A Parimutuel-like Mechanism for Information Aggregation: Forecasting Inside Intel*

Benjamin J. Gillen, Caltech
Charles R. Plott, Caltech
Matthew Shum, Caltech

November 14, 2013

Abstract

We evaluate the performance of an information aggregation mechanism (IAM) implemented inside Intel to forecast unit sales for the company. Developed, refined and tested in the laboratory using experimental methods, the IAM is constructed from institutional features that improve performance and overcome barriers to successful applications of other information aggregation methods. Its implementation at Intel provides a testbed for evaluating this new form of IAM's performance in a complex field environment. In contrast to prediction markets, which provide only a point forecast of future sales, the IAM characterizes the full distribution of participants' aggregated beliefs allowing a more detailed evaluation of its performance. We show this predictive distribution very closely matches the distribution over outcomes at short horizons while slightly underweighting low-probability realizations of unit sales at long horizons. Compared to Intel's "official forecast," the IAM forecasts perform well overall, even though they predate the official forecasts. The forecast improvements are most prominent at short forecast horizons and in direct distribution channels, where the effective aggregation of individually-held information drives the IAM to be more accurate than the official forecast over 75% of the time.

*The financial support of the Lee Center for Advanced Networking, the Gordon and Betty Moore Foundation, and the Laboratory of Experimental Economics and Political Science is gratefully acknowledged. We thank Dan Zhou for providing excellent research assistance, and Chew Soo Hong, Erik Snowberg, Allan Timmermann, Michael Waldman, and seminar audiences at Arizona State, EIEF Roma, MIT, Rice, UCSD, Universidad Carlos III Madrid, IIOC 2013, and the Econometric Society NASM 2013 for helpful comments.

1 Introduction

In many companies, internal forecasts of key financial and operational indicators provide a crucial performance metric and input into strategic planning decisions and managing market expectations. Typically, these forecasts are derived from the analysis of in-house experts, collecting dispersed information from disparate sources in a process consisting of as much art as science. In this paper, we study the use of a completely different type of procedure – an information aggregation mechanism (IAM) based on decentralized competition, motivated by economic theory, and refined and tested through experimental economics. The mechanism has been shown to work well in the simple and special cases of laboratory settings. The challenge is to test the robustness of the same mechanism when operating in the much more complex environment of a Fortune 500 company. Will it work at all? Will it be useful for improving the internal forecasting process used in the company?

The goal of an information aggregation mechanism (IAM) is to collect and aggregate the information held in the form of the subjective intuitions from a disperse collection of people. This task requires developing instruments to quantize information while setting proper incentives that balance the reward of revealing information with the hazards of free riding on others information and thus might lead to successful information aggregation. In a break from more traditional, theory-derived, approaches to mechanism design, the IAM was formulated and refined in a series of complex laboratory experiments. We discuss this break with theory and its econometric implications in Sections 3 and 4. The study of information aggregation in experimental economics laboratories has a long history, providing a valuable base on which to build. The ability of markets to perform the information collection and aggregation functions and also the sensitivity of such performance to the details of the market institution were first observed experimentally by Plott and Sunder (1982, 1988). Similarly, the possibility that markets might be designed to perform the aggregation function and implemented inside a business is well known (Chen and Plott (2002); Plott (2000)).

The mechanism studied in this paper shares institutional features with auctions, exchanges, and some betting processes. Because some features are also found in parimutuel betting systems, we call it a *Parimutuel Information Aggregation Mechanism (IAM)*. Indeed, major features of this mechanism were developed as a response to information aggregation shortcomings revealed in laboratory experiments of betting systems, which have a goal of entertaining participants as opposed to aggregating their information. At the same time, the design of the mechanism also reflects an attempt to avoid features that might inhibit the application of information aggregation mechanisms inside a business environment. As will become apparent, many features of the mechanism reflect an effort to draw on experiences derived from the application of markets to perform the same function.

We report results from a field experiment in which the IAM is implemented to forecast unit sales activity by Intel. As an international market leader in the hi-tech sector with annual revenues over $50 billion, Intel has one of the most recognizable brand names among American companies and its products are found in virtually all households in the country. Accurate forecasts of product sales are incredibly important both operationally, ensuring sufficient inventory is available for distribution, and financially, managing market expectations for shareholder value. With myriad distribution channels, forecasting product sales for the organization is an incredible task requiring analysts to aggregate information from sales reports, partner forecasts, and management guidance. As such, the requisite information for forecasting is dispersed through the firm among a variety of stakeholders. Adapting an IAM to this environment provides a more "scientific" approach to consolidating this information.

At Intel, we set up a collection of mechanisms to characterize uncertainty in future realizations of units sold for key products. The range of values that possible sale quantities can take is partitioned into a set of non-overlapping intervals, or "buckets." The analysts participating in each mechanism are asked to purchase "tickets" that pay off when the variable of interest takes a value within a given bucket. Analysts are allowed to buy as many tickets as they wish (up to a budget limit described below) and place them freely in any of the buckets. In this way, the distribution of tickets placed across the different buckets yields a natural measure of analysts' beliefs regarding the future realization of the variable of interest. The information aggregation mechanism automatically aggregates these beliefs across analysts, allowing decision makers to easily form "consensus" forecasts while also obtaining a glimpse into the uncertainty underlying these forecasts. In addition to the IAM's aggregated forecasts, we also have access to an internally-prepared "official" forecast that can serve as a benchmark against which to evaluate the mechanism's performance. Comparing the relative effectiveness of these mechanisms provides strong support for the IAM in an incredibly complex field setting.

Our empirical analysis uncovers two central findings confirming the effectiveness of Intel's implementation of the IAM. First, the degree to which the beliefs recovered from the IAM are consistent with rational expectations depends on the amount of noise and information in the forecasting environment. The IAM beliefs over unit sales matches the distribution of realized sales at short forecast horizons (up to three months) where individual information is relatively rich. At longer horizons (beyond six months), the mechanism's forecast distribution tends to understate the dispersion of uncertainty in sales, underweighting low probability events. While our information aggregation mechanism is designed to avoid key features of gambling processes, we find a "reverse favorite-longshot bias" in the beliefs derived from the mechanism similar to those found in empirical studies of betting markets. Second, the mechanism's expected outcome outperforms the official sales forecast at short horizons and in direct distribution product channels. This performance is highlighted by considering the ex-post optimal combination of the two forecasts, which heavily weights the IAM at short horizons and the official forecast at long horizons.

The appealing performance of the IAM is apparent not only in its empirical properties, but also in the degree to which Intel has expanded its utilization of the IAM in its forecasting and planning process. Starting from an initial pilot of the mechanism, Intel has continued to expand its implementation of the mechanism to several markets that target an important piece of the business. Further, the organization has explored using IAMs beyond forecasting sales itself, such as mechanisms suited to evaluating new ventures in research and development, project management risks, and a variety of other business problems that rely on information dispersed among a number of stakeholders. To our knowledge, the Intel IAM is now the longest-running implementation of an economically-motivated internal forecasting mechanism in industry, which we attribute to its unique features that build on experiences with the business applications of prediction markets.

2 Developing Information Aggregation Mechanisms

The purpose of an IAM is to quantify and collect information that might be held, in the form of vague and subjective intuitions, by dispersed individuals. The hope is that the collection and aggregation of this information produces a combined signal that has more information content than any single signal. A connection between markets and information transmission dates back to the foundations of economics (see Allen and Jordan (1998) for a review of this early development and general principles for the existence of rational expectations equilibrium). These theoretical results suggested that markets are capable of collecting and aggregating information, though exactly if and how that might happen was an open question.

2.1 Experimental Foundations for Information Aggregation

Motivated by theoretical suggestions of informationally efficient markets and rational expectations, Plott and Sunder (1982, 1988) looked to experiments as tools for examining the possibility. Plott and Sunder (1982) first demonstrated the ability of continuous double auction markets to transfer information from "insiders" who have information about the state to non-insiders who do not. Plott and Sunder (1988) builds on this initial finding, demonstrating further that the information transmission and collection can go beyond the simple transfer of information to a process of aggregating the information contained in multiple, independent sources. That is, market-based systems could effectively transfer "soft" information that exists in the form of intuitions into a quantitative signal consistent with Bayes Law. Of significance to the current design of an IAM, they demonstrated that the ability of markets to perform this task is dependent on the trading instruments available. In particular, markets perform the collection and aggregation well if populated by a complete set of Arrow-Debreu securities.[1]

The first application of a market based IAM inside a business was conducted by Chen and Plott (2002) inside Hewlett Packard Corporation. They implemented a complete set of Arrow-Debreu securities to aggregate information about future sales. The possible sales were divided into states, each state supporting an Arrow-Debreu security, and a continuous double auction market was opened for each of the securities. Since the payoff of the winning security was one and the payoff of losing securities was zero the prices of the complete set of securities could be interpreted as a probability distribution over the states. The mechanism was reported as successful but its use was limited due to difficulties related to coordinating and managing the mechanism. Many of the features of the IAM developed and tested here emerged in response to difficulties relating to deploying market-based IAM's inside businesses.

The design of the IAM reported and studied here shares some features with parimutuel betting processes - hence the reference to parimutuel incentives. In a parimutuel betting

[1]Information aggregation does not necessarily happen if the market has a single compound security and all agents do not have the same preferences. However the prices in a single compound security are related to the competitive equilibrium based on private information. This property, which is common to a private information equilibrium, was demonstrated experimentally by Plott and Sunder and expanded further by Berg et al. (2008). A recent theoretical literature (Ottaviani and Sørensen (2009, 2010)) explores equilibrium properties of parimutuel betting systems, with the thrust of these results indicating the potential for biased odds from parimutuel betting systems in equilibrium. These results underscore the importance of designing incentives that facilitate communication and learning among mechanism participants in a dynamic interaction.

system participants buy tickets on states of nature, such as the winner of a horse race, and tickets are sold at a fixed price. The revenue from all ticket sales are accumulated, called the purse, and paid to the holders of tickets on the winning bucket. The odds computed from this process reflect the number of tickets sold for a bucket divided into the size of the purse. There is a strong tendency for the odds to be related to the frequency with which the winner occurs. That tendency, which suggested a principle for a new type of IAM, was clearly established experimentally by Plott et al. (2003).

The parimutuel incentives in the IAM implemented at Intel differ from those in parimutuel betting systems in fundamental ways. First, tickets are not sold at a fixed price, but rather prices evolve with an exogenous trend, in order to encourage a timely completion of the process. Our specific timing setup is informed by the experiments in Axelrod et al. (2009) that demonstrate the importance of structuring the process to encourage participants to buy their tickets early rather than waiting until the last second in an attempt to free ride on information supplied by others.[2] Second, for purchasing the tickets, participants are allocated a fixed budget of a synthetic currency that had no value other than to buy tickets in the designated IAM. The use of a synthetic currency follows Plott and Roust (2009), and works to mitigate the negative impact of risk aversion on information aggregation.[3] Finally, the mechanism is not self-financing, with management providing a fixed cash prize distributed in proportion to the number of tickets in the winning bucket.

The IAM which we ran in Intel also features important differences from prediction systems based on markets, which have flourished in recent years.[4] Most importantly, the tickets

[2]Plott et al. (2003) demonstrated that the tendency to wait until the last second to buy tickets contributed to the creation of bubbles and retarded successful information aggregation. This property was replicated by Kalovcova and Ortmann (2009). In a very simple parimutuel betting system information is transferred through a process of observing betting and the importance of observing others in a parimutuel betting system context is examined by Koessler et al. (2012). Ottaviani and Sørensen (2004) present a theoretical analysis for the timing of bets in parimutuel betting systems, deriving results consistent with several of these experimental findings. The unique timing features in Intel's IAM help mitigate the impact of these incentives on information aggregation and provide an important differentiation between the IAM and parimutuel betting.

[3]Risk aversion has a tendency to inhibit participation even though an agent is informed and thus prevents information from getting into the system. Plott and Roust (2009) demonstrate that poor performance of the mechanism is closely related to poor information and to the extent that risk aversion diminishes the quality of information the removal of risk aversion is important.

[4]The Iowa Electronic Markets constituted the first "prediction markets" in the sense that the price of a binary security can be viewed as a probability and used to predict elections (see Berg et al. (2008) for a survey of these applications). Internal corporate prediction markets were broadly deployed at Google (Cowgill et al. (2009)) to gauge employees sentiments on everything from companys performance to general industry issues, though the information relayed by these markets were often biased by participation effects. Worth noting, though Google has ceased experimenting with their internal markets for business purposes, Intel continues to utilize our IAM in its sales forecasting process.

placed by IAM participants are not securities, and cannot be traded. Price speculation, which takes place in markets, cannot take place here. This payoff structure differs from prediction markets, where securities are traded by market participants over time. Manski (2006) discusses the difficulties in interpreting prediction market prices when participants may have heterogeneous beliefs, highlighting the issues not only in interpreting the data but also in the ability of the market mechanism to successfully aggregate information. The IAM is also less exposed to a "thin market" pheonomenon, as thin trading in a market can severely inhibit aggregation.[5] The timing features of our IAM (described above) were adopted to mitigate these problems. As such, our IAM is substantially removed from the features of an asset market. Indeed, the timing of the IAM is coordinated to be compatible with the busy schedules of participants. There is a fixed, pre-announced start and end time so that people know when to log in to actively participate. The sessions themselves are timed to hit key points of the Intel business cycle. By design, the output of the IAM is freshly available to other business processes that use it.

Another distinguishing feature of the Intel IAM is its freedom from self-selection and participation-induced bias. These selection issues are particularly prominent in the literature on prediction markets, which typically rely on participants selecting to engage in the mechanism, guided by the belief that increasing the size of the crowd maximizes its wisdom.[6] By contrast, the Intel IAM management invited participants chosen for the information to which they had access given their position in the organization.

Finally, we set up IAM's for sales forecasting that elicit participants' beliefs about variables (unit sales) which can take many (>>2) values. Specifically, we set up a complete set of simultaneous instruments, one for each value that the variable can take. This approach contrasts with many prediction markets, in which the outcome of interest is binary (or otherwise takes a small number of values); for instance, whether Obama or Romney would win the latest presidential election. Our approach operationalizes a general principle (see Plott (2000)) that the extent of information aggregation is limited by the dimensionality of the "message space" in which market participants operate. Taken as a whole, the activity in all these markets yields a complete probability distribution over the event space that, ideally,

[5] This was an issue encountered in the Hewlett-Packard IAM implemented by Chen and Plott (2002).

[6] This phenomenon may be accentuated in horse-racing parimutuel markets, in which individual decisions may be directed by the thrill of uncertainty and surprise rather than the desire to profit from exclusive information. While Woodland and Woodland (1994) and Gray and Gray (1997) find that thick betting markets for professional sports tend to satisfy market efficiency, a host of papers have explored potential cases of inefficiencies in recreational betting markets. Jullien and Salanie (2000) and Chiappori et al. (2009) discuss the identification and estimation of risk preferences using data from parimutuel markets.

will reflect the aggregation of private information about the various possible outcomes.

2.2 The IAM inside Intel

For description purposes we will consider a single variable, say unit sales for product i in quarter t, that we denote by $Y_{i,t}$. The positive real line is partitioned into K intervals, or "buckets," where each interval represents a range of possible values for sales that will be officially reported at the end of the sales period. The leftmost and rightmost buckets are, respectively, $[0, x_1)$ and $[x_{K-1}, \infty)$.

Participants interact with the mechanism in the form of an on-line interactive program. Mechanism organizers invite participants, who securely log in to their own account to access the IAM program. The mechanism makes "tickets" available for sale to participants, who spend an endowment of Francs (our synthetic experimental currency) on tickets and allocate them across the buckets. At the opening of each application all participants are given a fixed budget of 500 Francs for each of the predicted variables. The Francs cannot be transferred among participants, used in other applications, or assigned to buckets for another variable's IAM. As quality controls over the mechanism's operation, the IAM operates at a fixed time and only those invited are able to participate. The IAM program stores a wealth of data, including individual participant actions and time-stamps indicating when each of these actions took place.

The tickets for all buckets are priced the same and that price will move up at a pre-announced rate to ensure the mechanism closes in a reasonable time. For example, the opening price would be constant for fifteen minutes and then go up at a rate of one Franc per minute after that. These price changes discourage waiting until the last second to purchase, helping to offset individual incentives to hold back their private information and to improve their own information by learning from others' decisions. All participants are aware that their own information might be improved through seeing the purchases of others. They are also aware that their own information might be communicated by their own purchase of tickets. Inducing temporal discounting helps to mitigate these strategic incentives that otherwise hinder successful information aggregation. The price increase is constant but sufficiently substantial that by 40 minutes into the exercise the ticket prices are so high that the budget has little purchasing power. Notice, that this process is fundamentally different from betting processes.

Throughout the operation of the mechanism, participants have a continuously available record of the number of tickets that are currently placed in each of the buckets. At each instant during the application as well as at its termination, the placements of all tickets in all buckets are known. The individual participant also knows the proportion of tickets he or she holds in each bucket, which is particularly important because these proportions are the foundations for incentives. When the actual winning bucket becomes known those holding tickets in that bucket are given a part of a grand prize equal to the proportion of the winning bucket tickets that he or she holds. If participant n holds $z\%$ of the tickets sold for the winning bucket then participant n gets $z\%$ of the incentive prize. For example, if the incentive prize was \$10,000 and the individual held 10% of the tickets sold for that bucket then the payment to participant n would be \$1,000.[7]

Participants depend on the nature of the forecasting exercise. For forecasts of variables that have significant influence on financial performance, only insiders, those with access to limited financially relevant information, are permitted; forecasts that are not considered material to earnings reports may include a wider group. Typically, the forecasters are insiders with direct access to the most the information relevant to the forecasting problem, either directly involved in management or sales. Data already available (to the insiders), including current signals and historical results, are packaged for all participants to study in preparation for the IAM exercise, establishing a base of relevant information to provide an underlying distribution of common knowledge. As such, it is important to synchronize the start time so that those individuals with appropriate information could participate.

A typical IAM exercise involves forecasting for the current quarter plus the three upcoming quarters. The exercise takes place once a month and requires on the order of 30 minutes. Each participant is given a separate Franc budget for each item they forecast. All budgets are the same size and the budgets are not fungible across the items forecast. The number of participants varies from ten to twenty-five and each operates from a secure computer located wherever the participant happened to be located, home, office, traveling, etc. Typically the users are anonymous within the mechanism: both the list of participants and the winners are secret. Of course, the total of tickets purchased in each bucket of each forecast is public and known in real time as the tickets are purchased.

[7]The use of incentives inside a business reflected a belief and experience for experiments that incentives are central to the successful operations of information aggregation mechanisms. The performance of a mechanism without incentives (cheap talk) is explored by Bernnouri et al. (2011) as are the success of different measures of information aggregation. One might worry that these payments could provide a disincentive for employees to communicate information amongst each other during day-to-day operations, as they seek to exploit their information for advantage in the IAM. Contacts at Intel did not report any such behavior, but the possibility indicates the importance of balancing IAM incentives with other operational incentives.

3 Information Aggregation in the IAM

Since the IAM that was run inside Intel is practically complicated in many ways, a full-fledged dynamic equilibrium analysis is beyond the scope of our work here. To focus on the problem of assessing the information aggregation properties of the mechanism, we consider in this section some illustrative examples that provide intuition for the conditions under which information aggregation can or cannot occur. The fact that the information aggregation mechanism evolved through the application of laboratory methods, as opposed to pure theory, limits our access to a structural theory of behavior that informs our empirical analysis. Part of this barrier to testing the mechanism can be overcome in a laboratory context through careful design and knowledge of the experimental protocol. However, the central testable axiom supported in laboratory findings, that the IAM accurately reflects the conditional information contained in a disperse population's privately observed signals, arises largely as a consequence of rational expectations. Assuming this consequence extends to any information structure, the axiom provides the focus of our empirical study in the next section. In addition to characterizing successful information aggregation, these examples help illustrate differences between incentives in the IAM and those faced by participants in standard prediction markets.

Our first example, adapted from the "NOT SETS" example in Plott et al. (2003), examines how to evaluate the occurrence of information aggregation under laboratory experimental conditions. In this setting, the experimenter knows the underlying information generating process and can consequently evaluate whether the IAM was accurate without a complete theory of the process through with that accuracy is attained. The example illustrates the central tool used to improve the testing process in the absence of a complete theory and the role of communication in successful aggregation. However, extending the example also demonstrates that this simple test for aggregation is not available in the field, where we do not necessarily know the distributional and dependence structure for the latent information observed by individuals.

Our second example, which analyzes information aggregation in the context of a Dirichlet learning process, provides three lessons. First, it characterizes a tractable model of an information generating process that parallels the information aggregation mechanism in field applications. Second, the Dirichlet example highlights the theoretical challenges in developing a behavioral model connecting the information structure with its implications for the IAM through the principle of rational expectations. Third, and most importantly, the example provides a demonstration that rational expectations can be manifest in the accuracy of the information aggregation mechanism. That is, individual decision behavior and incentives do not preclude the possibility of information aggregation in the IAM, even though the successful outcome is not guaranteed by theory. Specifically, the Case 1 demonstrates that rational expectations will emerge if the underlying individual behavior is approximated by a naïve strategy. However, Case 2 demonstrates that reasonable deviations from the naïve strategy can result in a failure to aggregate information.

Despite the absence of a theoretical resolution, the issue of aggregation can still be approached empirically. We address this problem in Section 4, where Hypothesis 1 provides a formal statement of the testable empirical implications of IAM accuracy.

3.1 Example 1: The NOT SETS Experimental Condition

We begin illustrating the properties of information aggregation itself with the "NOT SETS" example adapted from Plott et al. (2003) and discussed further in Wit (1997)'s dissertation, which provided early successful tests of the IAM. Suppose the observed state of the world corresponds to one of three states, $Y \in \{A, B, C, D\}$, with a distribution that varies conditional on another variable $X \in \{1, 2, 3, 4\}$. The joint distribution over these two random variables is as follows:

$Y \setminus X$	1	2	3	4	$P\{Y\}$
A	$\frac{97}{400}$	$\frac{1}{400}$	$\frac{1}{400}$	$\frac{1}{400}$	$\frac{1}{4}$
B	$\frac{1}{400}$	$\frac{97}{400}$	$\frac{1}{400}$	$\frac{1}{400}$	$\frac{1}{4}$
C	$\frac{1}{400}$	$\frac{1}{400}$	$\frac{97}{400}$	$\frac{1}{400}$	$\frac{1}{4}$
D	$\frac{1}{400}$	$\frac{1}{400}$	$\frac{1}{400}$	$\frac{97}{400}$	$\frac{1}{4}$
$P\{X\}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	

Three players receive signals S_1, S_2, and S_3, each of which consist of the true realization of X and another, randomly drawn state. However, the signals are not independently generated so that the two randomly drawn states do not match. For example, if the true state is $X = 1$, then the signals generated by nature could be $(1, 2), (1, 3)$, and, $(1, 4)$. Consequently, if players were to reveal their signals to one another, they would know the true value of X. As such, the aggregated information about the distribution for Y from the observed signals, $Y|S_1, S_2, S_3$, corresponds to the conditional distribution for $Y|X$. In the example where $X = 1$, this distribution would place 97% of probability mass on the event $\{Y = 1\}$ and 1% of the probability mass on each of the remaining possible outcomes for Y.

If we remove the conditional structure that ensures players receive different "decoy" signals, then it would be possible for all three players to receive the same signal. For example, it would be possible for the true state to be $X = 1$ and the three signals generated by nature were each $(1, 2)$. In this case, the signals do not reveal the true value of X and so the aggregated information about the distribution for Y conditions only on $X \in \{1, 2\}$. As such, the expected distribution over Y places equal weight on the conditional distribution being generated by either the state $X = 1$ or $X = 2$. In the specific example, this distribution places 49% of the probability mass to each of the outcomes $\{Y = 1\}$ and $\{Y = 2\}$ and 1% probability mass to each of the outcomes $\{Y = 3\}$ and $\{Y = 4\}$.

3.2 Example 2: Individual Incentives with Dirichlet Beliefs

Our second example utilizes a Dirichlet sampling framework to illustrate the link between individual incentives and beliefs in the IAM.[8] Recall that we partition the set of feasible sales into a set of K ranges or "buckets," denoted by $x_1, \ldots, x_K$. Without loss of generality beyond this discretization, we can characterize the conditional probability that realized sales fall into a given bucket with a K-point multinomial distribution:

$$Y|\vec{\pi} \in \begin{cases} x_1, & \text{with prob. } \pi_1 \\ x_2, & \text{with prob. } \pi_2 \\ \cdots & \cdots \\ x_K, & \text{with prob. } \pi_K \end{cases} \tag{1}$$

The cell probabilities given by $\vec{\pi} = (\pi_1, \ldots, \pi_K)'$ are unknown quantities, about which agents are endeavoring to learn.

Suppose agents start off with a (common) prior that $\vec{\pi}$ follows a Dirichlet distribution with non-negative parameters $\vec{\alpha} = (\alpha_1, \ldots, \alpha_K)'$, supported on the K-dimensional unit simplex. That is, the prior distribution and expectation for the cell probabilities are:

$$\vec{\pi} \sim Dir\left(\alpha_1, \ldots, \alpha_K\right), \quad E\left[\pi_k\right] = \frac{\alpha_k}{\sum_{j=1}^{K} \alpha_j}$$

Suppose further that each agent updates her beliefs about $\vec{\pi}$ upon observing noisy signals

[8] The Dirichlet distrbution has also been used in other contexts (eg. Rothschild (1974)) to study learning and sampling in a discrete setting.

of $\vec{\pi}$. Specifically, an agent observes m signals $s_{n,1}, \ldots, s_{n,m}$, drawn independently from $MN(\vec{\pi})$. From these m signals, the agent can compute sample frequencies $p_{n,1}, \ldots, p_{n,K}$, where $p_{n,k} = \frac{1}{m} \sum_{j=1}^{m} \mathbb{1}_{s_{n,j}=k}$, the sample frequency that the signal falls into the k-th bucket. Given these conjugate distributional assumptions, the posterior distribution for $\vec{\pi}$ conditional on these signals will also be Dirichlet, corresponding to:

$$\vec{\pi}|\vec{s}_n, \vec{\alpha} \sim Dir\left(\alpha_1 + mp_{n,1}, \ldots, \alpha_K + mp_{n,K}\right)$$
$$E\left[\pi_k|\vec{s}_n, \vec{\alpha}\right] = \frac{\alpha_k + mp_{n,k}}{m + \sum_{j=1}^{K} \alpha_j} \equiv \overline{p}_{n,k}$$

In this setting, we want to examine whether information aggregation should occur – that is, whether the distribution of ticket placements coincides with the "true" underlying distribution of sales Y? We consider two cases in turn.

3.2.1 Case 1: Naive Behavior

First, we consider a scenario in which players follow a simple "naive" strategy where they place tickets proportional to their privately observed signals, so that $\nu_{n,k} \propto p_{n,k}$. Under the simplifying assumption that the signals are independent across agents, all of whom receive the same number of signals, we can obtain the aggregated posterior distribution, across all N agents:

$$\vec{\pi}|\vec{s}_1, \ldots, \vec{s}_N, \vec{\alpha} \sim Dir\left(\alpha_1 + m\sum_{n=1}^{N} p_{1,n}, \ldots, \alpha_K + m\sum_{n=1}^{N} p_{K,n}\right)$$
$$E\left[\pi_k|\vec{s}_1, \ldots, \vec{s}_N, \vec{\alpha}\right] = \frac{\alpha_k + m\sum_{n=1}^{N} p_{n,k}}{\sum_{j=1}^{K} \alpha_j + N * m} \equiv \tilde{p}_k$$

This last posterior distribution, $\vec{\pi}|\vec{s}_1, \ldots, \vec{s}_N, \vec{\alpha}$ represents the fully aggregated information regarding the distribution of the outcome variable Y available to participants. Intuitively, each individual's draw from the multinomial distribution corresponds to m "pieces" of information about the true distribution for Y and the Dirichlet distribution provides a convenient summary of the total information revealed to individuals. Now, if every player is equally informed via independent signals and prior beliefs are diffuse (so that $\vec{\alpha}$ is small), then, as $m \to \infty$, the law of large numbers ensures that that full-information posteriors converge to the true probabilities:

$$\frac{\alpha_k + m\sum_{n=1}^{N} p_{n,k}}{\sum_{j=1}^{K}\alpha_j + N * m}\tilde{p}_k \to \pi_k, \quad k = 1, \ldots, K. \tag{2}$$

At the same time, the assumption of naive behavior – that people place tickets proportional to the posterior beliefs, means that η_k, the total tickets in bucket k, will end up being proportional to the full-information posterior $\tilde{p}_k$: that is, information aggregation obtains: $\eta_k \propto \pi_k$.

However, as we see next, the parimutuel mechanism's incentives need not induce agents to place tickets in line with their posterior beliefs at any point in time. Moreover, even if players' behavior follows the naïve model proposed here, there are open questions concerning whether they will place tickets in line with their signals in the presence of heterogeneity or correlated information. These behavioral issues severely complicate the theoretical analysis of information aggregation.

3.2.2 Case 2: Strategic Ticket Placements

Next we consider more strategic behavior by the players in the IAM. Suppose the IAM is in a state where each bucket x_k has η_k tickets in it and denote the vector of states across buckets by $\vec{\eta}$; also, suppose further that agent n's posterior expected beliefs are given by $\vec{p}_n = [p_{n,1}, \ldots, p_{n,K}]'$. For a risk-neutral agent who has not yet placed any tickets, the marginal value of placing an additional ticket in bucket x_k is simply their posterior expectation of the realized outcome falling in that bucket divided by the number of tickets within the bucket:

$$V_n\left(x_k \middle| s_n, \vec{\eta}\right) = \frac{p_{n,k}}{1 + \eta_k}$$

Consequently, this agent would place their marginal ticket in the bucket that has the highest "odds" – that is, the largest posterior likelihood $p_{n,k}$ relative to the number of tickets in the bucket $(1 + \eta_k)$. As this discussion makes clear, an agent who places tickets strategically cares not primarily about what the most likely outcome is, but rather the outcome for which the empirical distribution of ticket placements differ most from his posterior likelihood. The two may be different, except in the case where posterior beliefs do not deviate too far from "consensus" (as reflected in the current distribution of ticket placements).

Similarly, if agent n has already placed $\nu_{n,k}$ tickets in bucket x_k, then his marginal expected payoff from placing an additional ticket in this bucket is now:

$$V_n\left(x_k|\nu_n, s_n, \vec{\eta}\right) = p_{n,k}\left(\frac{1+\nu_{n,k}}{1+\eta_k} - \frac{\nu_{n,k}}{\eta_k}\right)$$

which may be even further distorted away from $p_{n,k}$.

This possibility that agents may place tickets in a manner far from their posterior beliefs is an impediment to information aggregation in this market. Given the analysis above, there is no compelling reason to expect information aggregation to hold in this setting.[9] Still, it is ultimately an empirical question as to whether the IAM "works": whether it effectively aggregates information and leads to accurate forecasts of sales. While the many potential barriers to information aggregation indicate its success can be theoretically assured only in special circumstances, the following sections provide statistical tests of this hypothesis. These theoretical considerations, and their implied skepticism of the IAM performance, renders the IAM's effectiveness an all the more remarkable finding.

4 Testing the Information Aggregation Mechanism

The transition from experimental data to field data rests on the common, underlying hypotheses that forms the central empirical question of our study. Under the information aggregation hypothesis, the empirical distribution of ticket placements in the IAM should coincide with the "true" probabilities of sales. Recall $Y_{i,t}$ as the actual realization of unit sales and, denoting the information set of IAM participants at period $t-h$ by $\mathcal{F}_{i,t,t-h}$, we have:

Hypothesis 1. *Information Aggregation Mechanism Accuracy hypothesis:*

$$Y_{i,t}|\mathcal{F}_{i,t,t-h} \stackrel{d}{=} MN(\tilde{\eta}_{1|t-h}, \cdots, \tilde{\eta}_{K|t-h}), \quad \forall i, t, h. \tag{3}$$

where $\tilde{\eta}_{k|t-h}$ denotes the proportion of tickets placed in bucket k during the IAM at horizon

[9] In a series of papers, Ottaviani and Sørensen (2009, 2010) study betting in a parimutuel gambling environment, finding features of the betting system that cause market odds to differ substantially from the "full information" odds in equilibrium. An important feature of their results is the static, simultaneous, nature of the game. As demonstrated in Example 1 and in another example in Appendix A, communication among players and dynamic inference is important for information to be effectively aggregated.

$t - h$.

Why might we think this hypothesis holds a priori? The experimental evidence presented in Plott et al. (2003) certainly provide some empirical motivation. Similarly, Example 2 from the previous section illustrated both the possibility of information aggregation as well as the special conditions that are required for it to obtain theoretically. Finally, one could argue that, essentially by definition, an ex-post equilibrium supports information aggregation through a no-arbitrage argument.[10] Of course, multiple equilibria might exist, there is no complete theory of dynamics that supports convergence, and behavior in the field need not conform to theoretical models of equilibrium. Thus, the performance of the IAM is an empirical issue and, given the strong evidence for information aggregation in experiments, it makes sense to look for such evidence in the field.

Testing this hypothesis in the field application the analysis is more complicated because we lack information available to the experimenter in analyzing laboratory data. In the field, we require technical assumptions to compensate for the information differences between the field and the laboratory. In particular, within the laboratory experimental context, the designer knows exactly what the true conditional distribution for $Y_{i,t}$ given all information in the system. Consequently, the cell probabilities from experimental IAMs can be directly compared to these to test the Information Aggregation Mechanism Accuracy hypothesis. In the field, however, the econometrician lacks access to the series of conditional cell probabilities conditioning on the information available at different horizons. Rather, we only the realized value of $Y_{i,t}$ to test the forecast distributions across all horizons.

We operationalize the test by defining the cumulative conditional distribution $\hat{G}_{i,t|t-h}(y) = \sum_{k=1}^{max\{\kappa|x_\kappa \leq y\}} \tilde{\eta}_{k|t-h}$ corresponding to the ticket placements at horizon $t - h$. Then we transform the realized outcome $Y_{(i,t)}$ into its corresponding quantile in the conditional IAM distribution:

$$\hat{U}_{i,t,h} \equiv \hat{G}_{i,t|t-h}(Y_{i,t}) \sim_{H_0} U[0,1] \tag{4}$$

By translating the outcome into its conditional quantile from the IAM, we control for the

[10] Formally, suppose information has been aggregated through the interactions with the IAM and participants have arrived at a common posterior expectation for the cell probabilities characterizing the distribution over outcomes. As such, every player has the same posterior expected cell probabilities. Suppose every player then allocates their tickets proportionally to those common posterior expected cell payoff odds. By the payoff analysis from the Dirichlet specification in Example 2, these players would all be indifferent as to in which bucket they should place their tickets (within discretization errors). In a context with rational expectations, these posterior expected cell probabilities are accurate, and the fundamental relationship between the IAM and the empirical relationship between IAM probabilities and relative frequencies described by Hypothesis 1 would be satisfied.

heterogeneous conditional distributions from which $Y_{i,t}$ is being realized at different horizons. Accordingly, we can simply use a Kolmogorov-Smirnov test to evaluate whether we can reject that our sample of $\left\{\hat{U}_{i,t,h}\right\}_{t=1}^{T}$ are uniformly-distributed i.i.d. draws.[11] By analyzing the conditional quantile, such a test is robust to heterogeneity in the distributions across products, time, and the information available.[12]

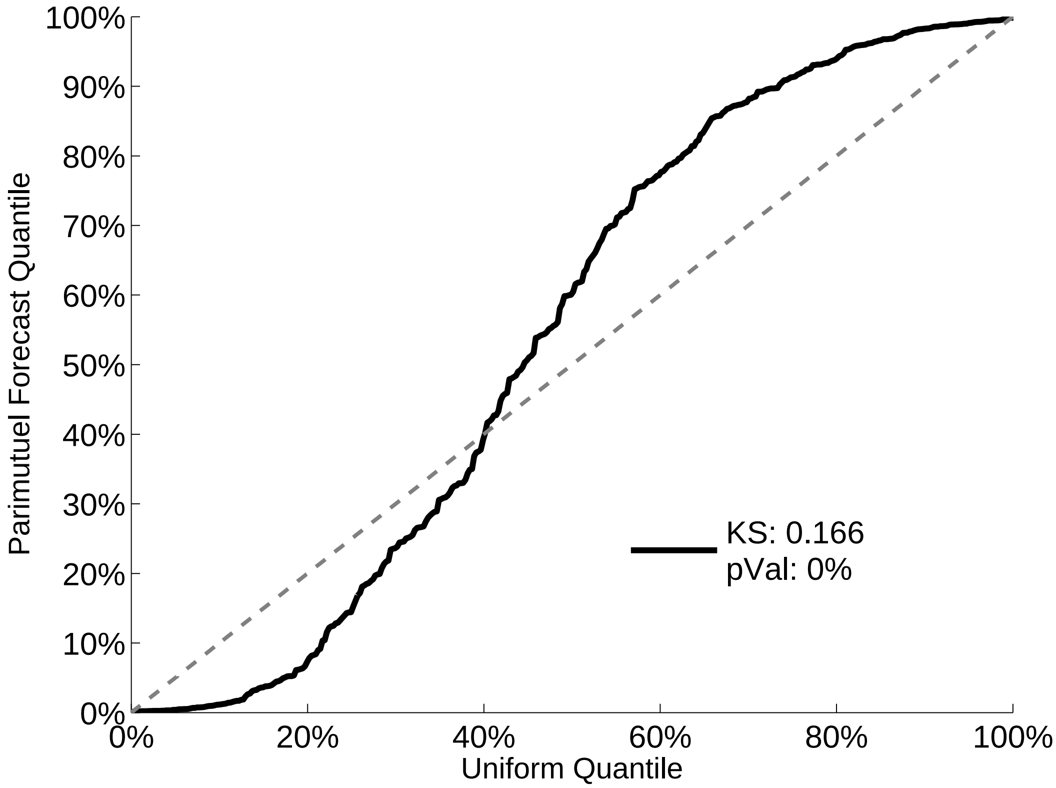


Figure 1: Quantile Plots for the Information Aggregation Mechanism

This figure presents the distribution of realized quantiles from the information aggregation mechanism in equation 4 against the theoretically accurate uniform distribution. The KS p-Value is reports the result for a Kolmogorov-Smirnov test of equality of the distributions.

[11]Since the IAM quantiles are only available for a discretized support, these quantiles are technically only identified within a range. As defined above, our reported results treat the probability mass in a bucket as lying entirely on the minimum of that bucket. However, our qualitative results are not sensitive to this treatment.

[12]While robust to heterogeneity, note that various features of our data, especially the panel structure coupled with multiple horizons, induce correlation across draws. As such, the p-Values of the Kolmogorov-Smirnov test are likely to be distorted with a downward bias. When the number of degrees of freedom for the Kolmogorov-Smirnov test is reduced by a factor of 4, the p-Value increases to 6.7%. While this correction is not valid, it does indicate a lack of robutstness for the results rejecting rational expectations. Unfortunately, allowing for correlated sampling structures in the Kolmogorov-Smirnov test is an intractable problem beyond our scope.

Figure 1 presents the empirical distribution for $\hat{U}_{i,t,h}$ plotted against the uniform distribution quantiles pooled across all products, periods, and forecast horizons. Visual inspection indicates two features the mechanism's distribution appears to distort in the realized distribution over outcomes. First, the S-shape of the graph indicates that the beliefs from the mechanism understate the likelihood of extreme outcomes. Second, the quantiles for the mechanism's distribution appear slightly to the left of the uniform, indicating the IAM distribution is a bit conservative at the median. Because of these two distortions, the Kolmogorov-Smirnoff test rejects the null hypothesis that the empirical distribution of outcomes matches the IAM's forecasted distribution for outcomes.

Given that we observe conditional forecast distributions at different horizons, we can empirically evaluate whether the forecast distribution accuracy improves as the forecast horizon shortens. To that end, instead of pooling the $\hat{U}_{i,t,h}$ across horizons, we can perform the Kolmogorov Smirnov test from equation 4 for each horizon sub-sample.[13] Figure 2 plots the quantiles of the forecast quantile distribution against the uniform distribution at horizons 1, 3, 5, 7, and 9. Notably, the plots appear to approach the 45 degree line as the horizon drops, as indicated by the pattern of Kolmogorov-Smirnov test statistics and p-Values. Hence, the reverse favorite-longshot bias described previously disappears as the forecasting horizon shrinks.

Our finding that, at least for long horizons, the beliefs elicited from the mechanism systematically understate tail probability events, is related to phenomena which have been much studied in the literature on betting markets. Specifically, the "favorite-longshot bias (FLB)" is an oft-reported empirical property in studies on betting markets.[14] In data patterns characterized by the FLB, the parimutuel odds on high probability events understate the realized probabilities (e.g., the odds on a horse "favored" to win the race understates the true odds of that horse winning). By contrast, we find a "reverse favorite-longshot bias," in which elicited beliefs understate the realized probabilities for *low-probability* (tail) events.

A number of explanations have been proposed for the FLB, when observed in betting environments, including probabilistic misperceptions, risk preferences, belief heterogeneity,

[13] Considering sub-samples also helps mitigate the effect of correlation in the sampling process for $\hat{U}_{i,t,h}$. In particular, one would expect a significant correlation in the conditional quantiles across horizons for a given product and period. Performing the test horizon-by-horizon prevents this distortion from affecting inference at the cost of weaker power due to fewer degrees of freedom. However, it doesn't fully control for the temporal autocorrelation in $\hat{U}_{i,t,h}$ at fixed horizon and across products for a fixed time period.

[14] There are four full chapters dedicated to its review alone in the *Handbook of Sports and Lottery Markets* (Hausch and Ziemba (2008a)).

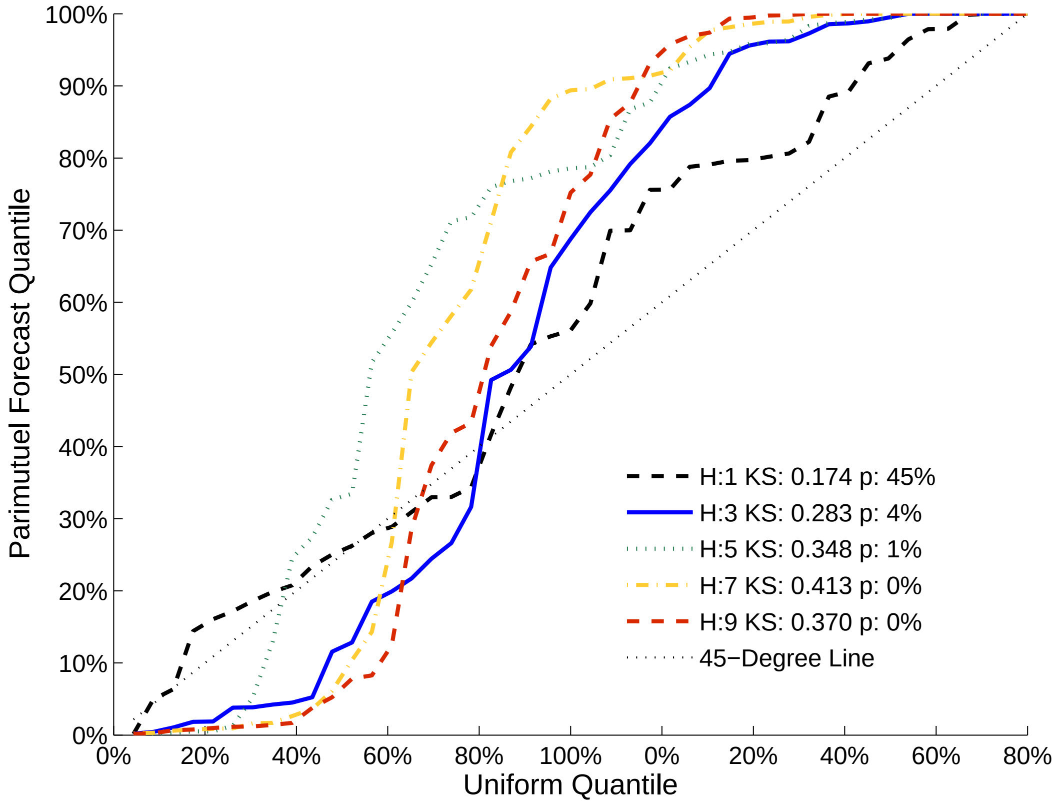


Figure 2: Quantile Plots for the Information Aggregation Mechanism by Horizon
This figure presents the distribution of realized quantiles from the information aggregation mechanism against the theoretical uniform distribution for horizon subsamples of the IAM. The KS p-Value reports the result for a Kolmogorov-Smirnov test of equality of the distributions.

and information incentives.[15] Experimental work has demonstrated that the distribution of tickets in a parimutual-type information aggregation mechanism is related to disequilibrium purchase of tickets.

The fact that the reverse-FLB disappears as the forecast horizon narrows (as in Figure 2) casts doubt on the risk explanation, because these preferences should not change with the forecast horizon. If risk aversion is driving the reverse-FLB, then the extent of the reverse-FLB would be primarily driven by the likelihood of the outcome itself. Importantly, the bias is driven by the likelihood of the event, rather than the uncertainty in the forecast distribution. That is, risk aversion would cause the likelihood of low-probability tail events

[15]See Ali (1977) for an early reference. Snowberg and Wolfers (2010) compare the relative likelihood of risk preferences and probabilistic misperception in betting markets, finding probabilistic misperceptions to be relatively more likely, but are silent on the role of strategic considerations. Gandhi and Serrano-Padial (2012) show that belief heterogeneity among racetrack bettors can also induce a longshot bias in prices.

to be understated in the forecast distribution, regardless of the forecast horizon – this is not what we see.

Ottaviani and Sørensen (2010) propose an alternative model of biases in parimutuel betting systems based on strategic betting rather than misperceptions or risk preferences. In this model, parimutuel betting participants are partially informed about the conditional distribution for a random variable in addition to aggregate uncertainty about the outcome itself.[16] Rational behavior in their model allows for a FLB or a reverse-FLB in the aggregated distribution depending on the ratio of privately-held information to noise in the forecast variable. Specifically, the reverse-FLB arises when information is very diffuse. To see the intuition from a mechanical perspective, consider the case of Lotto, a uniformly random parimutuel system. Since each number has an equal probability of being a winning number, any "favorites" which arise during the betting process must underpay, and "longshots" must overpay: that is, a systematic underweighting of low-probability events arises in the parimutuel odds.

This explanation implies that at long horizons, the noisiness of the uncertainty could itself lead to a reverse-FLB, which goes away as the horizon narrows and the quality of information about the conditional distribution improves, as we see in Figure 2. This explanation is further corroborated by limited analyst-level data which was available to us, in which tickets are placed later in long-horizon IAM's than in shorter-horizon mechanisms (an implication of information-driven reverse-FLB). For example, the average ticket placement time in the nine-month horizon mechanisms is 25% later than in the one-month horizon mechanisms. Further, while an average of 92% of tickets are placed before the first price increase at the one-month horizon, only 85% of tickets are placed before the first price increase at the nine-month horizon. This costly delay indicates participants' uncertainty increases their willingness to pay higher ticket costs in order to learn the information revealed within the IAM by earlier participants.

[16]Technically, the model analyzed by Ottaviani and Sørensen (2010) considers a simultaneous move game that differs from the dynamic IAM here with its increasing prices. Despite this tension in the direct application of Ottaviani and Sørensen (2010)s analysis, it seems reasonable to apply the intuition to the current setting.

5 Forecasting using the IAM

Having used distributional tests to evaluate the IAM's ability to aggregate information within the firm, we now examine how the mechanism can improve internal forecasts and decision-making. Many decisions, such as setting Wall Street expectations or sales incentive quotas, require a point forecast for unit sales. In these settings, we are less interested in the full conditional distribution of sales than in a central measure of the distribution. We now consider a horse-race between the IAM and the official forecast to evaluate their relative predictive accuracy and information content. In addition to establishing the organizational value of the IAM, these forecast tests allow us to evaluate the IAM in contrast to a concurrent, directly measured point expectation formed within the organization rather than solely relative to ex-post realized outcomes.

5.1 Forecast Data

We observe data from 2007 through 2010 across five major product lines. These forecasts are made in a dynamic business environment where expectations rapidly shifted from growth to stagnation. With the Lehman brothers crash occurring half-way through the sample, the forecasting exercise is particularly challenging at longer horizons. For proprietary reasons, Intel has requested we mask the actual values of units sold as well as the names of the products themselves. As all of our comparative analyses are insensitive to the numeraire, this masking has no effect on the results while allowing us to make our data publicly available.

To consolidate the distributional information from the mechanism into point forecasts, we propose two natural point-forecast estimates. The Mean forecast is taken as the expected value of the outcome under the forecasting distribution. The Mode forecast is taken as the outcome with maximal probability under the forecasting distribution. Due to the buckets in the mechanism, these forecasts are effectively interval-censored. To address this censoring, we take the mid-point of the bucket as representing the value for all forecast mass placed within that bucket.[17] Recalling the definition of actual unit sales for product line i in quarter t by $Y_{i,t}$, we refer to the official, mean, and mode forecasts at horizons $h \in \{1, 2, \ldots, 0\}$ months by $\hat{Y}^{(\text{Official})}_{i,t|t-h}$, $\hat{Y}^{(\text{Mean})}_{i,t|t-h}$, and $\hat{Y}^{(\text{Mode})}_{i,t|t-h}$ respectively. For each $k \in \{\text{Official, Mean, Mode}\}$, we can

[17] The first and last buckets representing ranges $[0, x_1)$ and $[x_{K-1}, \infty)$ are assigned values x_1, and x_{K-1}, respectively. While setting the label for the last bucket to ∞ would clearly be problematic for our results, we have considered several different specifications for these buckets, with little impact on our results from reasonable treatments.

then define the forecast error as $e^{(k)}_{i,t|t-h}$ from which we can compute forecast performance statistics such as the Root Mean Square Forecast Error (RMSFE).[18]

Table 1 reports summary statistics for the point forecasts and actual unit sales, including a break down by product lines. Realized quarterly sales, averaging 2.95 glorbs in the full sample, have a standard deviation over 1 glorb, indicating a highly variable forecasting environment. On average, the forecasts slightly overstate average sales, with the Mean and Mode average forecast improving upon the Official forecast bias by less than 0.01 units. Still, the IAM mean and mode forecasts deliver a root mean square error almost 10% lower than that of the the Official forecast. Within individual product lines, the RMSFE improvement is even more significant, ranging from 21% to 37%. By comparing the RMSFE with the standard deviation of sales within each product line, though, we see the forecasts in general perform quite poorly relative to the ex-post sample average.

Direct comparison of the forecasts is somewhat complicated by the timing of the IAM vis-a-vis the official forecasts. Since the mechanism takes place with a significant lag after the official forecast is released, the horizon h official forecast is most comparable to the horizon $h+1$ IAM forecast. Operationally, this implementation allows the analysts preparing the official forecast to observe the mechanism's distribution before announcing their forecast. When the last (most inaccurate) horizon is dropped from the official forecast and the nearest (most accurate) horizon is dropped from the IAM forecasts, the differential RMSFE is negligible in aggregate.

While there is some apparent heterogeneity across product lines, the magnitudes of realized unit sales are of the same order. Notably, though, all the tests we use are robust to product-line heterogeneity as we apply the appropriate clustering for the panel structure in computing standard errors. This robustness allows us to enhance the power of our tests by pooling the data across product lines, which is particularly helpful given the limited sample size we have available. In our statistical tests, we cluster our observations to correct for any autocorrelation introduced by the pooling and also analyze subsamples to ensure robustness of the findings from the pooled sample.

[18]From an organizational perspective, Intel uses the IAM as an input to its official forecast. As such, the horizon h IAM forecast is released shortly before the horizon $h-1$ official forecast, requiring an adjustment for the horizons when comparing forecasts.

Table 1: Summary Statistics for Forecast Data

This table presents summary statistics characterizing the average and standard deviation of unit sales, the target variable to be forecast, and the different forecasts available to decision makers. Due to variation in timing of the horizon forecasts relative to realized sales, the full-sample Root Mean Square Forecast Error is not directly comparable between the official and IAM forecasts. As such, the official RMSFE is also reported excluding the least-informed (9 month) horizon forecast and the IAM RMSFE is also reported excluding the most-informed (1 month) horizon forecast. The columns report results broken down by the product line for which forecasts are obtained to characterize heterogeneity.

	Full	Business Product Line				
	Sample	Disti 1	Disti 2	Direct 1	Direct 2	Direct 3
First Sales Period Quarter	2007Q1	2008Q1	2007Q2	2008Q3	2007Q1	2009Q1
Last Sales Period Quarter	2010Q1	2010Q1	2010Q1	2010Q1	2010Q1	2010Q1
Number of Forecast-Period-Product Obs	414	81	108	63	117	45
Number of Sales-Product Periods	46	9	12	7	13	5
Number of Forecasts per Product-Period	9	9	9	9	9	9
Average Unit Sales	2.95	1.94	1.97	3.77	3.69	4.00
Std Dev of Unit Sales	1.02	0.21	0.16	0.36	0.74	0.56
Average Official Forecast	2.99	2.01	2.00	3.84	3.71	4.07
Std Dev of Official Forecasts	1.03	0.18	0.18	0.43	0.79	0.39
Official Root Mean Square Fcst Error	0.48	0.26	0.22	0.54	0.63	0.65
Official RMSE excluding 9th Horizon	0.45	0.25	0.20	0.51	0.60	0.60
Average Parimutuel Mean Forecast	2.98	1.99	1.98	3.88	3.71	4.03
Std Dev of Parimutuel Mean Forecasts	1.04	0.18	0.16	0.41	0.78	0.39
Parimutuel Mean Root Mean Square Fcst Error	0.43	0.25	0.18	0.51	0.58	0.54
Mean RMSE excluding 1st Horizon	0.46	0.26	0.19	0.54	0.61	0.57
Average Parimutuel Mode Forecast	2.98	1.99	1.98	3.89	3.70	4.04
Std Dev of Parimutuel Mode Forecasts	1.04	0.19	0.16	0.42	0.78	0.41
Parimutuel Mode Root Mean Square Fcst Error	0.43	0.26	0.18	0.52	0.59	0.52
Mean RMSE excluding 1st Horizon	0.46	0.27	0.19	0.55	0.62	0.55

5.2 Predictive Accuracy with Quadratic Loss

In directly comparing the loss due to forecast errors, we find the IAM forecasts outperform the official forecast under quadratic loss in those settings where high-quality individual information may be dispersedly held. For example, the IAM performs particularly well at short forecast horizons and in direct distribution channels where internal participants are likely to have useful information for the forecasting exercise. In contrast, at long horizons and in channel distributor based product lines, the official forecast based on more structured models proves to be more informative.

Diebold and Mariano (1995), henceforth DM, tests provide the benchmark for directly comparing the predictive accuracy of two forecasts under a variety of possible loss functions. Here we simply define forecast loss as the square error of the forecast:

$$l^{(k)}_{i,t|t-h} = \left(\hat{Y}^{(k)}_{i,t|t-h} - Y_{i,t}\right)^2 \tag{5}$$

We can then compare the loss between two corresponding forecasts j and k:

$$\delta_{i,t,h}^{(j,k)} = l_{i,t|t-h}^{(j)} - l_{i,t|t-h}^{(k)} \tag{6}$$

The DM test statistic corresponds to the t-statistic for the average $\delta_{i,t,h}^{(j,k)}$, using a robust estimator of the variance allowing for auto-correlation of loss differentials within product lines and clustering for each revenue period. While DM's initial derivation of the test establishes its asymptotically normality, Harvey et al. (1997) show that Student's t distribution better controls for size.

We present the results of these tests in Table 2, comparing the loss of the horizon h IAM forecast with the horizon $h-1$ official forecast. Given the differential timing of the official forecast release and the mechanism, this treatment cedes a slight information advantage to the official forecast, which is always released *after* the mechanism has concluded. Despite this informational advantage, the full sample results indicate the point forecasts taken from the mechanism nonetheless deliver lower square loss than the official forecasts. This outperformance is especially surprising given that the offical forecasters know the IAM distribution *before* releasing their forecast. That is, the analysts' deviation from the IAM forecast actually worsens the forecast error.

Using DM tests to evaluate the mechanism's performance in subsamples, we find the IAM perform particularly well in exactly those contexts where individuals within the organization are likely to have disparate information. Conversely, the official forecast performs well in those settings that are conducive to high-level modeling and analysis based on historical and macroeconomic trends.

The root mean forecast improvement is monotonically declining in forecast horizon, with the official forecast outperforming the information aggregation mechanism at the 8 and 9 month horizons. The official forecast is also more likely to incorporate higher-level information regarding macroeconomic conditions that is especially relevant to long-horizon forecasting. Given that individual sales representatives and partner managers are unlikely to have materially relevant information about unit sales at these horizons, the official forecast's outperformance is to be expected here.

Similarly, the official forecast outperforms the mechanism's expectation in products distributed through channel sales operations. These distribution networks correspond to smaller-sized transactions and, as such, are driven primarily by trends in consumer spending best analyzed in aggregate. Further, these sales operations generate only about 10% of total

Table 2: Comparing Forecast Loss Across Mechanisms

This table presents Diebold-Mariano tests comparing the point forecasts from the official forecast and the mean and mode information aggregation mechanism forecast. Panel A uses the full sample of all forecasts and horizons, with Panels B and C reporting results for horizon and product subsamples. The Root Mean Δ Square Error reports the square root of the absolute average difference in the square error for the official and IAM forecasts, signed negatively for cases where the official forecast outperforms the IAM. The Outperformance Frequency captures the frequency with which the IAM forecast was more accurate than the official forecast. The DM-Statistic and p-Value report the Diebold-Mariano test statistic and p-Value using standard errors robust to autocorrelation up to the maximum horizon included in the sample, clustered by period and product.

Panel A: Pairwise Diebold Mariano Tests

	Mean Forecast				Mode Forecast				
	Root Mean Δ Square Error	Outperf Freq	DM Stat	p-Val	Root Mean Δ Square Error	Outperf Freq	DM Stat	p-Val	Num of Obs
Overall	0.45	61%	(2.20)*	3%	0.45	60%	(2.13)*	3%	368

Panel B: Pairwise Diebold Mariano Tests by Horizon

	Mean Forecast				Mode Forecast				
Horizon	Root Mean Δ Square Error	Outperf Freq	DM Stat	p-Val	Root Mean Δ Square Error	Outperf Freq	DM Stat	p-Val	Num of Obs
2 Months	0.69	78%	(2.18)*	3%	0.70	78%	(2.19)*	3%	46
3 Months	0.66	80%	(1.90)	6%	0.66	76%	(1.89)	6%	
4 Months	0.63	80%	(1.68)	10%	0.63	76%	(1.68)	10%	
5 Months	0.49	70%	(1.57)	12%	0.49	63%	(1.57)	12%	
6 Months	0.42	54%	(1.42)	16%	0.43	61%	(1.44)	16%	
7 Months	0.19	48%	(0.55)	59%	0.10	50%	(0.16)	88%	
8 Months	(0.22)	39%	1.12	27%	(0.22)	41%	1.24	22%	
9 Months	(0.25)	35%	1.48	15%	(0.27)	35%	1.67	10%	

Panel C: Pairwise Diebold Mariano Tests by Product Line

	Mean Forecast				Mode Forecast				
Product Line	Root Mean Δ Square Error	Outperf Freq	DM Stat	p-Val	Root Mean Δ Square Error	Outperf Freq	DM Stat	p-Val	Num of Obs
Disti 1	(0.21)	36%	1.03	31%	(0.22)	35%	1.07	29%	72
Disti 2	(0.07)	50%	0.47	64%	(0.07)	50%	0.47	64%	96
Direct 1	0.70	75%	(1.56)	12%	0.69	77%	(1.54)	13%	56
Direct 2	0.63	75%	(1.65)	10%	0.62	71%	(1.54)	13%	104
Direct 3	0.54	73%	(1.02)	32%	0.57	78%	(1.07)	29%	40

corporate profits. In contrast, for the direct channels that generate about 90% of firm profits, the IAM performs exemplary at aggregating the information held by individual agents.

On the whole, this analysis indicates the information aggregation mechanism performs well at forecasting in exactly those settings that economic theory indicates its usefulness. Overall, the mechanism outperforms the official forecast and it does so by dramatically reducing forecast loss in those settings where individually held information is likely to be particularly valuable.

5.3 Forecast Combination and Encompassing Tests

Instead of just choosing one forecast, we may wish to combine the information from the mechanism with the official forecasts into a single aggregated forecast. A robust literature considers optimal forecast combination, with the survey by Timmermann (2006) providing a good entry point. For the aggregated forecast to improve upon the IAM forecast, however, there must be some information in the official forecast that isn't already incorporated into the mechanism's expectations. As such, we'd like a test to make sure the mechanism's forecast doesn't encompass the official forecast before introducing it into a forecast combination exercise.

We follow the approach of Fair and Shiller (1990) in applying a regression-based test to evaluate the encompassing properties of the two forecasts. It's straightforward to show that the optimal weights with which to form a linear combination of forecasts can be calculated using the following regression.

$$Y_{i,t} = \gamma + \omega_{IAM}\hat{Y}^{(IAM)}_{i,t|t-h} + \omega_{Official}\hat{Y}^{(Official)}_{i,t|t-h} \tag{7}$$

In this context, one way to evaluate whether the official forecast encompasses the Mean forecast is to test the null hypothesis that $\gamma = 0$, $\omega_{IAM} = 0$, and $\omega_{Official} = 1$. If we reject this null hypothesis using an F-test, then we can be confident that the mechanism's forecast has additional information beyond that which is contained in the Official forecast. Similarly, if we reject the null hypothesis that $\gamma = 0$, $\omega_{IAM} = 1$, and $\omega_{Official} = 0$ then we will be able to say that the Official forecast contains information beyond that which is available from the mechanism's point forecast. Slightly less restrictive forms of these tests can also be formulated by dropping the condition on γ and still simpler tests can just evaluate the one-sided hypotheses that $\omega_{Official} \geq 0$ and $\omega_{IAM} \geq 0$.[19]

Table 3 summarizes the results of these regressions for the Mean and Mode forecast. These regressions show that the optimal combination of these forecasts negatively weight the official forecast. The F-Tests indicate that we can reject the null hypothesis that the

[19] These regression-based tests have been further generalized by Harvey et al. (1998), whose approach mirrors that of Diebold and Mariano (1995). We only present the results using the Fair and Shiller (1990) test here for brevity and ease of interpretation in relation to forecast combinations. In addition to Harvey et al. (1998), there are many other diagnostic tests we could apply to these two forecasting mechanisms. A full-blown forecast evaluation exercise is beyond the scope of the current paper, as it would distract from the focusing on the performance of the IAM itself. In the interests of evaluating multiple-horizon internal forecasting in dynamic panels, we present a more thorough analysis of forecasts themselves in the complementary paper Gillen et al. (2013).

Table 3: Forecast Combination Regressions

This table presents estimates from the forecast combination regressions 7. Panel A uses the full sample of all forecasts and horizons, with Panels B and C reporting results for horizon and product subsamples. The test F$(0, 1, 0)$ tests the hypothesis that $\alpha = 0$, $\omega_{IAM} = 1$, and $\omega_{Official} = 0$, similarly, F$(0, 0, 1)$ tests $\alpha = 0$, $\omega_{IAM} = 0$, and $\omega_{Official} =$. The tests F$(., 0, 1)$ and F$(., 0, 1)$ test the analogous restrictions without the zero-intercept condition. All tests use standard errors robust to autocorrelation up to the maximum horizon included in the sample, clustered by period and product.

			Panel A: Full-Sample Regressions					
		Parimutuel	Official					
	Intercept	Forecast	Forecast	F(0, 1, 0)	F(0, 0, 1)	F(., 1, 0)	F(., 0, 1)	
	St Dev	St Dev	St Dev	p-Value	p-Value	p-Value	p-Value	
Mean Fcst	0.26	0.76	0.14	12.80	144.82	17.63	215.78	368.00
	0.16	0.14	0.08	0%	0%	0%	0%	
Mode Fcst	0.27	0.75	0.15	14.08	145.06	19.67	216.13	368.00
	0.15	0.14	0.08	0%	0%	0%	0%	
			Panel B: Mean Forecast by Horizon					
		Parimutuel	Official					
Horizon	Intercept	Forecast	Forecast	F(0, 1, 0)	F(0, 0, 1)	F(., 1, 0)	F(., 0, 1)	
	St Dev	St Dev	St Dev	p-Value	p-Value	p-Value	p-Value	
h = 1	0.06	0.95	0.03	0.40	143.88	0.57	215.82	46.00
	0.11	0.04	0.02	75%	0%	57%	0%	
h = 2	0.10	0.90	0.07	0.68	73.79	1.01	110.67	
	0.12	0.07	0.04	57%	0%	37%	0%	
h = 3	0.12	0.88	0.08	0.68	45.96	1.01	68.93	
	0.18	0.12	0.05	57%	0%	37%	0%	
h = 4	0.22	0.85	0.07	0.74	20.01	1.06	29.47	
	0.24	0.21	0.11	54%	0%	36%	0%	
h = 5	0.29	0.71	0.18	1.52	12.52	2.16	18.35	
	0.28	0.26	0.15	22%	0%	13%	0%	
h = 6	0.37	0.48	0.38	3.11	5.29	4.33	7.85	
	0.31	0.26	0.13	4%	0%	2%	0%	
h = 7	0.48	0.07	0.75	6.38	3.25	8.75	4.39	
	0.31	0.30	0.25	0%	3%	0%	2%	
h = 8	0.54	(0.03)	0.82	6.48	3.06	8.75	4.23	
	0.34	0.25	0.17	0%	4%	0%	2%	
			Panel C: Mean Forecast by Product					
		Parimutuel	Official					
Product	Intercept	Forecast	Forecast	F(0, 1, 0)	F(0, 0, 1)	F(., 1, 0)	F(., 0, 1)	
	St Dev	St Dev	St Dev	p-Value	p-Value	p-Value	p-Value	
Disti 1	1.05	(0.23)	0.70	72.74	12.00	104.04	17.68	72
	0.13	0.10	0.13	0%	0%	0%	0%	
Disti 2	1.15	0.06	0.36	22.85	15.01	33.98	20.69	96
	0.29	0.26	0.18	0%	0%	0%	0%	
Direct 1	4.53	0.09	(0.29)	35.70	125.75	49.72	188.61	56
	0.28	0.09	0.06	0%	0%	0%	0%	
Direct 2	1.34	0.66	(0.02)	8.96	53.23	13.35	79.57	104
	0.80	0.23	0.09	0%	0%	0%	0%	
Direct 3	2.59	0.41	(0.06)	1.92	14.50	2.82	13.69	40
	3.34	0.47	0.49	14%	0%	7%	0%	

IAM forecast encompasses the official forecast, as the Official forecast contains information that can be used to moderate errors in the IAM forecast estimates. Indeed, even considering just the coefficient on the Official forecast, we can reject the minimally restrictive hypothesis that $\omega_{Official} \geq 0$.

The subsample results for Table 3 in panels B and C focus on just combining the mechanism's Mean forecast with the Official forecast at each horizon, though the results are similar although slightly weaker for the Mode forecast. The evidence in these subsamples is broadly consistent with the full-sample results. The significance of the results drops somewhat in the horizon-based subsamples, particularly at the longer horizons.

6 Conclusion

In this paper, we introduce a new information aggregation mechanism to a novel field application forecasting practical business information needs. This implementation provides a testbed for evaluating the theory, refining the practice of information aggregation, and exploring the intuition extracted from laboratory work. We perform these tests relative to realized outcomes in the field as well as measured conditional expectations held by informed decision makers. We find these mechanisms effectively reflect the uncertainty of forward business indicators such as future sales and yield forecasts that improve upon those generated by internal processes.

Further, by providing decision makers with a richer characterization of operational uncertainty, the mechanism can help them address problems with potentially asymmetric forecast loss and better control risks. As such, it is not surprising that Intel has extended its implementation beyond forecasting sales volumes but also to applications in evaluating research and development expenditures and tracking project management.

Field applications of information aggregation mechanisms also provide economists with valuable evidence for evaluating competing economic theories outside of tightly controlled laboratory settings. Our observed link between biases in the IAM and the forecast horizon provides suggestive evidence that information plays some role as a causal feature in affecting these biases. However, our findings here are certainly not conclusive, as many different causes could be postulated for the observed biases of the IAM in different circumstances. As such, a more narrow study would be required to provide more convincing evidence of these effects.

Our ability to develop the IAM as an accurate mechanism for measuring consensus probability beliefs in the absence of a unified theory provides an additional example of a well-established application for the experimental economics methodology. In an environment where theorists are confronted with an intractable problem, careful and principled experimentation can guide development of economic mechanisms that perform well despite their lack of clear theoretical foundation. The IAM's successful implementation here further confirms the broad experimental evidence that such a mechanism can effectively aggregate information. A long process of testing and refining information aggregation mechanisms in the lab has borne real-world validation and value.

References

Mukhtar M. Ali. Probability and utility estimates for racetrack bettors. *Journal of Political Economy*, 85:803–815, 1977.

Beth Allen and James Jordan. The existence of rational expectation equilibrium: a retrospective. In M. Majumdar, editor, *Organization with Incomplete Information: Essays in Economic Analysis*. Cambridge University Press, 1998.

Boris S. Axelrod, Ben J. Kulick, Charles R. Plott, and Kevin A. Roust. The design of improved parimutuel-type information aggregation mechanisms: Inaccuracies and the long-shot bias as disequilibrium phenomena. *Journal of Economic Behavior and Organization*, 69(2):170 – 181, 2009.

Joyce Berg, Robert Forsythe, Forrest Nelson, and Thomas Rietz. *Results from a Dozen Years of Election Futures Markets Research*, volume 1 of *Handbook of Experimental Economics Results*, chapter 80, pages 742–751. Elsevier, 2008.

Moez Bernnouri, Henner Gimpel, and Jacques Robert. Measuring the impact of information aggregation mechanisms: An experimental investigation. *Journal of Economic Behavior and Organization*, 78:302–318, 2011.

Kay-Yut Chen and Charles R. Plott. Information aggregation mechanisms: Concept, design, and implementation for a sales forecasting problem. *Mimeo*, 2002.

P. Chiappori, A. Gandhi, B. Salanie, and F. Salanie. Identifying preferences under risk from discrete choices. *The American Economic Review*, 99:356–362, 2009.

B. Cowgill, J. Wolfers, and E. Zitzewitz. Using prediction markets to track information flows: Evidence from Google. Working paper, 2009.

Francis X. Diebold and Roberto S. Mariano. Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263, 1995.

Ray C Fair and Robert J Shiller. Comparing information in forecasts from econometric models. *American Economic Review*, 80(3):375–89, 1990.

Robert Forsythe, Forrest Nelson, George R. Neumann, and Jack Wright. Anatomy of an experimental political stock market. *The American Economic Review*, pages 1142–1161, 1992.

A. Gandhi and R. Serrano-Padial. Are beliefs heterogeneous? the case of the longshot bias. working paper, 2012.

Ben Gillen, Charles Plott, and Matthew Shum. Bottom-up and top-down judgement forecasting at intel. working paper, 2013.

Philip K Gray and Stephen F Gray. Testing market efficiency: Evidence from the nfl sports betting market. *The Journal of Finance*, 52:1725–1737, 1997.

David Harvey, Stephen Leybourne, and Paul Newbold. Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13(2):281–291, 1997.

David I. Harvey, Stephen J. Leybourne, and Paul Newbold. Tests for forecast encompassing. *Journal of Business & Economic Statistics*, 16(2):254–259, 1998. ISSN 07350015.

Donald B Hausch and William T. Ziemba. *Handbook of Sports and Lottery Markets*. North-Holland, 2008a.

B. Jullien and B. Salanie. Estimating preferences under risk: The case of racetrack bettors. *Journal of Political Economy*, 108:503–530, 2000.

Katarina Kalovcova and Andreas Ortmann. Understanding the plott-wit-yang paradox. CERGE-EI Working Papers wp397, The Center for Economic Research and Graduate Education - Economic Institute, Prague, 2009.

Frederic Koessler, Charles Noussair, and Anthony Ziegelmeyer. Information aggregation and belief elicitation in experimental parimutuel betting markets. *Journal of Economic Behavior and Organization*, 83:195–208, 2012.

C. Manski. Interpreting the predictions of prediction markets. *Economics Letters*, 91:425–429, 2006.

Marco Ottaviani and Peter Norman Sørensen. The timing of bets and the favorite-longshot bias. FRU Working Papers 2004/12, University of Copenhagen. Department of Economics. Finance Research Unit, 2004.

Marco Ottaviani and Peter Norman Sørensen. Surprised by the parimutuel odds? *American Economic Review*, 99(5):2129–34, 2009.

Marco Ottaviani and Peter Norman Sørensen. Noise, information, and the favorite-longshot bias in parimutuel predictions. *American Economic Journal: Microeconomics*, 2(1):58–85, 2010.

C. R. Plott and K.A. Roust. The design of information aggregation mechanisms: A two-stage parimutuel market to avoid mirages bubbles. *Mimeo*, 2009.

Charles R. Plott. Markets as information gathering tools. *Southern Economic Journal*, 67 (1):1–15, 2000. ISSN 00384038.

Charles R. Plott and Shyam Sunder. Efficiency of experimental security markets with insider information: An application of rational-expectations models. *Journal of Political Economy*, 90(4):663–698, 1982. ISSN 00223808.

Charles R. Plott and Shyam Sunder. Rational expectations and the aggregation of diverse information in laboratory security markets. *Econometrica*, 56(5):1085–1118, 1988. ISSN 00129682.

Charles R. Plott, Jorgen Wit, and Winston C. Yang. Parimutuel betting markets as information aggregation devices: experimental results. *Economic Theory*, 22(2):311–351, 2003.

M. Rothschild. Searching for the lowest price when the distribution of prices is unknown. *Journal of Political Economy*, 82:689–711, 1974.

Erik Snowberg and Justin Wolfers. Explaining the favorite-long shot bias: Is it risk-love or misperceptions? *Journal of Political Economy*, 118(4):723–746, 2010.

Allan Timmermann. Forecast combinations. In C.W.J. Granger G. Elliott and A. Timmermann, editors, *Handbook of Economic Forecasting*, volume 1 of *Handbook of Economic Forecasting*, chapter 4, pages 135 – 196. Elsevier, 2006.

J. Wit. *Dynamics and Information Aggregation in Elections and Markets (Doctoral Thesis)*. Tinbergen Institute Research Series. Thesis Publ., 1997. URL http://books.google.com/books?id=X-grywAACAAJ.

J. Wolfers and E. Zitzewitz. Prediction markets in theory and practice. IZA Discussion Paper #1991, 2006.

Linda M Woodland and Bill M Woodland. Market efficiency and the favorite-longshot bias: The baseball betting market. *Journal of Finance*, 49:269–279, 1994.

Appendix A. Information Aggregation in The Conditional Normal Model

The approach to modeling agent information regarding the distribution of Y is somewhat at odds with common formulations of learning presented in asset market experiments. This example also illustrates the definition of our conditional information set as containing all information available to participants about the distribution of Y. We also use this example to illustrate the importance of coordinating behavior.

A.1. Information and Beliefs

In this example, consider learning about a random variable drawn from a normal distribution with unknown mean. In particular, suppose:

$$Y|\mu \sim N(\mu, 1)$$

with a diffuse prior for μ. Suppose that players 1 and 2, respectively, receive conditionally independent signals:

$$X_1|\mu \sim N(\mu, 2), \text{ and, } X_2|\mu \sim N(\mu, 2)$$

Player 1's expected posterior is $Y|X_1 \sim N(X_1, 3)$ and Player 2's expected posterior is $Y|X_2 \sim N(X_2, 3)$). Taking player 1 as an example, these distributions marginalize over the uncertainty in μ in the joint probability space for (Y, X_1, μ). In particular, player 1 knows that $Y|\mu \sim N(\mu, 1)$ and $\mu|X_1 \sim N(X_1, 2)$, so the probability that player 1 assigns to the event $\{Y \sim N(z, 1)\}$ comes from a $N\left([X_1 z], 2\right)$ pdf.

The key feature to note from this example is that the location of the ideally aggregated conditional distribution depends on the realized signals themselves because we don't know μ. In an IAM steady state with information aggregation, we will observe tickets distributed proportionally to a $N(0.5(X_1 + X_2), 2)$ distribution which will, on average, match a $N(\mu, 2)$ distribution. However, because of the sampling error in X_1 and X_2, a $N(\mu, 2)$ distribution is the ideal target for the IAM to deliver in expectation. Importantly, the IAM shouldn't deliver a distribution with unit variance because the IAM doesn't know μ and incorporates that uncertainty in the mean in characterizing its posterior. This case corresponds to a situation in which participants have very limited information about the distribution of Y.

If, instead, players receive the signals:

$$X_1|\mu \sim N(\mu, 0.0001), \text{ and, } X_2|\mu \sim N(\mu, 0.0001)$$

Player 1s posterior distribution for $Y|X_1 \sim N(X_1, 1.0001)$ and Player 2s posterior distribution for $Y|X_2 \sim N(X_2, 1.0001)$. The aggregated distribution is $Y|X_1, X_2 \sim N(0.5 * (X_1 + X_2), 1.00005)$. In this case, $0.5(X_1 + X_2)$ is going to be very close to μ and only the aggregate uncertainty in the variance will persist. This case corresponds to a situation in which participants have very good information about the distribution of Y.

A.2. Behavioral Strategies and Information Aggregation Failures

The Normal model presents a more challenging environment for information aggregation. For example, the aggregated distribution over Y must have a normal distribution. However, returning to the setting where players' signals about the mean have variance 2, suppose players simply allocate their tickets proportionally to the posterior distributions implied by their private information. In this case, $\nu_1 \sim N(X_1, 3)$ and $\nu_2 \sim N(X_2, 3)$. Since $\eta = \nu_1 + \nu_2$, the IAM would deliver a mixed-normal forecast distribution since the players, clearly failing to satisfy information aggregation about the full distribution.

Suppose alternatively that players move sequentially, allocate all their tickets in one move, and that player 1 first allocates his tickets according to his posterior beliefs. In this case, the IAM distribution that player 2 observes, $\nu_1 \sim N(X_1, 3)$, which serves as her prior distribution for Y. Further, player 2's prior for μ would no longer be diffuse, but rather normally distributed with mean X_1 and variance 2. On observing X_2, player 2's posterior corresponds to the aggregated information posterior beliefs. If player 2 simply places tickets in buckets with the highest expected payoffs, ignoring her past ticket placements, she will shift the IAM so that $\eta \sim N(0.5(X_1 + X_2), 2)$.

These examples indicate how tenuous information aggregation can be and the degree to which players must interact in order to effectively communicate and share information. As in the Multinomial-Dirichlet examples, correlation and heterogeneity in signal structure would require players' to deviate from the "private" strategies above to account for that structure in aggregating beliefs. However, it's important to note that throughout, these failures would be revealed in the distributional tests.

Experimental Instructions (Not for Publication)

Forecasting Instructions – (date and time)

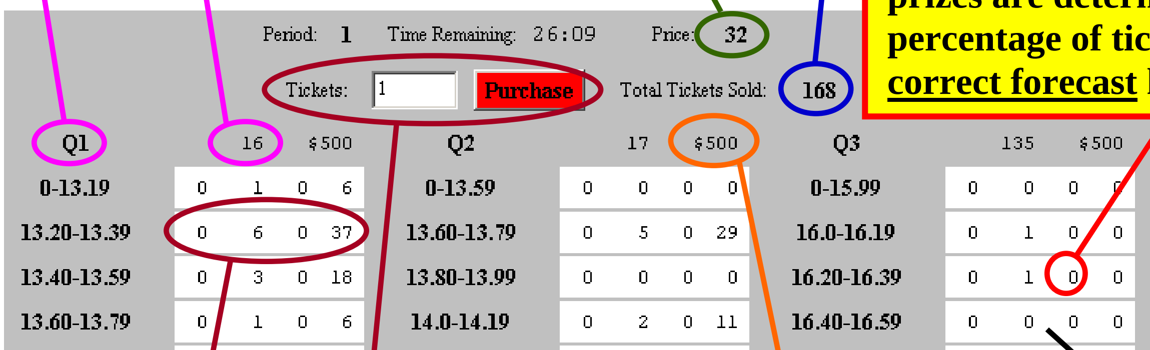
Forecasting (variable to be forecast)

Each column lists the set of forecasts for one quarter and total tickets sold

The price of a ticket will start to increase 15 minutes into the session

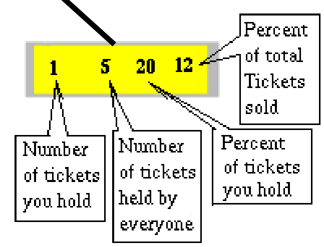
Total tickets sold to all participants for all quarters

Your chances of winning prizes are determined by the percentage of tickets <u>in the correct forecast</u> held by you



To purchase a ticket:
1. Click the white box of the range you choose
2. Enter the number of tickets
3. Click Purchase

Your unspent cash used to purchase tickets – compare to ticket price – separate budget for each quarter/column



Strategy:
1. You start with 500 units of house money for each quarter. <u>Spend it all</u> – but not on one forecast range unless you are certain.
2. Watch what others are doing. <u>The objective is to win money, not simply to record your beliefs.</u>
3. Prices will start at 5 units/ticket and not change for the first 15 minutes. Then they will go up by one unit per minute for 45 minutes. Do not wait too long to buy.

<u>Practice</u>:
http://location and time
<u>Real Deal</u>:
Time and location

Procedure

Step 1: Register

Register yourself in the system database. If you are not in the database the system will force you to register when you try to log into the Real Deal.

Go to (at any time including now) http://xxxx.caltech.edu/xxx
Select "Sign up as a new user". Choose an ID, a password, and enter a number into the "SS Number" field. We are not using real social security numbers – just pick a number with <u>9 digits</u> that you can remember (or write down). Part of a phone number might be a good idea.

Everyone should enter the following information. It will not be used for anything but is required in the stock application we are using.

University = "Company A" and Class = "Company A"
Street = "123 Main Street" City = "Anytown"
State = "CA" Zip = "12345" Country = "USA"
Enter your real e-mail address and phone number. (Enter area code "123" and then your real seven digit Intel phone number.)

Step 2: Practice

Go to the practice page http://xxxx.caltech.edu/Sales-practice/ prior to the Real Deal to become familiar with the forecasting application. Buy tickets for a few different forecasts and observe how the application responds.

Step 3: Get your secure ID

On the day of the Real Deal, ideally a few minutes before the start time, go to the Real Deal location, http://xxxxcaltech.edu/BusinessUnitYearQ#Date/. It will ask you for the user name and password that you used in Step 1. It will then give you your secure ID, which disguises your identity. Click the "Login" button to enter the Real Deal. You will not be able to use the application until the session begins.

Step 4: Participate in the Real Deal

The session will be held on <u>November 7 at 4:00 PM Pacific Time</u>. Be on time – a few minutes early would be wise. The trial will start exactly on time, allowing for clock differences, and move very quickly. It will likely be over in 30 minutes even though it will remain open for an hour.

Panics or problems: e-mail or call Mister X at ###-###-####. He will be working with Caltech to manage the trial and solve any problems.

We will put general announcements (if needed) on the Real Deal screens.

Determining Winners

Four prizes will be awarded for each of the three quarters forecast during the trial – see details below. We will know which forecast is correct once actual Q4 2006 and Q1, Q2 2007 Business Unit Billings are available. Prizes for each quarter will be awarded after the close of that quarter. All tickets in the correct forecast are considered winning tickets and will be entered into a drawing for prizes. After each prize drawing the winning ticket will be put back in the hopper, so each ticket may win more than one prize.

Q4 2006	**Q1 2007**	**Q2 2007**
Drawing 1: $100	**Drawing 1: $100**	**Drawing 1: $100**
Drawing 2: $100	**Drawing 2: $100**	**Drawing 2: $100**
Drawing 3: $50	**Drawing 3: $50**	**Drawing 3: $50**
Drawing 4: $50	**Drawing 4: $50**	**Drawing 4: $50**

These prizes will be distributed as an employee recognition award in the near term. Alternative payment methods may be developed in the long term.

Privacy

Participants will remain completely anonymous except to the research team at Caltech and to Mister X, the research manager at Company A. No one else participating in the trial will know for certain who is participating, so they certainly will not know which forecasts you choose. The final forecast generated by all participants will be published, but your personal forecast will be held in confidence by the research team. We will award prizes to the winners, but even the winners will not be announced.

We expect that participants will not share information with one another before, during or after the trial. Past research has shown that the best results are achieved when participants do not share information.