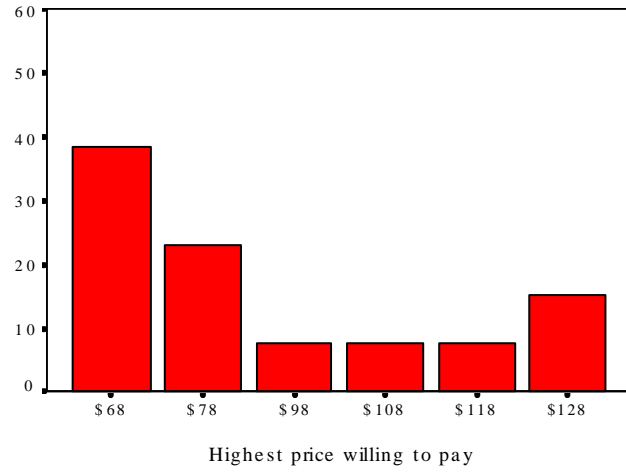
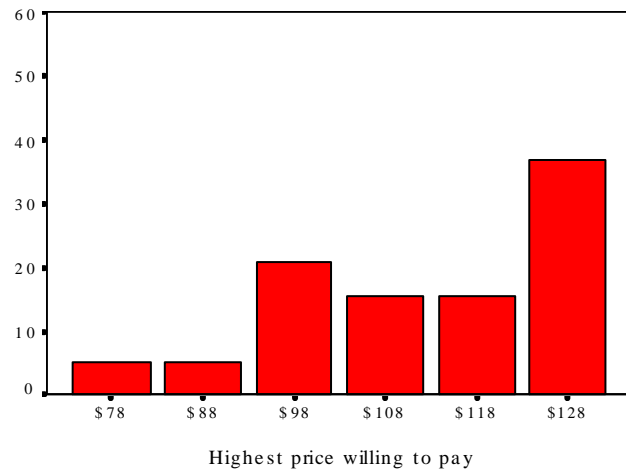


Figure 7.

Item #6

Retail Buyers (n = 19)



Appendix 1

Example of Photograph of Product in Online Survey



REFERENCES

Armstrong, J. Scott (1991), "Prediction of Consumer Behavior by Experts and Novices," *Journal of Consumer Research*, 18 (September), 251-256.

---- (2006), "Findings from Evidence-Based Forecasting: Methods for Reducing Forecast Error", forthcoming in the *International Journal of Forecasting*.

---- and Roderick J. Brodie (1999), "Forecasting for Marketing," in *Quantitative Methods in Marketing, 2nd ed.*, Graham J. Hooley and Michael K. Hussey, eds. London: International Thompson Business Press, 92-119.

Barbaro, Michael (2006). "Saks Restores Petite Sizes After Outcry," *New York Times online*, June 19, (accessed June 22, 2006), [available at <http://www.nytimes.com/2006/06/19/business/19cnd-saks.html>].

Blattberg, Robert C. and Stephen J. Hoch (1990), " Database Models And Managerial Intuition: 50% Model + 50% Manager," *Management Science*, Vol. 36 Issue 8 (August), 887-899.

Choi, Young-Eun and LuAnn Ricketts Gaskill (2000), "An Analysis of Mental Processes, Behaviors, and Job Satisfaction of Apparel Product Developers and Traditional Retail Buyers," *Journal of Business Research*, 49: 15-34.

Clemen, Robert T. (1989), "Combining Forecasts: A Review and Annotated Bibliography," *International Journal of Forecasting*, 5, 559-583.

Cox, Anthony D. Cox and John O. Summers (1987), "Heuristics and Biases in the Intuitive Projection of Retail Sales," *Journal of Marketing Research*, Vol. XXIV (August), 290-297.

Dandeo, Lisa M., Susan S. Fiorito, and Larry Giunipero, and Dawn H. Percy (2004), "Determining Retail Buyers' Negotiation Willingness for Automatic Replenishment Programs," *Journal of Fashion Marketing and Management*, Vol. 8, No. 1, 27-40.

Davies, Gary (1994), "The Delisting of Products by Retail Buyers," *Journal of Marketing Management*, 10, 473-493.

Derby, Meredith and Arthur Zaczekiewicz (2006), "Men's Biz Mediocre In March; Easter Shift Hurts Sales In All Categories Of Retail In Month," *Daily News Record*, 3, April 10.

Dermer, Jerry D. (1974), "Budgetary Motivation of Retail Store Managers and Buyers," *Journal of Retailing*, 50 (Fall), Issue 3, 23-30.

Ettenson, R.; and Wagner, J. (1986), "Retail Buyers' Saleability Judgements," *Journal of Retailing*; Vol. 62, 41-63.

Fairhurst, Ann.E. and Susan S. Fiorito (1990), "Retail Buyers' Decision-Making Process: An Investigation Of Contributing Variables," *International Review of Retail, Distribution and Consumer Research*, Vol. 1 Issue 1, 87-100.

-----, Sharron J. Lennon, and Kaekyung Yu (1996), "Retail Buyers' and Manufacturers' Sales Representatives' Perceptions of Market Show Services in Small Apparel Markets," *Clothing and Textiles Research Journal*, Volume 14, No. 3, 161-168.

Fitzgerald, Michael (2005), "Group Rethink: Can Technology Raise Society's IQ?" *Technology Review*, June: 80-81.

Hansen, Tommy Holm and Hans Skytte (1998), "Retailer Buying Behaviour: A Review," *The International Review of Retail, Distribution, and Consumer Research*, 8:3 (July): 279-301.

Hirschman, E. C (1981), "An Exploratory Comparison of Decision Criteria Used by Retail Buyers;" in *Retail Patronage Theory*, W.R. Darden and R. F. Lusch, editors. Proceedings of 1981 Workshop in Retail Patronage Theory, Norman, Oklahoma: 1-5.

---- and David Mazursky (1982), "A Trans-organizational Investigation of Retail Buyers' Criteria and Information Sources," New York University Institute of Retail Management Working Paper No 82-8.

Johnston, Robert D. (1995), "Listening To Customers Helps Buyers Pick The Best Merchandise." *Caribbean Business*, 23:32 (August 10), 42-46.

Kinkaide, Doris, Ginger A. Woodard, and Haesun Park (2002), "Buyer-Seller Relationships for Promotional Support in the Apparel Sector," *International Journal of Consumer Studies*, 26:4 (December), 294-302.

Kline, B. and J. Wagner (1994), "Information Sources and Retail Buyer Decision Making: The Effect of Product-Specific Buying Experience," *Journal of Retailing*, Vol. 70, 75-88.

Lazorchak, Shirley A. and Gwendolyn S. O'Neal (2001), "Department Store Apparel Buyers: Relationships Among Environmental Characteristics, Environmental Uncertainty, Boundary Spanning Activities, Managerial Discretion and Power," *Clothing and Textiles Research Journal*, Volume 19, No. 1, 11 – 21.

Lehmann, Donald R. (2004), "Metrics for Making Marketing Matter," *Journal of Marketing*, Vol. 68 (October 2004), 73-75.

Magi, Anne W. and Claes-Robert Julander (2005), "Consumers' Store-Level Price Knowledge: Why Are Some Consumers More Knowledgeable than Others?" *Journal of Retailing*, 81 (4), 319-329

McGinnis, Michael A. Gable, Myron, Madden, R. Burt (1984). "Improving the Profitability of Retail Merchandising Decisions-Revisited." *Journal of the Academy of Marketing Science*, 12 (1) (Winter), 49-57.

McIntyre, Shelby H., Dale D. Achabal, and Christopher M. Miller (1993), "Applying Case-Based Reasoning to Forecasting Retail Sales," *Journal of Retailing*, 69 (4), 372-398.

Morales, Andrea, Barbara Kahn, Leigh Mcalister, Leigh and Susan M. Broniarczyk, Susan M. (2005), "Perceptions Of Assortment Variety: The Effects Of Congruency Between Consumers' Internal And Retailers' External Organization," *Journal of Retailing*, 81 (2), 59-169,

Morwitz, Vicky (2001), "Methods for Forecasting from Intentions Data," in *Principles of Forecasting: Handbook for Researchers and Practitioners*, J.S. Armstrong ed., Norwell, MA: Kluwer Academic Publishers, 33-56.

Nocera, Joe (2006). "The Future Divined by the Crowd". *New York Times*, March 11, C1, C12.

Rajaram, Kumar (2001), "Assortment Planning in Fashion Retailing: Methodology, Application, and Analysis," *European Journal of Operational Research*, 129 (1), 186-209.

Ray, Russ (2006). "Prediction Markets and the Financial "Wisdom of Crowds." *Journal of Behavioral Finance*, 7 (1): 2-4.

Rust, Roland T., Tim Ambler, Gregory S. Carpenter, V. Kumar and Rajendra K. Srivastava, (2004), "Measuring Marketing Productivity: Current Knowledge and Future Directions," *Journal of Marketing*, 68 (4), 76-89.

Rust, Roland T., Katherine N. Lemon and Valarie A. Zeithaml (2004), "Return on Marketing: Using Customer Equity to Focus Marketing Strategy," *Journal of Marketing*, 68 (1), 109-127.

Sheth, J. N. (1973), "A Model of Industrial Buyer Behavior," *Journal of Marketing*, 37: 50-56.

_____ (1981), "A Theory of Merchandise Buying Behavior," in *Theory of Retailing and Non-traditional Sources*, R. Stampfl and E. Hirschman eds., , Chicago: American Marketing Association.

Shim, Soyeon and Antigone Kotsiopulos (1991), "Information-Seeking Patterns of Retail Apparel Buyers," *Clothing and Textiles Research Journal*, Volume 10, No. 1 (Fall), 20-30.

Surowiecki, James (2004). *The Wisdom of Crowds*. 2004, New York: Random House.

Van Herpen, Erica and Rik Pieters (2002), "The Variety of an Assortment: An Extension to the Attribute-Based Approach," *Marketing Science*, Summer, 21 (3), 331-341

Wagner, J., Ettenson, R., and Parrish, J. (1989), "Vendor Selection Among Retail Buyers: An Analysis by Merchandise Division," *Journal of Retailing*, 65, 58-77.

Wall, M., M. Sommers, and A. Wilcock (1994), "The Retail Buying of Fashion Goods: Underlying Themes of the Sourcing Process," *International Review of Retail, Distribution and Consumer Research*, 4 (1), 177-193.

Young, Vicki (2006), "Gap 1Q Profits Drop 16.8%," *Women's Wear Daily*, 2.