The Journal of Prediction Markets (2007) **1** 3, 189–208

CONDITIONAL PREDICTION MARKETS AS CORPORATE DECISION SUPPORT SYSTEMS – AN EXPERIMENTAL COMPARISON WITH GROUP DELIBERATIONS

Timm Sprenger, Paul Bolster† and Anand Venkateswaran‡*

College of Business Administration, Northeastern University, Boston, MA 02115

Predicting the future is an integral part of effective corporate decision making. Most firms face the critical challenge of aggregating information dispersed among its agents. These agents and thus the aggregation process are prone to judgmental biases. The primary research question we address is whether markets correct these biases better than group deliberations. Using an experimental setting, we find that information markets provide more accurate and less volatile forecasts than group deliberations. We also describe different sources of the behavioral biases we observe. For example, while a deliberating group can be led astray by an influential group member, traders tend to overweight personal preferences. Our results indicate that conditional prediction markets provide a more effective medium for aggregating information than group deliberations.

Keywords: Prediction Markets, Group Deliberation, Information Markets

I. INTRODUCTION

Predicting the future is an integral part of effective corporate decision making. Most firms face the critical challenge of aggregating information dispersed among its agents. As opposed to estimating the value or likelihood of some future event that will be verifiable with certainty, prediction markets can also be used to evaluate alternatives that may never materialize. The literature often refers to these markets as conditional markets. Hanson (1999) argues that conditional prediction markets could be used to directly guide decision making because they can "accurately estimate the consequences of important decisions".

The value of information markets to any firm depends on their direct comparison with existing information aggregation mechanisms. Several studies have compared the predictive power of markets with traditional mechanisms. Empirical studies have compared information markets to in-house experts (e.g. Chen and Plott (2002)), independent forecasting agencies (e.g. Wolfers and Zitzewitz (2004)), Delphi techniques (Berg and Rietz (2007)) and opinion polls (Berg *et al.* (2003) and Chen *et al.* (2005)). However, in many corporations the most relevant benchmark is arguably group deliberation. To our knowledge, this is the first study to compare prediction markets to this most widely used mechanism in the business world – found mostly in its form of the ubiquitous business meeting.[1]

---

*Tel.: +49-2823-98551. Email: sprenger.t@neu.edu
†Tel.: 617-373-5051. Email: p.bolster@neu.edu
‡Tel.: 617-373-7873. Email: anand@neu.edu

© The University of Buckingham Press (2007)

Besides its widespread use, there is another reason why group deliberation represents an interesting benchmark for prediction markets. While statistical groups and individuals are fundamentally different from market mechanisms, group deliberation has, at least in theory, the ability to replicate an efficient market. It involves multiple participants who exchange information and can thus update their priors to arrive at a consensus. Designed correctly, prediction markets can directly guide decision making. They are not limited to actual events, but can also evaluate conditional probabilities to provide a more informed picture of hypothetical future states of the world. In that sense, these markets are capable of assessing complex situations, a task traditionally reserved for deliberation.

Although the notion of informationally efficient markets assumes individual rationality, there are many studies which challenge this assumption. In light of these challenges, we address the following research question: Do markets correct behavioral biases better than group deliberations?

II. BEHAVIORAL BIASES IN GROUP DELIBERATIONS AND INFORMATION MARKETS

Both markets and group deliberations are composed of individuals. Much of economic theory is based on the notion that market participants act rationally and make utility optimizing decisions. This assumption has been questioned by numerous observations of behavioral biases.[2] Studies argue for the presence of behavioral biases in economic agents. For instance, (i) people are overconfident in their judgment and abilities (e.g. Russo and Shoemaker (1989)), (ii) individuals tend to overweight recent information and overreact to it (Tversky and Kahneman (1981)), and (iii) individual preferences depend on the way in which choices and outcomes are framed. Additional biases include forecasting error, regret avoidance and herding, and personal preferences.

The underlying assumption of the *Efficient Markets Hypothesis* (EMH) that markets fully and accurately reflect all available information appears unrealistic given a plethora of behavioral biases. For instance, evidence for continuing market inefficiencies can be found in the countless financial bubbles throughout history ranging from the tulip mania that swept through Holland in the 17th century to the US technology bubble in 2000. A variety of biases can help explain those inefficiencies.

One example of a behavioral bias is overconfidence. People are more optimistic about their own future than that of others (Weinstein (1980)). There are many other examples of behavioral biases. Most people overreact to unexpected and overweight anecdotal, recent information, a bias that is closely related to representativeness heuristics. De Bondt and Thaler (1985) have concluded that this behavior affects stock prices. It has also been found that regret avoidance may influence investment decisions and lead to suboptimal performance (Clarke *et al.* (1994)). Studies have also indicated that traders may trade according to their personal preferences as opposed to objective probability assessments. Surveys at the Iowa Elections Market (IEM) have found traders to be biased by their party preference and this bias appears to be reflected in both their trading activity and their portfolio holdings (Forsythe *et al.* (1999)). Despite the encouraging evidence supporting the benefit of prediction markets such behavioral biases may limit their effectiveness.

In firms, group deliberation remains one of the most widely used mechanisms to collect and aggregate information. Presumably business meetings and committees lead to better informed decisions. But there are a number of concerns regarding group deliberations and the fact that both practitioners and academics are continuously developing methods to better aggregate information shows that they are aware of certain failures within standard group discussions. Possible sources of deficiency in group deliberation include informational influences and social pressures. Informational influences arise due to failure of group members to disclose information because a public announcement by others has made them uncertain of their own beliefs. Examples include groupthink (Janis (1982)), overconfidence, polarization (Brown (1965)) and information cascades (Anderson and Holt (1997)). In the case of social pressures, people silence themselves for fear of disapproval from other group members even though they think that their information or belief is better. These pressures are also known as cost-benefit trade-offs, reputational cascades (Sunstein (2006)) and conformity (Asch (1963)).

A comparison of prediction markets and group deliberation must primarily focus on their ability to address these biases. Under the rational expectations hypothesis, systematic forecasting errors are impossible because they constitute valuable information. Due to the limitations of the rational expectations hypothesis, some alternative behavioral models of decision making have been developed, but most are designed to address only one specific bias.[3] The adaptive markets hypothesis (AMH) tries to reconcile behavioral biases and efficient markets in a more general framework. Lo (2004) argues that many of the biases that behaviorists cite as violations of rationality and inconsistent with market efficiency are, in fact, consistent with an evolutionary model of individuals adapting to a changing environment. The AMH suggests that

> "individuals make choices based on past experience and their best guess [...]. They learn by receiving positive or negative reinforcement from the outcome. [...] In this fashion, individuals develop heuristics to solve various economic challenges, and as long as those challenges remain stable, the heuristics will eventually adapt to yield approximately optimal solutions" (Lo (2004)).

Interestingly enough, the very emotions that may give rise to many behavioral biases may be intertwined with the ability to make rational choices. Lo concludes that "emotions are the basis for a reward-and-punishment system that facilitates the selection of advantageous behavior". So the emotional feedback mechanism helps humans to learn. Studies have documented that human judgment tends to be better calibrated when people perform "repetitive tasks with fast, clear feedback" (Odean (1998)). The arguably more focused informational input and the finite nature of most prediction markets may even give them an edge over financial markets and provide participants with a feedback loop which is better suited to calibrate the traders' judgment.

In a market environment the feedback mechanisms are not limited to emotions, but complemented by monetary profits. This feedback drives the most influential traders to be more rational and less prone to behavioral biases. A study of traders at the IEM supports this notion. Oliven and Rietz (2004) find that marginal traders are generally less biased. Price makers are those traders that enter new limit orders to buy or sell which are later accepted by others who can be referred to as price takers. These marginal traders who constitute roughly 15 percent of traders invest twice as much as others, trade more frequently, earn higher returns and make only one sixth of the errors. Highly biased traders, on the other hand, tend to buy and hold securities. So, the disproportionate influence of marginal traders is one reason why markets outperform opinion pools and polls. The distinction between marginal traders and price takers has an important implication. It explains why markets can be efficient and their prices can be used as a predictive tool, while at the same time, behavioral biases affect the majority of participants and can provide additional inputs to support decision making.

Group deliberations often lack this direct feedback mechanism and provide less incentive for members to learn. Extensive empirical evidence suggests that individuals exert less effort in groups than when alone, a tendency known as social loafing (Ingham *et al.* (1974)). In addition, groups often fail to hold individuals accountable for their contribution. Without this feedback loop individuals are less likely to correct their errors.

Information markets however, not only correct behavioral biases through individual feedback but also provide a superior mechanism to correct others' biases. If an individual recognizes a judgmental error, she is rewarded for correcting it through the price mechanism. The fear of social sanctions is limited because trading is mostly anonymous. Deliberative groups often fail to correct their members' judgmental errors, even if they could. In experiments using questions with definite answers, deliberating groups have done only slightly better than their average group member, but far from their best (Gigone and Hastie (1997)). Groups have a better chance to outperform their individual members when the outcome is verifiable (MacCoun (2002)). In group deliberations, the cost-benefit trade-off and the desire to conform can limit error correction (Cannon and Edmondson (2001) and Edmondson (1996)).

Additionally, there are reasons to believe that markets provide a better incentive system for the discovery of new information. While traders can seek new information continuously to confirm or refute the current market prices, the participants of group deliberations are usually "trapped" for the duration of the deliberation round. Also participants in information markets are often self-selected and the intensity of their beliefs is captured in the volume they are willing to trade.

III. EXPERIMENTAL COMPARISON OF MARKETS AND GROUP DELIBERATIONS

To examine the relative merits of prediction markets and group deliberations, we develop and execute a laboratory experiment which simulates conditional prediction markets. Information markets and group discussions are used to make predictions about the success of five new product concepts for cellular phones. The markets are conditional prediction markets with student participants forecasting the relative future market share among these five products in the college population. The use of information markets for the evaluation of product concepts is limited to two studies. Chan *et al.* (2002) introduced the idea of securities trading of concepts (STOC) and conducted two laboratory experiments to evaluate new product introductions for air pumps and sport utility vehicles. Soukhoroukova and Spann (2005) have replicated this experiment design with mp3-players.

The experiment is based on two treatments: 12 markets with trading groups (Traders) and 12 deliberative groups (Talkers). Each group has four members, so that a total of 96 subjects participate in the experiment. The small size was also chosen to make the markets more directly comparable to the average group discussion. Both Traders and Talkers are assigned the task of predicting the future market share among the college population of five cell phone models labeled A, B, C, D and E.

All phones including their attributes are actual product concepts from four different manufacturers which had not yet been released at the time of the experiment. Prices are set close to the manufacturers' recommended retail prices minus a \$150 to \$200 discount representative of the markdown that wireless providers usually give with the purchase of a new service contract. Figure 1 shows the salient features of the five choices.

There are various reasons for the choice of cell phones. Cell phones have a limited number of important attributes which simplifies the decision process and does not draw too much attention away from the forecasting task. Further, almost every college student owns a cell phone and most have gone through the purchasing process at least once. Many students' cell phone purchases are subject to behavioral biases and tend to be emotional.

Traders trade the shares of the five cell phone concepts {A, B, C, D, E}. The value of every share is linked directly to the future market share of the product. The payoff structure is simple. For example, a share of a phone with a 20 percent market share is worth \$20.

To ensure that the trading mechanism provides sufficient liquidity, a market board is chosen as the trading mechanism for this experiment.[4] The board functions as a market maker with infinite liquidity that changes prices according to demand. The board stands ready to buy and sell securities in \$1 intervals. Thus, there is a bid-ask spread of \$1 in every one of the five markets. Traders can either buy securities for the lowest possible price available on the market board or sell securities by exchanging them for the highest possible price that the market

	Product A	Product B	Product C	Product D	Product E
Price	$250	$200	free	$185	$300
Size/ Weight	4.1x1.7x0.6 in/ 3.2oz (91gr)	4.2x2.0x0.5 in/ 4.0oz (115gr)	4.0x1.7x0.7 in/ 3.0oz (85gr)	3.5x1.8x0.9 in/ 3.5oz (98gr)	4.6x3.5x0.9 in/ 4.1oz (118gr)
Talk Time/ Standby	4hrs/9d	3hrs/9d	7hrs/15d	5hrs/11d	4hrs/6d
Camera	Pics & Video (2 mpx)	Pics & Video (2 mpx)	Pics (0.3 mpx)	Pics&Video (1.3 mpx)	Pics&Video (6 mpx)

FIGURE 1. Product concepts underlying the experiment.

board offers. By buying securities traders essentially increase the stock price by $1. The inverse holds for security sales. The trading mechanism is thus very similar to more complicated designs that have been created to address the thin market problem typical for many information markets, but it is simple enough for subjects without prior trading experience.

Since the total market share of all 5 phones is 100 percent by definition, the value of a complete bundle of securities must be $100. Therefore, the market maker also exchanges a complete bundle {A, B, C, D, E} for $100 at any time and provides a unit portfolio in exchange for $100 in cash. This operation takes place at no risk to the market board, but provides increased flexibility for possible trading strategies. If the combined prices of {A, B, C, D, E} exceeds $100, Traders could take advantage of this arbitrage opportunity by buying a bundle from the board for $100 and selling it to the market at a riskless profit. If the combined prices fell below $95, traders could buy a bundle and sell it to the marker maker for more than what they paid for it on the market board. Because of the bid-ask spread of $1 across 5 securities, arbitrage opportunities only arise when prices are outside of the efficient bounds of $95 and $100. Bundle trades could increase or decrease the supply of shares according to demand. This also gives traders a chance to sell securities short by buying a bundle for $100 and selling only selected shares. Traders did take advantage of all of these trading strategies.

The Traders in each group are rotated to ensure that every participant has an equal chance to trade. At every turn, Traders are allowed to make up to 5 trades (not considering any exchanges of complete bundles). A round is defined as all four traders getting a chance to trade. Trading continues until no one wants to trade anymore assuming that this would indicate an equilibrium state.

The starting prices for all five phones are set to $20 and subjects are provided with an initial endowment of five bundles and $300 in play-money.

The initial cash endowment is intended to provide sufficient liquidity to move prices, but prevent speculative bubbles. We developed an Excel-based market board, illustrated in Appendix B, to simplify trade execution.

The experiment is conducted in four sessions which last about 40 minutes each. At the start, all subjects are given a questionnaire and pictures of the phones and product attributes. Participants are informed that they are taking part in an experimental study of prediction markets and group deliberations and asked to individually fill out the first part of a questionnaire regarding their personal preferences and individual market share estimates. Appendix A contains a copy of the administered questionnaire. Subjects are informed that all phones are actual models to be released in the near future and that the future sales data would reveal the true market share distribution among these five competitors. A brief introduction to the basics of prediction markets and their use in a corporate setting is followed by an explanation of the securities used in this experiment and the mechanism of the market board. The concepts of selling and buying securities as well as arbitrage are illustrated with a physical market board that resembles the software version. To provide incentives, the best Trader in each trading group and the deliberative group with the best group estimate are rewarded with prizes.

Participants are then randomly assigned the role of either a Trader or a Talker. They receive a label with a group number and a personal ID. The deliberative groups meet in separate rooms to discussion and create consensus estimates. A computer terminal is provided for every trading group. Traders are allowed to communicate.

The accuracy of the predictions can only be assessed vis-à-vis the actual market share for the five products. Some of the phones will not be released until the end of 2007 and even then official market data may not serve as a reliable proxy for the relative market share due to other factors such as the marketing campaign and the number of wireless providers offering the phone.

Therefore we conduct an independent online survey to provide a proxy for the actual market share. We receive 134 responses from participants who are invited by e-mail and provided with the same information as the subjects. A simple first choice model is employed. Results of the response are reported in Table 1. Since the margin of error is substantial, those results that use this market share proxy as a benchmark need to be interpreted with caution.

TABLE 1
SURVEY RESULTS THAT SERVE AS THE BENCHMARK MARKET SHARE ESTIMATE

	A	B	C	D	E	Total
First Choice	31	27	39	17	20	134
Percentage	23.1%	20.1%	29.1%	12.7%	14.9%	100%
Margin of Error (95%)	*7.1%*	*6.8%*	*7.7%*	*5.6%*	*6.0%*	

IV. EXPERIMENT RESULTS

We divide our results into three categories; Overall accuracy and precision, behavioral biases and forecast accuracy, and market efficiency.

Overall Accuracy and Precision

All existing empirical studies that address the predictive power of information markets have focused on measuring the accuracy of predictions, with accuracy mostly defined as the mean or median result across multiple observations. We propose that accuracy is only a partial measure of predictive power and should be complemented by precision, defined as the variance of observations. The existing prediction market research has so far failed to recognize precision as an equally important component.

Arguably the most conspicuous attribute of a forecast is its accuracy. Both prediction markets and group deliberations are used because they are supposed to improve the accuracy of individual judgment. As a statistical group, the average of all individual subjects can serve as a reference point for both Traders and Talkers (Figure 2). We find that the statistical mean of all market share estimates is, indeed, close to the actual outcome. Except for phones D and E which end up being close together, the ranking of the models is accurate. The estimates are well calibrated with respect to the actual results as can be seen from a comparison of the benchmark (dashed line) with the line of best fit through the five estimates (solid).

While accurate on average, the estimates lack precision as evidenced by their large standard deviations. In other words, asking any one of the individuals could easily yield a very different and inaccurate picture. In most corporate settings, average accuracy by itself is not sufficient to justify decisions. A firm cannot hold

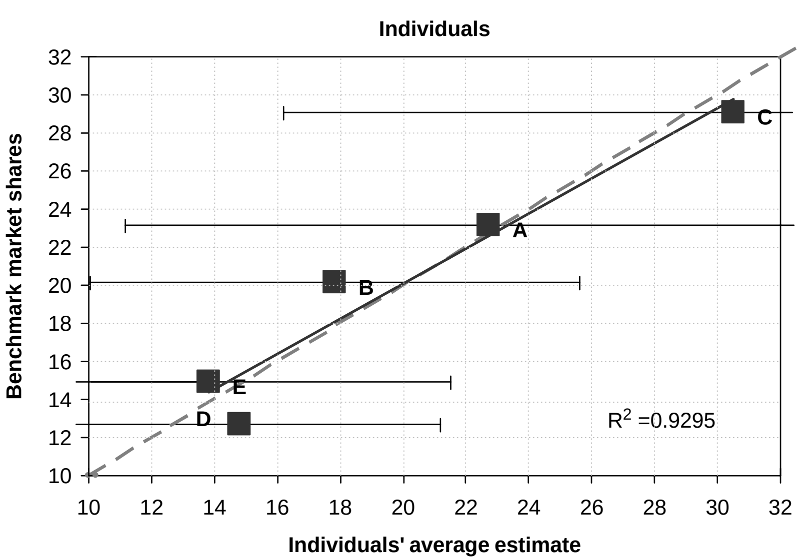


FIGURE 2. Average and standard deviation of individual estimates.

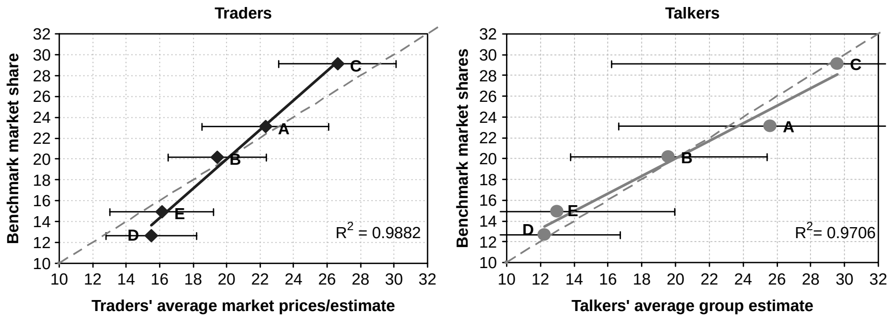


FIGURE 3. Average and standard deviation of Traders' and Talkers' market share estimates.

twelve different group meetings or run twelve separate prediction markets to ensure accuracy. Thus the variance among individual estimates becomes critical.

We compare the average estimate of the 12 trading and deliberative groups, and find that both mechanisms generate an accurate prediction that is again very close to the benchmark (Figure 3). On average, both Traders and Talkers are able to rank all five phones correctly according to their market share. While Traders overestimate high and underestimate low market shares, Talkers' estimates are slightly skewed in the opposite direction.

With respect to the variance of estimates, trading groups outperform group discussions substantially. Group deliberation usually reduces the variance of their members' judgment (Brown (1965)). This appears to be the case in this experiment, too, where the standard deviation of estimates is lower for Talkers than for individuals. However, the direct comparison of the two group mechanisms shows that the variation across the 12 estimates is much lower for the Traders with an average standard deviation of 3.2 percent compared to 7.9 percent for the Talkers.[5] In fact, if one takes this variation into account, it is clear that the 12 deliberative groups forecast all five phones to be within one standard deviation of the average predictions. The 12 trading groups, on the other hand, have a lower variance in their prediction. Of course, the lower dispersion holds irrespective of the estimated benchmark.

Considering differences in precision, the results across the 12 trading groups are substantially better than for their deliberative counterparts. The overall mean average error (MAE) of 2.87 percent and the mean average percentage error (MAPE) of 16 percent are roughly twice as good. The average correlation between market prices and the benchmark market share is 0.83 whereas it is only 0.64 for the Talkers. So while both Traders and Talkers are, on average, very accurate in their predictions, Traders outperform Talkers in terms of precision.

One cautionary note deserves mention. The lower variation in Trader estimates may be the result of a behavioral bias. All five shares are initially valued at \$20. Tversky and Kahneman (1974) find that the starting point in a measurement process has a strong influence on the median response. This behavioral bias is often referred to as "anchoring". The magnitude of the most prominent dimension of a decision object serves as a reference point for the decision. Its value is modified upward or downward, but the adjustment is often insufficient. Since the products in Chan *et al.* (2002) had previously been used in physical and web-based assessments of market shares, there was a direct comparison for the market based results. In fact, the market share forecasts deviated substantially from previous results and the variations between products were much lower in the prediction markets. This may provide further evidence for a potential "anchoring effect" in prediction markets which can distort results.

While, on average, both Traders and Talkers predict market shares with high accuracy, Traders outperform Talkers both in relative and absolute terms. Traders also exhibit a much lower variance than Talkers.

Behavioral biases and accuracy

We have seen that the variation of results is much greater across the Talkers relative to the Traders. This confirms some of the limitations regarding group deliberations mentioned previously. In most group discussions, participation and influence are unequal. Often informal leaders emerge (Brown (1965)). Variance can also be due to polarization where group members adopt a more extreme version of the opinion they held prior to deliberation.

After Talkers finalize their market share estimates, individuals are separately asked to name who they thought was the most talkative and influential member in their group. Figure 4 shows the relationship between group performance and the individual performance of the most influential group member. This analysis only includes groups that clearly identified a most influential group member by simple majority. The regression of the MAPE of those influencers and the group performance indicates a positive correlation between bad judgment on part of the influencer and an inferior group judgment.[6] This relationship could be due to an inaccurate influencer, or due to the fact that people with bad judgment happened to be in one deliberative group. However, if we compare the relationship between the group members' average accuracy and the group estimate, we find no statistically significant relationship (Figure 5). This means

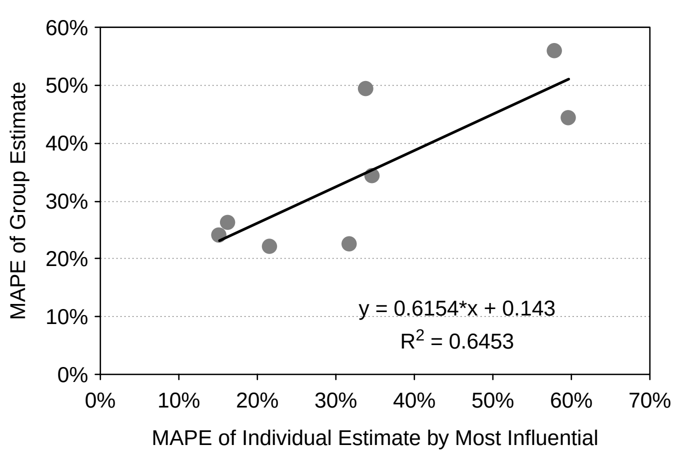


FIGURE 4. Corr. of influencer and group MAPE. * Denotes statistical significance at the 5 percent level.

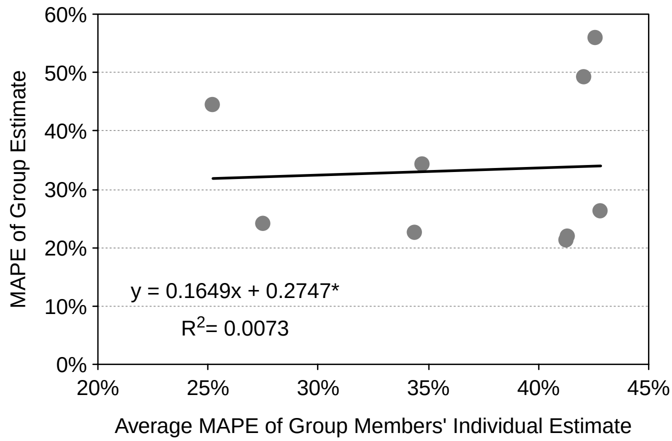


FIGURE 5. Corr. of MAPE and group estimate. * Denotes statistical significance at the 5 percent level.

that while group members with bad individual estimates were able to correct each other and maybe cancel out opposing errors, they were unable to correct the errors of the most influential group member.

We have shown that an inaccurate and influential member of a Talker group diminishes the accuracy of consensus forecast. We next examine the influence of individual personal preferences on trading activity among Traders. Individual traders often tend to be biased by their personal preferences. These biases may also yield useful information in a business context. If traders at the Hollywood Stock Exchange, for instance, trade primarily in contracts related to their favorite movies, the portfolio holdings could allow insights into their preferences. Transferred to this experiment, there may be a bias among traders to trade in their favorite phones. Prior to trading, Traders are asked which phone they personally prefer. If there is no bias in the choice of securities in which Traders trade, subjects should be just as likely to trade in their favorite phone as they are to trade in any of the other four. Assuming no bias, the random chance that Traders make a trade in their favorite security should be one out of five, or 20 percent.

Out of our overall sample of 985 trades, 28 percent (z-score: 6.05; p-value: 0.00) involve subjects' favorite phones (Table 2). Thus, Traders show a significant preference to trade in their favorite phones. Further, 68 percent of favorite phone trades are buys. Participants are even more likely to make their very first trade in their favorite phone (33 percent; z-score: 2.31; p-value: 0.02) and 100 percent of first trades are decisions to buy. Conversely, trades in the least favorite phone are below 20 percent, and most of those trades are decisions to sell. This data supports the notion that there is a bias among traders to trade according to their preferences.

While, we are able to confirm disproportionate trading in favorite phones, we actually may be capturing Traders' interest in phones expected to achieve the highest market share. There may be some overlap between the two subgroups. Thirty-three percent (z-score: 9.80, p-value: 0.00) of all trades and 46 percent (z-score: 4.47, p-value: 0.00) of first trades involve the phone traders

Table 2
Trading activity by traders' preferences and individual estimates

	Favorite	Least Favorite	High Estimate	Low Estimate
First Trades	33.3%*	16.7%*	45.8%*	0%*
	0.02	*0.56*	*0.00*	*0.01*
Total	16	8	22	7
Buy	100%	50%	100%	29%
Sell	0%	50%	0%	71%
All Trades	27.7%*	16.6%*	32.5%*	15.7%*
	0.00	0.01	0.00	0.01
Total	273	164	320	106
Buy	68%	29%	77%	12%
Sell	32%	71%	23%	88%

* Denotes statistical significance at the the 5 percent level.
p-value in italics.

individually forecast with the highest market share. The vast majority of those trades are buys.

Apart from trading activity, the resulting portfolio may reveal a bias, too. To test the portfolios in this experiment for a potential bias, we use the following method. We construct a typical market portfolio by averaging the number of shares that all traders hold in each of the five securities. We then benchmark all individual portfolios against this average portfolio to see if it is over- or underweighted with respect to a certain security. Since the chances for any one portfolio to be over- or underweighted should be about half, we test the portfolio composition analogous to this trading behavior.

The 48 portfolios of individual Traders do not show a statistically significant overweighting of favorite securities. Seventy-one percent (z-score: 2.89; p-value: 0.00) of Traders did, however, underweight their least favorite phone. There is also a tendency to overweight those models that are highest in the individual estimates (67 percent; z-score: 2.31; p-value: 0.02) and underweight the lowest estimate (84 percent; z-score: 3.89; p-value: 0.02). Thus, the portfolio composition allows us to infer the traders' least favorite securities as well as their individual beliefs.

To what extent do individual traders update their private estimates when they assess the value of securities? Traders may condition their beliefs on market signals or discard their personal preferences and interpret market information (Chan *et al.* (2002)). Spann and Skiera (2004) argue that prediction markets can be used by management to identify employees with superior forecasting ability. For this to hold there must be a relationship between individual forecasting ability and trading success. It is also possible that successful traders do not have superior private information, but are simply good processors of market signals. In this experiment, the prior individual estimate and the trading success of 48 Traders are collected. OLS estimates

TABLE 3
FINAL PORTFOLIO WEIGHTING BY TRADERS' PREFERENCES AND ESTIMATES

	Favorite	Least Favorite	High Estimate	Low Estimate
Overweight	52.1%	29.2%*	66.7%*	15.6%*
	0.77	*0.00*	*0.02*	*0.00*
Underweight	47.9%	70.8%*	33.3%*	84.4%*
	0.77	*0.00*	*0.02*	*0.00*

* Denotes statistical significance at the the 5 percent level.
p-values in italics.

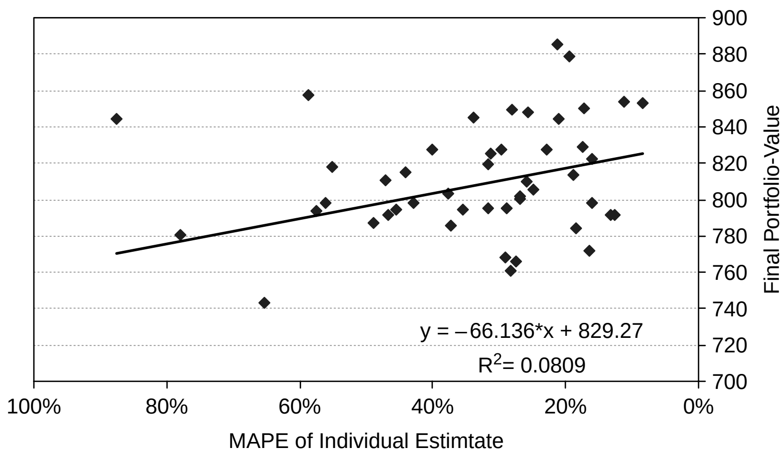


FIGURE 6. Individual accuracy and trading success. * Denotes statistical significance at the 5 percent level.

confirm a positive relationship between a better individual estimate and trading success measured by the final value of the portfolio (Figure 6).

To summarize our estimates of behavioral biases and consensus estimates, Talkers can be led astray by an inaccurate but influential group member. Traders show a tendency to disproportionately trade in their most favorite phones and are more likely to buy than sell those securities. Traders also tend to overweight (underweight) phones expected to garner the highest (lowest) market share in their final portfolio. Trading success can be an indicator of individual forecasting ability.

Market Efficiency

An efficient market incorporates all relevant information into prices. Next, we examine equilibrium or price stabilization behavior in these experimental markets. The results from our experiment suggest that trading patterns vary substantially across trading groups with the number of trades ranging from 42 to 155. Some groups traded for three rounds whereas others did not reach a conclusion until round nine. While some traders exploited the maximum number of permitted trades in each round, others averaged only 2.6 trades per

TABLE 4
TRADING STATISTICS AND MAPE FOR THE TRADING GROUPS

Trading Group	Trades	Rounds	Trades/Round	Avg. Trades/Trader	MAPE
1	129	7	18.4	4.6	5.6%
2	74	4	18.5	4.6	19.7%
3	155	9	17.2	4.3	17.5%
4	92	5	18.4	4.6	17.4%
5	60	4	15.0	3.8	10.3%
6	47	3	15.7	3.9	12.2%
7	79	5	15.8	4.0	24.9%
8	59	3	19.7	4.9	17.0%
9	93	6	15.5	3.9	17.1%
10	100	5	20.0	5.0	9.2%
11	55	3	18.3	4.6	18.5%
12	42	4	10.5	2.6	10.7%
Average	82.1	4.8	16.9	4.2	15.0%

turn. Since none of these statistics correlate with the accuracy of the market estimates, we can exclude them as factors for efficient information aggregation.

In an efficient market, investors cannot systematically outperform the market. The initial prices of all five securities are equal. Although all phones are modeled closely after real product concepts and are competitively priced, one might argue that it is easier to make a profitable trade early by buying low or selling high. Such a trade is profitable for the trader and moves prices in the right direction. We define such a trade as "right". On the other hand, a "wrong" trade is one where a trader trades in a security that is already overvalued or sells a security that is undervalued. In the first round of trading, more than 65 percent of trades are "right". This is the highest proportion of right trades in any round. It is still favorable in the second round, and by the third round, the proportion of right trades is virtually equal to the proportion of wrong trades. One would expect to see this behavior in an efficient market. Overall, 56% and 44% of trades represent right and wrong trades respectively. This 12 percent differential generates market prices which are substantially better than the deliberative estimates. This confirms that a limited number of marginal traders or trades are sufficient for efficient markets, leaving significant room for behavioral biases.

Although each trading session resulted in equilibrium, the final market prices of eight trading groups sum to more than $100 and only two groups close trading below $100. Two of the eight groups overprice the bundle {A, B, C, D, E} by more than ten percent. Thus, while traders could have made a riskless profit by exploiting this arbitrage opportunity, they did not. In fact, eight of twelve markets closed with an arbitrage opportunity. This phenomenon is not unusual. Rietz (2005) has observed that state-contingent Arrow-Debreu contracts, similar to the shares used in this experiment, were consistently overpriced by 14.5 percent to 20 percent in laboratory markets with two securities that represented a bundle of all possible outcomes.[7]

TABLE 5
PERCENTAGE OF RIGHT AND WRONG TRADES PER ROUND

Cumulative Round	1	2	3	4	5	>5
Right	65.1%	61.7%	58.0%	60.7%	56.3%	56.0%
Wrong	34.9%	38.3%	42.0%	39.3%	43.7%	44.0%
Round by Round						
Round	1	2	3	4	5	>5
Right	65.1%	58.4%	50.5%	51.3%	53.8%	53.5%
Wrong	34.9%	41.6%	49.5%	48.7%	46.2%	46.5%

Taken together, our results indicate that there are a number of significant biases in traders' behavior. Performance and non-performance related trader information can be used to draw conclusions about the traders' preferences and abilities. Despite these behavioral biases the markets are able to make predictions that are better than most individuals and even group deliberations. This underscores the ability of markets to overcome and correct individual biases effectively.

V. CONCLUSION

This study directly compares information markets and group deliberations. Both prediction markets and group deliberations are subject to numerous judgmental biases. A theoretical examination suggests that the feedback mechanism provided by markets makes them better suited to correct those biases. We argue that, in the case of information markets, average accuracy alone is insufficient to appropriately measure predictive power. In a direct comparison with other decision support tools, precision should also be assessed to reach a more meaningful conclusion.

Our experimental design is straightforward. We provide small groups of individuals with identical information and ask them to forecast market shares of five competing cell phones. We then compare the output from business meetings with that from information markets. Our results strongly indicate that information markets provide more accurate and less volatile estimates than group deliberations. However, both prediction markets and group deliberations are subject to numerous judgmental biases. While some of these biases have been found in other prediction markets, this study confirms their existence in the context of new product introductions. We also confirm different sources of behavioral biases in the two approaches. For example, Talkers can be led astray by an influential group member while Traders tend to overweight purchases of their favorite phone. Apparently the structure of information markets alleviates Traders' individual biases more effectively resulting in more accurate and precise forecasts.

Of course, prediction markets are subject to certain limitations and cannot fully replace meetings, opinion polls or outside consultants. They do, however, have considerable potential to supplement these traditional methods used to collect and aggregate information. It is a promising field for further research that needs to be tested outside of an experimental setting.

NOTES

1. There is a limited amount of research concerned with the direct theoretical comparison of prediction markets and group deliberations. Sunstein (2006) is one example.
2. For a more comprehensive overview refer to Kahneman *et al.* (1982).
3. Prospect theory (Kahneman and Tversky (1979)), which generates behavior consistent with loss aversion, but not overconfidence or regret, is one prominent example.
4. The original market board was invented by Robin Hanson who has used it for a limited number of winner-take all markets where traders had to identify the culprit in a murder mystery. A number of modifications to Hanson's market board were made for this laboratory market. Prior to this experiment, it had not been used in any formal study of information markets.
5. Appendix C contains results of individual and group estimates.
6. This analysis eliminates one extreme outlier. The most influential decision maker in one group had a MAPE of 98 percent which was the worst individual estimate among all subjects. Obviously, the group was able to correct this extremely conspicuous error.
7. Chen and Plott (2002) also observed that in all of their prediction market experiments at Hewlett-Packard, prices violated the no-arbitrage conditions and summed to be greater than the winning payoff (2002).
8. Questions 12 and 13 were posed orally after the group discussion. They asked subjects to identify the most influential or most talkative (question 12) and the least influential or least talkative (question 13) group members.
9. This is only a partial screenshot. Participants were able to scroll up and down and trade every price between \$0 and \$100.

REFERENCES

Lisa Anderson and Charles Holt 'Information cascades in the laboratory' *American Economic Review* (1997) 87(5) 847–862.

Solomon Asch 'Effect of group pressure upon the modification and distortion of judgments' in Goetzkow Harold (ed) *Groups, Leadership and Men* (New York, NY, Russell & Russell, 1963) pp 177–190.

Joyce Berg, Forrest Nelson and Thomas Rietz Accuracy and Forecast Standard Error of Prediction Markets, Working Paper, University of Iowa, Henry B. Tippie College of Business Administration (2003). Retrieved on March 30, 2007, from http://www.biz.uiowa.edu/iem/archive/forecasting.pdf

Joyce Berg and Thomas Rietz Aggregating Information: A Comparison of Prediction Markets and Delphi Techniques, University of Iowa, Henry B. Tippie College of Business Administration, Unpublished (2007).

Roger Brown *Social Psychology* (New York, NY, The Free Press, 1965).

Mark Cannon and Amy Edmondson 'Confronting failure: Antecedents and consequences of shared beliefs about failure in organizational work group' *Journal of Organizational Behavior* (2001) 22(2) 161–177.

Chen, Kay-Yut and Plott Charles Information Aggregation Mechanisms: Concept, Design and Implementation for a Sales Forecasting Problem, Working Paper 1131, California Institute of Technology, Division of the Humanities and Social Sciences (2002).

Chen, Yiling, Chu, Chao-Hsien, Mullen, Tracy and Pennock, David Information Markets vs. Opinion Pools: An Empirical Comparison (2005). Retrieved on March 30, 2007, from http://dpennock.com/papers/chen-ec-2005-market-vs-pool.pdf

Roger Clarke, Scott Krase and Meir Statman 'Tracking errors, regret, and tactical asset allocation' *Journal of Portfolio Management* (1994) 20(3) 16–24.

Werner De Bondt and Richard Thaler 'Does the stock market overreact?' *Journal of Finance* (1985) 40(3) 793–805.

Amy Edmondson 'Learning from mistakes is easier said than done: Group and organizational influences on the detection and correction of human error' *Journal of Applied Behavioral Science* (1996) 32(1) 5–28.

Robert Forsythe, Thomas Rietz and Thomas Ross 'Wishes, expectations and actions: price formation in election stock markets' *Journal of Economic Behavior and Organization* (1999) 39(1) 83–110.

Daniel Gigone and Reid Hastie 'Proper analysis of the accuracy of group judgments' *Psychological Bulletin* (1997) 121(1) 149–167.

Robin Hanson 'Decision markets' *IEEE Intelligent Systems* (1999) 14(3) 16–19.

Alan Ingham, George Levinger, James Graves and Peckham Vaughn 'The ringelmann effect: Studies of group size and group performance' *Journal of Experimental Social Psychology* (1974) 10(4) 371–384.

Irving Janis *Groupthink* (Boston, MA, Houghton Mifflin, 2nd ed., 1982).

Andrew Lo 'The adaptive markets hypothesis: Market efficiency from an evolutionary perspective' *Journal of Portfolio Management* (2004) 30(1) 15–29.

MacCoun 'Comparing micro and macro rationality' in Rejeev Gowda and Fox Jeffrey (eds) *Judgments, Decisions, and Public Policy* (Cambridge University Press, 2002) pp 116–137.

Terrance Odean 'Volume, volatility, price and profit when all traders are above average' *Journal of Finance* (1998) 53(6) 1887–1934.

Kenneth Oliven and Thomas Rietz 'Suckers are born but markets are made: Individual rationality, arbitrage, and market efficiency on an electronic futures market' *Management Science* (2004) 50(3) 336–351.

Rietz, Thomas Behavioral Mis-pricing and Arbitrage in Experimental Asset Markets, Working Paper, University of Iowa, Henry B. Tippie College of Business Administration (2005). Retrieved on March 30, 2007 from http://www.biz.uiowa.edu/faculty/trietz/papers/market.pdf

Edward Russo and Schoemaker *Decision Traps: Ten Barriers to Brilliant Decision Making and How to Overcome Them* (New York, NY, Simon&Schuster, 1989).

Arina Soukhoroukova and Martin Spann New Product Development with Internet-based Information Markets: Theory and Empirical Application. In 13th European Conference on Information Systems (ECIS), Regensburg, Germany (2005).

Martin Spann and Bernd Skiera 'Einsatzmöglichkeiten virtueller Börsen in der Marktforschung' *Zeitschrift für Betriebswirtschaft* (2004) 74(EH2) 25–48.

Cass Sunstein 'Deliberation and information markets' in Robert Hahn and Paul Tetlock (eds) *Information Markets – A New Way of Making Decisions* (Washington, DC, The AEI Press, 2006) pp 67–100.

Amos Tversky and Daniel Kahneman 'Judgment under uncertainty: Heuristics and biases' *Science* (1974) 185 1124–1131.

Amos Tversky and Daniel Kahneman 'The framing of decisions and the psychology of choice' *Science* (1981) 211 453–458.

Neil Weinstein 'Unrealistic optimism about future life events' *Journal of Personality and Social Psychology* (1980) 39(5) 806–820.

Justin Wolfers and Eric Zitzewitz 'Prediction markets' *Journal of Economic Perspectives* (2004) 18(2) 107–126.

APPENDIX

A. Questionnaire used for the experimental comparison[8]

Prediction Markets – A Classroom Experiment

Questionnaire:

1. Name: ____________________ 2. E-mail: __________________

3. Which phone would *you* most likely buy (first choice)? ___ (A-E)

4. Rank the 5 phones from *your* most likely to least likely purchase!
A: ___ B: ___ C: ___ D: ___ E: ___

5. Estimate the future market share *among college students*!
A: ___% B: ___% C: ___% D: ___% E: ___% (must sum to 100%)

6. How confident are you that your estimate is correct, from 0 (not all all) to 10 (very confident)? _____

7. Role: ☐ Trader ☐ Talker

8. Group #: ______ (assigned) 9. Your group member ID#: ____ (1-4)

Traders Only:

10. Your Portfolio After Trading Ends:
Shares of Product: A: ____ B: ____ C: ____ D: ____ E: ____
Cash Position: $ _________

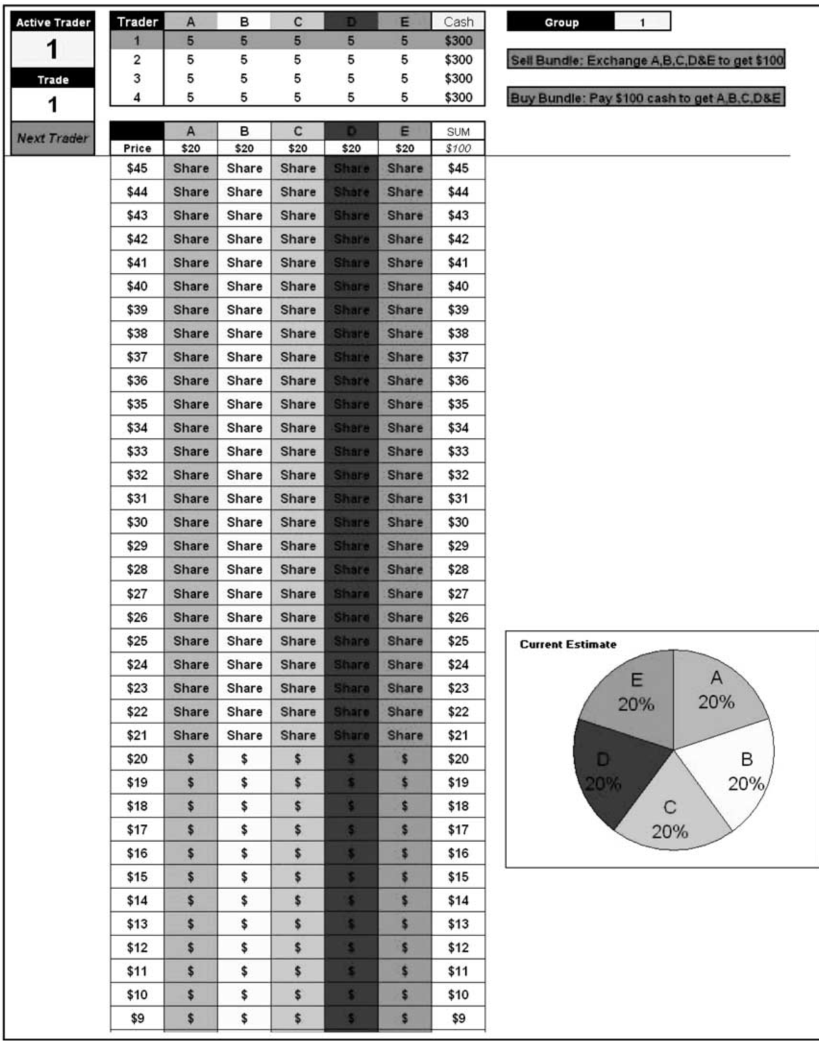
11. How confident are you that your group's market price estimates are correct, from 0 (not at all) to 10 (very confident)? _____

Talkers Only (remember ID# of your group members):

10. Group estimate of future market shares:
A: ___% B: ___% C: ___% D: ___% E: ___% (must sum to 100%)

11. How confident are you that your group's estimate is correct, from 0 (not at all) to 10 (very confident)? _____
12. Question 12 (will be asked orally after discussion) ? _____
13. Question 13 (will be asked orally after discussion) ? _____

B. Screenshot of the market board trading software[9]



C. Market share estimates by Traders, Talkers and Individuals

Traders (closing market prices)

Group	A	B	C	D	E	SUM	MAE	MAPE	COR.
1	*21*	*21*	*32*	*14*	*15*	103	1.45	7%	0.97
2	*20*	*17*	*36*	*19*	*18*	110	4.51	25%	0.78
3	*24*	*27*	*22*	*16*	*14*	103	3.81	19%	0.64
4	*31*	*17*	*29*	*9*	*14*	100	3.15	17%	0.90
5	*24*	*20*	*24*	*16*	*16*	100	2.10	11%	0.94
6	*24*	*23*	*24*	*14*	*20*	105	3.04	16%	0.84
7	*19*	*17*	*28*	*16*	*25*	105	4.36	26%	0.53
8	*19*	*18*	*30*	*19*	*17*	103	3.11	19%	0.79
9	*26*	*19*	*24*	*19*	*17*	105	3.50	20%	0.78
10	*25*	*20*	*27*	*14*	*12*	98	1.67	9%	0.95
11	*23*	*23*	*30*	*22*	*21*	119	3.85	26%	0.87
12	*21*	*20*	*26*	*16*	*13*	96	2.13	12%	0.94
MEAN	**23**	**20**	**28**	**16**	**17**	**104**	**3.06**	**17%**	**0.83**
STDEV	**3.4**	**3.0**	**4.0**	**3.4**	**3.7**				

Traders (normalized market prices)

Group	A	B	C	D	E	SUM	MAE	MAPE	COR.
1	*20*	*20*	*31*	*14*	*15*	100	1.24	6%	0.97
2	*18*	*15*	*33*	*17*	*16*	100	3.86	21%	0.78
3	*23*	*26*	*21*	*16*	*14*	100	3.63	18%	0.64
4	*31*	*17*	*29*	*9*	*14*	100	3.15	17%	0.90
5	*24*	*20*	*24*	*16*	*16*	100	2.10	11%	0.94
6	*23*	*22*	*23*	*13*	*19*	100	2.61	13%	0.84
7	*18*	*16*	*27*	*15*	*24*	100	4.57	26%	0.53
8	*18*	*17*	*29*	*18*	*17*	100	2.94	18%	0.79
9	*25*	*18*	*23*	*18*	*16*	100	3.32	18%	0.78
10	*26*	*20*	*28*	*14*	*12*	100	1.69	9%	0.95
11	*19*	*19*	*25*	*18*	*18*	100	3.41	20%	0.87
12	*22*	*21*	*27*	*17*	*14*	100	1.87	11%	0.94
MEAN	**22**	**19**	**27**	**15**	**16**	**100**	**2.87**	**16%**	**0.83**
STDEV	**3.8**	**2.9**	**3.5**	**2.7**	**3.1**				

Talkers (group estimates)

Group	A	B	C	D	E	SUM	MAE	MAPE	COR.
13	*35*	*15*	*25*	*10*	*15*	100	4.78	23%	0.73
14	*40*	*20*	*15*	*20*	*5*	100	9.67	49%	0.28
15	*20*	*20*	*40*	*10*	*10*	100	4.36	21%	0.95
16	*20*	*25*	*35*	*12*	*8*	100	4.30	22%	0.91
17	*20*	*30*	*30*	*10*	*10*	100	4.30	24%	0.83
18	*20*	*15*	*30*	*10*	*25*	100	4.39	26%	0.67
19	*17*	*20*	*30*	*20*	*13*	100	3.28	20%	0.71
20	*30*	*15*	*20*	*10*	*25*	100	6.78	35%	0.38
21	*20*	*15*	*40*	*15*	*10*	100	5.28	26%	0.87
22	*30*	*25*	*10*	*15*	*20*	100	7.64	34%	-0.11
23	*15*	*10*	*60*	*10*	*5*	100	12.36	56%	0.84
24	*40*	*25*	*20*	*5*	*10*	100	8.69	44%	0.62
MEAN	**26**	**20**	**30**	**12**	**13**	**100**	**6.32**	**32%**	**0.64**
STDEV	**9.0**	**5.8**	**13.4**	**4.5**	**6.9**				

Individuals (96 subjects)

							Based on Average Estimate		
	A	B	C	D	E	SUM	MAE	MAPE	COR.
MEAN	**23**	**18**	**31**	**15**	**14**	**100**	**6.06**	**8%**	**0.96**
STDEV	**11.6**	**7.8**	**14.3**	**6.3**	**7.7**				

Based on Individual Estimates

MAE	MAPE	COR.
9.77	**37%**	**0.54**