DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

A Parimutuel-like Mechanism for Information Aggregation: A field Test Inside Intel

Benjamin J. Gillen
California Institute of Technology

Charles R. Plott
California Institute of Technology

Matthew Shum
California Institute of Technology

SOCIAL SCIENCE WORKING PAPER 1367

September, 2014

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

Benjamin J. Gillen
Caltech

Charles R. Plott
Caltech

Matthew Shum
Caltech

September 27, 2014

Abstract

Field tests of a new Information Aggregation Mechanism (IAM) developed via laboratory experimental methods were implemented inside Intel Corporation for sales forecasting. The IAM, which incorporates selected features of parimutuel betting, is uniquely designed to collect and quantize as probability distributions any dispersed, subjectively held information that might exist. The tests demonstrate the robustness of experimental results and the practical usefulness of the IAM. The IAM yields predicted distributions of future sales that are very accurate at short horizons; indeed, more accurate than Intel's official in-house forecast 59% of the time. A symmetric game model suggests why the IAM works.

*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 production 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 successfully aggregate organizational information about the uncertainty surrounding potential business outcomes? Does the mechanism reveal information not already apparent in the company's internal forecasting process? Is its operation understandable from the point of view of available theory and supported by robust econometric analysis of the data?

The purpose 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 quantify this information while setting proper incentives that balance the reward to revealing information against the hazards of free riding on others' information. The study of information aggregation in experimental economics laboratories has a long history, providing a valuable history of successful mechanisms based on theoretical principles but refined through practical testing. The ability of markets to perform the information collection and aggregation functions and 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)). How well do these mechanisms work in a large-scale field setting? Are the results from the laboratory robust in the field? This paper addresses these questions

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 betting systems, which have a goal of entertaining participants as opposed to aggregating their information. At the same time, the design of the mechanism 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 a similar function.

We report results from a long-running 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. Forecasts of product sales are 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. At Intel, we set up mechanisms to collect and aggregate these pieces of information about future realizations of units sold for key products. In these mechanisms, the range of values that possible sale quantities can take is partitioned into a set of non-overlapping intervals, or "buckets." The participants in each mechanism are given an opportunity to purchase "tickets" that pay off when the variable of interest takes a value within a given bucket. Participants 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 potential measure of participants' beliefs regarding the future realization of the variable of interest. The information aggregation mechanism automatically aggregates these beliefs across participants, allowing the construction of "consensus" forecasts while also obtaining a glimpse into the underlying uncertainty.

We approach the evaluation of performance by posing two questions. The first question focuses on proof of principle. "Does the process do what it was intended to do?" The second question focuses on the consistency of the behavior with the principles on which the design was constructed. "Does it do what it does for the right reasons?" The answers to these questions are known for tests in laboratory experimental environments. The questions are now posed for the mechanism performance in a field environment.

The process is intended to aggregate information so the first question asks if it is successful. Evaluating its success requires assessing the amount of information that might exist in the environment and characterizing how that information is reflected in the IAM. Section 3 presents a model of information and how the uncertainty in this environment, both related to outcomes and to odds of outcomes, is represented in the IAM. Given this conceptual link, we can then test the IAM's performance by comparing its predictions with realized results. As success requires accuracy of both the predicted outcome and the likelihood the IAM assigns to that outcome, the IAM's accuracy depends both on the process working and on the existence of information to be collected and analyzed. Excellent prediction performance might not be impressive if the prediction is trivial in some sense and not likely to be repeated in other applications. On the other hand, poor prediction performance could reflect the lack of information to be aggregated as opposed to any failures of the mechanism's collection of the meager information available. Our empirical tests indicate that the IAM reliably characterizes the uncertainty regarding sales in those settings where Intel has good information about that uncertainty and that the information reflected by the IAM is not trivially available through other internal information sources.

The second question addresses the power of the underlying theoretical principles to explain and interpret the performance of the IAM. The IAM was not developed from first principles as suggested by the growing literature on implementation theory, so a complete theory of its construction is not available. Instead the IAM was built on insights drawn from the behavior of markets and entertainment devices together with features of the rules used for their operation. While the development of the IAM through experimental testing departs from a pure theory approach, the controlled exploration of the IAM's empirical properties yield theoretical insights into why it works and its sensitivity to details of the environment. The model of the information and environment developed in the section is extended to a game theoretic model of individual and systems behavior and information aggregation in the environment. Clearly the simple theory abstracts from important features of real-world interaction that could complicate information aggregation and underline theoretical skepticism about possible mechanism success. Section 3 addresses many complicating issues. Nonetheless, even in the presence of complicating features, the IAM may successfully aggregate information in accordance with well-understood economic principles. Establishing this possibility motivates the notion of information aggregation outcome as a null hypothesis, the observable implications of which we develop in the first half of Section 4. Importantly, we develop tests for this null hypothesis in a reduced form model to ensure their robustness to structural misspecification.

Testing the performance of Intel's IAM implementation is complicated by the fact that we don't observe the true unknown distribution over sales in the field. In addition to the IAM's aggregated forecasts, we also have access to an internally-prepared "official" forecast that can serve as a contemporaneous benchmark providing an indicator of the quality of information that exists within the organization at the time of the IAM run. Beyond six months before the quarter ends, the forecast errors from both the IAM and the official forecast are as large as the standard deviation of sales themselves. However, forecast errors become much smaller as the horizon narrows, indicating that information improves substantially, especially in the last month before the quarter ends. Together, these findings support the natural hypothesis that Intel's information about sales uncertainty is much better at shorter horizons than longer horizons. This benchmark will allow us to verify that the IAM accurately reflects the uncertainty in sales in contexts where Intel has good information about that uncertainty, in comparison with those situations in which good information did not exist.

Our empirical analysis uncovers two central findings confirming the effectiveness of Intel's IAM implementation. First, in terms of aggregating information, we find that the IAM accurately characterizes uncertainty in sales at those forecast horizons when the participating individuals have good visibility into that uncertainty. In a simple reduced form test, the IAM beliefs over unit sales matches the distribution of realized sales up to three months before the end of the quarter. The result establishes that, when the information existed to be aggregated, the IAM effectively gathered and aggregated it. At forecast horizons beyond six months, the mechanism's forecast distribution reflects a tendency to overstate expected sales consistent with the official forecast's bias and also tends to understate the dispersion of uncertainty, underweighting low probability events and illustrating a "reverse favorite-longshot bias."

Second, we show that the IAM reveals information about sales not available in the official internal forecast. We find the expected outcome reported by the mechanism robustly outperforms the official sales forecast, delivering lower forecast error in 59% of the IAM's runs. The relevance of this performance is highlighted by considering the ex-post optimal combination of the two forecasts. This optimal combination of forecasts heavily weights the IAM positively and assigns a negative weight to the official forecast across all forecasting horizons. Rather than questioning the value of the IAM as a forecasting tool, these findings question the information value of the official forecast relative to the IAM. Though this robust outperformance may be driven by other, non-informational, purposes the official forecast serves in the organization, the information value of the IAM forecast's is clearly substantial.

2 The Information Aggregation Mechanism Structure

A connection between markets and information transmission dates back to the foundations of economics (The intuition is clearly seen in Hayek (1945) and Fama (1970) begins to be formalized with Green (1973, 1977) and see Allen and Jordan (1998) for a review of this early formal development and general principles for the existence of rational expectations equilibrium). The theoretical results suggested that markets are capable of collecting and aggregating information, though exactly how that might happen remains an open question. The existence of such possibilities suggests that it is possible to design an Information Aggregation Mechanism with the purpose 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.

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 in 1996 by Chen and Plott (2002) inside Hewlett Packard Corporation. They implemented a complete

[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).

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 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 entertainment-based partimutuel betting systems in fundamental ways. First, tickets are not sold at a fixed price, but rather prices increase at a pre-announced rate 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

[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.

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 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 the self-selection and participation-induced bias. Successful information aggregation balances an individual's strategic influence in the IAM, as measured by their impact on the final distribution over tickets, with the quality of their private information. This balance is often upset by agents' participation decisions, which exacerbate the influence of heterogeneous risk aversion, outside wealth, and behavioral biases. These selection issues arise prominently in prediction markets for entertainment purposes, which typically rely on participants selecting themselves to voluntarily engage in the mechanism, guided by the belief that increasing the size of the crowd

[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 company's performance to general industry issues, though the information relayed by these markets were often biased by participation effects.

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

maximizes its wisdom.[6] By contrast, Intel management invited IAM participants chosen for the information to which they had access given their position in the organization. This intervention ensures the population consists of relatively homogeneous participants, mitigating potential imbalances that might complicate information aggregation.

Finally, we set up IAM's for sales forecasting that elicit participants' beliefs about variables (unit sales) that 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 intuition (see Plott (2000)) that 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, 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

[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.

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.

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.

3 A Model of the Information Environment and Aggregation

We present a model for extending the application of the mechanism to field environments, facilitating the interpretation of the data and subsequent testing in terms of information aggregation. The modeling effort proceeds in three parts. We begin by characterizing the nature of the randomness in the field environment in 3.1. Here we depart from a parametric tradition in which uncertainty is represented by an unknown expectation for an unknown state variable. As opposed to this single point that might be identified with certainty, the information we which to aggregate is intended to capture a minimally parametric representation of uncertainty in the full distribution of the state. In analyzing these distributions over distributions, we propose a natural signal technology for how participants acquire information. With a precise definition of information in 3.2, we evaluate the implications of individual decisions when that behavior is structured from rule-of-thumb ticket purchases. We show that, if individuals follow this rule of thumb, the mechanism will aggregate information and the relative ticket allocation in the IAM will match the conditional uncertainty in sales. We then show, in 3.3, that information aggregation is attained in the unique, symmetric Nash equilibrium any IAM in which participants arrive at common posterior beliefs. Imagination leads to many reasons why the IAM might not work. Such possibilities are discussed in Section 3.4.

3.1 A Model of Information Aggregation in the IAM

We start with a simple model using a Dirichlet sampling framework to represent the links among the environment, individual beliefs, and information aggregation in the IAM. 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:

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

The cell probabilities given by $\pi = (\pi_1, \ldots, \pi_K)'$ are unknown quantities, about which agents are endeavoring to learn. Their uncertainty about these quantities corresponds to a distribution over distributions, a modeling environment for which the Dirichlet is particularly well-adapted.[7]

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

$$\pi \sim Dir(\alpha_1, \ldots, \alpha_K), \quad E[\pi_k] = \frac{\alpha_k}{\sum_{j=1}^{K} \alpha_j} \tag{2}$$

Suppose further that each agent updates her beliefs about π upon observing noisy signals of π. Specifically, an agent observes m_n signals $s_{n,1}, \ldots, s_{n,m_n}$, drawn independently from $MN(\pi)$. From these m_n signals, the agent can compute sample frequencies $\hat{p}_{n,1}, \ldots, \hat{p}_{n,K}$, where $\hat{p}_{n,k} = \frac{1}{m_n} \sum_{j=1}^{m_n} \mathbb{1}\{s_{n,j} = k\}$, the sample frequency with which the signal falls into the k-th bucket. Given these conjugate distributional assumptions, the posterior distribution for

[7] The Dirichlet distribution has also been used in other contexts (eg. Rothschild (1974)) to study learning and sampling in a discrete setting. It features prominently in the examples from Ottaviani and Sørensen (2010).

π conditional on these signals will also be Dirichlet, corresponding to:

$$\pi|s_n,\alpha \sim Dir\left(\alpha_1 + m_n\hat{p}_{n,1},\ldots,\alpha_K + m_n\hat{p}_{n,K}\right) \tag{3}$$
$$E\left[\pi_k|s_n,\alpha\right] = \frac{\alpha_k + m_n\hat{p}_{n,k}}{m_n + \sum_{j=1}^{K}\alpha_j} \equiv \tilde{p}_{n,k}$$

In this setting, we want to examine whether information aggregation should occur – that is, whether the distribution of ticket placements coincides with the "true" or expected probabilities for the underlying distribution of sales Y based on all agents' information. Add the simplifying assumption that the signals are independent across agents, let $M = \sum_{n=1}^{N} m_n$ denote the total number of signals and let $\hat{p}_k = \frac{1}{M}\sum_{n=1}^{N}\sum_{j=1}^{m_n}\mathbb{1}\left\{s_{n,j}=k\right\}$ be the proportion of all signals in bucket k. We obtain the aggregated posterior distribution across all N agents:

$$\pi|s_1,\ldots,s_N,\alpha \sim Dir\left(\alpha_1 + \sum_{n=1}^{N} m_n\hat{p}_{1,n},\ldots,\alpha_K + \sum_{n=1}^{N} m_n\hat{p}_{K,n}\right) \tag{4}$$
$$E\left[\pi_k|s_1,\ldots,s_N,\alpha\right] = \frac{\alpha_k + \sum_{n=1}^{N} m_n\hat{p}_{n,k}}{M + \sum_{j=1}^{K}\alpha_j} \equiv \tilde{p}_k$$

This last posterior distribution, $\pi|s_1,\ldots,s_N,\alpha$ represents the fully aggregated information regarding the distribution of the outcome variable Y available to participants. Intuitively, each individual's draws from the multinomial distribution correspond to m_n "pieces" of information about the true distribution for Y and the Dirichlet distribution provides a convenient summary of the total information revealed to individuals.

Note that the posterior beliefs still allow for aggregate uncertainty in the cell probabilities themselves to persist in the populations' information set. That is, while the expected cell probabilities are fixed, the realized cell probabilities remain random with a positive variance. Nonetheless, these expected cell probabilities represent all the information available about the uncertainty in how realized sales will turn out as opposed to a principle that rests on the possibility that no uncertainty exists. This allows us to define "successful" information aggregation in the mechanism in expectation as:

Definition 1 (Information Aggregation in Expectation)**.** We say the IAM aggregates information in expectation if the expected cell probabilities are proportional to the allocation of tickets within the IAM.

Now, given a large number of independent signals, then as $M \to \infty$, either because each agent receives a lot of information (m_n becomes large) or many agents receive information (N becomes large), the law of large numbers ensures that full-information posteriors converge to the true probabilities:

$$\tilde{p}_k = \frac{\alpha_k + \sum_{n=1}^{N} m_n \hat{p}_{n,k}}{M + \sum_{j=1}^{K} \alpha_j} \to \pi_k, \quad k = 1, \ldots, K. \tag{5}$$

This allows us to define what "exactly successful" information aggregation in the mechanism means as:

Definition 2 (Exact Information Aggregation)**.** We say the IAM aggregates information exactly if the true unobserved cell probabilities, conditional on available information, are proportional to the allocation of tickets within the IAM.

These two definitions contrast the aggregate uncertainty in outcomes, characterized by the cell probabilities π, and the aggregate uncertainty in the distribution over outcomes, characterized by the Dirichlet posterior distribution. Clearly, as these definitions do not explicitly rely on the Dirichlet structure, they can be readily interpreted as applying to any measurable setting.

3.2 Individual Behavior and Information Aggregation

We begin our analysis by demonstrating that a simple strategy of ticket placement can lead the distribution over tickets in the IAM to match the posterior expected cell probabilities. Suppose player n follows a simple "naive" or "straight-forward" strategy that places $\nu_{n,k}$ tickets in bucket k proportionally to their privately observed signals, so that $\nu_{n,k} \propto \hat{p}_{n,k}$. Suppose further that every player is equally endowed with C units of currency and equally informed via (m) independent signals and prior beliefs are diffuse (so that α is arbitrarily small). Then, trivially, the distribution over tickets will be proportional to the average signal.

$$\nu_k \propto \hat{p}_k = \frac{\left(M + \sum_{j=1}^{K} \alpha_j\right)}{M} \tilde{p}_k - \frac{1}{M} \alpha_k \approx \tilde{p}_k \tag{6}$$

We note that this simple strategy underweights the prior beliefs of agents, so if individuals follow this behavioral rule, the distribution reported by the IAM will display a small distortion from failing to incorporate prior information. This distortion could be remedied by seeding the IAM with the prior distribution, which is public and common knowledge by assumption. However, if the prior is genuinely diffuse (i.e., as $\alpha \to 0$), then the distortion becomes arbitrarily small and the naïve strategy will lead the IAM to aggregate information in expectation. Further, if the number of participants becomes large or the participants become arbitrarily well informed about (i.e., as $M \to \infty$) then the IAM will aggregate information exactly. These results are summarized in the following proposition.

Proposition 1. *[Information Aggregation from Naïve Behavior]*

- *Suppose the information aggregation environment is characterized by the distributional assumptions embedded in equations 1 - 5.*
- *Suppose further that players reveal their private information by following a naïve strategy that places their tickets proportionally their private signals, so that $\nu_{n,k} \propto \hat{p}_{n,k}$.*

Then as $\alpha \to 0$, the IAM aggregates information in expectation:

$$\nu_k \propto \hat{p}_k \underset{\alpha \to 0}{\rightarrow} \tilde{p}_k = E\left[\pi_k | s_1, \ldots, s_n, \alpha\right]$$

Further, as information accumulates, with either $N \to \infty$ or $m \to \infty$, then the IAM aggregates information exactly:

$$\nu_k \propto \hat{p}_k \underset{M \to \infty}{\rightarrow} \pi_k.$$

This simple result follows immediately from equation 6 and highlights the complex balancing act that gives rise to successful information aggregation. The aggregated information set balances each player's private signal content with their ticket purchasing capacity. With agents that are completely unresponsive to the external environment, information aggregation occurs so long as these features are exogenously balanced, achieved here by the assumption that all players are equally informed and have equal budgets. This feature highlights the importance of management's role in determining IAM participants, the synthetic currency, and the dynamic interactions among participants. Participant selection ensures the subjects in the IAM are as well-informed as possible. The synthetic currency encourages participants to spend all of their budget, mitigating risk aversion's impact on participation as a complicating feature. Finally, the dynamic interactions of subjects in the IAM are designed to facilitate communicating both the content of their information and their relative informedness to achieve this balance.[8] The following result summarizes this finding:

Result 1 (Naïve Behavior Supporting Information Aggregation)**.** *If individuals follow the rule of thumb of purchasing tickets in proportion to their observed information, the mechanism will aggregate information and the relative ticket allocation in the IAM will match the conditional uncertainty in sales.*

3.3 Information Aggregation and Agreement as an Equilibrium Property

Suppose all information in the system is publicly revealed, so that every participant in the IAM agrees on the posterior distribution for π and the posterior expected cell probabilities, so that $\tilde{p}_{n,k} = \tilde{p}_k$ for all players. Given a common prior and common knowledge of rationality, the result from Aumann (1976) stipulates this sort of agreement would be a necessary feature for any equilibrium in the mechanism. This property allows the model to abstract from the complications induced by strategic communication, providing a clear and tractable perspective on the possibility of informative equilibria in the IAM in settings with rational expectations.

In this environment, an obvious symmetric Nash equilbrium exists, namely one where individuals place their tickets proportionally to the jointly agreed upon posterior expected cell probabilities, $\nu_{n,k} \propto \tilde{p}_k, \forall n, k$. In fact, this equilibrium is the unique symmetric, simultaneous equilibrium in the IAM, which we establish in the next Proposition.

Proposition 2. *[Information Aggregation as an Equilibrium Property]*

[8]Without considering the dynamic elements of the IAM, one might be tempted to suggest this result indicates players' endowed budgets be proportional in some sense to the quantity of private information they have available. If players have heterogeneous endowments of information and follow the naïve strategy proposed here, such a heterogeneous endowment would be necessary for unbiased information aggregation. However, this analysis ignores the possibility that players can choose to place tickets in an "uninformative manner. In particular, players may place a portion of their tickets according to their private information and, after a while, place their remaining tickets in a manner that preserves the relative distribution over tickets in the IAM. This consideration highlights the role dynamic interactions regarding "quantity" of ticket placements play in allowing participants to modulate their position according to their "quantity" of private information

- *Suppose the information aggregation environment is characterized by the distributional assumptions embedded in equations 1 - 5. Suppose further that all private signals are publicly revealed, so that* $\tilde{p}_{n,k} = \tilde{p}_k = E\left[\pi_k | s_1, \ldots, s_n, \alpha\right], \forall n, k$

- *Suppose tickets are infinitely divisible and each player places their tickets proportionally to the posterior expected cell probabilities, so that* $\nu_{n,k} \propto \tilde{p}_k, \forall n, k$.

This behavioral strategy represents the unique symmetric equilibrium outcome with agreement, under which the IAM aggregates information in expectation. Further, as information accumulates and $M \to \infty$, *then* $\tilde{p}_k \to \pi_k$ *and the IAM aggregates information exactly in this equilibrium.*

Proposition 2 indicates that, if private information can be effectively communicated, the IAM's incentives will lead to information aggregation through equilibrium regardless of subjects' initial endowments of wealth and information. Further, the IAM stabilizes in a state where it aggregates information, in the sense that players do not have an incentive to disrupt the IAM. Since the conditions assumed here are essentially equivalent to the definition of an ex-post equilibrium, the result should not be terribly surprising. However, it highlights the simple link between rational expectations, the IAM's incentives, and information aggregation.

Result 2 (Symmetric Equilibrium with Agreement Supports Information Aggregation)**.** *Information aggregation within the IAM can be supported as a unique, symmetric Nash equilibrium.*

3.4 Complicating Features and the Non-Necessity of Information Aggregation

The theoretical development thus far illustrates the principles underlying the IAM process to conceptually motivate the transition from the laboratory to the field environment. Clearly, the theoretical arguments above do not establish information aggregation as a necessary feature of the IAM, but just as a possible feature with firm theoretical grounding that one might reasonably expect to carry over from the lab to the field. Many unmodelled features in the real world complicate the simple model of incentives and information here and might lead the IAM to fail in aggregating all information. While the equilibrium model supports an interpretation of information aggregation as guided by incentivized behavior, it also suggests conditions under which information aggregation might not be observed. Further, many alternative models of behavior exist which might not support information aggregation as an outcome of the IAM. These theories have potential for helping identify limitations to the successful operation of the IAM and highlight the degree to which successful information aggregation presents a surprising result

Many environmental features might complicate information aggregation even in the context of equilibrium models with perfectly rational agents. Thin markets can operate to reduce information expressed in prices. For example, an individual who knows a ticket will win with near certainty has no incentive to buy more than one ticket if it is the only ticket sold. Thus, without sufficient competition the available information might not find its way into the relative volumes of ticket sales. As further illustrated in Ottaviani and Sørensen (2010), strategic behaviors could also degrade the information contained in market odds. Though costly ticket purchases cannot be viewed as cheap talk,[9] participants are aware of the possibility that others are watching purchases and thus have an incentive to mislead others as well as hide what they know. These incentives might cause the agreement property on which information aggregation is based to not be satisfied. Another theoretical complication arises from the existence of multiple, possibly asymmetric, equilibria, which are likely to arise in more realistic models.

Beyond theory, practical features might complicate the successful operation of the IAM. While the starting time of the IAM is well known, the participants must operate on a business schedule. Because ticket prices go up as time progresses, a late arrival might have limited purchasing power and the participant's information would not get into the system. The timing difficulties could be correlated with information with those with the best information facing the most constrained time. The result would be a bias of the information available to the IAM.

Any number of individual behavioral properties could work to degrade or remove the information aggregation capacities of the IAM. How individuals quantize their beliefs or

[9]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.

whether intuitions can be captured by subjective probabilities are open research questions. Individuals could ignore the information contained in IAM odds, thus failing to update public information to their private information when making ticket purchases. A "long shot" bias may be an inherent property of individual decisions under uncertainty. Similarly, individuals could ignore their private information and simply mimic the decisions of others in a Banerjee (1992) and Bikhchandani et al. (1992) style cascade. Manski (2006) shows that prices from prediction markets need not reflect the mean beliefs of traders in those markets, even when participants observe price-related information, due to essential breakdowns in the balance highlighted by Proposition (1). Wolfers and Zitzewitz (2006) present conditions under which prediction market signals correspond to mean beliefs, underscoring the importance of accounting for risk aversion in individual behavior. As these risk attitudes could create a wedge between purchases and subjective probabilities, the degree to which the reported prices match the frequency of observed outcomes is an empirical question.

4 Testing the Information Aggregation Mechanism

If subjects possess information about the distribution of sales and if the mechanism works, then realized sales outcomes will reflect the aggregated distribution of tickets within the IAM. This section focuses on evaluating the degree to which the information reported by the IAM accurately reflects the uncertainty in sales. Our key finding is that the IAM performs well in those settings where the organization had good visibility into the uncertainty in sales, an outcome that would not be feasible if the IAM did not successfully aggregate information. In settings where the organization has very limited visibility into future sales, there is little information for the IAM to aggregate as reflected in the accuracy of its reported distributions over sales.

There are two ways that we could empirically reject the hypothesis that the IAM works to aggregate subjectively held information about sales. First, it could produce an unreasonably accurate prediction of sales data when no information existed to be aggregated. Second, it could produce a poor prediction of sales data when information existed to be aggregated. The composite nature of our hypothesis in a natural observational environment precludes a structural test of the motivating theory from the previous section. As such, we consider reduced-form tests of the information aggregation property's observable implications in Section **??** that compare the full distributions reported by the IAM to the realized outcomes. These tests would allow us to reject the hypothesis that the mechanism aggregates information efficiently for any number of reasons. However, given results that the aggregated beliefs reported by the IAM effectively match the empirical distribution of outcomes, it is inconceivable that such a complex outcome could arise simply by random chance.

4.1 Mechanism Outcomes, Sales, and Forecast Evaluation Data

We observe data on actual sales, the official forecast, and the outcome of the IAM from 2006 through 2013 across five major product lines.[10] To extract a point forecast from the IAM, we define the IAM Mean as the expected value of the outcome under the distribution over tickets within the IAM.[11] Recalling the notation for actual sales of product i in quarter t as $Y_{i,t}$, we refer to the official and IAM forecasts at horizons $h \in \{1, 2, \ldots, 9\}$ months by $\hat{Y}^{(\text{Official})}_{i,t|t-h}$ and $\hat{Y}^{(\text{IAM})}_{i,t|t-h}$, respectively. For each $k \in \{\text{Official, IAM}\}$, we denote the forecast error as $e^{(k)}_{i,t|t-h} = Y_{i,t} - \hat{Y}^{(k)}_{i,t|t-h}$ to evaluate forecast performance.

To characterize the forecasting environment, Table 1 reports summary statistics for the point forecasts together with the actual unit sales, including a break down by product lines. Realized quarterly sales are normalized by product line to average one unit quarterly sales with a one unit standard deviation. On average, the official and IAM forecasts slightly overstate average sales, but the bias is less pronounced in the IAM Mean, which improves upon the Official forecast bias by approximately 5%. Overall, the IAM forecasts deliver a root mean square error almost 8% lower than that of the Official forecast in the full sample, with that outperformance being quite stable across all forecast horizons.

The content of Table 1 can be summarized by the following:

Result 3 (Information about Sales Improves as the Forecast Horizon Narrows)**.** *The accuracy of point forecasts for both the IAM and the official forecast improve as the forecast horizon narrows. Conversely, the quality of this information degrades as the forecast horizon*

[10] 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. Each product was normalized to have an average of one unit sold with a unit standard deviation of sales.

[11] Alternatives, such as the modal forecast are largely the same. 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. 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.

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 we consider. 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 forecast horizon at which forecasts are generated.

	Full		Forecast Horizon		
	Sample	Last Mth	1-3 Mth	4-6 Mth	7-9 Mth
First Period	200602	200602	200602	200603	200604
Last Period	201303	201303	201303	201303	201303
Num of Obs	930	114	339	312	279
Num of Qtrs	30	30	30	28	26
Average Sales	1.00	0.96	0.97	1.02	1.01
Std Dev Sales	0.99	1.00	0.99	0.99	1.00
Average Official	1.19	0.99	1.01	1.20	1.40
Std Dev Official	1.04	1.01	1.02	1.04	1.04
Official RMSE	0.96	0.38	0.63	0.97	1.24
Average IAM Mean	1.14	0.93	0.99	1.16	1.32
Std Dev IAM Mean	1.02	1.01	1.00	1.02	1.00
IAM Mean RMSE	0.88	0.29	0.55	0.89	1.15

lengthens. At these long horizons, there may not be substantial information for the IAM to collect and report.

4.2 Empirical Implications of Information Aggregation

Whether or not the successful information aggregation observed in experimental implementations of the IAM can be achieved in a field context presents the central empirical question of our study. Investigating information aggregation in the field is complicated by the econometricians' inability to access the detailed information used by experimenters in analyzing laboratory data. The experimental designer knows exactly what the true conditional distribution for $Y_{i,t}$ given all information in the system, providing a host of observable restrictions with which to test each market. In the field, however, the econometrician only observes the realized value of $Y_{i,t}$, a severely restricted view into the data generating process that requires leveraging the panel of forecast distributions along with structural assumptions to test whether it exhibits successful information aggregation.According to definitions 1 and 2, successful information aggregation is characterized by a match between the empirical distribution of ticket placements in the IAM should and the "true" probabilities of sales (whether in expectation or exactly). Letting $F_{i,t|t-h}$ denote the information set of IAM participants at period $t-h$, the hypothesis of successful information aggregation can be stated:

Hypothesis 1. *Information Aggregation Mechanism Accuracy hypothesis:*

$$Y_{i,t}|\mathcal{F}_{i,t,t-h} \stackrel{d}{=} MN(\tilde{\eta}_{1|t-h}, \cdots, \tilde{\eta}_{K|t-h}), \quad \forall i, t, h. \tag{7}$$

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

Why might we think Hypothesis 1 holds a priori? Clearly this is a very strong hypothesis constructed from a set of maintained hypotheses outlined in Section 3 that cannot be tested directly and have the capacity to lead to the rejection of Hypothesis 1. The equilibrium model developed in Section 3 reflects principles often found useful for interpreting economic data but it also rests on a series of "as if" propositions that cannot be examined directly. At base, Hypothesis 1 is not inconsistent with a body of accepted theory. Furthermore, Experimental results indicate that the IAM performs as suggested by theory. However, as highlighted in Section 3.4, there are many instances in which the hypothesis may fail. In the last analysis, whether or not the hypothesis holds in a given field setting is an empirical matter.

We can test Hypothesis 1 without additional structural assumptions 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:

$$\hat{Q}_{i,t,h} \equiv \hat{G}_{i,t|t-h}(Y_{i,t}) \sim_{H_0} U[0,1] \tag{8}$$

By translating the outcome into its conditional quantile from the IAM, we control for the heterogeneous conditional distributions from which $Y_{i,t}$ is being realized at different horizons.

Using the well-know property that the quantiles of a random variable are sampled uniformly by the true distribution, in cases where the IAM distribution matches the true conditional distribution for $Y_{i,t}|\mathcal{F}_{i,t|t-h}$, then $\hat{Q}_{i,t,h} \sim U[0,1]$. Accordingly, we can simply use a Kolmogorov-Smirnov test to evaluate whether we can reject that our sample of quantiles is drawn from a set of uniformly-distributed random variables.[12] Analyzing the conditional quantiles renders the test robust to heterogeneity in the distributions across products, time, and information sets.[13] We summarize this finding with the following result:

Result 4 (Testability of the Full IAM Distribution)**.** *A Kolmogorov-Smirnov test can be used to measure the degree to which the probabilities of sales in a quarter match the relative frequencies of actual sales.*

4.3 Exact Information Aggregation in Periods with Good Information Available

Panel A of Figure 1 shows that the empirical distribution for almost perfectly matches the uniform distribution quantiles for the last IAM run for the quarter, which occurs around one week before the quarter ends, but several weeks before sales are finalized. As indicated by the RMSFE in Table 1, uncertainty in final sales persists even in the last run of the IAM, though we might expect participants in the IAM to have good information about this uncertainty. As a setting in which we can be confident that information exists to be aggregated, this context provides the ideal test for our hypothesis.

The results in Figure 1 support the hypothesis, with a mean absolute deviation of less than 3% and the Kolmogorov-Smirnoff test statistic corresponding to a p-Value of 86%, indicating that we cannot distinguish the IAM quantiles from the uniform distribution. That is, the probabilities predicted by the IAM match the relative frequencies of sales. Figure 1's Panel B extends the sample to include IAM runs to include the two and three month horizon forecasts, essentially all IAMs run in the quarter for which sales are being forecast.

[12]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.

[13]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. Unfortunately, analyzing correlated sampling structures in the Kolmogorov-Smirnov test is an intractable problem beyond our scope.

Visual inspection again verifies that the IAM quantiles match up well with the uniform distribution. The apparent distortion in the tails does not register as statistically significant, as the Kolmogorov-Smirnov test still lacks significance with a p-Value of 9%.

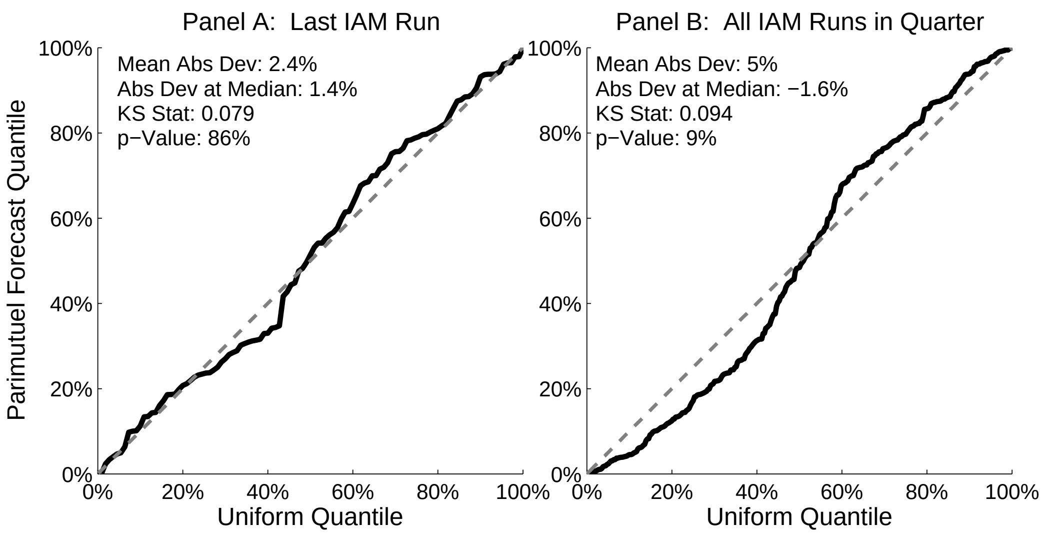


Figure 1: Quantile Plots for the Information Aggregation Mechanism
This figure presents the distribution of realized quantiles from the information aggregation mechanism defined in equation (8) against the theoretically accurate uniform distribution. Panel A reports the results for the last run of the IAM before the quarter ended. Panel B reports the results for all IAM runs during the quarter for which sales are being forecast. Mean Abs Dev reports the mean absolute deviation between the two distributions while the Abs Dev at Median reports the absolute difference between the two distributions at the median of the uniform distribution. The KS Stat and p-Value report correspond to a Kolmogorov-Smirnov test of equality of the distributions.

4.4 Information Aggregation with Limited Information Available

If participants have very little information about the uncertainty in sales, the IAM's performance should deteriorate. This setting provides a sort of placebo test for the IAM, since if there is no information to aggregate, how could the IAM aggregate information exactly? To evaluate this effect, we consider the IAM's performance at long-forecast horizons. In these markets, we might expect the IAM to aggregate information in expectation, but not exactly. What are the observational differences between this outcome and exact information aggregation?

Panels A and B in Figure 2 plot the quantiles of the forecast quantile distribution against the uniform distribution at short medium (4-6 Months) and long horizons (7-9 Months), respectively. Notably, the IAM in this setting no longer plots along the 45 degree line, but rather follows a distinct S pattern. We also see evidence of the bias in the location of these forecasting distributions, reflecting optimistic expectations for business performance.

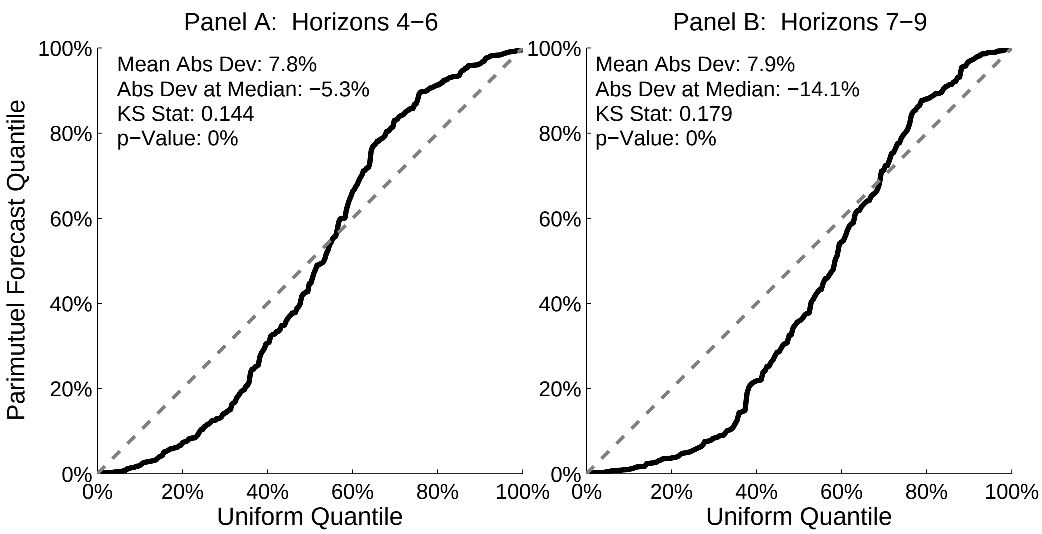


Figure 2: Quantile Plots for the Information Aggregation Mechanism
This figure presents the distribution of realized quantiles from the information aggregation mechanism defined in equation (8) against the theoretically accurate uniform distribution. Panel A reports the results for the IAM's run 4-6 months before the quarter ended. Panel B reports the results for IAM runs 7-9 months before sales are finalized. Mean Abs Dev reports the mean absolute deviation between the two distributions while the Abs Dev at Median reports the absolute difference between the two distributions at the median of the uniform distribution. The KS Stat and p-Value report correspond to a Kolmogorov-Smirnov test of equality of the distributions.

Our finding that,the IAM systematically understate tail probability events in the absence of quality information 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 gambling environments, including probabilistic misperceptions, risk preferences, belief heterogeneity,

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

and information incentives.[15] Ottaviani and Sørensen (2010) propose an alternative model of biases in entertainment-based parimutuel betting systems based on strategic behavior 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 an 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.

The results at these horizons underscore the adage that making predictions is hard, especially about the future. Due to the many composite hypotheses being tested in this evaluation, there are any number of reasons the IAM performance could break down here. In the field, we do not have sufficient data to determine what structural model best characterizes participant behavior. However, the hypothesis that the IAM aggregates information exactly provides the only plausible explanation for the data's support of Hypothesis 1 at the 1-3 month horizon.

Result 5 (IAM Accuracy Hypothesis Supported at Short Horizons)**.** *Within a two or three month horizon the probabilities predicted by the IAM are the same as the relative frequencies of sales. For longer horizons the probabilities of extreme events are understate the actual probabilities.*

[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.

[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 The Information Value of the IAM for Forecasting

We now examine the degree to which information available through the mechanism can help improve the internal, official forecasts of product sales. Our analysis reveals two key findings. First, natural forecasts extracted from the IAM provide lower mean square predictive loss than the official forecast. Second, the IAM reports valuable information relating to expected sales that is not already reflected in the official forecast. These findings underscore the information reported by the aggregated information reflected by the IAM is not merely a restatement of information readily available elsewhere within the organization.

In these tests, we draw on techniques from the forecast evaluation literature (see, for instance, West (2006) and the references therein) to compare the accuracy of the point prediction forecast from the IAM with the official forecast. Of course, the feedback between the official forecast and the IAM induces potential correlation in forecast errors both between and across forecasts, so we implement our tests to be robust to any correlations that might arise through this feedback cycle. Direct comparison of the forecasts is somewhat complicated by the timing of the IAM vis-a-vis the official forecasts. To provide comparable information sets, we evaluate the IAM forecast relative to the official forecast prepared immediately after the IAM. Our conservative treatment cedes a slight information advantage to the official forecast, allowing us to be sure that any outperformance by the IAM is attributable to its aggregation of dispersed beliefs.

5.1 Predictive Accuracy and Forecast Performance

We begin with a direct horse-race between the IAM Mean forecast and the Official forecast in terms of forecast error. 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. Treating forecast loss as the square error of the forecast:

$$l_{i,t|t-h}^{(k)} = \left(\hat{Y}_{i,t|t-h}^{(k)} - Y_{i,t}\right)^2 \tag{9}$$

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

$$\delta_{i,t,h}^{(j,k)} = l_{i,t|t-h}^{(j)} - l_{i,t|t-h}^{(k)} \tag{10}$$

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. 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.

	Num of Obs	Freq IAM Outperforms	Avg Abs Δ Loss(*100)	Diebold-Mariano Test t-Statistic	p-Value
Full Sample	930	59%	0.56	(2.78)	0.6%
1 Mth	114	66%	0.34	(2.99)	0.3%
2-3 Mths	225	56%	7.09	(1.81)	7.1%
4-6 Mths	312	58%	6.43	(1.86)	6.4%
7-9 Mths	279	59%	5.22	(1.95)	5.2%

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.

Table 2 shows the IAM forecast clearly outperforms the official forecast in the overall sample, with the outperformance especially pronounced at the one month horizon where the IAM Mean outperforms the official forecast 66% of the time. 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. As such, the IAM's systematic outperformance is especially surprising given that the official forecasters know the IAM distribution *before* releasing their forecast.

As Table 2 illustrates, the official forecast deviates from the IAM forecast in the wrong direction about 59% of the time. The Diebold-Mariano tests provide clear evidence of the IAM Mean outperforming the official forecast in the full sample and in the 1 month horizon. Looking at other forecast horizons, the test's significance drops with the smaller number of observations, but the IAM Mean is a better forecast than the Official forecast over 55% of the time across all horizons.

5.2 Forecast Combination and Encompassing Tests

Beyond its consistent outperformance of the official forecast, we now show that the IAM contains relevant information about sales that's not already reflected in the official forecast. To this end, we analyze how best to combine the information from the IAM forecast with the official forecasts into a single aggregated forecast to perform a series of encompassing tests.[17] If the IAM encompasses the Official forecast, a decision maker who learns the official forecast after having already observed the IAM would not update their beliefs about potential sales. If the Official forecast does not encompass the IAM, then the information from the IAM would be valuable to a decision maker who only has access to that official forecast.

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.

$$Y_{i,t} = \gamma + \omega_{IAM}\hat{Y}^{(IAM)}_{i,t|t-h} + \omega_{Official}\hat{Y}^{(Official)}_{i,t|t-h} \tag{11}$$

If the IAM encompasses the official forecast, then the weight assigned to the official forecast in the optimal forecast combination would be zero (i.e., $\omega_{Official} = 0$), which can be evaluated using the standard t-Statistic based on robust standard errors. A sharper test of encompassing could evaluate the joint hypothesis that $\omega_{IAM} = 1$ and $\omega_{Official} = 0$, while an even sharper test could add the restriction that $\gamma = 0$ to the null hypothesis. These tests are all readily evaluated using F-statistics with robust standard errors.

Rather than the official forecast encompassing the IAM Mean, the results in Table 3 indicate the IAM Mean is more likely to encompass the official forecast than the reverse. In Panel A's full sample estimates, the Official forecast is actually negatively weighted, though this negative weight is not statistically distinguishable from zero. Looking at the horizon subsamples, we can always reject the hypothesis that the IAM is encompassed by the Official forecast, that is, the data clearly show that $\omega_{IAM} \neq 0$. However, we cannot always reject the null hypothesis that $\omega_{Official} = 0$, since the optimal forecast weights the official forecast negatively to control variability in the IAM Mean. We also reject all of the composite encompassing F-tests, providing statistical evidence of the information value to be gained by combining the forecasts.

[17] A robust literature considers optimal forecast combination, with the survey by Timmermann (2006) providing a good entry point.

Table 3: Forecast Combination Regressions

This table presents estimates from the forecast combination regressions 11. 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.

	Intercept	IAM Mean Weight	Official Fcst Weight		F(0, 1, 0)	F(0, 0, 1)	F(., 1, 0)	F(., 0, 1)
Full Sample	0.32	94%	-33%	F-Stat	10.05	22.84	14.97	33.03
Std. Error	(0.09)	(17%)	(18%)	p-Value	0%	0%	0%	0%
1 Month	0.08	98%	-3%	F-Stat	5.73	22.57	5.60	31.47
Std. Error	(0.03)	(13%)	(13%)	p-Value	0%	0%	0%	0%
2-3 Months	0.17	115%	-36%	F-Stat	5.65	10.53	8.47	14.96
Std. Error	(0.09)	(29%)	(30%)	p-Value	0%	0%	0%	0%
4-6 Month	0.35	85%	-26%	F-Stat	7.59	13.03	10.69	18.04
Std. Error	(0.10)	(22%)	(24%)	p-Value	0%	0%	0%	0%
7-9 Month	0.53	79%	-40%	F-Stat	8.07	16.13	11.95	24.19
Std. Error	(0.17)	(29%)	(26%)	p-Value	0%	0%	0%	0%

Result 6 (IAM Reports Information Not Included in the Official Forecast)**.** *The IAM is more accuracte than the official forecast. An analysis of the optimum combination of frecasts reveals that the official forecast carries a negative weight, meaning that the predicition of the IAM can be improved by adding information that the official forecast is systematically wrong relative to the IAM.*

6 Discussion of Conclusions

This paper analyzes a field test of the Information Aggregation Mechanism that was developed and refined in laboratory experimental environments. The purpose of the mechanism is to quantize, collect, and aggregate information that is held in the form of intuition by a dispersed collection of people. Having been adopted to generate monthly predictions of quarterly sales inside Intel corporation for over seven years, the field test establishes the robustness of the IAM's information aggregation properties to applications in uncontrolled environments,. The results underscore that the IAM performs well in both absolute terms and relative to other forecasts and that it does so for understandable reasons.

Results 1 & 2: A Theory of Behavior Bridges the Gap Between Lab and Field Environments

Our analysis encounters several major differences between field tests and research conducted under controlled laboratory procedures worthy of mention. A mechanism developed through the process of laboratory experimentation need not be understood in terms of an overarching general theory that would be necessary to develop the mechanism following the practice of mechanism design. Partial theoretical principles, e.g. aspects of rational expectations and strategic behavior, can be applied and tested in the laboratory because the controlled setting ensures the experimenter knows all relevant features of incentives and information. Directed by such partial equilibrium principles, trial and error guides the experimenter through refining the mechanism in response to observed shortcomings. A general theory explaining behavior is unnecessary in the special cases evaluated in lab settings, but its absence challenges our confidence when assuming that systems behavior will persist when transferring from the lab to a field environment. As we no longer know and control the variables driving individual behavior in that setting, the field environment introduces unmodelled features that complicate drawing general conclusions regarding a theory of behavior in this setting.

We bridge the laboratory and field setting by presenting a simple theoretical model with precise assumptions that characterizes the essential observational implications of successful information aggregation as an outcome of the IAM. Considering information and incentives, we present theoretical equilibrium properties that motivate the possibility of successful information aggregation as driven by robust behavioral principles. Though these behavioral principles are not individually testable in the field, they individually have support from experimental contexts.

Results 3, 4, & 5: IAM Accurately Reflects Uncertainty in Environments with Good Information

Our evaluation of the IAM's performance focuses on the aggregate hypothesis that the probabilities emerging from the IAM should match the relative frequencies of the events predicted. This reduced-form hypothesis imposes no additional structure on the data generating process beyond the simplest conditions for basic measurability, so the statistical tests are both powerful and robust. Evaluating all moments of the complete distribution reported by the IAM, we verify this hypothesis in settings where participants have good information about the uncertainty surrounding potential outcomes. We observe the IAM's performance degrade with the quality of information, as measured by external characteristics of the information environment. In particular, some systematic error appears in the IAM distribution as the forecast horizon increases and the bias becomes worse as the official forecast accuracy becomes poor.

Result 6: IAM Captures Unique Information Unavailable Elsewhere

We also verify that the IAM reports more information about expected sales than is available from the official forecast. The IAM forecasts can be improved by incorporating information from the official forecast, but the optimal forecast combination assigns a negative weight to the official forecast. Though the official forecast serves many purposes within the organization, the results demonstrate that the success of the IAM is not exclusively due to a dependence on the official forecast.

Overall Results: Empirical Support for Information Aggregation in the Field

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. Together the collection of results suggest that the IAM performance reflects much of the information available on which to base sales predictions. A tempting working conjecture is that the IAM captures all information available that can be used to predict future sales, but the data tell us that possibilities for improvement exist. The fact that predictions can be improved by adding the official prediction as a contrarian predictor indicates one approach to forecast improvements. Of course, by adding key people, better data to individuals, and decision aids, the prediction might be further improved.

The results demonstrate that the model performs as predicted. Does it do it for the right reasons? We modestly admit the answer to that question cannot be known with any confidence through field testing. Many maintained assumptions that lie at the heart of our working model could fail and the resulting inaccuracies and errors could be offsetting. We do know the essential features of the model are supported by experimental evidence. We also know that key implications of the model exhibit powerful observable restrictions in both experiments and in applications. These implications, motivated by the underlying model and theory, are supported a successful field test. Thus, the information aggregation implications revealed by the model in experiments appear robust in their application to uncontrolled environments.

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.

Robert J Aumann. Agreeing to disagree. *The Annals of Statistics*, 4(6):1236–1239, Nov 1976.

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.

Abhijit V Banerjee. A simple model of herd behavior. *Quarterly Journal of Economics*, 107 (3):797–817, 1992.

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.

Sushil Bikhchandani, David Hirschleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5): 992–1026, 1992.

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.

Eugene Fama. Efficient capital markets: A review of thoery and empirical work. *Journal of Finance*, 25:381–417, 1970.

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

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.

Jerry Green. Information efficiency and equilibrium. *Harvard Institute of Economic Research Discussion Paper*, (284), 1973.

Jerry Green. The non-existence of informational equilibria. *The Review of Economic Studies*, 44(3):451–463, Oct 1977.

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

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

F.A. Hayek. The use of knowledge in society. *American Economic Review*, 35:519–530, 1945.

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. 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.

Kenneth D. West. Forecast evaluation. In C.W.J. Granger G. Elliott and A. Timmermann, editors, *Handbook of Economic Forecasting*, volume 1 of *Handbook of Economic Forecasting*, chapter 3, pages 99 – 134. Elsevier, 2006.

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. Proof of Proposition 2

Proposition 2. *[Information Aggregation as an Equilibrium Property]*

- *Suppose the information aggregation environment is characterized by the distributional assumptions embedded in equations 1 - 5. Suppose further that all private signals are publicly revealed, so that* $\tilde{p}_{n,k} = \tilde{p}_k = E[\pi_k | s_1, \ldots, s_n, \alpha], \forall n, k$
- *Suppose tickets are infinitely divisible and each player places their tickets proportionally to the posterior expected cell probabilities, so that* $\nu_{n,k} \propto \tilde{p}_k, \forall n, k$*.*

This behavioral strategy represents the unique symmetric equilibrium outcome with agreement, under which the IAM aggregates information in expectation. Further, as information accumulates and $M \to \infty$*, then* $\tilde{p}_k \to \pi_k$ *and the IAM aggregates information exactly in this equilibrium.*

Part 1: Best Response Along Equilibrium Path

Given expected cell probabilities and other players' ticket placements, it is optimal for a player to place tickets according to the expected cell probabilities. This partial-equilibrium result establishes that the above assumptions suffice for $\nu_{n,k} \propto \tilde{p}_k, \forall n, k$, to be a best response.

Consider the decision problem faced by the n-th player, conditioning on the players' beliefs $\tilde{p}_n, k = \tilde{p}_k$ and the assumption that all other players are placing their tickets proportionally to the aggregate posterior beliefs. Player n's payoff from any ticket allocation is:

$$E[u_n(\nu) | s_1, \ldots, s_N, \alpha] = \sum_{k=1}^{K} \frac{\nu_{n,k}}{(N-1)\tilde{p}_k + \nu_{n,k}} E[\pi_k | s_1, \ldots, s_N, \alpha] = \sum_{k=1}^{K} \frac{\nu_{n,k}}{(N-1)\tilde{p}_k + \nu_{n,k}} \tilde{p}_k$$

Taking first order conditions of the Lagrangian that incorporates a shadow cost (λ) for the constraint that tickets be fully allocated:

$$\frac{\partial}{\partial \nu_{n,k}} E[u_n(\nu) | s_1, \ldots, s_N, \alpha] = \frac{(N-1)\tilde{p}_k^2}{((N-1)\tilde{p}_k + \nu_{n,k})^2} - \lambda = 0$$

$$\sum_{k=1}^{K} \nu_{n,k} = 1$$

The budget constraint enforces these first order conditions to balance across each of the K cells, so player n's utility maximizing strategy accords with the equilibrium prediction that the players allocate tickets according to the posterior expected cell probabilities.

$$\frac{(N-1)\tilde{p}_k^2}{((N-1)\tilde{p}_k + \nu_{n,k})^2} = \frac{(N-1)\tilde{p}_j^2}{((N-1)\tilde{p}_j + \nu_{n,j})^2} \implies \frac{\nu_{n,k}}{\nu_{n,j}} = \frac{\tilde{p}_k}{\tilde{p}_j}$$

Part 2: Uniqueness of Equilibrium Outcome

We now establish uniqueness of the equilibrium outcome. First, we show that at least one player has a profitable deviation if the IAM's distribution of tickets is not proportional to the agreed-upon posterior odds. Second, we show that asymmetric ticket allocations are not supportable with agreement.

(a) Suppose the IAM's distribution of tickets is not proportional to $\tilde{p}$, then at least one player has a profitable deviation.

Without loss of generality, suppose $\tilde{p}_1 > \eta_1$ and order the indices so that $\frac{\tilde{p}_1}{\eta_1} \geq \frac{\tilde{p}_2}{\eta_2} \geq \cdots \geq \frac{\tilde{p}_K}{\eta_K}$. Choose as player 1 a subject that weakly underallocates tickets to bucket 1, so that $\nu_{1,1} \leq \eta_1 < \tilde{p}_1$ and select bucket k so that $\nu_{1,k} \geq \eta_k$. Consider the gains and losses to player 1 from shifting ϵ tickets from bucket k to bucket 1.

$$\begin{aligned} \text{Gains from Increasing } \nu_{1,1}\text{:} \quad & \left(\frac{\nu_{1,1}+\epsilon}{N\eta_1+\epsilon} - \frac{\nu_{1,1}}{N\eta_1}\right)\tilde{p}_1 = \frac{N\eta_1-\nu_{1,1}}{N\eta_1+\epsilon}\frac{\tilde{p}_1}{\eta_1}\frac{\epsilon}{N} \\ \text{Cost of Decreasing } \nu_{1,k}\text{:} \quad & \left(\frac{\nu_{1,k}-\epsilon}{N\eta_k-\epsilon} - \frac{\nu_{1,k}}{N\eta_k}\right)\tilde{p}_k = \frac{N\eta_k-\nu_{1,k}}{N\eta_k-\epsilon}\frac{\tilde{p}_k}{\eta_k}\frac{\epsilon}{N} \end{aligned}$$

We want to show that this deviation is profitable for some $\epsilon > 0$, for which it will be sufficient to show:

$$\frac{N\eta_1 - \nu_{1,1}}{N\eta_1}\frac{\tilde{p}_1}{\eta_1} = \left(1 - \frac{\nu_{1,1}}{N\eta_1}\right)\frac{\tilde{p}_1}{\eta_1} > \left(1 - \frac{\nu_{1,k}}{N\eta_k}\right)\frac{\tilde{p}_k}{\eta_k} = \frac{N\eta_k - \nu_{1,k}}{N\eta_k}\frac{\tilde{p}_k}{\eta_k}$$

This inequality holds by the assumptions of our construction:

$$\frac{\tilde{p}_1}{\eta_1} - \frac{\tilde{p}_k}{\eta_k} \geq \underbrace{\frac{\tilde{\nu}_{1,1}}{\eta_1}}_{\leq 1}\frac{\tilde{p}_1}{\eta_1} - \underbrace{\frac{\tilde{\nu}_{1,k}}{\eta_k}}_{\geq 1}\frac{\tilde{p}_1}{\eta_1}\frac{\tilde{p}_k}{\eta_k} \implies \frac{\tilde{p}_1}{\eta_1} - \frac{\tilde{p}_k}{\eta_k} > \frac{1}{N}\left(\frac{\tilde{\nu}_{1,1}}{\eta_1}\frac{\tilde{p}_1}{\eta_1} - \frac{\tilde{\nu}_{1,k}}{\eta_k}\frac{\tilde{p}_k}{\eta_k}\right)$$

(b) Suppose the IAM's distribution of tickets is proportional to $\tilde{p}$, so that $\frac{\tilde{p}_1}{\eta_1} = \frac{\tilde{p}_2}{\eta_2} = \cdots = \frac{\tilde{p}_K}{\eta_K}$, but two players are not playing the same strategy. At least one player has a profitable deviation.

Suppose player 1's allocaiton differs from the IAM odds. Let $\nu_{1,1} = \eta_1 - \xi$, $\nu_{1,2} = \eta_2 + \xi$, and consider the gains and losses to player 1 from shifting $\epsilon = \xi/N$ tickets from bucket 2 to bucket 1.

$$\begin{aligned} \text{Gains from Increasing } \nu_{1,1}\text{:} \quad & \frac{N\eta_1-\nu_{1,1}}{N\eta_1+\epsilon}\frac{\epsilon}{N} \\ \text{Cost of Decreasing } \nu_{1,2}\text{:} \quad & \frac{N\eta_2-\nu_{1,2}}{N\eta_2-\epsilon}\frac{\epsilon}{N} \end{aligned}$$

We will show this deviation is profitable by verifying that:

$$\frac{N\eta_1 - \nu_{1,1}}{N\eta_1 + \epsilon} > \frac{N\eta_2 - \nu_{1,2}}{N\eta_2 - \epsilon}$$

This inequality can be established by direct substitution:

$$\tfrac{N\eta_1 - \nu_{1,1}}{N\eta_1 + \epsilon} = \tfrac{(N-1)\eta_1 + \xi}{N\eta_1 + \xi/N}; \quad \tfrac{N\eta_2 - \nu_{1,2}}{N\eta_2 - \epsilon} = \tfrac{(N-1)\eta_2 - \xi}{N\eta_1 - \xi/N}$$

Then:

$$\frac{(N-1)\,\eta_1 + \xi}{(N-1)\,\eta_2 - \xi} > \frac{N\eta_1 + \xi}{N\eta_2 - \xi} > \frac{N\eta_1 + \xi/N}{N\eta_2 - \xi/N} \implies \frac{N\eta_1 - \nu_{1,1}}{N\eta_1 + \epsilon} > \frac{(N-1)\,\eta_2 - \xi}{N\eta_1 - \xi/N}$$

Part 3: Information Aggregation Properties

By the agreement assumption and the results of Parts (1) and (2), the IAM ticket allocation represents rational expectations for $E\left[\pi|s_1, \ldots, s_N, \alpha\right]$. Clearly, if every player places tickets proportionally to $\tilde{p}$, then the aggregated distribution of tickets in the IAM will match this distribution. As in Proposition 1, information accumulates with either a large number of players or with players' signal counts, allowing for the application of a Law of Large Numbers to show that the IAM aggregates information exactly.

Experimental Instructions (Not for Publication)

Forecasting Instructions – (date and time)

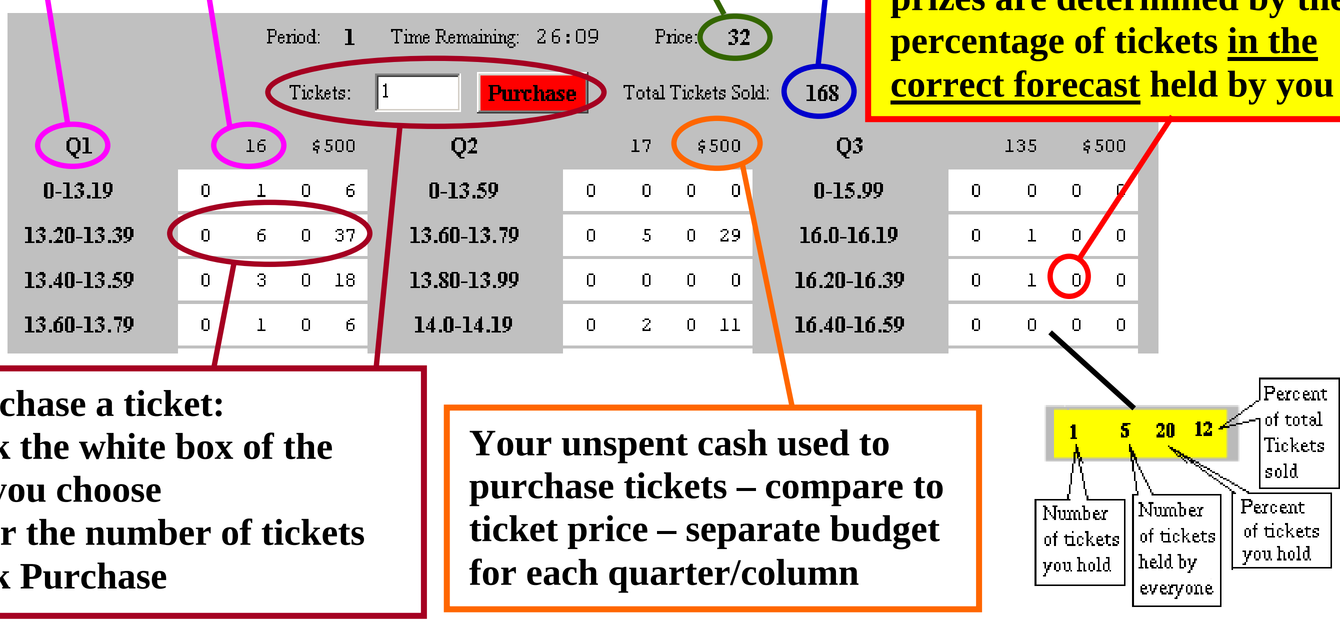
Forecasting (variable to be forecast)

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>:
<u>http://location</u> 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 9 digits 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 November 7 at 4:00 PM Pacific Time. 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.