Association for Information Systems
AIS Electronic Library (AISeL)

ICIS 2010 Proceedings

International Conference on Information Systems (ICIS)

2010

PREDICTIVE MODEL MARKETS: DESIGN PRINCIPLES FOR MANAGING ENTERPRISE-LEVEL ADVANCED ANALYTICS

Sule Balkan
Arizona State University, sule.balkan@asu.edu

Michael Goul
Arizona State University, michael.goul@asu.edu

PREDICTIVE MODEL MARKETS: DESIGN PRINCIPLES FOR MANAGING ENTERPRISE-LEVEL ADVANCED ANALYTICS

Completed Research Paper

Sule Balkan
W. P. Carey School of Business
Arizona State University
Main Campus
PO BOX 874606
Tempe, AZ 85287-4606
Sule.Balkan@asu.edu

Michael Goul
W. P. Carey School of Business
Arizona State University
Main Campus
PO BOX 874606
Tempe, AZ 85287-4606
Michael.Goul@asu.edu

Abstract

As advanced analytics penetrate a wide range of business applications, companies face the challenge of managing analytics-based assets, such as predictive models. Tasks ahead include model selection, scoring and deployment planning. One way to optimize model selection is to tap the combined knowledge of company staff through a "prediction market," a virtual market designed to reveal participants' aggregate wisdom by seeing where people "invest" their money. In the context of predictive-model selection, this paper refers to such devices as predictive-model markets. This paper examines design possibilities for building experimental markets that can ultimately be used to test whether predictive-model markets will improve model selection and deployment. The researchers test two types of incentives for participation: economic and social. Study results indicate that such markets can effectively work using either; a surprising finding is that social incentives did not improve effectiveness when added to economic incentives.

Keywords: Advanced analytics, prediction markets, predictive models

Introduction

The realm of advanced analytics encompasses predictive modeling, data mining, scoring, text analytics, forecasting, optimization, simulation and experimental design. According to The Data Warehousing Institute, 38% of organizations currently are practicing advanced analytics, and 85% say they will be practicing advanced analytics within three years [Russom, 2009]. Constantly changing business environments, combined with organizational imperatives to cut costs and generate sales, are spurring interest in these business-intelligence (BI) tools. Another driver is intense competition leading to tight margins that need to be closely monitored and controlled.

In addition, there is an increasing need to better harvest knowledge from growing volumes of collected data. That data-volume growth is fueled by more efficient and cost-effective technological capabilities that support tracking and tagging, as well as remote data capture using sensors, and data capture that accrues from fielding an increasingly mobile workforce.

Furthermore, there is a widespread shift to pervasive business intelligence using embedded analytics that are integrated into processes. Increasingly, these are key to campaigns conducted across a wide variety of business application areas [e.g., Watson and Wixom, 2007]. Because advanced analytics are becoming pervasive, and they now are commonly deployed enterprise-wide, there is an important need for new strategies to manage analytics-based assets.

Recent work sheds light on new analytics-based asset-management approaches. Some researchers assert that as the number of models increases to support more and more business objectives, so does the requirement to manage these models reliably and securely as valuable corporate assets [Chu, 2009 and Wei, et al., 2009]. Managing such assets will require the following design properties:

1. There needs to be a secure, centralized repository for storing and organizing models based on a company's business objectives.
2. Within that repository, the assets need to be organized by the various company campaigns in which they are deployed.
3. Methods for managing that repository need to support a variety of different types of advanced analytics, e.g., predictive models, classification models, segmentation and rules-based models.
4. There needs to be a way to delineate Champion Models (those that are in use in campaigns) from Challenger Models (those that may be applied in a campaign in the event the Champion model's performance decays).
5. The repository needs to provide a governance structure so that models can be tracked in terms of a lifecycle: where they are in their development, whether they have been validated, and where they've been deployed. There is also a need for information about each model that can assist in determining when a model might need to be retired and what new models should be used instead.

Few, if any, design principles for advanced analytics repositories have been addressed in the literature. New research is needed to examine and expand on achieving the five repository properties identified above. For example, extensions could support the direct linkage of model scoring to both batch and on-demand applications. Likewise, an extension could enable comparisons of Champion and Challenger Models on a variety of performance dimensions. And, extensions could support the staging and testing of model workflows whereby model dependencies can be examined and improved. Perhaps the repository should support management of model versions. In statistical theory, there are models referred to as fused models whereby multiple models can be combined (or their main attributes combined) in the design of a new model [e.g., Bonissone, et al., 2008]. The construction of such meta-models could be supported. Finally, since there are many enterprise stakeholders who are involved in building models, their roles and levels of engagement with such a repository represent areas that are ripe for new research. Overall, however, the bottom-line objective of a repository is to support campaigns that meet business objectives. Any extensions must be targeted to this directive.

In this paper, we advance design principles for enterprise-level management of advanced analytics to deal with several of the extensions discussed above. We focus on what has been referred to as "information markets" or "prediction markets," and we use these terms interchangeably throughout the paper. In prediction markets, stakeholders contemplate and make transactions within a virtual market, and they place bets on known contracts with unpredictable outcomes. These bets then are examined by policymakers and decision makers who leverage the collective knowledge that is expressed through the distribution of bets.

In the next section of the paper, we discuss prediction markets with an eye towards the design-principle possibilities such markets might facilitate for enterprise-level management of advanced analytics assets, specifically those possibilities related to the management of predictive models. We examine the potential value of adapting the prediction-market approach, the nature of betting (and the associated contracts) that might be embedded in such an adaptation, the different market signaling that would be central to the adaptation, and the incentives for stakeholders to participate truthfully. Of these, one of the most important issues to address first is the matter of incentives. We contrast economic motivators with a combination of economic and social catalysts in an empirical study. The study methodology and results are described and then discussed in subsequent sections of the paper.

Overall, we have applied a storyboarding design methodology to build dashboards for conducting our predictive-market research in the enterprise-level advanced-analytics context. We conclude with a series of challenges for what our findings suggest in terms of next steps in what will most certainly be an important business-intelligence research stream: the evolution of design elements for advanced-analytics repositories intended to serve the enterprise. In our research, we call such repositories predictive-model markets. A predictive-model market exhibits design essentials one through five noted above, and it is a prediction market in which enterprise stakeholders can contribute to the repository of models, examine the performance of these models and bet on future performance in an effort to fine-tune model selection and development.

Related Work: Tools for Optimizing Advanced Analytics

Three areas of related work lay the foundations for our research on predictive model markets: prediction markets themselves, dashboards as informational devices, and collaboration in decision-making efforts. Combined, the areas of research present a set of design possibilities for building experimental markets that can ultimately test the most important research question: Can predictive-model markets improve the enterprise-level deployment of the most effective advanced analytics?

Prior work has established that model deployment is an under-investigated area of predictive-modeling research, and this is most likely due to the inadequacies of many historical business-intelligence ecosystems to provide sufficient and suitable technological support for wider-scale use of advanced analytics [Balkan and Goul, 2010]. A design-science approach is a viable method for advancing early research. Our first objective is to create a viable predictive-model market within which to conduct experiments on important design principles.

Prediction Markets

Prediction markets gained notoriety when a policy-analysis market was created by the Defense Advanced Research Projects Agency (DARPA) to determine if the approach could help to predict future geopolitical events [Looney, 2003]. The market was unique in that it was designed specifically for information aggregation and revelation. Similar predictive markets have since evolved, and many reference Surowiecki's The Wisdom of Crowds, a book maintaining that decisions made using the aggregate wisdom of a group often are better than decisions that individuals in the group could have made alone [2004].

Berg and Rietz [2003] define prediction markets as "markets that are designed and run for the primary purpose of mining and aggregating information scattered among traders and subsequently using this information in the form of market values in order to make predictions of the future."

Tziralis and Tatsiopoulos surveyed the prediction-market literature from 1991-2006, and identified twenty-seven theoretical papers, seventy-two on applications, twenty law and policy reviews and thirty-three descriptive papers, theses, dissertations or conference proceedings directly related to prediction markets [2007]. Today, The Journal of Predictive Markets published by Buckingham Press covers recent research related to this topic, and it featured a special issue in 2009 devoted to the study of predictive markets in corporate settings.

Several popular predictive markets exist today and are accessible through the Internet. These include the Hollywood Stock Exchange (www.hsx.com), Iowa Electronic Markets (http://www.biz.uiowa.edu/iem/index.cfm) and Foresight Exchange (www.ideosphere.com). Many prediction markets rely on virtual currency, and many, such as the Hollywood Stock Exchange, front bettors with a play-money stake. Researchers have found that there are no differences in the accuracy of prediction markets whether they are based on play or real money [Servan-Schreiber, et al., 2004].

Key design features of predictive markets that have a potential bearing on predictive-model markets include three areas where design options have been documented in the literature. The first area relates to the type of contract that is used as the basis for the market. The second area relates to the incentives that are central to the engagement of market participants. The third category involves the mechanisms that make the market operate. For each of these, the design options are briefly described as follows:

Contract choices – Design choices relevant to the payoffs associated with the outcomes of future events. Each contract is intended to reveal the market's expectation in order to support decision making. Basic design options include:

- Winner-take-all
- Index (the payoff varies over some continuous distribution based on a number that rises or falls)
- Spread (betting on a cutoff that determines whether an event occurs; as in sports betting where the spread of a score is the basis for a bet)
- Exotic (betting on unique combination of events)

Incentive choices – Design choices relevant to the nature of incentives for participants to make truthful bets and to participate and engage in the prediction market (in both performance and effort). Basic design options include:

- Fixed payments
- Performance-based incentives
- Incentives using real money
- Incentives using play-money
- Economic incentives alone
- Economic and social incentives combined
- Plausible combinations of the incentives mentioned above

Mechanism choices – Design choices relevant to the scoring rules applied to market events, the appropriate signals the market uses to convey information, performance of the market over time, the state of the market at some reporting-time interval, the possibility of informing and facilitative arbitrage, and determinations of how to reflect market and environmental volatility. As an example of design choices, a market might use a mechanism that focuses on stock price. In such a market, scoring rules would be required for adjusting and changing stock prices based on demand, performance and other factors.

Dashboards

Processes related to building and deploying dashboards are relevant to predictive-model markets because they are well understood in the business intelligence community, the information conveyed in a predictive-model market can be organized as a dashboard, and alternative dashboard configurations can be leveraged to conduct experiments to investigate various design principles. Dashboards also provide a lens into additional design options in the construction of a predictive-model market. For example, four areas of design choices related to building dashboards provide a viable way to consider different operationalizations of predictive-model ideals in the advanced-analytics area.

First, considerations of key performance indicators (KPIs) are a major concern in dashboard-development design [Eckerson, 2006]. Second, dashboard granularity is vital. Granularity addresses level of detail provided in a dashboard [Nichols, et al., 2009]. Third, a key dashboard consideration is how to best make use of visualization [Bellamy, et al., 2007]. Finally, dashboards, like prediction markets, rely on mechanisms that facilitate user control of representations and processes. For each of these, the design options for a predictive model market are listed below:

Performance measures (KPIs) choices – Design choices include the nature and type of the key indictors to be included in the market tracking system that are transparent to market participants. There are numerous examples related to both the market and to the domain in which the dashboard applies. From a market perspective, these could include stock prices, trading volume, historical profits and more. From a domain perspective, these could include Champion model performance as measured by lift values by decile, Gini coefficients, Gini coefficient decay and model type (e.g., decision tree, neural network, or logistic regression).

Granularity choices – Design choices related to the level of detail to provide in a dashboard, the hierarchical options for presenting dashboard information and the facilitation of operations like drill-down. The granularity decision is influenced by users' cognitive styles, the possibility of information overload, and the speed by which information needs to be acted upon. This also relates to the design of controls needed to navigate the dashboard.

Visualization choices – Design choices relevant to dashboard usability. These choices are made through usability studies involving stakeholders, matching visual representation appropriateness to different types of data, etc. The nature of KPIs is also important for the visualizations that are used.

Mechanism choices – Design choices related to whether to enrich dashboards with added functionality and information. For example, should you facilitate the sharing of user comments, or do you want to display rules and alerts? And, how much metadata should be included in representations including visualizations? Mechanism choices such as these can be significantly affected by the intended role of the dashboard (e.g., to serve as a change agent, to control operations, to support root-cause analysis, to do straightforward monitoring, etc.).

Collaboration and Social Incentives

While dashboard-mechanism design choices include selecting whether users can share comments, there are many other design options that could support collaboration and provide for social incentives in a predictive model market. For example, Davenport and Prushak address many of the incentives that may encourage knowledge sharing from a knowledge-market perspective (e.g., reciprocity, street credentials, reputation, and altruism). Others have discussed how Web 2.0 and 3.0 capabilities might influence knowledge sharing. Malone et al. have included prediction markets in the broader context of approaches that harness collective intelligence and address collaboration in this context, as well [2010]. While it is beyond the scope of this paper to discuss all design choices related to collaboration and social incentives that could be relevant to predictive model markets, such design features could be an essential component. In the next sections of this paper, we deal directly with this issue.

Constructing Predictive-Model Markets

For our experiments covered in this paper, we have adapted the approach of Weinhardt, Holtman et al. [2003] and Weinhardt, Neumann, et al. [2006] who described the development of markets through a method called market engineering. The first several phases of their approach include environmental analysis, design and evaluation. In evaluation, there is typically consideration of the conceptual model of a market system. The model is to be examined via economic tests to determine the outcome performance of the market and assess the business model. This phase is supported by analytical and experimental evaluation methods. The tests are then used to determine if the business model is acceptable, if scoring rules need to be adjusted, etc. In the following, we explain the design features of the dashboard we ultimately adopted in the context of the design options previously discussed in the "Related Work" section of this paper. These design options were operationalized in a companion case study called The Wales Market, which is excerpted and attached as Appendix 1. In the Wales Market case, we recount how a fictional BI executive named Yolanda Wales creates a market in which participants bet on the performance of various predictive models by buying shares in those models. The models in the case were used to guide casino-marketing efforts. In the following, we discuss this case in terms of design option choices from above:

Contract Type: The Wales Market reflects payoffs in the form of stock prices that rise or fall based on the performance, supply, and demand of shares for a particular market over a particular decile. A decile is a demarcation of ten percent; for example, ten percent of targets as scored by a predictive model.

Incentives: The Wales Market is designed to investigate two hypotheses related to incentives with fixed design features. We examine a market scenario with only economic incentives and a market scenario with economic and social incentives combined. For the Wales Market, we fixed the incentive type to performance-based incentives, we did not allow for arbitrage, we did not include signals regarding external market volatility, and we held a number of other design options constant.

Mechanisms (Market): We included many market signals in the Wales Market case to reflect a campaign-based scenario with both a Champion model and multiple Challenger Models. We included the application of a Champion model to what we refer to as an action window, and we limit the size of that window to three deciles. This means the campaign only made promotional offers to subjects in the first three deciles. We include information about the model type, the stock price, the cost of the model in a campaign, the revenue from the application of a model, and we show what we refer to as the dispersed revenue. Dispersed revenue refers only to Challenger Models. Since they weren't actually used in the campaign, the Challenger Models' real revenues can only reflect the revenues of individual targets who were scored within in a decile of the action window associated with the Champion model. We also used a fixed betting scheme where a stated amount for the total of all possible bets was fixed. There were no required rules for the way bets could be allocated.

Performance Measures: Challenger models have unknown revenues for their targets who weren't scored into the action window deciles of the target model. This unknown reflects one approach to assessing the risk associated with the adoption of a Challenger Model over the Champion model. Other approaches to assessing risk include the fact that model performance is fairly well known to deteriorate in deciles beyond the first two or three. The performance measures therefore reflect the overlap of the Challenger Models with the Champion model by using color codes. There are also tables to show the overlap. Stock prices are included as is the number of prior times a model has been used as a campaign's Champion model.

Granularity: We included a single prior performance of the Champion model. The dashboard is a single-level view of that performance for this preliminary experimentation. We also chose an experimental context familiar to subjects who had already studied the Harrah's case from the Teradata University Network.

Visualizations: We used both tables and graphs to display the relatively complex dispersed-revenue information. In our early analysis, interpreting this information was the most cognitively demanding task of the Wales Market. Color coding facilitates fast visual comparison of a Challenger Model by decile to the deciles of the Champion model.

Mechanisms (Dashboard): We did not include collaboration support as a component of the dashboard designed for the Wales Market case. However, as will be explained below, this was addressed in the experimental design. We used a paper-based version of the case, dashboard and the betting mechanism. It is important to note that we included only one dashboard in what would realistically be a complex set of dashboards for an enterprise-level predictive-model market. More complex dashboards would need to take into account different campaign types, different instantiations of campaigns using shared models, etc.

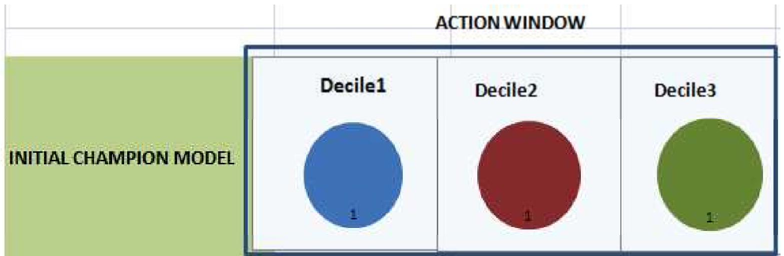
In the next section, we discuss the experimental design with additional explanation of how design options were used in the experimental phase of examining design criteria for predictive-model markets.

Experimental Design

In order to advance design principles for predictive-model markets, we examine several properties. Our objective is to empirically test the validity of the information-market approach and understand the governance structure of enterprise-level predictive-model management. In an organization practicing advanced analytics with a centralized repository for predictive models, there needs to be a way to delineate Champion Models from Challenger Models. We have designed an experiment to test several dimensions based on the literature on information markets and incentive mechanisms. We created a dashboard of models (one Champion and three Challenger Models) with several market signals, and asked our subjects to invest a total of $1000 among different stocks, and each stock refers to a specific model decile. The subjects were informed that the winner would be rewarded. (Appendix 1 includes the case and the details). The experimental design dimensions are as follows:

Market Signals

We have designed a dashboard to report the performance results from a campaign that used the Champion model across the deciles of the action window, where the phrase, "action window" refers to the number of deciles that were used in a completed campaign cycle. The dashboard included actual performance information at decile level for the Champion model (Figure 1).

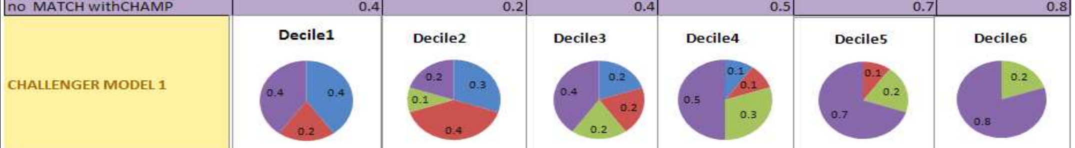


	Decile1	Decile2	Decile3
Revenue	$ 900.00	$ 900.00	$ 700.00
Cost	$ 50.00	$ 50.00	$ 50.00
Share Price	$ 65.00	$ 60.00	$ 55.00
Model Builder	Bob Hamilton		
Prior Championships	5		
Creation Date	Jan-08		
Total Profit	$ 2,350.00		

Figure 1. Action Window

Next, we added Challenger Models to the dashboard and reported dispersed performance metrics. Since we cannot fully observe the performance of the Challenger Models, the objective of the dashboard is to provide information about the dispersion of the observed outcomes across the Challenger Models and their deciles based on their overlap with the Champion model. The dispersion is the only way market participants can infer performance about the Challenger Models (Figure 2). Our objective is to understand if, in the information market we created for predictive-model selection for the next campaign cycle, the market signals we incorporated to the design are actually sufficient to support arriving at a viable bet. As can be seen from Figures 1 and 2, subjects could view various market signals, such as pie charts, (dispersed) revenue, cost, share price, model builder, number of championships, creation date and total (dispersed) profit.

Model 1 overlap	Decile1	Decile2	Decile3	Decile4	Decile5	Decile6
ovelap with Champ DECILE 1	0.4	0.3	0.2	0.1	0	0
ovelap with Champ DECILE 2	0.2	0.4	0.2	0.1	0.1	0
ovelap with Champ DECILE 3	0	0.1	0.2	0.3	0.2	0.2
no MATCH withCHAMP	0.4	0.2	0.4	0.5	0.7	0.8



	Decile1	Decile2	Decile3	Decile4	Decile5	Decile6
Dispersed Revenue	$ 540.00	$ 980.00	$ 500.00	$ 210.00	$ 150.00	$ 140.00
Cost	$ 50.00	$ 50.00	$ 50.00	$ 50.00	$ 50.00	$ 50.00
Share Price	$ 67.00	$ 66.00	$ 59.00	$ 50.00	$ 40.00	$ 30.00
Model Builder	NeuralNet Consulting					
Prior Championships	1					
Creation Date	Jan-10					
Total Dispersed Profit	$ 2,220.00					

Figure 2. Challenger Model Performance

Granularity

Most companies use one Champion Model and, when significant model decay becomes evident in performance results, managers retire and replace the Champion with a new model [Chu et. al. 2009]. With the introduction of a dashboard design that includes the performance of Champion and various Challenger Models, we wanted to provide information to aid in assessing the use of such a new Champion Model. This assessment would occur at the beginning of each campaign cycle. Also, by introducing each model decile performance as a separate stock, we have introduced the concept of being able to deploy fused models, in which a final fusion will be informed by the market bets and the winner stocks. This enables us to test whether the incorporation of individual model deciles as stocks is a valid design approach.

Market Structure: Economic and Social Markets

Based on the literature on economic theory, the economics of information markets, and the impact of incentives, our market design took into consideration different aspects of the market structure. According to traditional economic theory, rational workers will choose to improve their performance in response to a scheme that rewards such improvement with financial gain (Lazear, 2000). Therefore, in our design we have included some incentives for the winners.

Based on the literature on information economy, on the other hand, information is concurrently traded as an economic good and given away free for all. Therefore, there exists a unique tension between purely social and purely economic market-design conventions (Raban, 2008). There are two different outcomes from previous research on the impact of social activity on economic activity. First, it has been found that social activity catalyzes economic-market approaches to the design of prediction markets (Raban, 2008). Also, markets containing signals of both social-market relationships and money-market relationships will be perceived and traded much like money-market relationships (Heyman and Ariely, 2004). While these results are important, they do not directly inform the design of predictive-model market-incentive structures. In order to empirically test the impact of social markets on economic markets, we have included a social-market dimension to our experimental design. In one setting, we have provided only monetary incentives to subjects as a reward and no interaction was allowed. In another setting, along with monetary incentives, we allowed subjects to cooperate and exchange information. In order to measure the information exchange, we provided subjects with candy to exchange for information. As this represents the social market, the number of candies subjects were left with did not create an advantage or a disadvantage so long as particpants completed their decision. Heyman and Ariely used candy in this manner in a prior study to examine social exchange [2004].

Rational Behavior

Based on Surowiecki's Wisdom of Crowds, markets are made up of diverse people with different levels of information and intelligence [2004]. Yet, when you put all those people together and they start buying and selling, they come up with generally intelligent decisions. In order to see if deep knowledge and understanding about predictive models made a difference, we ran our experiments on undergraduate sophomore IS students, as well as students who are in their last week of finalizing their Master's Degree in Information Management. The latter have been taught more sophisticated material on business intelligence and predictive modeling, but both sets of students know BI concepts, vocabulary, as well as how predictive models are developed and assessed. An important test is whether the results from the information market we created for predictive-model management will yield different or homogeneous outcomes across different groups.

Incentives

We had four groups of students participating in information market to make a decision on the model combinations to be used in a fictitious next-campaign cycle. All four groups were exposed to the same dashboard, as well as the same market signals, and they were asked to make decisions on which models and model deciles should be used in the next cycle of the campaign. "Winners" earned $50 gift certificates, but we didn't announce how the winners would be chosen. At the end of the experiment in each group, we gave the $50 awards to those individuals who came closest to their group's collective predictions.

Results

Table 1 depicts the final experimental design to test the validity of employing an information market for model selection. Subjects in all four groups were given a $1,000 in play money to bet on twenty different stocks. Stocks consist of model deciles. The experiment used the Model Dashboard that is included in Appendix 2.

The hypotheses we are testing can be summarized as follows:

- Does the predictive-model market that we have designed have necessary and sufficient attributes to inform decisions that enable viable bets?
- Will introduction of a social-market dimension change the market and change its outcomes?
- Does experience/education level impact the market and change its outcomes?

Table 1. Experimental Design

	Economic Market No Interaction	Economic & Social Market Interaction
Undergraduate IS Students	Economic Incentive UG-E	Economic Incentive (+ 5 Candies) UG-E+S
Graduate MSIM Students	Economic Incentive G-E	Economic Incentive (+ 5 Candies) G-E+S

Figures 3 and 4 show the betting results from our all four groups (UG-E, UG-E+S), (G-E, G-E+S) respectively:

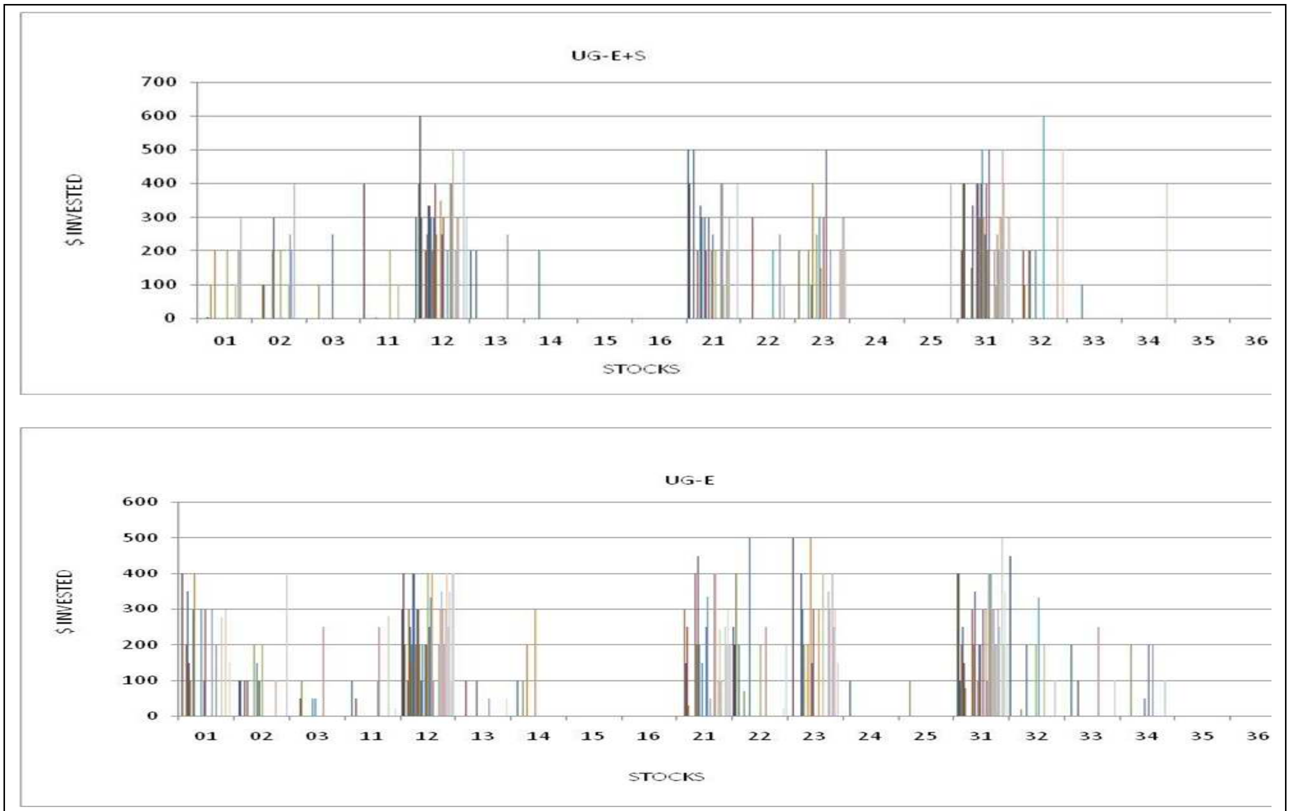


Figure 3. Betting Results-Undergraduate

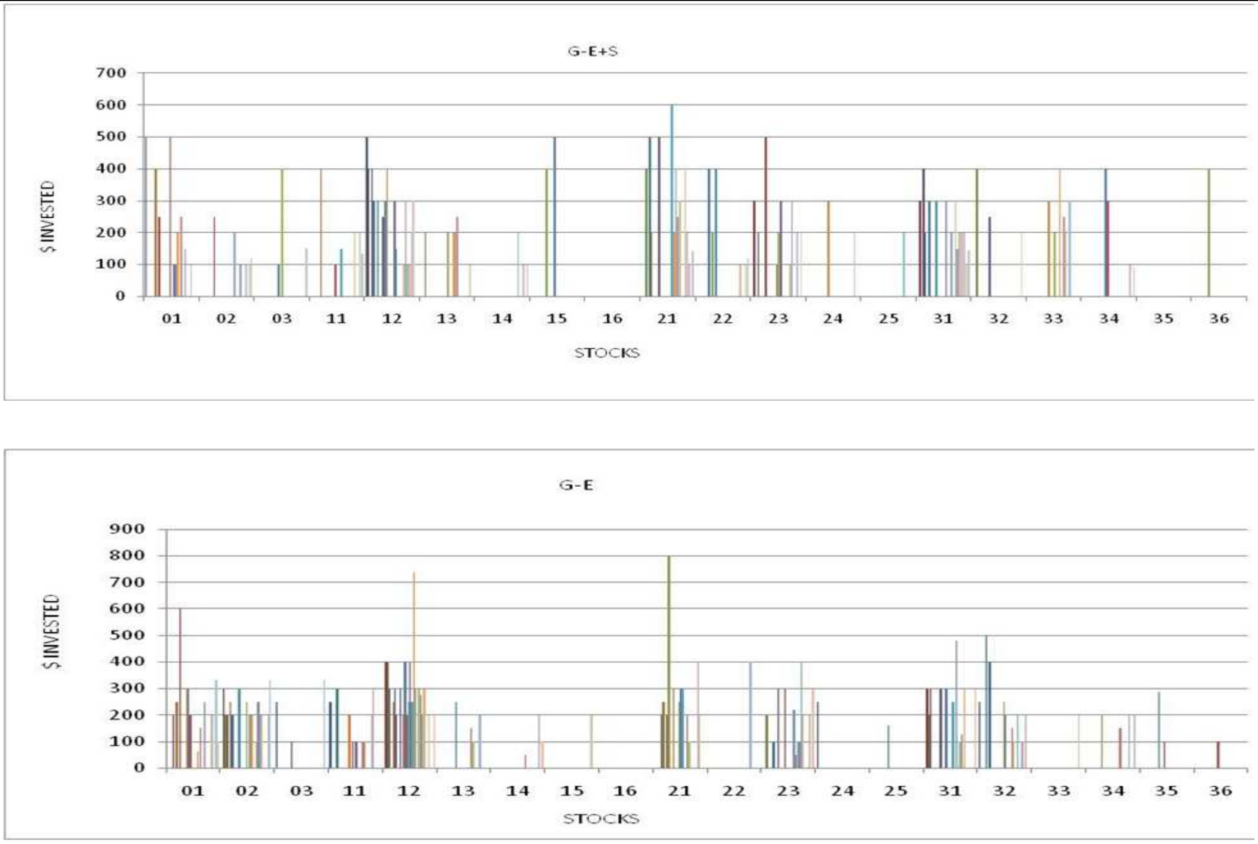


Figure 4. Betting Results-Graduate

Table 2 shows the count and average of bets on the top five stocks across four groups:

Table 2. Betting Results-Top 5 Socks

STOCKS	TOTAL	count UG-E	count UG-E+S	count G-E	count G-E+S	avg UG-E	avg UG-E+S	avg G-E	avg G-E+S
12	95	34	23	21	17	$ 274.50	$ 327.52	$ 302.86	$ 317.65
31	79	28	24	12	15	$ 268.93	$ 311.79	$ 254.58	$ 232.93
21	66	22	18	13	13	$ 250.41	$ 299.11	$ 276.92	$ 322.31
23	54	17	15	12	10	$ 300.00	$ 253.33	$ 205.83	$ 240.00
01	44	14	6	14	10	$ 251.79	$ 166.67	$ 228.43	$ 255.00

As can be seen from Figures 3 and 4 and Table 2, the top three stock choices were the same (12, 31, and 21) across the four different treatment groups. The market structure that we set up yielded the same outcome in four different settings. The overall sum of bets follow the same rank ordering, in spite of a spike on stock 02 for the G-E group, which has 16. Statistical test results (t-test at .05 significance level) indicated that there was no statistically significant difference in the average dollar amount bet in each stock across economic and social markets for both graduate and undergraduate groups.

Figure 5 addresses the count of bets across the four different treatment groups for the top 5 stocks:

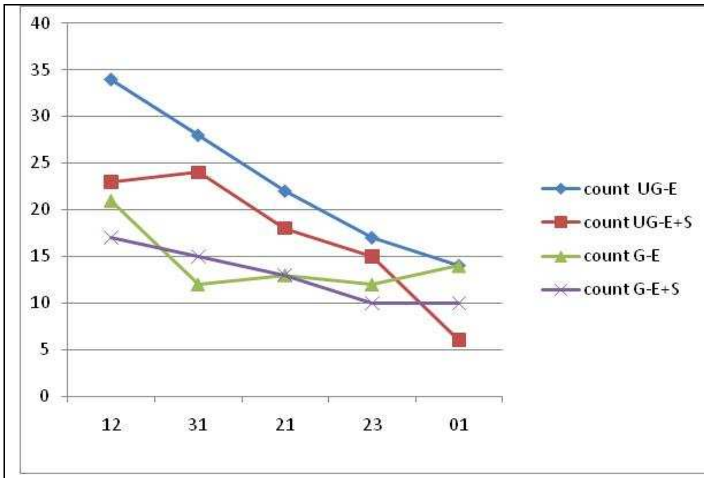


Figure 5. Count of Bets

As can be seen from Figure 5, the total market participants vary across groups (40, 33, 31, and 29 for UG-E, UG-E+S, G-E, G-E+S respectively), but rank ordering overall was consistent with the majority. For the economic and social groups who were allowed to interact in exchange for candies, we have observed that undergraduate students exchanged more candies (2.13) than graduate students (0.72).

Discussion

This empirical study has shown that different groups in a predictive-model market, given the same dashboard, demonstrated surprisingly consistent betting behaviors. This was not because there were clear winners to bet on; realistic uncertainty was conveyed in the market constructs. We therefore infer that the market operates as expected for a prediction market. We demonstrated how a predictive-model market can be used to test design conventions like incentives (economic and social). The addition of social-market conventions in the form of interaction did not change the results.

The market we designed and tested evidenced the ability of multiple stakeholders to engage and exert appropriate effort. Both graduate and undergraduate students had very similar cumulative betting results when provided with the same dashboard and same design principles. This suggests applicability of the design to stakeholders, such as database administrators or analysts. Another interesting outcome we found was that some people did not bet the total amount allowed. Their records were eliminated from the results. In addition, although the results were not impacted, economic and social markets were livelier than purely economic markets.

The survey we conducted right after the betting was over showed that, on average, both graduates and undergraduates ranked revenue, pie charts and stock prices as the top three considerations when making their decision. Model-builder and creation-date information were ranked lowest in usage by both groups. (Students were not warned in advance that a survey would follow on the usage of dashboard KPIs). This provides evidence that the dashboard contained representations that had different priorities in the decision support role they played for those participating in the market.

Conclusions and Next Steps

In this paper, we have identified an approach to managing advanced analytics at the enterprise level in which stakeholder knowledge about predictive models and their performance can be leveraged to support model-selection decisions. We have conducted experiments to investigate design principles for the approach, and our main finding is that there is no significant difference in stakeholder betting behavior based on economic or combined economic and social incentives. This is an important finding because it advises that predictive-model markets can be designed with either orientation.

In addition, we found that the market representations and the level of granularity in the betting processes we developed made sense to market participants, and they were able to make informed bets. The bets placed were remarkably consistent across treatments, a result that has been documented in other prediction market settings. This is important because it provides evidence that the design conventions used enable results similar to those reported in the literature for prediction markets in general.

Our study is limited in that we used student groups, but the Masters level subjects did have an average of more than ten years of experience in the information systems field. All subjects had BI training — including training in the development of predictive models — and they were familiar with the domain of the case through prior exposure to BI in the gaming industry.

An important future research stream relates to advancing dashboard design conventions to better represent enterprise-level realities. Two alternatives include hierarchical dashboards and campaign-related dashboards that are instantiated at the initiation of each campaign. An initial set of results related to predictive model markets are important before moving to advance this more complex set of design issues. In fact, aligning predictive-model market notions with new approaches to agile and pervasive analytics will most certainly require additional research. However, if the approach can facilitate the coordination of multiple stakeholders in the predictive-model management process, there are interesting possibilities for increased collaboration resulting in improved enterprise capabilities in the area of advanced analytics. And, as our experiments have demonstrated, the possibilities for including personnel who have a variety of perspectives and play different BI roles can be accommodated in a prediction-model-market setting.

Our next steps in this research stream are to focus on complete campaign deployment cycles by extending the market model to multiple periods and different model configurations. To that end, we have embraced a portfolio theoretic approach that enables the formulation and validation of an index that can be tracked over time [Balkan and Goul, 2010]. That index enables the evaluation of model performance in the multi-period setting. We are working to establish mechanisms for revaluing stock prices based on performance, supply, and demand. Our longer-term intent is to build a web-enabled marketplace that can be shared with other researchers interested in using a predictive-model market in evaluating design conventions and in the teaching of BI. We envision a configurable predictive-model market that can be used to replicate findings and explore new design options, as well as a market context that will ultimately help to build an accumulating set of design-convention knowledge.

Of course, knowledge about initial design possibilities will need to continue to be expanded. To understand the true efficacy of the approach on bottom-line campaign performance, we will seek opportunities for field and case studies. Being able to demonstrate a web-enabled predictive-model market prototype is an important step in building foundations for this next level of evaluation. It is our hope that as advanced analytics proliferate in organizations through the ideal of pervasive BI, new enterprise-level approaches can be evolved through collaborative academic and industry research streams. What our evidence provides is an important first step: Predictive-model markets can be designed based on either economic or combined economic and social incentives. Much more work is needed to ascertain if predictive-model markets have a place in the evolution of enterprise-data warehouse architectures.

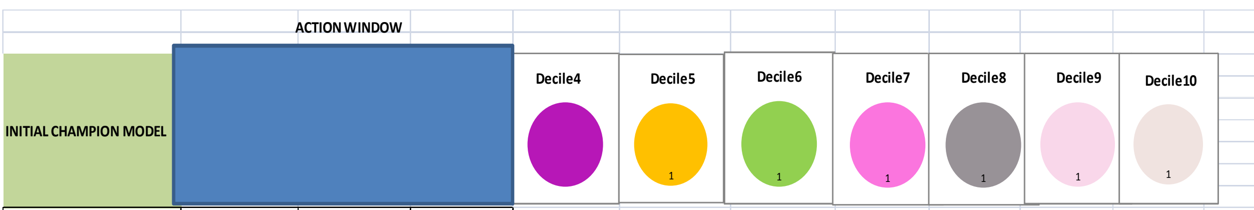
Appendix 1: Excerpts From The Wales Market Case

Yolanda Wales is Vice-President of the Business Intelligence (BI) Solutions Area of an international conglomerate named “The Sanders.” The Sanders is composed of several divisions doing business in the gaming industry. Business intelligence is the primary enabler of customer-relationship management aimed at getting customers to enjoy cross-casino experiences. Using the company’s trademarked “Play-To-Win” card, customers’ preferences, gaming patterns and expenditures are constantly being analyzed to build predictive models designed to help score customer profitability given certain types of special offers.

Because the company's BI capabilities have advanced to the point where analysts are creating numerous high-quality predictive models, Yolanda needs to come up with a new model dashboard to help all of her analyst teams, the marketing group and other managers quickly ascertain which models are the best to use in particular marketing campaigns. When a model is first used in a campaign — such as a mailing to encourage customers to come to a casino during a week that is known to be slow given the normal business cycle — that model is known as the Initial Champion Model. It is difficult to compare the predicted performance of a non-Champion model (called a "Challenger Model") with the known performance of the Champion model. For example, if the Champion model for a campaign targeted 30 customers and 10 actually came and spent money using their Play-To-Win card, then for that Champion model, we know with certainty the profitability of that Champion model. In contrast, a Challenger Model might have targeted a different set of customers, and we could only speculate on their profitability if we tried to compare the Champion and non-Champion models. However, there are likely to be some customers that were targeted in both models. This is known as the model overlap.

Yolanda has come up with a unique strategy for the dashboard she envisions. She explains: "Why not create a market of models where each model can be described and its performance can be shown to all of the main stakeholders in a campaign? We'll stake them each $1,000 — in real money — and ask them to bet on each campaign's model performance by purchasing shares of different models and deciles within those models. After we know the results of the campaign, we'll let everyone know the actual and predicted performance. Then, we'll start another round of betting. The market will react by changing the share prices of the models. New share prices will reflect model performance and demand from the prior cycle of the campaign."

Ms. Wales selects her junior executive, Mora Modeles to develop the idea. Mora has the difficult task of figuring out how to determine the performance of a Challenger Model from one cycle of a campaign to another – and, more importantly, how a Challenger Model would pay off when the results weren't known with certainty. On one hand, the Challenger Model might have performed better on the most profitable subjects that it had in common with the Initial Champion Model. On the other hand, since the Challenger Model might have targeted different subjects, the profitability of those subjects would never be known. Mora decides that what she can do is provide the market with information only about what was actually known:



Revenue	$ 900.00	$ 900.00	$ 700.00
Cost	$ 50.00	$ 50.00	$ 50.00
Share Price	$ 65.00	$ 60.00	$ 55.00
Model Builder	Bob Hamilton		
Prior Championships	5		
Creation Date	Jan-08		
Total Profit	$ 2,350.00		

This set of graphs and numbers in the dashboard reflects the first stage of a campaign where the Initial Champion Model was used. This model scored all of the possible people to be targeted in the campaign (i.e., the population). For example, if the model was a good model, it would assign the highest scores to those people from whom Sanders could reap the most profit. The action window shows the graphs in the dashboard that are relevant to the first stage of the campaign.

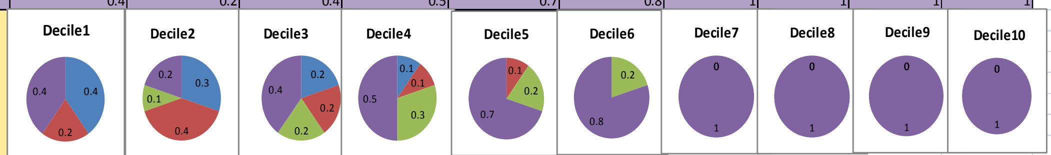
After the first cycle of a campaign, Mora knows the profitability of the Initial Champion Model with certainty. By color-coding the percentage of respondents conforming to each of what are referred to as scored deciles, she can compare the Champion to Challenger Models and easily see how dispersed the scored subjects might be amongst the different scoring capabilities of each model. A decile is a division of the population into ten equal groups. The top decile (Decile1, for example) refers to the top 10% of subjects who received the highest scores by the first model.

In the Initial Champion Model, the color blue signifies those who were targeted in the campaign (because they are inside the action window), and it indicates that this set of 'blue' subjects were those in the top 10% of all subjects as scored by the initial Champion model. Similarly, the color burgundy signifies those scored by the Initial Champion Model to be in the second 10% of the scores given to all subjects. The cost of a model by decile refers to promotion cost, and that promotion cost is fixed by decile. The share price refers to the bet that Yolanda's stakeholders can make. A higher share price reflects the amount of profit that a model has generated in its lifetime, and it takes into account the demand from the bettors for that model's performance for the particular decile. Other important information about a model also is provided. Some bettors rely on this information in making their decisions almost as much as they rely on the pure profits a model produces. That profit is shown for the Initial Champion Model at $2550. This is calculated by considering only the action window and the amount of revenue by decile, less the fixed costs for the deciles.

To look at how the Initial Champion Model compared with Challenger Models (e.g., Challenger Model 1), Mora and her colleagues need to be able to see the dispersion of scored subjects as shown in the following:

Model 1 overlap	Decile1	Decile2	Decile3	Decile4	Decile5	Decile6	Decile7	Decile8	Decile9	Decile10
ovelap with Champ DECILE 1	0.4	0.3	0.2	0.1	0	0	0	0	0	0
ovelap with Champ DECILE 2	0.2	0.4	0.2	0.1	0.1	0	0	0	0	0
ovelap with Champ DECILE 3	0	0.1	0.2	0.3	0.2	0.2	0	0	0	0
no MATCH withCHAMP	0.4	0.2	0.4	0.5	0.7	0.8	1	1	1	1

CHALLENGER MODEL 1



	Decile1	Decile2	Decile3	Decile4	Decile5	Decile6
Dispersed Revenue	$ 540.00	$ 980.00	$ 500.00	$ 210.00	$ 150.00	$ 140.00
Cost	$ 50.00	$ 50.00	$ 50.00	$ 50.00	$ 50.00	$ 50.00
Share Price	$ 67.00	$ 66.00	$ 59.00	$ 50.00	$ 40.00	$ 30.00
Model Builder	NeuralNet Consulting					
Prior Championships	1					
Creation Date	Apr-10					
Total Dispersed Profit	$ 2,220.00					

Note that in the first decile of the comparison, there are three different colors in the pie chart. The blue and burgundy correspond to those subjects scored in the Initial Champion Model in their respective deciles, and the charts show that there was a 40% overlap of subjects in Decile 1 between the two models. This overlap means the Initial Champion Model scored subjects in the first decile, and so did Challenger Model 1. The purple area depicts those that were scored only by the second model, but there is no overlap with the Initial Champion Model. In other words, the results are unknown for this 40% - because they were not targeted in the campaign. However, there is a number referred to as the "dispersed revenue." That shows the actual amount of revenue made by the people targeted in Challenger Model 1 in the corresponding decile of that model.

Mora would explain: "The customers we make the most money on are scored into multiple deciles by the Initial Champion model. However, a Challenger Model may make more profit within a decile because it happened to score those most profitable customers in the particular decile." For example, consider the revenue of the second decile of Challenger Model 1. It is $980.00 – while the second decile for of the Initial Champion Model is $900.00. The reason it made more revenue is because more of what they call "whales" or big spenders in the gaming industry happened to be scored in the second decile of Challenger Model 1. As Mora told her market participants, "Your challenge in this market is to take your stake and allocate it to deciles for the next round of the campaign. For example, you may choose to invest in the second decile of Champion Model 1 because it made more profit than the Initial Champion Model. But since you are making an investment, you need to take the share price into account as well. You also have to remember that this is a market, and anything can happen in the next cycle. Investing everything on one outlier can get you in the long run I know, because I used to run a craps table. I've seen some big losers bet on a number just because it came up last time."

Mora has prepared the attached sheet to represent the model market information she can provide. This snapshot of the markets describes the actual performance of the Champion model and the integrated performances of the non-Champion models. There are three Challenger Models in the dashboard and, for each; the dispersion of revenues is shown as based on actual expenditures from the first cycle of the campaign when the Initial Champion Model was used. The total profit for each of the Challenger Models is lower than that of the Initial Champion Model because to capture all of the subjects that were targeted by the Initial Champion Model, would take more mailings. Additional fixed costs are associated with each of these mailings.

In the second and subsequent cycles of a campaign, Yolanda uses what is referred to as a 'fused model.' That is, she can use the subjects targeted by some "Model A" as ranked in the first decile, some "Model B" as ranked in the second decile, and so forth. In her experience, the bettors' bets provide the best information on which models and their associated deciles to include in the next cycle.

Mora Modeles' assistant, Winnie Whiner, is always assigned to give the new model builders a training class on investing in the Wales Market. When Winnie explains to the newbies how it all works and how to make an investment decision, she makes it plain and simple: "Look, you've got a thousand bucks and you're going to buy shares in the particular deciles of models. Nothin' more, nothin' less. You're hoping your share price will go up after the next cycle of the campaign. If the share prices go up, you've made some money in the market. That will happen if there's lots of demand for those shares, and there's usually lots of demand where there's the highest profitability from what you know from the prior campaign cycle. You've got to figure out which model deciles are worthy of your investment."

Once Winnie finishes her instructions, the market opens and the betting begins.

Wales Market Betting Sheet

Campaign: Weekday Promotion Event

Offer: Free Night Stay on Tuesday or Wednesday

Campaign Manager: Mora Modeles

Bettor:

REMEMBER: YOU HAVE A TOTAL AMOUNT OF $1000 TO INVEST

Initial Champion Model

Share Price:	$65.00	$60.00	$55.00
	Decile 1	Decile 2	Decile 3
Investment		**200**	

Challenger Model 1

	$67.00	$66.00	$59.00	$50.00	$40.00	$30.00
	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6
Investment						

Challenger Model 2

	$70.00	$60.00	$50.00	$45.00	$40.00
	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5
Investment	**400**				

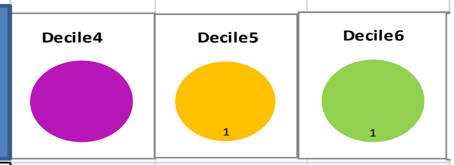
Challenger Model 3

	$48.00	$47.50	$46.00	$45.00	$20.00	$15.00
	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6
Investment			**400**			

Appendix 2: Wales Market Dashboard

ACTION WINDOW

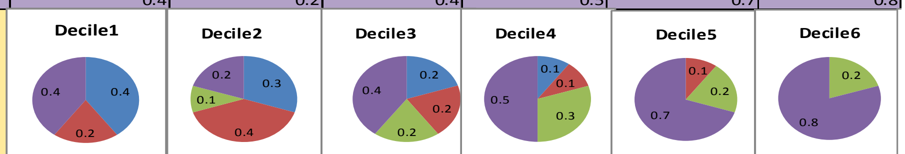
INITIAL CHAMPION MODEL



Revenue	$ 900.00	$ 900.00	$ 700.00
Cost	$ 50.00	$ 50.00	$ 50.00
Share Price	$ 65.00	$ 60.00	$ 55.00
Model Builder	Bob Hamilton		
Prior Championships	5		
Creation Date	Jan-08		
Total Profit	$ 2,350.00		

Model 1 overlap	Decile1	Decile2	Decile3	Decile4	Decile5	Decile6
ovelap with Champ DECILE 1	0.4	0.3	0.2	0.1	0	0
ovelap with Champ DECILE 2	0.2	0.4	0.2	0.1	0.1	0
ovelap with Champ DECILE 3	0	0.1	0.2	0.3	0.2	0.2
no MATCH withCHAMP	0.4	0.2	0.4	0.5	0.7	0.8

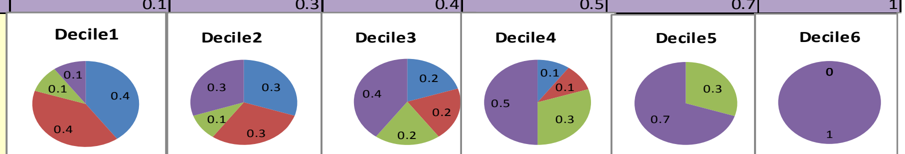
CHALLENGER MODEL 1



Dispersed Revenue	$ 540.00	$ 980.00	$ 500.00	$ 210.00	$ 150.00	$ 140.00
Cost	$ 50.00	$ 50.00	$ 50.00	$ 50.00	$ 50.00	$ 50.00
Share Price	$ 67.00	$ 66.00	$ 59.00	$ 50.00	$ 40.00	$ 30.00
Model Builder	NeuralNet Consulting					
Prior Championships	1					
Creation Date	Apr-10					
Total Dispersed Profit	$ 2,220.00					

Model 2 overlap with Camp	Decile1	Decile2	Decile3	Decile4	Decile5	Decile6
ovelap with Champ DECILE 1	0.4	0.3	0.2	0.1	0	0
ovelap with Champ DECILE 2	0.4	0.3	0.2	0.1	0	0
ovelap with Champ DECILE 3	0.1	0.1	0.2	0.3	0.3	0
no MATCH	0.1	0.3	0.4	0.5	0.7	1

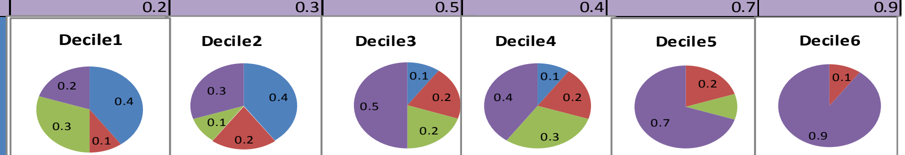
CHALLENGER MODEL 2



Disbursed Revenue	$ 790.00	$ 610.00	$ 810.00	$ 60.00	$ 260.00	
Cost	$ 50.00	$ 50.00	$ 50.00	$ 50.00	$ 50.00	
Share Price	$ 70.00	$ 60.00	$ 50.00	$ 45.00	$ 40.00	
Model Builder	New-DelhiAnalytics					
Prior Championships	1					
Creation Date	Jun-09					
Total Profit	$ 2,280.00					

Model 3 overlap with Camp	Decile1	Decile2	Decile3	Decile4	Decile5	Decile6
ovelap with Champ DECILE 1	0.4	0.4	0.1	0.1	0	0
ovelap with Champ DECILE 2	0.1	0.2	0.2	0.2	0.2	0.1
ovelap with Champ DECILE 3	0.3	0.1	0.2	0.3	0.1	0
no MATCH	0.2	0.3	0.5	0.4	0.7	0.9

CHALLENGER MODEL 3



Disbursed Revenue	$ 860.00	$ 610.00	$ 410.00	$ 480.00	$ 50.00	$ 90.00
Cost	$ 50.00	$ 50.00	$ 50.00	$ 50.00	$ 50.00	$ 50.00
Share Price	$ 48.00	$ 47.50	$ 46.00	$ 45.00	$ 20.00	$ 15.00
Model Builder	David Herman					
Prior Championships	1					
Creation Date	Jan-09					
Total Profit	$ 2,200.00					

Reference

Balkan, S. and M. Goul, 2010 “Advances in Predictive Modeling: How In-Database Analytics will Evolve to be a Game Changer,” *Business Intelligence Journal*, Volume 15, Number 2.

Balkan, S. and M. Goul, 2010 “A Portfolio Theoretic Approach to Administering Advanced Analytics: The Case of Multi-Stage Campaign Management,” Working Paper.

Bellamy, R.K.E., Erickson, T., Fuller, B., Kellogg, W.A., Rosenbaum, R., Thomas, J.C., Vetting Wolf, T. “Seeing is Believing: Designing Visualizations for Managing Risk and Compliance,” *IBM Systems Journal*, Vol. 46, No.2.

Bonissone, P.P., F. Xue and R. Subbu, 2007, “Fast Meta-Models for Local Fusion of Multiple Predictive Models,” *Applied Soft Computing Journal*, 2008.

Conitzer, V. *, 2007* “Prediction Markets, Mechanism Design, and Cooperative Game Theory,” *25th Conference on Uncertainty in Artificial Intelligence.*

Chu, C. R., 2009, “Computer-implemented Systems and Methods for Updating Predictive Models,” United States Patent Application Publication, US 2009/0106178 A1, April 23.

Davenport, T. and L. Prusak, 1998, *Working knowledge: Managing What Your Organization Knows,* Harvard Business School Press, Boston, MA.

Eckerson, W.,2006, *Performance Dashboards: Measuring, Monitoring and Managing Your Business,* Wiley and Sons, New Jersey.

Heyman, J., Ariely, D., 2004, “Effort for Payment: A Tale of Two Markets,” *Psychological Science*, Vol. 14, No.11.

Howson, C., 2008, “The Road to Pervasive BI,” *Intelligent Enterprise*, February, http:/www.intelligententerprise.com/showArticle.jhtml?articleID=206801408.

Lazear, E.P., 2000, “Performance Pay and Productivity,” American Economic Review, 90, 5, December, 1346-1361.

Looney, Robert, 2003, “DARPA’s Policy Analysis Market for Intelligence: Outside the Box or Off the Wall,” *Strategic Insights,* Vol. II Issue 9, September.

Malik, Shadan, 2005, *Enterprise Dashboards: Design and Best Practices for IT*, Wiley.

Malone, Thomas W., Laubacher, Robert and Dellarocas, Chrysanthos N., 2010, “Harnessing Crowds: Mapping the Genome of Collective Intelligence,” *MIT Sloan Research Paper* No. 4732-09.

Nichols, J., Demirkan, H., Goul, M., Keith, M., 2009, “Towards a Theory of Agile Dashboards for Service Dashboards for Service Oriented Organizations,” *Americas Conference on Information Systems (AMCIS)*, AMCIS 2009 Proceedings.

Raban, D. R., 2008, “The Incentive Structure in an Online Information Market,” *Journal of American Society for Information Science and Technology*, 59(14): 2284-2295.

Russom, Philip, 2009, “Next Generation Data Warehouse Platforms,” *TDWI Best Practices Report*.

Servan-Schreiber, E., Wolfers, J., Pennock, D. M., Galebach, B., 2004, “Prediction Markets: Does Money Matter?” *Special Issue: Innovative Auction Market, Electronic Markets,* Volume 14(3):243-251.

Surowiecki, James, 2004, *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations* Little/Brown.

Tziralis G. and Tatsiopoulos I., 2007, “Prediction Markets: An Extended Literature Review,” *The Journal of Prediction Markets*, Vol. 1.

Watson, H., Wixom, B., 2007, "The Current State of Business Intelligence," *Computer*, vol. 40, no. 9, pp. 96-99, August.

Wei, J., E. Gao, F. Wang and R. Chu, 2009, “Dashboard Reports for Predictive Model Management,” Paper 045-2009, SAS Global Forum.

Weinhardt, C., Neumann, D., Holtmann, C., 2006, “Computer-aided Market Engineering,” *Communications of the ACM,* Vol. 49, No. 7.

Weinhardt, C.,Schnizler, B., Luckner, S., 2003, “Market Engineering,” *Wirtschaftsinformatik,* 45(6): 635-64.