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

Benjamin J. Gillen
Caltech

Charles R. Plott
Caltech

Matthew Shum
Caltech

November 8, 2015

Abstract

A new Information Aggregation Mechanism (IAM), developed via laboratory experimental methods, is implemented inside Intel Corporation in a long-running field test. The IAM, incorporating features of pari-mutuel betting, is uniquely designed to collect and quantize as probability distributions dispersed, subjectively held information. IAM participants' incentives support timely information revelation and the emergence of consensus beliefs over future outcomes. Empirical tests demonstrate the robustness of experimental results and the IAM's practical usefulness in addressing real-world problems. The IAM's predictive distributions forecasting sales are very accurate, especially for short horizons and direct sales channels, often proving more accurate than Intel's internal forecast.

*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, Columbia, EIEF Roma, MIT, Rice, UCSD, Universidad Carlos III Madrid, IIOC 2013, and the Econometric Society NASM 2013 for helpful comments.

1 Introduction

Many companies rely on internal forecasts of key financial and operational indicators, both as an input into production decisions and a leading indicator for managing market expectations about firm performance. 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 a completely different type of procedure – an information aggregation mechanism (IAM) based on decentralized competition, motivated by economic theory, and refined through testing by experimental economic methods. In relatively simple and special laboratory settings, the mechanism has been shown to effectively quantize, collect, and aggregate information. Here, we take up the challenge of testing the broader applicability and robustness of the same mechanism when operating in the much more complex environment of a Fortune 500 company. How well do information aggregation mechanisms work in a large-scale field setting? Are the results from the laboratory robust in the field? 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?

The purpose of an information aggregation mechanism (IAM) is to collect and aggregate information from a disperse collection of people. This task requires developing instruments to quantify participants' information and providing incentives that reward an individual for truthfully reporting private information rather than free-riding on others' information or misrepresenting that information for speculative purposes. The study of information aggregation in experimental economics laboratories has a long track record of successful mechanisms based on theoretical principles that have been refined through practical testing. The ability of markets to perform the information collection and aggregation functions and the sensitivity of this performance to the details of the market institution were first observed experimentally by Plott and Sunder (1982, 1988). Similarly, the possibility that markets could be specifically designed to perform the aggregation function and then implemented inside a business is well known (Chen and Plott (2002); Plott (2000)). The mechanism we study here shares institutional features with auctions, exchanges, and some betting processes with several substantial differences. Because some features are also found in pari-mutuel betting systems, we call it a *Pari-mutuel Information Aggregation Mechanism (IAM)*. Indeed, major features of this mechanism were developed to address challenges encountered when applying betting systems, traditionally designed for entertainment purposes, to information aggregation. In its field implementation, the IAM's design reflects an attempt to mitigate environmental complications that might inhibit the application of information aggregation mechanisms inside a business environment.

We report results from a long-running field experiment in which the IAM is implemented to forecast sales activity by Intel. As an international market leader in the hi-tech sector with annual revenues over $50 billion, Intel has an instantly recognizable brand name and its products are found in virtually all households in the country. Forecasts of product sales are important both operationally, in managing inventory and production for distribution, and financially, in managing market expectations for shareholder value. With myriad distribution channels, forecasting product sales for the organization is an incredible task requiring analysts to consolidate 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, ripe for collection and aggregation through economic mechanisms. At Intel, we applied the IAM to collect this information. In these mechanisms, the range of values that 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 actual sales take a value within a given bucket. Participants are allowed to buy tickets (up to a budget limit given below) and place them in any of the buckets. In this way, the distribution of tickets placed across the different buckets yields a measure of participants' "consensus" beliefs regarding the future realization of sales.

The main question we address is whether the IAM successfully aggregates information in a complicated real-life environment, providing an empirical validation of the mechanism's efficacy. This question is naturally motivated by the experimental evidence presented in section 2, where we also detail the implementation of the IAM at Intel. In an online appendix, we provide theoretical motivation for this inquiry by presenting a simplified model of the IAM in which information is successfully aggregated in the unique ex-post symmetric equilibrium. Allowing the possibility of information aggregation as a systematic outcome of the IAM, we proceed with our empirical investigation in Section 3 by presenting nonparametric (distribution-free) tests for the hypothesis of information aggregation.

Our results show that the IAM reliably characterizes the uncertainty regarding sales in "information-rich" environments where Intel has reliable and accurate information about potential sales. Specifically, when forecasting unit sales up to three months into the future, the IAM forecast distribution for sales matches the distribution of actual sales very closely. At horizons beyond four 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 presenting a "reverse favorite-longshot bias." Together, these findings support the hypothesis that Intel's information about sales uncertainty is much better at shorter horizons than longer horizons.

A similar pattern emerges when contrasting forecasts of products sold directly to large customers, for which Intel management is actively engaged, against those sold indirectly through distributors, for which Intel management has more limited information and influence. We find that forecasts of sales through direct channels are consistently more accurate and reliable than forecasts for sales through indirect channels. In short, these findings suggest that in relatively information-rich forecasting environments (as in forecasting near-term events, or Intel's direct sales), then the IAM effectively harvests and aggregates it.

We also find that the information reflected by the IAM is not readily available through other internal information sources. In addition to the IAM's aggregated forecasts, we also have access to an internally-prepared "official" forecast that serves as a contemporaneous benchmark providing an indicator of the quality of information within the organization at the time of the IAM run. Our second empirical result establishes that the IAM reveals considerable 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 63% 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. Though this robust outperformance may be partly attributable to other, potentially non-informational, purposes the official forecast serves in the organization, the information value of the IAM's forecasts is clearly substantial.

2 The Information Aggregation Mechanism Structure

A connection between markets and information transmission dates back to the foundations of economics.[1] Theoretical results suggest that markets are capable of collecting and aggre-

[1] The intuition connecting markets and information is clearly seen in Hayek (1945) and Fama (1970), with with Green (1973, 1977) providing some of the first steps towards its formal presentation. See Allen and Jordan (1998) for a review of this early formal development and general principles for the existence of rational expectations equilibrium.

gating information, though the price-formation process by which this result arises remains a vibrant research agenda surveyed in Vives (2010). The occurrence of information aggregation in markets suggests that it is possible to design an Information Aggregation Mechanism with the specific purpose of quickly and efficiently quantifying and collecting information that might be held by dispersed individuals in the form of vague and subjective intuitions.

2.1 Experimental Foundations for Information Aggregation

Experimental investigations of information aggregation began with the study of experimental markets. 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 information transmission and collection can go beyond the simple transfer of information to a process of quantizing and aggregating information distributed among multiple, independent sources. That is, market-based systems could effectively transfer "soft" information into a quantitative signal consistent with Bayes Law. Moreover, they pointed out that markets perform the collection and aggregation well if populated by a complete set of Arrow-Debreu securities.

The first application of a market based IAM inside a business was conducted in 1996 by Chen and Plott (2002) at Hewlett Packard Corporation to obtain forecasts about future sales. Possible sales levels were partitioned into states, and an Arrow-Debreu security was introduced for each state. Simultaneous continuous double auction markets were then introduced 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. Though this implementation was reported as successful, its use was limited due to difficulties in coordinating participation and managing the mechanism. Many features of the IAM discussed here emerged in response to difficulties encountered deploying market-based IAM's inside businesses.

The design of the IAM reported and studied here shares some features with pari-mutuel betting processes - hence the reference to pari-mutuel incentives. In a pari-mutuel 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 value of tickets sold for a bucket divided into the size of the purse, which display 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 pari-mutuel incentives in the IAM implemented at Intel are fundamentally different from those in partimutuel betting systems for entertainment purposes. 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), which demonstrated the importance of providing incentives for participants to buy their tickets early rather than waiting until the last second in an attempt to free ride on information supplied by others. 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 pari-mutuel betting.[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) as a device to mitigate the negative impact risk aversion might have on participation and information aggregation. Finally, the mechanism subject to neither self-financing nor individual rationality constraints, with Intel management funding the cash prize and dictating subject participation.

2.2 Implementing the IAM inside Intel

The IAM inside Intel differs in important ways from other types of prediction mechanisms, including market forms of organization that have flourished in recent years.[3] Many prediction

[2] Plott et al. (2003) (replicated in Kalovcova and Ortmann (2009)) demonstrated that the tendency to wait until the last second to buy tickets contributed to the creation of bubbles and retarded successful information aggregation. In a very simple pari-mutuel betting system information is transferred through a process of observing betting; this is examined by Koessler et al. (2012). Ottaviani and Sørensen (2004) present a theoretical analysis for the timing of bets in pari-mutuel betting systems, deriving results consistent with several of these experimental findings. Hu and Wallace (2013) present a two-stage mechanism for aggregating information that evokes features similar to dynamic pari-mutuel betting systems and, especially, the staged price increases in the IAM.

[3] 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 and other firms (Cowgill et al. (2009); Cowgill and Zitzewitz (2014)) to gauge employees' sentiments on everything from a company's performance to general industry issues.

markets function as security exchanges where participants buy and sell contracts that pay off depending on the realized value of the outcome of interest. In contrast, the tickets placed by IAM participants cannot be traded. The price speculation driving bubbles in markets, based on potential profits from changes in asset prices, cannot arise in the Intel IAM. At the same time, the literature studying prediction markets and experimental tests have shown how "thin markets," arising from a lack of coordinated participation, can inhibit information aggregation by slowing the market's convergence to equilibrium. In order to avoid thin markets and expedite equilibrium convergence, the timing of the IAM was adapted to be compatible with the schedules of participants for contemporaneous interaction.

The Intel IAM participants were selected according to the information they had about Intel operations that were relevant to the variable being predicted. Potential participants were informally assigned one of four classifications for the level their information: (i) Street, (ii) Intel general, (iii) Intel specific, and (iv) Intel technical. The last two of these classifications indicated participants with access to detailed information about the variable to be predicted, so "Prediction" groups typically consisted of 5-10 people from (iii) and (iv). (In contrast, groups containing many participants from (i) or (ii) tended to produce predictions that mirrored public knowledge, drowning out the informed signals from insiders.)

While participation was rewarded according to performance in the IAM, the levels of incentives did not compare with the pressures of busy schedules. Intel discovered early that it was important for management to tell selected participants that their participation in the IAM was valuable to Intel. Thus, management fiat is able to ensure self-selection of participants and individual rationality constraints do not distort the IAM's performance. Such selection issues are particularly prominent in public prediction markets for entertainment purposes, in which participants select themselves and voluntarily engage in the mechanism.[4] 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 not publicly announced.

We set up IAM's for sales forecasting that elicit participants' beliefs about variables (unit

[4] This phenomenon may be accentuated in horse-racing pari-mutuel 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 and using data from pari-mutuel markets to learn about these preferences.

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 2010 presidential election. 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. 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 $[x_{K-1}, \infty)$. Figure 1 presents an instruction sheet given to participants, which shows a partial screenshot illustrating the buckets and ticket information available to them during the IAM.

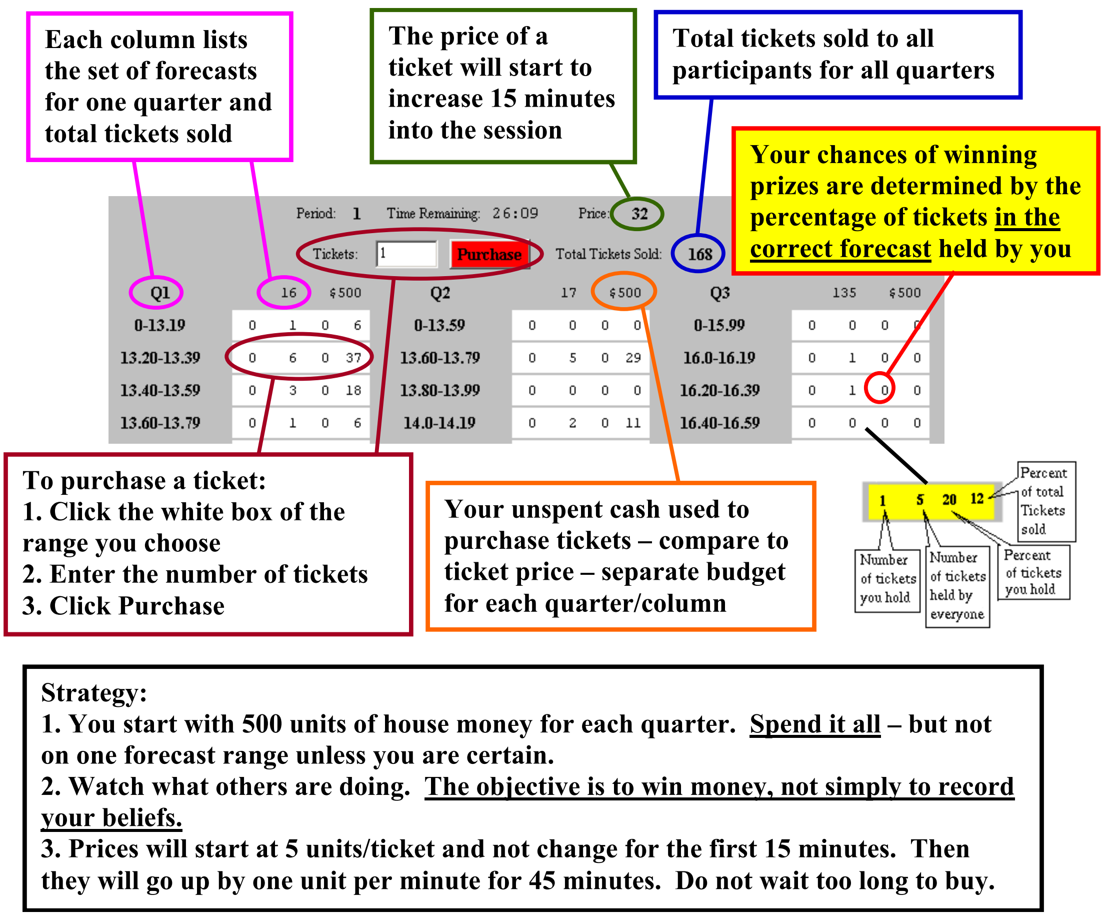

Figure 1: Screen Shot from Instruction Sheet for IAM Participants

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

The 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. The temporal discounting helps to mitigate these strategic incentives that otherwise hinder successful information aggregation.

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. 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.

3 Testing the Information Aggregation Mechanism

Departing from the mechanism design tradition of constructing a mechanism from first principles, the IAM's experimental development integrated modified features of familiar exchange and entertainment processes to create a system that facilitates information aggregation. This departure complicates presenting a complete theory of its construction, dynamics, and strategic foundations. However, an understanding of how information aggregation takes place emerges from the controlled exploration of the IAM's empirical properties.

In an online appendix, we present a tractable model based on the Dirichlet sampling process that accommodates a variety of empirical settings and provides some theory about the mechanism's behavior. The model characterizes how information aggregation arises in pari-mutuel-style mechanisms and how the IAM might respond in different environments. Even in this simplified setting, the characterization is approximate and a multiplicity of potentially exotic and asymmetric equilibria could arise in which information may or may not aggregate. Our analysis doesn't completely characterize this set of equilibria and a more thorough analysis could likely identify settings in which information aggregation does not occur. However, we do identify the unique symmetric ex-post Nash equilibrium in which information does successfully aggregate. Importantly, the results demonstrate that information aggregation is clearly not theoretically ruled out and establish that its existence is an empirical question.

Using data drawn from the Intel experience, we now turn to investigating the consistency between the IAM's reported distribution over sales and the revealed uncertainty in realized sales. There are two ways that we could empirically reject the hypothesis that the IAM successfully aggregates 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. We consider nonparametric (distribution-free) tests that compare the full distributions reported by the IAM to the frequency of realized outcomes. The distribution-free feature of our testing approach ensures that our results do not depend on potentially wrong specifications of the data-generating process.

As a preview of our results, the IAM performs well in markets where we expect the organization to have accurate and reliable information about future sales. For sales realized within the next three months and for products sold through direct distribution channels, the IAM accurately reflects the realized distribution over sales. In other settings where we find the information to be less accurate or reliable, the IAM performs less well; this is not surprising, as in these cases, there is "less information" for the IAM to aggregate.

3.1 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 broken down by the channel through which the products were being sold. 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.[5] 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. Realized quarterly sales are normalized by product line to average one unit quarterly sales with a one unit standard deviation.[6] 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 forecast 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 columns of Table 1 break down forecast accuracy by forecast horizon and by the product's sales channel. Several clear patterns emerge in these results. First, the accuracy of point forecasts for both the IAM and the official forecast improve as the forecast horizon narrows. Conversely, the accuracy of this information degrades as the forecast horizon lengthens. Second, a similar pattern distinguishes information quality across sales channels. In the

[5]Alternatives, such as the median or modal forecast yield largely the same results as reported in an online appendix. Due to the buckets in the mechanism, these forecasts are effectively interval-censored. We take the mid-point of the bucket as 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. We have considered different specifications for these buckets with little impact on our results.

[6]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 aggregated data publicly available. For each product, we normalized our data by subtracting the average realized sales, scaling by the standard deviation so that realized sales would have unit variance, and adding one to the normalized value. With this normalization, realized quarterly sales ranged from -1.59 to 2.66.

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 (RMSFE) is not directly comparable between the official and IAM forecasts. To address this timing issue, we report the official RMSFE excluding the least-informed (9 month) horizon forecast and the IAM RMSFE excluding the most-informed (1 month) horizon forecast. The columns report subsample results broken down by the forecast horizon at which forecasts are generated and the channel through which the sales were delivered.

	Full	Subsamples by Forecast Horizon				Sales Channel	
	Sample	Last Mth	1-3 Mth	4-6 Mth	7-9 Mth	Indirect	Direct
First Quarter	2006Q2	2006Q2	2006Q2	2006Q3	2006Q4	2006Q4	2006Q2
Last Quarter	2013Q3	2013Q3	2013Q3	2013Q3	2013Q3	2011Q4	2013Q3
Num of Markets	979	113	339	328	312	339	640
Num of Quarters	30	30	30	29	28	21	30
Average Sales	1.00	0.97	0.97	1.00	1.03	1.00	1.00
Std Dev Sales	1.00	1.02	1.01	1.00	0.98	1.00	1.00
Average Official	1.22	1.00	1.04	1.23	1.41	1.22	1.22
Std Dev Official	1.05	1.01	1.04	1.05	1.03	0.91	1.12
Official RMSFE	0.97	0.37	0.65	0.98	1.22	1.20	0.82
Average IAM Mean	1.16	0.96	1.01	1.15	1.33	1.11	1.19
Std Dev IAM Mean	1.00	1.02	1.00	1.00	0.97	0.86	1.07
IAM Mean RMSFE	0.87	0.27	0.55	0.89	1.11	1.08	0.74

Direct channel, forecasts are more accurate than those for sales through Indirect channels. Finally, at long forecast horizons and in indirect sales channels, there may not be substantial information for the IAM to collect and report. In direct sales channels and especially at short forecast horizons, Intel has high quality information about sales to aggregate.

3.2 Nonparametric Tests 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 $Y_{i,t}$ that, in comparison, is a severely restricted view into the data generating process that requires relatively large samples to provide a rich perspective of mechanism performance. We define successful information aggregation as characterized by a match between the empirical distribution of ticket placements in the IAM and the "true" probabilities of sales (whether in expectation or exactly). Letting $\mathcal{F}_{i,t|t-h}$ denote the information set of IAM participants regarding $Y_{i,t}$ in period $t-h$:

Hypothesis 1. *Information Aggregation Mechanism Accuracy hypothesis:*

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

where MN refers to the multinomial distribution, $\tilde{\eta}_{k,t|t-h}$ denotes the proportion of tickets placed in bucket k for the IAM run at time $t-h$ to generate the h horizon forecast for the t^{th} period's sales.

To test Hypothesis 1, we define the IAM's reported cumulative conditional distribution of $Y_{t|t-h}$ corresponding to the ticket placements at horizon $t-h$ by $\hat{G}_{i,t|t-h}(y) = \sum_{k=1}^{max\{\kappa|x_\kappa \leq y\}} \tilde{\eta}_{k,t|t-h}$. Then we transform the realized outcome $Y_{(i,t)}$ into its corresponding quantile in the conditional IAM distribution:

$$\hat{Q}_{i,t,h} \equiv \hat{G}_{i,t|t-h}(Y_{i,t}) \tag{2}$$

By translating the outcome into its conditional quantile from the IAM, we allow for the possibility that, for different horizons' forecasts of the same product quarter, the actual sales $Y_{i,t}$ may derive from different (conditional) distributions. Using the well-known property that the quantiles of a random variable are sampled uniformly, if the IAM distribution matches the true conditional probability distribution for $Y_{i,t}|\mathcal{F}_{i,t|t-h}$, then $\hat{Q}_{i,t,h} \sim_{H_0} U[0,1]$. That is, when we evaluate the IAM distribution at an actual realization of sales, that variable will be uniformly distributed on the unit interval under Hypothesis 1. This result allows us to consolidate observations across instances of the IAM's for different periods, horizons, and products to evaluate the hypothesis.

Accordingly, we can use a Kolmogorov-Smirnov test to evaluate whether we can reject the hypothesis that our sample of quantiles is drawn from a set of uniformly-distributed random variables. Analyzing the conditional quantiles renders the test robust to heterogeneity in the distributions across products, time, and information sets.[7]

3.3 Information Aggregation in Information-Rich Settings

Panel A of Figure 2 shows that the empirical distribution of IAM quantiles almost perfectly matches the uniform distribution 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 2 support the hypothesis, with a mean absolute deviation of less than 4% and the Kolmogorov-Smirnoff test statistic corresponding to a p-Value of 72%, 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 2'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. Visual inspection again verifies that the IAM quantiles match up well with the uniform distribution with a slight distortion in the tails. This distortion registers as statistically significant, with the Kolmogorov-Smirnov test reporting a p-Value of 2%.

3.4 Information Aggregation in Information-Scarce Settings

If participants have very little information relevant for predicting 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. Panels A and B in Figure 3 plot the quantiles of the forecast quantile distribution against

[7] Since the IAM quantiles are only available for a discretized support, these quantiles are technically only identified within a range. Our reported results treat the probability mass in a bucket as lying entirely on the minimum of that bucket, though our qualitative results are not sensitive to this treatment. While robust to heterogeneity, we 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.

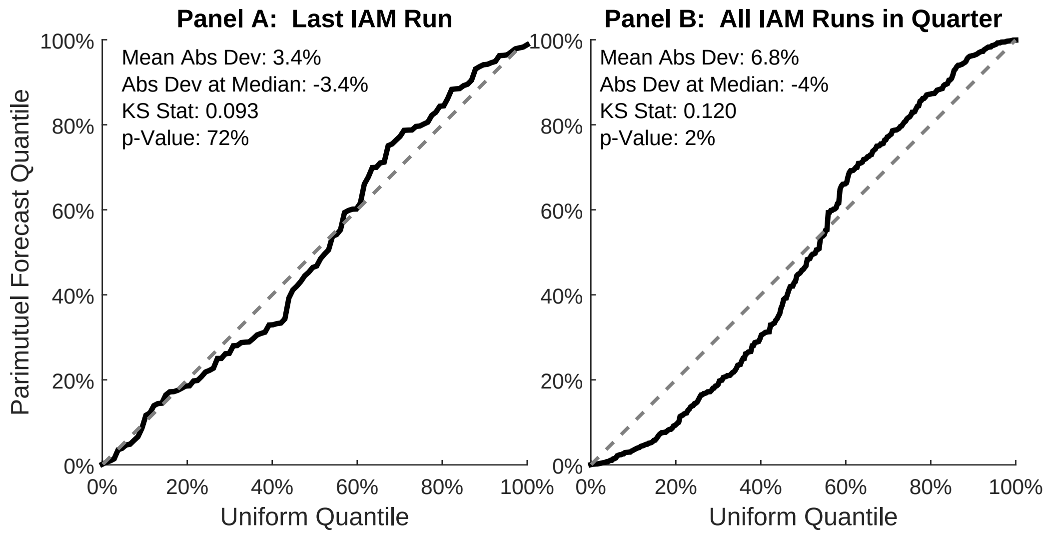

Figure 2: Quantile Plots for the Information Aggregation Mechanism
This figure presents the distribution of realized quantiles from the information aggregation mechanism defined in equation (2) against the theoretically accurate uniform distribution. Panel A reports the results for the last run of the 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.

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 heading into the great recession.

The S-pattern we observe indicating the IAM systematically understates tail probability events in the absence of useful 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.[8] In data patterns characterized by the FLB, the pari-mutuel 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

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

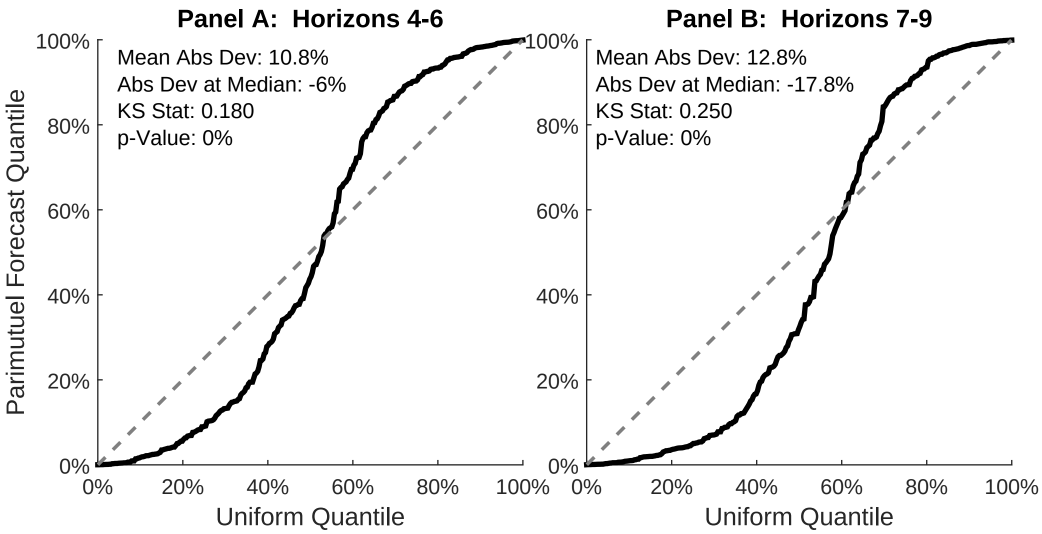

Figure 3: Quantile Plots for the Information Aggregation Mechanism
This figure presents the distribution of realized quantiles from the information aggregation mechanism defined in equation (2) against the theoretically accurate uniform distribution. Panel A reports the results for the IAM's run 4-6 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.

environments, including probabilistic misperceptions, risk preferences, belief heterogeneity, and information incentives.[9] Ottaviani and Sørensen (2010) present a strategic analysis of these biases in simultaneous, single-shot, pari-mutuel betting systems with persistent aggregate uncertainty in the distribution over states. In this model, pari-mutuel betting participants are partially informed about the conditional distribution for a random variable in addition to aggregate uncertainty about the outcome itself. 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 pari-mutuel 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

[9]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 (2015) show that belief heterogeneity among racetrack bettors can also induce a longshot bias in prices.

systematic underweighting of low-probability events arises in the pari-mutuel odds. This result indicates that the dynamic features of the IAM, which are not considered in Ottaviani and Sørensen (2010)'s analysis of a simultaneous betting, do not completely resolve these biases in information-scarce environments.

3.5 Information Aggregation across Sales Channels

As noted above, forecasts for sales made through direct channels are more accurate than those through indirect channels. Having observed the IAM's accuracy supported in short horizons, we'd expect similar support in the direct sales channel. Further, if the IAM's reverse-FLB at long horizons is driven by information quality, the reverse-FLB should present more clearly in the indirect channel than in direct-channel sales.

Figure 4 presents the quantile plots of the IAM against the uniform distribution broken down by sales channel. Panels A and B present the quantiles of the IAM for all forecasts within three months of the quarter end, breaking down the results from Figure 2's Panel B into direct and indirect sales. Panel A indicates minimal distortion of IAM quantiles in Direct Sales, which cannot be statistically distinguished from the uniform distribution at the 9% level. Panel B suggests the bias observed in Figure 2's Panel B is mainly driven by forecasts for indirect product sales, which generates a 1% p-Value in the Kolmogorov-Smirnov test. Panels C and D confirm the results from Figure 3 indicating the reverse-FLB is present in both direct sales forecasts and indirect sales forecasts at these long forecast horizons. Comparing the Mean Absolute Deviation across these environments indicates the reverse-FLB is slightly sharper for indirect sales at long horizons, but this marginal effect is swamped by the overall presentation of the reverse-FLB.

The results across these subsamples underscore both the difficulty of making predictions in information-scarce environments and the potential for consolidating information to make predictions in information-rich environments. Due to the many composite hypotheses being tested in this evaluation, there are any number of reasons the IAM performance could break down in information-scarce environments. 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 support the data provides to Hypothesis 1 in information-rich environments.

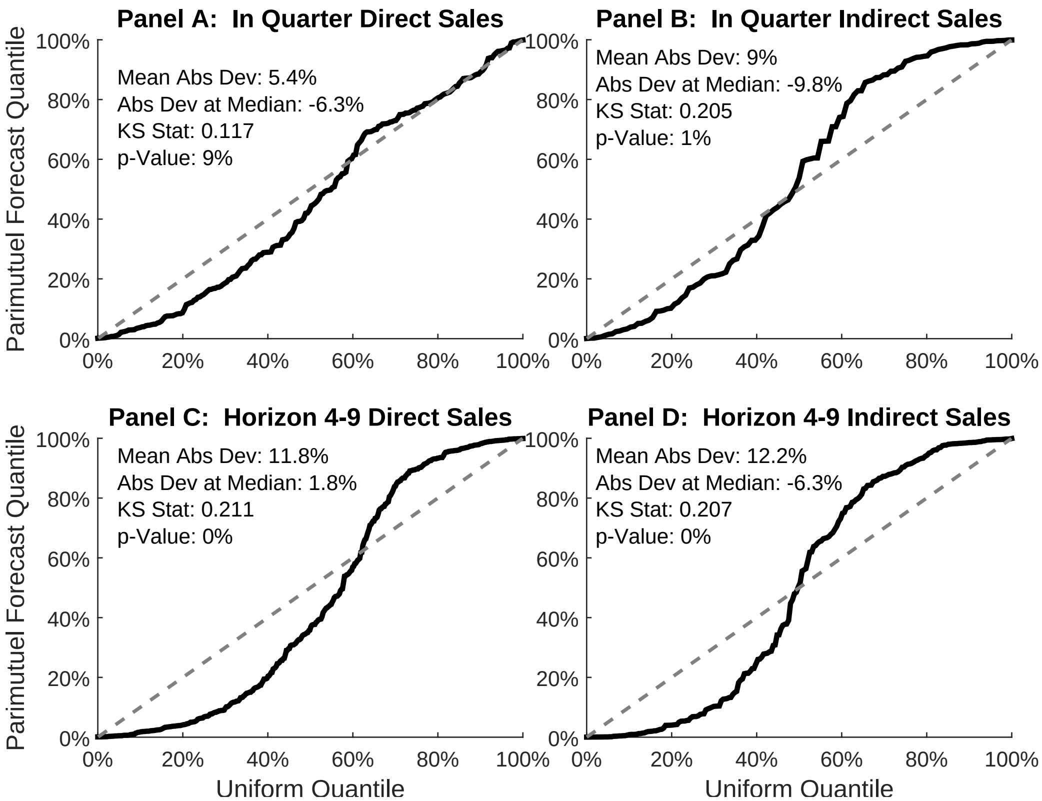

Figure 4: IAM Distributional Accuracy Across Sales Channels

This figure presents the distribution of realized quantiles from the information aggregation mechanism defined in equation (2) against the theoretically accurate uniform distribution. Panel A reports the results for all IAM runs predicting sales through the Direct channel during the quarter for which the sales are being forecast. Panel B reports results for IAM runs for predicting Indirect sales during the same period. Panels C and D provide the results for direct and indirect sales, respectively, at forecast horizons of four to nine months. 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 The Information Value of the IAM for Forecasting

We now examine the degree to which information available through the mechanism can help improve internal 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 confirm that 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

Table 2: Comparing Forecast Loss Across Mechanisms

This table presents Diebold-Mariano tests comparing the point forecasts from the official forecast and the mean information aggregation mechanism forecast. The Freq IAM Outperforms captures the frequency with which the IAM was more accurate than the official forecast. The Avg Abs Δ Loss reports the square root of the absolute average difference in the square error for the two forecasts, signed negatively when the official forecast outperforms the IAM. The DM-Statistic and p-Value report the Diebold-Mariano test statistic and p-Value using three-way clustered standard errors by product, period, and horizon.

	Num of Obs	Freq IAM Outperforms	Avg Abs Δ Loss(*100)	Diebold-Mariano Test t-Statistic	p-Value
Full Sample	979	63%	17.57	(2.61)	0.9%
Forecast Horizon					
1 Mth	113	65%	6.64	(2.44)	1.6%
2-3 Mths	226	62%	15.41	(2.04)	4.3%
4-6 Mths	328	62%	15.87	(2.32)	2.1%
7-9 Mths	312	65%	24.88	(2.74)	0.6%
Sales Channel					
Indirect	339	59%	28.79	(2.44)	1.5%
Direct	640	65%	11.63	(3.23)	0.1%

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 relative to 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.

4.1 Predictive Accuracy and Forecast Performance

We begin with a direct horse-race between the IAM Mean 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:

$$l^{(k)}_{i,t|t-h} = \left(\hat{Y}^{(k)}_{i,t|t-h} - Y_{i,t}\right)^2 \tag{3}$$

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

$$\delta_{i,t,h}^{(j,k)} = l_{i,t|t-h}^{(j)} - l_{i,t|t-h}^{(k)} \tag{4}$$

The DM test statistic corresponds to the t-statistic for the average $\delta_{i,t,h}^{(j,k)}$, using a robust variance estimator.[10]

Table 2 shows the IAM Mean 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 65% of the time. 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 63% of the time. The DM tests clearly show that the IAM Mean outperforms the official forecast in the full sample and in the one-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 in over 59% of the observations across subsamples. Contrasting performance across sales channels, the IAM more consistently outperforms the official forecast in Direct sales channels than in indirect sales channels.

4.2 Forecast Combination and Encompassing Tests

Beyond its consistent outperformance of the official forecast, the IAM also contains relevant information about sales not already reflected in the official forecast. To show this, we analyze the optimal combintation of the IAM forecast with the official forecasts into a single aggregated forecast.[11] This investigation provides a mechanism for a series of encompassing tests.

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.

[10] Given the myriad sources of dependence in our data, we apply Cameron et al. (2011)'s three-way clustering strategy on product-quarter, sales quarter, and IAM session, similarly to two-way clustering strategies for panels proposed by Petersen (2009) and Thompson (2011). While DM's initial derivation of the test establishes its asymptotically normality, Harvey et al. (1997, 1998) establish the expected result that a Student's t distribution better controls for size.

[11] Timmermann (2006) provides a good survey of the substantial literature on optimal forecast combination.

Table 3: Forecast Combination Regressions

This table presents estimates from the forecast combination regressions 5. 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 $\omega_{Official} = 1$. The tests $F(., 0, 1)$ and $F(., 0, 1)$ test the analogous restrictions without the zero-intercept restriction. All tests use three-way clustered standard errors by product, period, and horizon.

	Constant	IAM Mean Weight	Official Fcst Weight	Encompassing F-Tests	F(., 1, 0)	F(., 0, 1)	F(0, 1, 0)	F(0, 0, 1)
Full Sample	0.29	111%	-48%	F-Stat	3.69	32.44	5.50	47.49
Std. Error	(0.14)	(15%)	(16%)	p-Value	1.2%	0%	0.4%	0%
Forecast Horizon								
1 Month	0.07	116%	-20%	F-Stat	2.25	35.90	3.19	51.94
Std. Error	(0.06)	(16%)	(14%)	p-Value	8.6%	0%	4.5%	0%
2-3 Months	0.15	131%	-50%	F-Stat	2.96	19.89	4.10	27.36
Std. Error	(0.08)	(19%)	(20%)	p-Value	3.3%	0%	1.8%	0%
4-6 Month	0.31	88%	-27%	F-Stat	3.11	6.59	4.64	9.80
Std. Error	(0.13)	(28%)	(30%)	p-Value	2.7%	0%	1%	0%
7-9 Month	0.52	107%	-65%	F-Stat	4.54	16.17	6.19	23.62
Std. Error	(0.24)	(27%)	(24%)	p-Value	0.4%	0%	0.2%	0%
Sales Channel								
Indirect	0.65	104%	-65%	F-Stat	6.02	22.35	8.26	32.53
Std. Error	(0.22)	(23%)	(21%)	p-Value	0.1%	0%	0%	0%
Direct	0.16	111%	-39%	F-Stat	2.41	26.01	3.61	36.79
Std. Error	(0.1)	(17%)	(16%)	p-Value	6.6%	0%	2.8%	0%

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.

$$Y_{i,t} = \gamma + \omega_{IAM} \hat{Y}^{(IAM)}_{i,t|t-h} + \omega_{Official} \hat{Y}^{(Official)}_{i,t|t-h} \tag{5}$$

If the IAM encompasses the official forecast, 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 (multi-way clustered) 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.

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 the full sample estimates, the Official forecast is actually negatively weighted, though this negative weight is not statistically distinguishable from zero. Looking at subsamples of these results, 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.

5 Conclusions

This paper analyzes a field test of the Information Aggregation Mechanism that was developed and refined in laboratory experimental environments. The results underscore that the IAM performs well both in absolute terms and relative to other forecasts and that it does so for understandable reasons. **To our knowledge, the Intel IAM is now the longest-running implementation of an economically-motivated internal forecasting mechanism in industry. We attribute its long-running success to the unique features built on experimental development and experiences with the business applications of prediction markets.**

Intel's motivation for running the IAM derived from a desire to explore new approaches to producing the forecasts required for key strategic decisions within the company. Economics-based approaches (such as the IAM) are attractive because of the idea that market incentives might reduce potential biases and distortions in the existing forecasting procedure. From a practical perspective, the IAM has provided Intel with greater insight into its business conditions and improved the organization's ability to forecast sales. Overall, Intel's management has reacted to the findings in this paper with interest and surprise. Accordingly, the scale and scope of the IAM inside Intel may be expanded and broadened and its output incorporated into the official forecast reports.

At the same time, Intel management identified some challenges interpreting its results that apply to automated prediction systems generally. For changes in forecasts, management wanted a story, an explanation relating the numbers to the operational environment to guide their reaction to those changes. Management currently develops this insight through discussions about IAM results with key participants and business stakeholders, suggesting ways to further enhance prediction systems' value to an organization. Future implementations of the IAM at Intel could be supplemented with qualitative surveys of participants to provide a glimpse into the mechanisms and explanations for the observed forecasts. In response to observed changes in forecasts for an aggregated product's sales, another approach might tactically deploy the IAM for geographic- or market-specific segments of the product to identify whether changes are driven by a specific segment amenable to management intervention. Further development could apply the IAM to evaluate probabilities for joint events, for example, characterizing the expected completion date of a project conditional on the date at which a specific milestone is reached.

References

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

B. Allen and J. 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.

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

B. Axelrod, B. Kulick, C. Plott, and K. 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.

J. Berg, R. Forsythe, F. Nelson, and T. 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.

A. Cameron, J. Gelbach, and D. Miller. Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29(2):238–249, 2011.

K. Chen and C. 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 and E Zitzewitz. Corporate prediction markets: Evidence from google, ford, and firm x. *Mimeo*, 2014.

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

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

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

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

A. Gandhi and R. Serrano-Padial. Does belief heterogeneity explain asset prices: The case of the longshot bias. *The Review of Economic Studies*, 82(1):156–186, 2015.

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

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

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

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

D. Harvey, S. Leybourne, and P. Newbold. Tests for forecast encompassing. *Journal of Business & Economic Statistics*, 16(2):254–259, 1998.

D. Hausch and W. Ziemba. *Handbook of Sports and Lottery Markets*. North-Holland, 2008a.

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

T. Hu and N. Wallace. Information aggregation in a large two-stage market game. *Mimeo*, 2013.

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

K. Kalovcova and A. Ortmann. Understanding the plott-wit-yang paradox. CERGE-EI (Prague) Working Papers wp397, 2009.

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

M. Ottaviani and P. Sørensen. The timing of bets and the favorite-longshot bias. University of Copenhagen, FRU Working Papers 2004/12, 2004.

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

M. Petersen. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies*, 22(1):435–480, 2009.

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

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

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

C. Plott and S. Sunder. Rational expectations and the aggregation of diverse information in laboratory security markets. *Econometrica*, 56(5):1085–1118, 1988.

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

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

S. Thompson. Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics*, 99(1):1 – 10, 2011.

A. 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.

X. Vives. *Information and Learning in Markets: The Impact of Market Microstructure*. Princeton University Press, 2010.

K. 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.

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

Online Appendices (Not for Publication)

A A Model of the Information Environment and Aggregation

Here we present a simple model of learning based on the Dirichlet sampling process to characterize how information aggregation might arise in a pari-mutuel-style mechanism similar to our 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 probability that realized sales fall into a given bucket with a K-point multinomial distribution:

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

In this environment, agents are endeavoring to learn about the entire distribution of sales, as described by the set of unknown parameters $\pi = (\pi_1, \ldots, \pi_K)'$. The fact that agents must learn about an entire probability distribution distinguishes this learning environment from typical univariate learning models in economics. In this environment, agents' evolving beliefs about the unknown probabilities π corresponds to a "distribution over distributions," a modeling environment for which the Dirichlet is particularly well-adapted.

Suppose agents start off with a (common) prior that π follows a Dirichlet distribution with non-negative concentration 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:

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

Each agent updates her beliefs about π upon observing signals about the likelihood of different states. Specifically, an agent observes m_n signals $s_{n,1}, \ldots, s_{n,m_n}$, drawn independently from a multinomial distribution $MN(\pi)$. From these m_n signals, the agent can compute sample frequencies $\hat{p}_{n,1}, \ldots, \hat{p}_{n,K}$, where $\hat{p}_{n,k} = \frac{1}{m_n} \sum_{j=1}^{m_n} \mathbb{1}\left\{s_{n,j} = k\right\}$, the sample frequency with which the signal falls into the k-th bucket. Given these conjugate distributional assumptions, the posterior distribution for π conditional on these signals will also be Dirichlet:

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

In this setting, we want to characterize precisely what information aggregation means in terms of the underlying distribution of sales Y based on all agents' information. Adding 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 $\hat{p}_k = \frac{1}{M}\sum_{n=1}^{N}\sum_{j=1}^{m_n}\mathbb{1}\left\{s_{n,j}=k\right\}$ be the proportion of all signals in bucket k. By conjugacy, the aggregated posterior distribution across all N agents will again be Dirichlet:

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

This last posterior distribution, $p\left(\pi|s_1,\ldots,s_N,\alpha\right)$, 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 "bits" 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.

Given a large number of independent signals, so that $M \to \infty$, 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:

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

This allows us to define what "exactly successful" information aggregation 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.

A.1 Incentives for Information Revelation

Given our definitions above, we now examine how incentives might guide individual behavior to reveal their posterior beliefs in the IAM. Suppose the IAM at time t is in a state where each bucket x_k has $\eta_k^{(t)}$ tickets in it and denote the state vector of tickets across buckets by $\eta^{(t)}$. Denote agent n's interim posterior expected beliefs at time t, conditional on the IAM state and history up to t, by $p_n^{(t)} = \left[p_{n,1}^{(t)}, \ldots, p_{n,K}^{(t)}\right]'$. For a risk-neutral agent who has not yet placed any tickets, the marginal value of placing an additional ticket in bucket x_k is simply their posterior expectation of the realized outcome falling in that bucket divided by the number of tickets within the bucket:

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

In this case, the agent would maximize their payoff by placing their marginal ticket in the bucket that has the highest "odds" – that is, the largest posterior likelihood $p_{n,k}$ relative to the number of tickets that would be placed in the bucket $(1+\eta_k^{(t)})$. As this discussion makes clear, an agent who places tickets strategically primarily cares not about what the most likely outcome is, but rather the outcome for which the distribution implied by the IAM state differs most from his interim posterior. The most likely outcome is only preferred when a player's beliefs are consistent with the consensus, so that $p_{n,k}^{(t)} \propto \eta_k^{(t)}$, in which case the player has a slight preference for placing tickets in the bucket with highest probability. This preference slightly distorts incentives because of the finite number of tickets placed in the IAM, which results in agents evaluating incentives according to the number of tickets in the bin "plus one." This effect could induce a slight reverse-favorite longshot bias, but since this distortion clearly becomes negligible as the number of tickets grows.

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

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

which may also be distorted by the player's existing holdings, particularly if $\nu_{n,k}/\eta_k$ is large (i.e., if player n has placed a large share of the tickets in bucket k). However, in markets where information is spread diffusely across players, $\nu_{n,k}/\eta_k$ would be small and this distortion becomes negligible.

Note that incentives in the IAM are structured so that participants gain nothing from misrepresenting their beliefs. This feature presents an important element of the IAM's design that distinguishes it from continuous double auctions and facilitates its objective of aggregating information. When players disagree about the likelihood of events, the placement of tickets within the IAM ebbs and flows until a consensus forms. The IAM provides an intuitive and accessible mechanism for participants to communicate their beliefs quickly and efficiently. As long as two players disagree, they will be able to express that disagreement by placing more tickets. As this dynamic highlights the disagreement, players update their beliefs until they converge on a consensus distribution over states implied by the IAM.

Result 1 (Incentive Compatibility of Reporting Information in the IAM)**.**
Incentives in the IAM encourage participants to place tickets in the bins for which they most disagree with the probabilities implied by the distribution of tickets already placed within the IAM.

A.2 Information Aggregation as an Equilibrium Property

The incentives identified in the previous section indicate that subjects are encouraged to place tickets in a manner that most expresses their disagreement with the state of the IAM. We now consider how those incentives interact with the nature of equilibrium in the IAM. Suppose all information in the system is publicly revealed, so that every participant in the IAM agrees that the posterior distribution for π is given by equation (A.4). 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 us to abstract from the complications induced by strategic communication and the Folk theorem for dynamic settings, 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 1. *[Information Aggregation as an Equilibrium Property]*

- *Suppose the information aggregation environment is characterized by the distributional assumptions embedded in equations A.1 - A.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.*

The proof of this proposition is given at the end of this section. Proposition 1 indicates that, if private information can be effectively communicated, the IAM's incentives will lead to information aggregation through equilibrium once agreement obtains. 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.*

The main challenge in establishing information aggregation as a necessary property of equilibrium lies in the complexity of dynamic behavior and determining beliefs off the equilibrium path. Given the dynamic nature of the IAM and the complexity of beliefs and strategies, we'd expect a multiplicity of potentially exotic and asymmetric equilibria that we could not hope to characterize completely. That said, the purpose of our analysis is not to provide an exhaustive characterization of all equilibria in the IAM nor do we intend to relate these equilibria to properties of mechanisms typically associated with prediction and financial markets. Rather, our analysis is restricted to motivating the expectation that Intel's experience in implementing the IAM can inform other complex organizations designing systems to address information aggregation problems. To this end, though the equilibrium analysis above is incomplete, we do establish the possibility of information aggregation and characterize the incentives by which it presents a natural outcome of the IAM.

We conclude this section with a remark on the role of robust theoretical foundations for information aggregation in our application. The mechanism design literature has a rich tradition of mechanisms in which information aggregates in the unique equilibrium satisfying individual rationality and incentive compatibility constraints. Despite the robustness of these devices' theoretical properties, their implementation in complex settings is constrained by their complexity as well as their lack of familiarity to mechanism participants. More complex mechanisms have yet to pass through the crucible of experimental validation by which the rules and incentives of the IAM have been refined and, in so doing, these mechanisms will almost certainly require practical refinements as a result. The pari-mutuel incentives and dynamic elements of the IAM may complicate a complete characterization of its equilibria, but they do present established institutions that facilitate participant interactions. This feature is not exclusively a consequence of theoretical intuition, but relies critically on the experimental foundations for the IAM's design.

A.3 Proof of Proposition 1

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:

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

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

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

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

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.

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

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 allocation 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. If information accumulates with either a large number of players or with players' signal counts, so that a Law of Large Numbers applies, the IAM would aggregate information exactly.

B Forecast Evaluation for Other Predictive Measures from the IAM

This appendix replicates the forecast evaluation results from the main body of the paper using the Median and Mode from the IAM in place of the Mean of the IAM's Distribution.

Table B.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 Median and the Mode of the IAM Distribution. The columns report results broken down by the forecast horizon at which forecasts are generated and the channel through which the sales were delivered.

	Full	Forecast Horizon				Sales Channel	
	Sample	Last Mth	1-3 Mth	4-6 Mth	7-9 Mth	Indirect	Direct
First Period	200602	200602	200602	200603	200604	200604	200602
Last Period	201303	201303	201303	201303	201303	201104	201303
Num of Obs	979	113	339	328	312	339	640
Num of Qtrs	30	30	30	29	28	21	30
Average Sales	1.00	0.97	0.97	1.00	1.03	1.00	1.00
Std Dev Sales	1.00	1.02	1.01	1.00	0.98	1.00	1.00
Average Median	1.16	0.96	1.01	1.16	1.34	1.11	1.19
Std Dev Median	1.01	1.02	1.00	1.01	0.98	0.86	1.08
Median RMSE	0.88	0.27	0.55	0.89	1.12	1.08	0.75
Average IAM Mode	1.15	0.97	1.01	1.13	1.33	1.09	1.18
Std Dev IAM Mode	1.01	1.01	1.01	1.01	0.98	0.88	1.07
IAM Mode RMSE	0.89	0.29	0.57	0.89	1.13	1.10	0.75

Table B.2: Comparing Forecast Loss Across Mechanisms

This table presents Diebold-Mariano tests comparing the point forecasts from the official forecast and the median 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 three-way clustered standard errors by product, period, and horizon.

Panel A: Median IAM Forecast

	Num of Obs	Freq IAM Outperforms	Avg Abs Δ Loss(*100)	Diebold-Mariano Test t-Statistic	Diebold-Mariano Test p-Value
Full Sample	979	60%	16.86	(2.53)	1.2%
Forecast Horizon					
1 Mth	113	67%	6.71	(2.46)	1.6%
2-3 Mths	226	58%	15.24	(1.88)	6.1%
4-6 Mths	328	58%	15.77	(2.29)	2.2%
7-9 Mths	312	61%	22.84	(2.40)	1.7%
Sales Channel					
Indirect	339	56%	27.86	(2.27)	2.4%
Direct	640	63%	11.02	(3.62)	0%

Panel B: Mode IAM Forecast

	Num of Obs	Freq IAM Outperforms	Avg Abs Δ Loss(*100)	Diebold-Mariano Test t-Statistic	Diebold-Mariano Test p-Value
Full Sample	979	59%	15.32	(2.42)	1.6%
Forecast Horizon					
1 Mth	113	61%	5.50	(2.28)	2.4%
2-3 Mths	226	59%	12.88	(1.65)	10%
4-6 Mths	328	58%	15.64	(2.26)	2.4%
7-9 Mths	312	59%	20.31	(1.87)	6.2%
Sales Channel					
Indirect	339	52%	24.69	(1.97)	4.9%
Direct	640	63%	10.36	(5.25)	0%

Table B.3: Forecast Combination Regressions

This table presents estimates from the forecast combination regressions 5. 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 $\omega_{Official} = 1$. The tests $F(.,0,1)$ and $F(.,0,1)$ test the analogous restrictions without the zero-intercept restriction. All tests use three-way clustered standard errors by product, period, and horizon.

Panel A: IAM Median Forecast

	Intercept	IAM Median Weight	Official Fcst Weight		F(., 1, 0)	F(., 0, 1)	F(0, 1, 0)	F(0, 0, 1)
Full Sample	0.30	106%	-43%	F-Stat	4.05	43.40	5.86	58.02
Std. Error	(0.14)	(14%)	(14%)	p-Value	0.7%	0%	0.3%	0%
Forecast Horizon								
1 Month	0.07	114%	-18%	F-Stat	2.91	47.78	3.96	66.33
Std. Error	(0.06)	(16%)	(14%)	p-Value	3.8%	0%	2.2%	0%
2-3 Months	0.15	122%	-41%	F-Stat	3.58	28.95	4.73	38.04
Std. Error	(0.08)	(16%)	(16%)	p-Value	1.5%	0%	1%	0%
4-6 Month	0.32	90%	-29%	F-Stat	3.25	7.54	4.87	11.29
Std. Error	(0.13)	(25%)	(28%)	p-Value	2.2%	0%	0.8%	0%
7-9 Month	0.53	98%	-58%	F-Stat	4.35	12.70	5.79	18.21
Std. Error	(0.24)	(31%)	(28%)	p-Value	0.5%	0%	0.3%	0%
Sales Channel								
Indirect	0.65	93%	-56%	F-Stat	5.50	21.23	7.59	30.75
Std. Error	(0.22)	(23%)	(20%)	p-Value	0.1%	0%	0.1%	0%
Direct	0.16	111%	-40%	F-Stat	3.17	36.24	4.60	46.46
Std. Error	(0.1)	(17%)	(15%)	p-Value	2.4%	0%	1%	0%

Panel B: IAM Mode Forecast

	Intercept	IAM Mode Weight	Official Fcst Weight		F(., 1, 0)	F(., 0, 1)	F(0, 1, 0)	F(0, 0, 1)
Full Sample	0.30	72%	-11%	F-Stat	2.66	17.98	3.93	26.93
Std. Error	(0.15)	(14%)	(15%)	p-Value	4.7%	0%	2%	0%
Forecast Horizon								
1 Month	0.04	95%	1%	F-Stat	0.31	52.63	0.44	67.17
Std. Error	(0.06)	(9%)	(9%)	p-Value	81.6%	0%	64.6%	0%
2-3 Months	0.17	96%	-17%	F-Stat	2.28	19.32	3.38	25.69
Std. Error	(0.09)	(20%)	(17%)	p-Value	8.1%	0%	3.6%	0%
4-6 Month	0.33	68%	-7%	F-Stat	3.99	6.39	5.95	9.45
Std. Error	(0.14)	(21%)	(25%)	p-Value	0.8%	0%	0.3%	0%
7-9 Month	0.53	50%	-12%	F-Stat	3.95	7.51	5.15	9.75
Std. Error	(0.24)	(27%)	(27%)	p-Value	0.9%	0%	0.6%	0%
Sales Channel								
Indirect	0.67	69%	-34%	F-Stat	4.90	19.46	6.76	27.60
Std. Error	(0.23)	(21%)	(18%)	p-Value	0.2%	0%	0.1%	0%
Direct	0.17	70%	1%	F-Stat	1.81	27.46	2.21	40.75
Std. Error	(0.12)	(16%)	(11%)	p-Value	14.3%	0%	11%	0%

C Experimental Instructions

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.