2011-02-03

Innovations in Defense Acquisition: Asymmetric Information, Mechanism Design and Prediction Markets

Coughlan, Peter; Arkes, Jeremy; Gates, William

NPS-AM-11-006

ACQUISITION RESEARCH SPONSORED REPORT SERIES

Innovations in Defense Acquisition: Asymmetric Information, Mechanism Design and Prediction Markets

3 February 2011

by

Dr. Peter Coughlan, Associate Professor

Dr. William Gates, Associate Professor, and

Dr. Jeremy Arkes, Associate Professor

Graduate School of Business & Public Policy

Naval Postgraduate School

Prepared for: Naval Postgraduate School, Monterey, California 93943

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **03 FEB 2011**	2. REPORT TYPE	3. DATES COVERED **00-00-2011 to 00-00-2011**

4. TITLE AND SUBTITLE
Innovations in Defense Acquisition: Asymmetric Information, Mechanism Design and Prediction Markets

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School,Graduate School of Business & Public Policy,555 Dyer Road, Room 332,Monterey,CA,93943-5103

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Prediction markets, sometimes called information markets, idea markets or event futures, are similar to financial stock markets where the ?stocks? and their prices reflect the consensus view regarding the outcomes of specifically defined future probabilistic events. Prediction markets quickly and efficiently gather and summarize information from a disparate and diverse group of people, providing a two-way information flow; individual traders are informed by the consensus opinion and their market decisions inform the aggregate consensus. The remarkable accuracy of prediction markets in forecasting election results economic outcomes, and other variables has defense managers intrigued by the possibility of applying these markets as a managerial decision tool. Overall implementing prediction markets is straight forward, but in practice the devil is in the details, including security or contact design, trading rules, participation incentives and the number and characteristics of the traders. Small changes in any design element can significantly affect prediction market performance. This research highlights the implementation issues involved in designing and running prediction markets. If improperly designed, prediction markets will be confusing and uninformative, at best. Poorly designed early pilots can portray prediction markets as a flawed concept as opposed to a useful concept with a flawed implementation.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**67**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

The research presented in this report was supported by the Acquisition Chair of the Graduate School of Business & Public Policy at the Naval Postgraduate School.

To request Defense Acquisition Research or to become a research sponsor, please contact:

NPS Acquisition Research Program
Attn: James B. Greene, RADM, USN, (Ret.)
Acquisition Chair
Graduate School of Business and Public Policy
Naval Postgraduate School
555 Dyer Road, Room 332
Monterey, CA 93943-5103
Tel: (831) 656-2092
Fax: (831) 656-2253
e-mail: jbgreene@nps.edu

Copies of the Acquisition Sponsored Research Reports may be printed from our website www.acquisitionresearch.org

Abstract

Prediction markets, sometimes called information markets, idea markets or event futures, are similar to financial stock markets where the “stocks” and their prices reflect the consensus view regarding the outcomes of specifically defined future probabilistic events. Prediction markets quickly and efficiently gather and summarize information from a disparate and diverse group of people, providing a two-way information flow; individual traders are informed by the consensus opinion and their market decisions inform the aggregate consensus.

The remarkable accuracy of prediction markets in forecasting election results, economic outcomes, and other variables has defense managers intrigued by the possibility of applying these markets as a managerial decision tool. Overall, implementing prediction markets is straight forward, but in practice the devil is in the details, including security or contact design, trading rules, participation incentives, and the number and characteristics of the traders. Small changes in any design element can significantly affect prediction market performance.

This research highlights the implementation issues involved in designing and running prediction markets. If improperly designed, prediction markets will be confusing and uninformative, at best. Poorly designed early pilots can portray prediction markets as a flawed concept as opposed to a useful concept with a flawed implementation.

Keywords: Prediction markets, information aggregation, information markets, idea markets or event futures

THIS PAGE INTENTIONALLY LEFT BLANK

About the Authors

Peter J. Coughlan, is an Associate Professor of Business and Economics at the Graduate School of Business and Public Policy, Naval Postgraduate School. He received his PhD in Economics from the California Institute of Technology in June 1999. Prior to joining the Naval Postgraduate School in 2004, Dr. Coughlan was an Assistant Professor of Business Administration in the Strategy Group at the Harvard Business School. Dr. Coughlan's research interests include experimental economics and game theory/mechanism design. He also has research interests in strategic and competitive dynamics, particularly applied to technological change.

Peter J. Coughlan
Graduate School of Business and Public Policy
Naval Postgraduate School
Monterey, CA 93943-5000
Fax: (831) 656-3407
E-mail: pjcoughl@nps.edu

William R. Gates is an Associate Professor of Economics and Dean of the Graduate School of Business and Public Policy at the Naval Postgraduate School. He received his PhD in Economics from Yale University in 1984. Before joining the Naval Postgraduate School in 1988, Dr. Gates work as an economist at the Jet Propulsion Laboratory in Pasadena, CA. Dr. Gates' research interests include game theory (matching games and auctions), incentives and asymmetric information (incentive contracting) and public goods (economics of defense alliances).

William R. Gates
Graduate School of Business and Public Policy
Naval Postgraduate School
Monterey, CA 93943-5000
Phone: (831) 656-2754
Fax: (831) 656-3407
E-mail: bgates@nps.edu

Jeremy Arkes is an Associate Professor of Economics at the Graduate School of Business and Public Policy, Naval Postgraduate School. He received his

PhD in Economics from the University of Wisconsin in 1997. Prior to joining the Naval Postgraduate School in 2007, Dr. Arkes worked as a Research Analyst at the Center for Naval Analyses and as an Economist at Rand Corporation. Dr. Arkes' research interests include Posttraumatic Stress Disorder (PTSD), military manpower, sports economics and betting markets, the effects of the recessions on social behavior such as drug use, and how divorces affect children.

Jeremy Arkes
Graduate School of Business and Public Policy
Naval Postgraduate School
Monterey, CA 93943-5000
Phone: (831) 656-2646
Fax: (831) 656-7633
E-mail: jaarkes@nps.edu

NPS-AM-11-006

ACQUISITION RESEARCH
SPONSORED REPORT SERIES

Innovations in Defense Acquisition: Asymmetric Information, Mechanism Design and Prediction Markets

3 February 2011

by

Dr. Peter Coughlan, Associate Professor

Dr. William Gates, Associate Professor, and

Dr. Jeremy Arkes, Associate Professor

Graduate School of Business & Public Policy

Naval Postgraduate School

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

1. Introduction

Prediction markets (PMs), sometimes called information markets, idea markets or event futures, are essentially small-scale securities markets. They are similar to financial stock markets, however the ultimate value of “stocks” or “shares” traded depends upon the outcome of specifically defined future events, rather than on the future earnings of a publicly traded company. Similarly, whereas a company’s stock price reflects the real-time consensus view of that company’s future earnings, stock prices in prediction markets reflect the real-time consensus view about the expected outcome of the associated future events.

In their review of prediction markets, Tziralis and Tatsiopoulos (2007) described prediction markets based upon a definition by Berg, Nelson, and Rietz (2008):

> Prediction markets are defined 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 about specific future events. (p. 1)

Prediction markets are an excellent way to quickly and efficiently gather and summarize information from a disparate and diverse group of people. Prediction markets aggregate knowledge in a unique way. From the perspective of an individual trader, the current market price represents the collective consensus among other market participants. Viewing this consensus allows an individual trader to make his or her own assessment by combining the market’s aggregated information with his or her own private information. Thus, prediction markets provide a two-way information flow; individual traders are informed by the consensus opinion and their market decisions inform the aggregate consensus.

Research has demonstrated that the collective judgment of a large group, none of whom may be “experts” on a particular issue, will usually be more accurate than the judgment of individual experts (or even a small group of experts). (Wolfers and Zitzewitz, 2004) Each participant in a prediction market, ranging from high-level managers to “assembly-line” workers, can provide a unique perspective and valuable information about the future outcome in question. The trading prices in prediction markets provide management with a timely, accurate, and continuously updated picture about the likelihood of future events—enabling them to evaluate risk and provide an early warning of issues requiring management attention.

Both the understanding of PMs and the range of market applications have been growing in recent years. For example, PMs have been employed to predict election outcomes and have been found to do so more accurately than existing polling mechanisms (Berg et al., 2008). In addition, corporations have used PMs to predict new product (or project) sales, launch dates, regulatory approval, and achievement of development milestones.

What PMs do that other methods of gathering dispersed personal information cannot, is aggregate many opinions into a single, collective, market-based forecast of future events. A PM also allows participants to express what may be unpopular opinions in an anonymous fashion.

The U.S. Navy, like corporations, is interested in many important but uncertain future outcomes, such as recruiting and retention success and milestone or delivery dates for certain acquisitions. Given these uncertainties, a well-designed prediction market applied to forecasting such relevant future outcomes could provide the Navy with valuable information. Greater information on these outcomes could help the Navy plan better and allocate or manage resources more efficiently. For example, having more accuracy in predicting the final fiscal-year retention numbers would help the Navy plan for their recruiting goal for the following year. Similarly, more accurately predicting the date at which an acquisition program will achieve a program milestone would enhance program management.

2. From "Stock" Price to Prediction

Before delving further into the promise and pitfalls of prediction markets for military forecasting, it is critical to understand what PM prices actually predict (or what they are intended to predict). The nature of these predictions depends critically on the definition or design of the *stocks* or *securities* being traded.

In private-sector stock markets, for example, it is important to understand that what is actually being traded are company *shares*, which are each, in essence, a contract which entitles the shareholder to certain future rights (such as a claim on that company's future earnings). Traders in prediction markets are similarly buying and selling contracts which specify how the payoffs to holders of these contracts are tied to future events.

Common types of contracts traded in prediction markets include:

(1) "winner-take-all" contracts, which pay off a fixed amount if and only if a specific event occurs;

(2) "index" contracts, which pay off a variable amount which is tied to a specific future measure; and

(3) "spread" contracts, which combine aspects of the winner-take-all and index contract types, paying off a fixed amount if and only if a specific future measure is above or below a threshold, which is adjusted by the market-maker to balance the two sides of the market (just as the "point-spread" is adjusted in football betting).

What any particular PM is actually designed to predict is very much a function of the type of contract being traded. In particular, the market price of a winner-take-all contract reflects the market expectation of the *probability* of a specific future event. The market price of an index contract reflects the market expectation of the *mean value* of a future measure. Finally, the market price of a spread contract reflects the market expectation of the *median value* (or any other *percentile value* of interest) for some future measure.

3. An Illustration of a Prediction Market in Action

To understand how a prediction market actually *aggregates* information, consider this simple example: Suppose that there are three possible (mutually-exclusive) outcomes of some future event. For simplicity, let us label these outcomes A, B, and C. Further suppose that there is a prediction market in which traders buy and sell shares associated with each outcome (A shares, B shares, and C shares, respectively).

Suppose that the shares in the prediction market are winner-take-all contracts, such that each outcome's associated shares pay off $100 each if that particular outcome occurs and pay off $0 otherwise. Thus, if outcome A occurs, A shares will pay off $100 each, while B shares and C shares will pay off nothing ($0 each). If outcome B occurs, B shares will pay off $100 each, while A shares and C shares will pay off nothing ($0 each). Finally, if outcome C occurs, C shares will pay off $100 each, while A shares and B shares will pay off nothing ($0 each).

Prediction Market Prices as Consensus Opinion

Suppose that each outcome is initially considered equally likely by all traders. Thus, all traders initially believe that there is a 33.33% chance that outcome A will occur, a 33.33% chance that outcome B will occur, and a 33.33% chance that outcome C will occur. Therefore, each trader estimates that there is a 33.33% chance that any particular outcome's shares will be worth $100 each and a 66.67% chance that the shares will be worthless. Hence, each trader has an initial expected value of any share of any of the three outcomes of $33.33 (33.33% x $100).

In a well-functioning prediction market, therefore, we would expect A shares, B shares, and C shares to each quickly achieve a market price around $33.33 each. No trader should be willing to buy shares above this price, and no trader should be willing to sell shares below this price. Thus, the market prices accurately reflect the consensus opinion regarding the likelihood of each future outcome. Of course, the consensus opinion is not very helpful at this point, as it considers all outcomes equally likely.

Information Dissemination in the Prediction Market

Now let us see what might happen if limited, but valuable, information is revealed to one or a few traders. In particular, suppose that, over time, it becomes apparent to some traders that certain outcomes are no longer possible. In other words, at some point in time, some trader(s) may discover that a particular outcome is definitely not going to occur. While such information will only be revealed to one or a few market traders at a given time, we assume all traders are aware that other traders may receive such definitive “outcome-excluding” information. While unrealistic, this certainty condition simplifies the illustration; without such certainty outcomes would not be as definitive.

For example, suppose trader Alan learns at some point that outcome A will not occur, but that outcomes B and C are still considered equally likely. How might Alan act on this information? First of all, Alan now knows that A shares are worthless, thus he would be willing to sell A shares for any positive price. Similarly, Alan now considers outcomes B and C to each have a 50% chance of occurring, so he would be willing to buy B or C shares for any price less than $50.

Now consider what would happen if Alan did, in fact, take the action of selling A shares in the prediction market. As noted above, none of the other traders are currently willing to pay more than $33.33 for A shares, so Alan would have to offer to sell at a price below this amount. Suppose Alan offers to sell A shares at $32 each. He might get some initial takers at this price, but the other traders will begin to realize that somebody in the market must have learned that outcome A is no longer possible. Thus, the market price of A shares will quickly decline to $0.

Suppose Alan instead decided to act upon his inside information by buying B shares or C shares, which he now values at $50 each. As noted above, none of the other traders are currently willing to sell B or C shares for any price below $33.33

each, so Alan will have to offer to buy at a price above this amount. Suppose Alan offers to buy B shares (or C shares) at $35 each. He might again get some initial takers at this price, but the other traders will begin to realize that somebody in the market must have learned that outcome B (or C) has suddenly become more likely (as the result of another outcome being eliminated). Thus, traders will gradually (if not immediately) realize that B shares (or C shares) now must be worth at least $50 each, and thus the market price of these shares will quickly rise to $50, at which point no trader should be willing to pay a price above this amount.

Furthermore, note that the pricing dynamics for each of the three share types will be mutually reinforcing: As the price of A shares declines, this indicates that B shares and C shares must be more valuable, and vice versa. Ultimately, Alan's effort to benefit from his private or inside information would result in market prices of $0 for A shares, $50 each for B shares, and $50 each for C shares.

In the scenario described, note that only a single trader actually had any information of value, but the nature of the prediction market quickly disseminated this inside information to all other traders, as if they had the same knowledge first-hand. This is the fundamental characteristic of prediction markets: The market creates an incentive for individual traders to act on (or take advantage of) their own private or inside information (or assessment), yet acting on this information necessarily disseminates this previously private information to others. Hence, prediction markets make private information public.

Information Aggregation in the Prediction Market

Now let us see what might happen in the prediction market when further valuable information is revealed to traders. Suppose, in particular, that trader Bill now discovers that outcome B is no longer possible and that only outcome A or outcome C can possibly occur. Based on this private information alone, Bill would consider B shares to be worthless and would value A and C shares at $50 each.

Alan's previously private information (that outcome A will not occur) has already been disseminated to all traders in the market; however, Bill can fully deduce that neither outcome A nor B are possible, so outcome C is going to occur. Thus, Bill considers A and B shares to be worthless and values C shares at $100 each.

Now consider what would happen if Bill were to act on his new private information. As noted above, after the dissemination of Alan's private information, the prevailing market prices would be $0 for A shares, $50 each for B shares, and $50 each for C shares. Thus, Bill could profit by selling B shares for any positive price and could profit by buying C shares for any price less than $100.

However, to sell B shares Bill will have to offer to sell at some price below $50 which, by the same dynamic described above, will gradually (if not immediately) reveal to other traders that further inside information regarding outcome B has been revealed to some trader(s). Thus, other traders will begin to realize that B shares are worthless and the market price will quickly decline to $0.

Suppose Bill instead decided to act upon his inside information by buying C shares, which he now values at $100 each. Because none of the other traders would be willing to sell C shares for any price below $50 each, Bill would have to offer to buy at a price above this amount. However, in doing so other traders will again begin to realize that somebody in the market must have learned that outcome C will, in fact, occur. Thus, traders will gradually (if not immediately) realize that C shares now must be worth $100 each, and thus the market price of these shares will quickly rise to $100.

Again, note that the opposite pricing dynamics for B and C shares will be mutually reinforcing: As the price of B shares declines, this indicates that C shares must be more valuable, and vice versa. Ultimately, Bill's effort to benefit from his private or inside information would result in market prices of $0 for A shares, $0 for B shares, and $100 each for C shares.

Information Aggregation vs. Information Averaging

While the example above was hypothetical, note that this same scenario has been simulated in laboratory prediction markets using human subjects, and the result is the same: The prediction market fully aggregates the information available to all traders and accurately predicts the correct outcome (Plott & Sunder, 1982).

It is significant to note that, in the scenario described, no single trader had private information that revealed the true eventual outcome. Instead, the vast majority of traders had absolutely zero private information, while two traders had limited private information that individually eliminated only one of the three possible outcomes. Thus, information about the future outcome was limited, fragmented, dispersed, and private. Yet, despite these problematic information conditions, the prediction market fully aggregated the information to formulate an accurate consensus forecast.

It is also important to distinguish how information aggregation is different—and more accurate—than information *averaging*. One might think, for example, that similar prediction accuracy could be achieved if decision-makers simply gathered the individual assessments of many (or select) individuals and averaged the result. This is incorrect for several reasons.

First of all, prediction markets create an incentive to reveal privately held information: market gains can be achieved only by acting on this information. In contrast, there is little incentive to reveal valuable private information in a simple survey, particularly if the individuals have some stake in the future outcome (which is often the case).

In addition, it would be very difficult to identify *which* individuals should be surveyed. In a prediction market mechanism, the individuals with valuable information self-select: Those who have valuable private information have an incentive to proactively trade based on this information, while those who do not have valuable private information simply reactively trade in response to market trends. In the scenario described above, for example, surveying all traders would incorporate numerous poorly informed assessments of 33% chance for each outcome, mitigating the impact of the assessments of those individuals who actually do have valuable information (Alan and Bill in the example).

Finally, even if the actual “experts” or “insiders” with valuable private information could be identified, averaging their individual assessments would still not produce as accurate a forecast as the aggregation of assessments achieved via the prediction market. In the scenario above, for example, Alan and Bill each had valuable private information. If asked their assessment of the likelihood of each of the three outcomes, Alan would assign 0% probability to outcome A, 50% to outcome B, and 50% to outcome C. Bill, on the other hand, would assign 50% probability to outcome A, 0% to outcome B, and 50% to outcome C. Averaging these two experts assessments would assign a combined probability of 25% to outcome A, 25% to outcome B, and 50% to outcome C. Thus, while averaging the expert assessments would, in this case, identify the most likely outcome, it would not match the much more precise (and correct) forecast of the prediction market: 100% chance of outcome C.

4. Private Sector Prediction Market Applications and Results

Beyond the theoretical or conceptual appeal of the prediction market concept, the potential Navy benefit of employing this forecasting tool is also supported by numerous “real-world” private-sector applications, many of which have been publicly touted as highly successful endeavors. Current private-sector applications of internal prediction markets include sales forecasting, project execution, product design, trend forecasting, and resource allocation.

Sales forecasting predicts the likely volume of sales in dollars or units. Project execution predicts when projects will reach their planned milestones. Product design forecasts which product features or enhancements customers will prefer. Trend forecasting reveals new or existing market, technology, or customer trends. Finally, resource allocation enables business units to trade resources according to their needs and can be used to support objectives such as corporate social responsibility (Corporate Executive Board, 2006).

The applications and experiences of some corporate prediction market early-adopters are summarized in Table 1. In what follows, we will describe in more detail the specific design, usage, and results of prediction markets as employed at Hewlett-Packard (HP), Siemens, and Microsoft.

Table 1. Experiences of Early Prediction Market Adopters
(Kiviat, 2004; Malone, 2004)

Application	**Practitioner**	**Description**
Sales Forecasting	Hewlett-Packard	Hewlett-Packard used an internal market system to forecast printer sales with considerable accuracy. Front line sales employees exchanged contracts representing the future sales volume based on their predictions of future printer sales. When trading ended, the contract valued most highly represented the most likely sales range. HP's official forecast erred by 13%, while the market erred by 6%. In further trials, the market performance exceeded the accuracy of official forecasts 75% of the time.
Product Development	Eli Lilly	Eli Lilly applied internal markets to predict correctly which of six potential new drugs would have the greatest success in passing product development hurdles. Employees involved in different stages of drug development traded market contracts based on their information. The market aggregated information with accuracy and opinion detail that would not have emerged had traders responded to a poll.
General Forecasting	Google	Google uses internal markets to forecast events such as new product launch dates and new office openings. The company applies market predictions to determine the likelihood that an event will occur on a specific date.
Project Milestones	Siemens	Siemens used internal prediction markets to predict software project milestones. On one occasion, traditional methods suggested a software project would be delivered on time, but the prediction market suggested it would be 2-3 weeks delayed. The project turned out to be 11 workdays late.
Project Milestones	Microsoft	Microsoft uses internal markets to predict whether projects will meet milestones articulated in their project plans.

Hewlett-Packard

An internal prediction market was designed and implemented at HP with the hope of producing more accurate printer sales forecasts than the firm's internal processes (Chen & Plott, 2002). A total of twelve predictions were performed over a period of three years. The prediction markets at HP included predictions for eight products. In some cases dollar sales were predicted and in other cases the number of units sold was predicted.

The market design employed at HP was the web-based double auction market of Marketscape software, developed at the Laboratory of Economics and Political Science at Caltech (Chen & Plott, 2002). From the web interface, participants could enter a buy offer, a sell offer, or acceptance of an offer. If a trade was possible, it was executed and if not, the order was placed in an order book.

In order to predict future sales of a product, HP established a prediction market with multiple securities, each associated with a particular sales volume interval. For example, if intervals of 100 units were used, there would be a security for 0-100 units sold, 101-200 units sold, and so on. Depending on the interval in which the final outcome falls, the corresponding security pays one dollar per share; all other securities pay nothing. Thus, the HP prediction markets used winner-take-all contracts.

The payoff for HP markets involved real money in which the "winning" security paid off a fixed amount; all other securities paid nothing. HP had issues engaging employees to participate in an activity in which they may lose money, thus HP supplemented participants with money at the beginning of the market sessions to ensure participation and minimize the potential employee loss (Chen & Plott, 2002). The markets at HP typically included 20-30 people, mostly from the marketing and finance divisions (Chen & Plott, 2002). Additionally, about five participants were from HP Labs, who had little or no information about the predicted event, but provided additional market liquidity.

The internal prediction market forecasts at HP were closer to the actual sales outcomes than the official forecasts in six out of eight events (Chen & Plott, 2002). These results sparked interest at other private-sector firms for using prediction markets to help forecast future sales or other market outcomes rather than relying on traditional forecasting methods alone.

Siemens

Siemens has used prediction markets for software projects. Ortner (1998) describes an implementation at Siemens in which an internal prediction market correctly forecasted that the firm would fail to deliver a software project on time even when traditional planning tools suggested the deadline could be met. The Siemens market, like HP, used a fully computerized double auction market with a software product called FX, developed by Kumo Inc. (Ortner, 1998).

For this software project Siemens created two separate prediction markets. One asked a simple question: Can the project be finished in the planned time horizon? The payoff rule was a simple winner-take-all design with *Yes* and *No* shares. Hence, the prevailing market price for a Yes security predicted the probability of meeting the planning time horizon, while the price of the No security predicted the converse probability.

The second market was designed to predict the length of the possible delay. This market included two shares called *Early* and *Late*, with Early shares yielding a greater payoff if the project was on-time or only a few weeks late, while Late shares yielded a greater payoff if the project was four or more weeks late. In particular, the payoff structure for this second market was set up in a linear fashion such that Early shares paid the maximum of (1 - 0.2* weeks late) or zero, while NO shares paid the minimum of (0.2* weeks late) or one. This payoff structure is illustrated in Figure 1. Thus, if the market consensus predicted a one week delay, we would expect EARLY shares to trade at a price of 0.8 while LATE shares traded at 0.2.

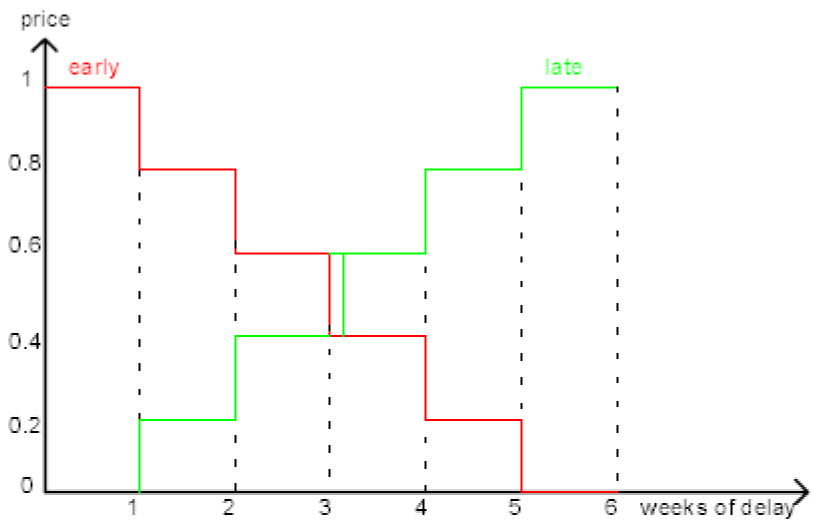

Figure 1. Pay-off structure at Siemens.
(Ortner, 1998)

Siemens opened the market to all people working in the project except upper level management. 63 traders joined the market and about 50 became active traders. Of the participant pool, 67% were developers, 31% group managers, and 2% project managers (second level—first level managers were not allowed to join the experiment because of their manipulating power; Ortner, 1998). The Siemens market did not use any uninformed traders.

Results at Siemens: Initially, after opening the two markets, the winner-take-all market YES shares approached a price of 0.43 and fluctuated between 0.43 and 0.40 for approximately six weeks. About one month prior to the deadline, the YES shares for the winner-take-all market plummeted indicating the market did not believe the project would reach its planned milestone, although it was still possible according to the traditional project plan used by the management team (Ortner, 1998). In the end, the market was closed when the project manager announced the milestone time limit was not reached. So each YES share paid 0 and the NO shares paid 1. In the second market, used to predict the time delay, after only 1 month of trading and more than 3 months before the scheduled deadline, the market predicted a delay of 2–3 weeks. In the end, the actual delay turned out to be 11 workdays.

HP and Siemens' experiences suggest that motivating employees to trade may be a major challenge, but the results of active trading can produce very valuable forecasts. Both firms ran real money exchanges with a relatively small trading population (20–60 people) and subsidized market participation by either providing traders with a portfolio or matching initial deposits. Even with the subsidies and small trading population, the predictive performance of these markets was remarkable.

Microsoft

Microsoft uses internal markets to predict whether projects will meet milestones articulated in their project plans. Microsoft's markets rely on an automated market maker that enables traders to access the market at their convenience to buy and sell contracts. By using a market maker, traders can exchange contracts without relying on others' willingness to buy or sell (the role of market-maker mechanisms is discussed in more detail later in this report).

Microsoft's prediction markets use multiple contracts, each representing a different predicted date on which a project will reach a certain milestone. Microsoft also has run test markets in the past involving naïve and informed traders. "Naïve traders did not impact the accuracy of market predictions because informed traders corrected market price fluctuations caused by naïve participants" (Corporate Executive Board, 2006, p. 10). However, Microsoft currently limits participation to informed traders because uninformed traders are less likely to participate. Microsoft selects its traders by targeting employees who have enough information to make educated trades and by selecting traders from different corporate functions to aggregate different types of information, giving more accurate results (Corporate Executive Board, 2006). The only concern is excluding someone from the market who has relevant information, but has been overlooked by management. Microsoft encourages participants to trade when they think they can contribute to the market.

5. Other Real World Prediction Market Applications and Results

The Iowa Electronic Market, run by the University of Iowa, is probably the best known prediction market amongst economists. The Iowa Electronic Market uses a double auction market with winner-take-all and index contract types. It is a real money market with no endowment, and participation is open to anyone interested, but likely only attracts those particularly intrigued and aware of the market's existence. In 1988, the original Iowa experiment allowed trades in a contract that paid 2½ cents for each percentage point of the popular vote in the presidential election received by Bush, Dukakis, and others. More recently, it has run prediction markets based on the 2008 presidential election, the 2008 congressional elections, and the 2008 Minnesota senate election.

The Iowa Electronic Markets have yielded very accurate predictions which outperform the forecasts from large-scale polling organizations (Berg, Forsythe, Nelson, & Rietz, 2001). Figure 2 shows data from the four U.S. presidential elections between 1988 and 2000. The horizontal axis shows the number of days until the election and the vertical axis displays the average absolute error between the prediction market price (linked to the two-party share of the popular vote) and the actual popular vote percentage earned in the election. In the last week before the elections, the prediction markets have predicted vote shares with an average absolute error of approximately 1.5 percentage points, compared to the final Gallup poll forecasts that differed by 2.1 percentage points (Wolfers & Zitzewitz, 2004).

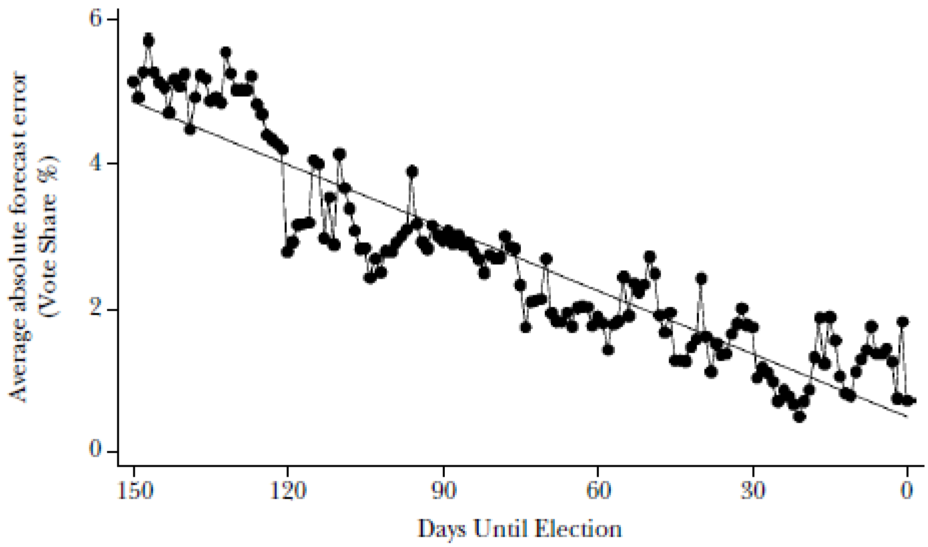

Figure 2. Information Revelation Through Time
(Wolfers & Zitzewitz, 2004)

The superior performance of the Iowa markets may be attributable to the fact that "traders are self-selected with a clear interest in predicting what will actually happen, rather than what they hope will happen" (Corporate Executive Board, 2006. P. 5). In a poll, respondents predict events without any context of others' beliefs. In a prediction market, each participant knows the current consensus and factors this information into decision-making.

Another example of the relative performance of a prediction market comes from the Economic Derivatives market established by Goldman Sachs and Deutsche Bank. This market is tied to macroeconomic outcomes, such as non-farm payrolls, retail sales, levels of the Institute for Supply Management's manufacturing diffusion index, and initial unemployment claims (Gürkaynak & Wolfers, 2005). The market mechanism is a pari-mutuel system where all bets that the specified outcome either will or will not occur are pooled for a given strike price; this pool is then distributed to the winners in proportion to the number of options purchased. The Economic Derivatives market uses multiple contracts, allowing traders to take a position on specified ranges in which the data will fall. The outcome results in a probability density function, which prior to this market was unavailable.

Figure 3 compares the performance of the Economic Derivatives market with a survey of economists in predicting economic outcomes based on data gathered by Gürkaynak and Wolfers (2005). This figure shows that the market-based forecast approximates the information in the survey-based forecasts. Additionally, the markets' response to data releases are better captured in the market-based expectations than survey-based expectations, suggesting that the markets perform and react better than survey-based forecasting (Gürkaynak & Wolfers, 2005).

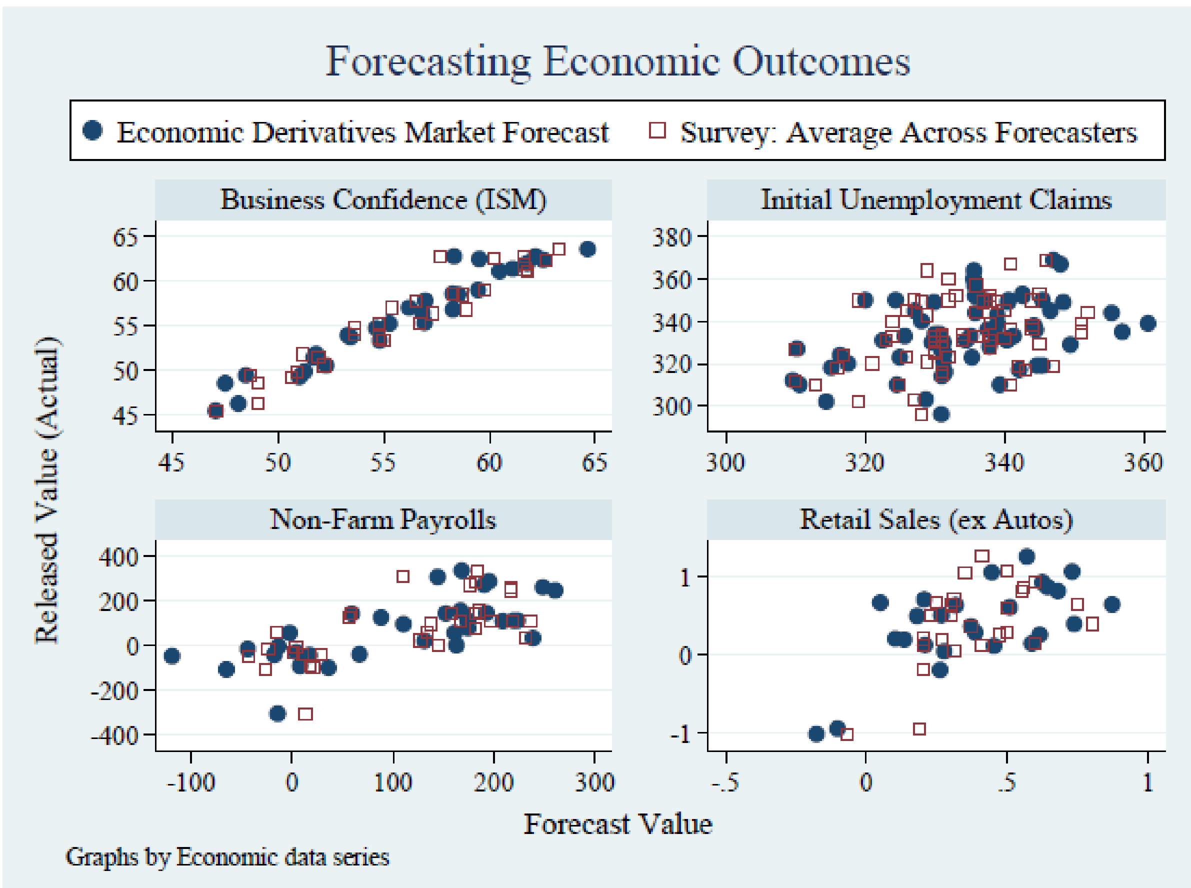

Figure 3. Forecasting Economic Outcomes
(Gürkaynak & Wolfers, 2005)

Table 2 summarizes some of the more popular prediction markets available for public trade.

Table 2. Popular Prediction Markets

Market	Focus
Iowa Electronic Markets <www.biz.iowa.edu/iem> Run by University of Iowa	Small-scale election markets.
Centrebet <www.centrebet.com> For profit company	Northern Territory bookmaker, offering odds on election outcomes, current events, sports, and entertainment.
TradeSports <www.tradesports.com> For profit company	Traded in political futures, financial contracts, current events, sports, and entertainment. No longer in business.
Economic Derivatives <www.economicderivatives.com> Run by Goldman Sachs and Deutsche Bank	Large-scale financial market trading in the likely outcome of future economic data releases.
Newsfutures <www.newsfutures.com> For profit company	Political, finance, current events and sports markets. Also technology and pharmaceutical futures for specific clients.
Foresight Exchange <www.ideosphere.com> Non-profit research group	Political, finance, current events, science and technology events suggested by clients.
Hollywood Stock Exchange <www.hsx.com> Owned by Cantor Fitzgerald	Success of movies, movie stars, and awards. Data used for market research.
Intrade <www.intrade.com> For profit company	Political, financial, current and similar event futures.

6. Prediction Market Design: The Devil is in the Details

The preceding sections of this report provide strong evidence of prediction markets' power and potential when they are "done right." However, the problem is that it can be very difficult to design and implement prediction markets in the right way. Much of the remainder of this report, in fact, will detail a number of the significant pitfalls and concerns in prediction market design.

Overall, it is a fair characterization to say that prediction markets may be very straightforward in basic principle, but in practice the devil is certainly in the details, and there are many details with which to be concerned. Everything from security or contact design, to trading rules, to incentives, to the number and characteristics of the traders, can influence the value and overall performance of any prediction market.

Small changes in any of these market design elements can have significant effects on overall prediction market performance. Thus, when deciding how best to design and implement any prediction market, it always pays to follow the carpenter's motto: Measure twice, cut once.

THIS PAGE INTENTIONALLY LEFT BLANK

7. Measuring Prediction Market Performance

Of course, before we can even discuss the impact of specific market design elements on prediction market performance, it is important to understand that measuring prediction market performance is itself a potential quagmire. What does it mean to say that a prediction market is "working well" or "not working well?" How would you even know how well it is actually working?

Prediction Market Prices as Predictors

Because the "prediction" provided by any prediction market is primarily reflected in the market price of the asset or contract being traded, it is essential to understand the microeconomics driving this price determination. Most fundamentally, the price of any contract sold in a prediction market is determined by how much the buyer is willing to pay for that contract as well as how much the seller is willing to accept for that same contract. These willingness-to-pay and willingness-to-accept amounts are determined by each trader's perceived "value" for that contract.

However, the perceived value of a contract traded in a prediction market is actually comprised of two elements: arbitrage value and intrinsic value. Arbitrage value is the potential financial benefit to the holder of a contract from reselling the contract at a higher price at some later point in time. Arbitrage value is the simplest interpretation of the "buy low, sell high" adage.

The intrinsic value of an asset, on the other hand, is the expected financial benefit of holding the contract indefinitely, or until the market closes. In financial stock markets, for example, the intrinsic value of shares in a company would be equivalent to the net present value of dividends or other future disbursements or financial benefits to shareholders.

Recognize that many people think about the value of shares in the stock market primarily in terms of their arbitrage value. The common thinking is something along the lines of, "I expect the stock price to rise significantly in the future, so I should buy shares now and sell them when the price goes up." While certainly a profitable endeavor if such a forecast of the future stock price is true, the logic falls short of answering the key questions: Why should we expect the price to rise or fall in the future? What drives future price changes?

If all traders focus only on the arbitrage value of contracts, then prices in prediction markets have little predictive value. In this case, the price in a prediction market would reflect only trader expectations about other traders' price expectations, which in turn reflect expectation about other traders' price expectations, creating a circular flow of expectations without any core foundation.

Therefore, for markets to have any predictive power this flow of expectations of expectations of expectations must, at some point, end with an expectation regarding the intrinsic value of the contract. As long as the flow of expectations ends with some trader expectation of the actual financial benefit of holding the contract indefinitely, then the market price has some predictive value. In particular, the market price of a contract traded in a prediction market could be seen to have some value in terms of predicting the future events which will ultimately determine that contract's true intrinsic value.

Note that each layer of expectation which drives the market price creates additional potential for miscalculation or error. In other words, trader expectations of the intrinsic value of a contract are likely to be more accurate (collectively) than trader expectations of trader expectations of trader expectations of the intrinsic value. Thus, the more that trades in prediction markets are driven by expectations of intrinsic value (as opposed to expectations of future prices), the more reliable market prices are as predictors of future events.

Thus, as we proceed to discuss the issue of measuring market performance, it is important to note that one criterion to apply is to determine whether a particular market design or practice fosters trading based more on perceived arbitrage value (generating less reliable predictions) or trading based more on perceived intrinsic value (generating more reliable predictions).

Prediction Market Performance: A Relative Measure

It is also important to note that prediction market performance (at least in terms of accuracy) cannot really be captured by any absolute measure. Instead, the only reliable way to measure prediction market accuracy is by using relative measures. In other words, any meaningful statement about the accuracy of a prediction market must be made relative to the accuracy of some other forecasting method.

To understand this, suppose we wanted to evaluate the accuracy of a prediction market designed to predict the outcome of a coin flip. In particular, suppose the contracts in this particular prediction market pay off exactly $1.00 if the outcome of the coin flip is heads, but pay off nothing if the outcome of the coin flip is tails. At what price should we expect contracts to trade in such a prediction market? The answer, of course, is 50 cents, which translates into a forecast of 50% chance of a heads outcome.

Of course, we already knew that there was a 50/50 chance of heads versus tails, so would it be fair to say that the prediction market "failed" to provide a valuable forecast? Of course not. The prediction market not only gave a forecast which was as accurate as any alternative, but it in fact produced the most accurate forecast possible.

As a further example, consider the forecasting accuracy of TV news meteorologists. Suppose that the TV weatherman in Las Vegas is far more accurate at predicting his city's annual precipitation than the TV weatherman in Seattle. Does this mean that the Las Vegas meteorologist is a better forecaster? Of course not. It is simply an easier task to forecast the weather in Las Vegas than in Seattle. This example draws to mind the Los Angeles TV weatherman portrayed by Steve Martin in the movie *L.A. Story* who simply pre-tapes a week's worth of weather forecasts so he can go on vacation.

Thus, measuring the performance of any prediction market requires considering not only the nature and difficulty of the particular prediction challenge but also considering the performance of alternative forecasting methods. This makes reliably measuring prediction market performance a particularly difficult task, as the complexity of the prediction task may be poorly understood, and there may not be viable alternative forecasting methods to serve as benchmarks for comparison.

Probabilistic Predictions: The Limited Observations Problem

A cursory examination of prediction markets now in practice suggests that many, if not a majority, of these markets rely on trading winner-take-all contracts which, as discussed previously, pay off a fixed amount if and only if a specific event occurs. Recall that the market price of a winner-take-all contract reflects the market expectation of the probability of a specific future event.

However, such probabilistic predictions introduce additional difficulty in terms of measuring accuracy. To see this, consider this statement from Todd Proebsting at Microsoft, who drew further on the weatherman analogy introduced above, noting that

> There's a common 'weatherman' misunderstanding about prediction markets, especially in the press. Perhaps counter-intuitively, a weatherman is not wrong if the sun comes out after a 90 percent forecast for rain because there was still a 10 percent chance of sunshine. Instead, the weatherman is a good predictor if it rains 90 percent of the time when he gives a 90 percent chance of rain—any more or less would be poor predictions. Prediction markets work the same way. (Corporate Executive Board, 2006. P. 9)

Thus, we cannot reliably measure the accuracy of winner-take-all prediction markets without a large sample of predictions to evaluate. Unfortunately, however, it takes time to generate such a large portfolio of predictions, and managers desiring to apply prediction markets as a decision tool rarely have the money or patience for extended "field testing" prior to actual reliance upon the market predictions. Thus, any prediction market design would need to be tested and evaluated extensively in a controlled "laboratory" environment prior to actual implementation.

Measuring Performance by Measuring Information Aggregation

It is also important to note that the basic premise of measuring prediction market performance based on forecast accuracy is itself inherently flawed. Decision-makers, of course, ultimately are most concerned about the accuracy of any forecast upon which they rely. However, a perfectly designed and well-functioning prediction market may often produce fairly imprecise forecasts while, in contrast, a very badly designed and poorly functioning prediction market may produce comparatively more precise forecasts.

How is this possible, you ask? The answer is rooted in the fact that the predictive power of prediction markets is fundamentally based on the ability of such markets to efficiently aggregate collective knowledge. In essence, prediction markets only reveal what is already known, but these markets add value by collecting and integrating knowledge that may be atomized, dispersed, and otherwise hidden. However, the accuracy of any prediction market forecast is ultimately limited by the accuracy of collective trader knowledge once it is gathered and integrated. In other words, garbage in, garbage out.

Consider once again the example from Section 3 of this report, in which there are three possible (mutually exclusive and a priori equally likely) future outcomes, labeled A, B, and C, along with a prediction market in which traders buy and sell shares associated with each outcome (A shares, B shares, and C shares, respectively). Recall that the shares in the prediction market are winner-take-all contracts, with each outcome's associated shares paying off $100 each if that particular outcome occurs and paying off $0 otherwise. Hence, a prevailing market price of $X for a particular outcome's shares indicates a market predicted probability of that outcome equivalent to X%.

Now suppose we were trying to evaluate two competing prediction market designs—labeled design #1 and design #2—each with different rules or procedures, information or communication, endowments, market-maker algorithms, and so on. Consider what might happen if we were to compare the performance of these two competing market designs in terms of their ability to accurately forecast which outcome (A, B, or C) will ultimately occur in the scenario above.

In testing these two competing market designs, of course, a different population of traders would have to participate in each market design. If some traders were to participate in both market designs, the "cross-contamination" of information would make it unclear which of the two designs should really be given credit (or blame) for the accuracy (or inaccuracy) of any given forecast.

The two distinct populations of traders engaged in design #1 and design #2, however, would presumably have different collective bases of knowledge. So, suppose some traders participating in market design #1 know with certainty that outcome A will not occur while the remaining traders in this population have no special knowledge and thus consider outcomes A, B, and C all equally likely to occur. For the population of traders participating in design #2, in contrast, suppose that some of the traders know with certainty that outcome A will not occur, some of the traders in this same population know with certainty that outcome B will not occur, while the remaining traders in this population have no special knowledge and thus consider outcomes A, B, and C all equally likely to occur.

With these knowledge conditions in mind, consider the implications of the following outcome:

(1) Market design #1 produces final prices for each outcome of $0 for A, $50 for B, and $50 for C;

(2) Market design #2 produces final prices for each outcome of $0 for A, $30 for B, and $70 for C; and

(3) Outcome C does, in fact, occur.

What might be concluded regarding the two market designs based on these observations?

The easy answer would be to conclude that design #2 must be a better prediction market design (i.e., it has better rules, procedures, endowments, algorithms, and so on) because it forecasted a 70% probability of the correct eventual outcome (C), whereas design #1 only forecasted a 50% probability of the correct outcome. However, the truth is actually quite the opposite: Design #1 outperformed design #2 in the basic dimension of prediction market success—information aggregation.

In producing final prices of $0 for A, $50 for B, and $50 for C, market design #1 achieved 100% information aggregation. The best estimate of the probability of outcome C occurring given the entire collective knowledge of the design #1 trader population was, in fact, 50%. However, in contrast, the final prices of $0 for A, $30 for B, and $70 for C generated by market design #2 reveal a failure to fully aggregate the collective knowledge of the trader population. In particular, a fully efficient prediction market using the trader population of design #2 would have produced final prices of $100 for $C and $0 for the other two outcomes. Instead, however, market design #2 essentially "left knowledge on the table" and fell short of full information aggregation.

Of course, an outside observer would be unable to recognize the superior performance of market design #1 in this example because the knowledge or information that is to be aggregated is hidden from observation. Therefore, in real world implementations of prediction markets one can only observe the final forecasts produced by the market but cannot ultimately know whether the particular market design was, in fact, doing its job by fully aggregating the collective base of knowledge within the trader population.

Thus, the only way to truly evaluate the performance of prediction market design elements is to do so within a controlled setting in which you can observe the information conditions prior to implementing the prediction market. In particular, a researcher must control the information that is possessed by each individual trader by, in fact, being the sole provider of that information himself or herself. In sum, effectively evaluating potential prediction market designs requires the researcher (or evaluator) to *induce* the initial information conditions by providing certain information to each trader. The traders then buy and sell contracts using the market design under evaluation. Ultimately the researcher measures the degree to which the final market prices fully aggregate the information initially disseminated across the trader population.

For this reason, it is strongly recommended that any future evaluation of prediction markets as a defense management tool include careful and comprehensive laboratory experimentation to evaluate and compare alternative market designs. The bottom line is that such controlled experiments, in which the degree of information aggregation can actually be measured, are the only true means for testing the basic performance (at least in principal) of any prediction market design.

8. Limitations of Prediction Markets as a Decision Tool: The Endogeneity Problem

The remarkable accuracy of prediction markets in forecasting election results, economic outcomes, and other variables has managers in both the private sector and public sector intrigued by the possibility of applying these markets as a managerial decision tool. Some imagine using prediction markets to forecast consumer response to new product alternatives and then adjusting marketing decisions in accordance with these forecasts. Others envision using prediction markets to forecast the future success of the various elements within an organization's research, acquisition, or investment portfolio and then allocating resources where the potential returns are most promising. The possibilities in which prediction markets might be able to guide decision-making seem endless.

Unfortunately, however, there is an inherent limitation in using prediction markets for many decision-making applications. In particular, if the decision which the prediction market is intended to assist will itself impact the future variable(s) the market is designed to forecast, there exists an endogeneity problem that will limit the accuracy and/or usefulness of any such prediction market. In particular, decision-makers may not be able to decipher whether the prediction market is providing a forecast of the future without any managerial action or whether the prediction market is, in fact, anticipating certain managerial action and therefore is forecasting a future which already incorporates the impact of the very decision the market is intended to assist.

For example, consider a potential application of prediction markets to aid defense manpower decision-making. Suppose, in particular, that the Navy implements a prediction market forecasting retention outcomes (i.e., whether or not certain goals will be met or even the actual number of sailors retained). If this market forecasts low retention levels, the Navy might consider responding with corrective action, such as increasing retention bonuses. The problem, of course, is that an efficient and effective prediction market should have already anticipated such action by the Navy, so the bonus increase could have zero or minimal impact on the retention levels forecast by the prediction market. Observing this, the Navy might then decide to increase retention bonuses even further, but rational traders in the prediction market would have anticipated this action as well. The bottom line is that Navy decision-makers would never be able to know whether the prediction market was forecasting retention levels given the current bonus amounts or assuming some adjusted (unknown) bonus amount. Hence, the value of such a prediction market for guiding manpower decision-making would be seriously compromised.

This endogeneity problem was one of the critiques of the Defense Advanced Research Projects Agency (DARPA) initiative, which had prediction markets for terrorist acts and assassinations of certain world leaders. For example, suppose that some intelligence analyst believed that a certain country's leader was going to be assassinated. That analyst would then buy shares in that outcome, driving the price up. Officials would observe the increased market probability of that assassination occurring, so they may take steps to deter it, which would bring the probability and the price down. Thus, the analyst's well-informed investment would then lose money (or at least gain little to no return). If the analyst had rationally anticipated these deterrent actions, however, he or she would never have made the investment in the first place, understanding that there is little opportunity for profit.

So, how does one address this endogeneity problem? Our research has so far uncovered two different approaches in the area of prediction market contract design and trading:

(1) decision-independent contracts, and

(2) conditional contracts.

A *decision-independent contract* is a contract whose prediction market price provides valuable information for a particular managerial decision under consideration, yet whose outcome will be unaffected by that decision. Consider, for example, a decision-supporting prediction implemented by Google to determine how many people would sign up for their e-mail services (Dye, 2008). The purpose of the market was not to adjust marketing activities or other variables which could influence the outcome being forecasted, but rather to estimate how much bandwidth/storage to allocate to the e-mail services. Thus, unlike the example of the prediction market for retention outcomes described above, the policy response to the Google prediction market would not affect the outcome forecasted by the market.

Designing and trading *conditional* contracts is a second approach to overcoming the endogeneity problem when a prediction market is to be used as a decision support tool. Conditional contracts are contracts whose payoffs depend not only on the forecasted outcome, but also upon the presence of certain assumed prior conditions. For example, a conditional contract designed to forecast retention outcomes might be formulated around the following question: If a retention bonus of $10,000 is offered, how many Service members in this specialty area will be retained? Similar conditional contracts might be designed around bonus levels of $5,000, $15,000, or any other amounts under consideration.

Such conditional contracts have the benefit that they are unaffected by the managerial decision they are intended to inform, and they also allow management to measure the direct variable they may be hoping to influence with their decision (in this case, retention levels). The primary drawback of such conditional contracts is that there must be a clear and non-distorting rule for how much these contracts pay off if the antecedent condition does not occur. In the example above, for instance, what is the contract worth if the retention bonus is ultimately not set at $10,000 or whatever amount was pre-supposed by the contract conditions? Note that, if the answer to this question is that such contracts become worthless, then their predictive value is destroyed because the value (and, thus, market price) of such contracts will always be downwardly influenced by traders' estimated probability that the pre-supposed condition will change.

Instead, to maintain the predictive value of conditional contracts, the prediction market designer must formulate and publicize a “terminal value” rule that will not (or, at least, should not) have any effect on trader value for conditional contracts, and thus not create any market price distortion. For example, the terminal value rule could state that if the particular circumstance pre-supposed by a conditional contract does not ultimately occur, then each share of that contract will pay off an amount equal to the average market price over some period of time prior to market closing, or the end of trading for that particular contract. Such a rule for paying off “orphan” or “lame duck” conditional contracts should exert no pressure, upwards or downwards, on market prices, thus preserving the predictive power of these contracts. The only impact of such a rule should be to reduce trading and stabilize the market price, at a level indicative of the final prediction, as it becomes clear that the prerequisite condition will not become reality.

9. Incentive Issues in Prediction Markets with "Play Money"

An additional concern that is particular to applying prediction markets to defense management is the issue of trader incentives. In particular, due to legal or cultural restrictions related to pay for performance or even gambling, there is an aversion to using real monetary incentives for prediction markets involving government employees as traders. This raises a concern because the best performing prediction markets, whether as managerial decision-aids or forecasting elections and financial outcomes, have all used the potential for real money profits (or losses) as the traders' incentive.

Whereas there has been some research to indicate the potential value of prediction markets that employ only "play money" incentives (Servan-Schreiber, Wolfers, Pennock, & Galebach, 2004), these results must be interpreted with some skepticism. More specifically, this particular research investigated trading to predict the outcome of sporting events, for which many individuals have an inherent incentive to participate. For example, millions of people join fantasy sports leagues (which essentially involve predicting how individual athletes are going to perform) without any financial incentive, sometimes even paying to participate. The fact that individuals can be motivated to trade in sports prediction markets with only play money is thus not surprising. Such inherent motivation is far less likely to be present in prediction markets associated with outcomes of managerial interest, such as technological progress or cost inflation of a new product design.

Some have argued that bragging rights or a sense of competition should be sufficient motivators for active trading in play-money prediction markets. Whether or not this is true, these non-financial incentives create their own problems. First of all, so-called bragging rights require non-anonymity. After all, one can't brag about superior performance if others don't know how well you have done. However, in many cases the potential anonymity of trading in prediction markets is a key advantage. In particular, a trader may have inside information that a particular project is "doomed to failure" but, for political reasons, is uncomfortable revealing this information publicly. However, an anonymous prediction market would allow such a person to act on, benefit from, and disseminate this valuable inside information without speaking out publicly as a whistle-blower.

There's also a fundamental question regarding how trader performance would even be measured in such play-money markets. While it is a common practice to rank or recognize traders at any point in time based on the current value of their overall portfolio (cash-on-hand plus current market-value of all contracts held), such a measurement system has inherent problems. In particular, given that the purpose of the prediction markets is to forecast future outcomes, you would presumably like to recognize traders based on the accuracy and advance foresight of their predictions regarding these outcomes. However, ranking traders or measuring performance based on current portfolio value fails to do this, instead rewarding or recognizing traders for their ability to predict market prices (or other trader's predictions) rather than the actual outcome of events in question.

At any point in time, the true intrinsic value of any individual's portfolio in a prediction market is unknown, and can only be evaluated after the forecasted events have occurred. To provide a concrete example, consider an individual who may have held an enormous amount of Enron stock before that company collapsed amid scandal. Prior to the Enron crash, such an individual would have been rated or ranked very highly by the portfolio value criterion discussed above when, in fact, he had actually been a very unwise investor.

Sometimes play-money prediction markets are enhanced by offering a prize to the top trader or traders. Besides the performance measurement problem discussed above, this practice of rewarding (or even just recognizing) only the top performers creates additional market-distortion problems. In particular, such a reward-only-the-top approach disproportionately diminishes the value to traders of small gains and increases the value of large gains.

For example, consider the decision-problem of a trader who is near the “back of the pack” as the prediction market approaches closure. Small or medium gains are worthless to the trader at this point, making the only trades of value those that have the potential for large gains that could immediately vault him into the group of leading traders who will be rewarded or recognized. Thus, the reward-only-the-top incentive approach drives traders to go for “home runs,” focusing on high-risk/high-reward investments to the detriment of low-risk/low-reward investments. This could cause the market price of low-probability events to be overstated and the market price of high-probability events to be understated.

THIS PAGE INTENTIONALLY LEFT BLANK

10. Security or Contract Design Issues

While both public and private sector leaders are often easily intrigued by the potential benefit of prediction markets as a forecasting tool, our research indicates that they often find it far more difficult to actually identify future events to predict. In particular, the design requirements for a viable (and tradable) prediction market contract can prove surprisingly difficult to meet.

For example, to produce reliable and valuable forecasts, a prediction market contract must be designed to measure something that is:

- *Important:* First and foremost, the contracted event must be something for which the outcome is relevant to managerial decision-making.
- *Quantifiable:* Prediction market contracts should not be based on subjective assessments of “success” or some other measure. Instead, there must be a specific quantifiable outcome measure on which the contract payoff depends (even if this measure is a zero/one yes or no measurement).
- *Clearly defined:* The event in question and the potential outcomes of that event must be clearly understood by all traders. Moreover, when there are different potential measurements of the event it must be 100% clear upon which measurement the contract pay off will ultimately depend.
- *Contingency covered:* The contract pay off under all possible scenarios must be determined, specified, and clearly communicated. For example, what happens if the event in question does not actually occur? What happens if the measurement in question doesn’t take place? What happens in the case of a tie or other knife-edge outcome?
- *Decision-independent:* As discussed earlier, the event being forecast cannot be affected by the managerial decision which the forecast is intended to support.
- *Subject to alternative forecasts:* While not essential, it is valuable to employ prediction market contracts for events which are also forecast via other means (expert analysis, official estimates, opinion polls, etc.). Doing so allows decision makers to determine whether prediction market forecasts do indeed outperform alternative methodologies.

In principle, each of these contract design requirements is intuitively justifiable. However, in practice these restrictions may significantly limit the pool of potential events to be predicted. We discovered this firsthand when in a pilot prediction market study conducted for The Chief of Naval Personnel (N1).

11. Student Pilot Study

Two Naval Postgraduate School (NPS) students, LT Michael Chinn and LT Leslie Huffman, implemented a prediction market to explore prediction market applicability for manpower related outcomes. The prediction market took place from early August 2009 through September 2009 and the results are reported in their MBA Joint Applied Project Report (Chinn & Huffman, 2009). In conducting this experiment, Chinn and Huffman addressed five prediction market design issues: claim definition, claim structure, participation incentives, market participants, and trading mechanism. The pilot test was funded through the Chief of Naval Personnel.

Claim Definition

The pilot study was initially planned to involve several prediction market questions related to manpower outcomes and a few “fun” questions to complement and encourage participation. N1 had offered guidance on which manpower outcomes we were going to use. However, as Chinn and Huffman (2009) detail, claim definition proved more difficult than expected; outcome measurements for some of the proposed claims were ambiguous or poorly defined, other proposed claims had largely predetermined outcomes or were likely targets for management intervention if they diverged from their final target values. The final portfolio of claims included five with Navy relevance, either because they touched directly upon a Navy manpower outcome or because they involved general economic conditions that significantly impact Navy manpower issues:

- What will be the Navy's end-strength (for officers and enlisted personnel) for FY2009?
- On September 30, 2009, what will the Navy's FY10 enlisted accession goal be?
- What will be the official September 2009 national seasonally-adjusted unemployment rate (per the U.S. Department of Labor)?

- Will the Dow Jones Industrial Average (INDU) close above 9,400 by Close of Business (COB) on Friday, August 14, 2009?
- Will the FY 2010 Defense Appropriation bill be signed into law before October 1, 2009?

There were four other questions on baseball, football, and the Emmy awards. These were intended to add a little fun to the market and hopefully spur more involvement in the military-related outcomes. The fun questions were introduced periodically throughout the pilot study to maintain participant interest throughout.

Claim structure

All claims were winner-take-all contracts, though some questions involved multiple securities with outcomes defined over a specific range. As such, the market price for all contracts reflected the probability for that specific outcome.

Participation Incentives

Due to Legal and cultural restrictions, the pilot market was restricted to play money with the primary motivation being bragging rights for the top performers. In this pilot it was possible to determine a final ranking because all contracts closed at the end of the pilot. However, the value of bragging rights was limited because most traders registered with pseudonyms and could not be individually identified. Pseudonyms were used to provide participants anonymity as desired.

Market Participants

The pilot market sponsor identified 58 potential participants, 26 of whom were from within the sponsor's organization, five were from the Naval Postgraduate School (the students and project advisors), and the others were from manpower-related offices from throughout the Navy. Participants were individually invited to participate in the pilot prediction market. All participants were considered informed; uninformed participants were not invited. With a play money market and bragging rights as the only incentive, uniformed bidders would not likely participate or provide any meaningful liquidity.

Participation in the prediction market was very low, as feared. Of the 53 non-NPS potential participants, only 19 created an account and had at least one trade. Trading also fell sharply as time progressed, from 57 non-NPS trades in the first week, to 15 or less in the third week and all subsequent weeks. The introduction of the fun questions had no apparent effect on the trading in the Navy manpower-relevant questions.

Prediction Market Trading Mechanism

The pilot prediction market was implemented through Inkling Market's website (http://inklingmarkets.com/). The owner of Inkling provided the market platform free of charge; there is normally a charge per user. The Inkling owner met with LT Chinn and LT Huffman, and was very helpful in establishing the market. The Inkling platform involves an automatic market maker. Thus, one person's wish to buy a stock does not depend on someone else wanting to sell the stock for an agreed-upon price. This is a critical design consideration in thin markets with limited participants and trading activity. In addition, the Inkling interface is relatively easy for a user to understand, an important consideration for the inexperienced participants involved in this pilot.

Lessons Learned: Implementing prediction markets is difficult

The primary implementation issue in the pilot prediction market involved limited participation. As mentioned above, participation was low and fell off dramatically. In addition, adding fun questions had little impact on overall participation. One potential explanation for low participation may involve the limited number of participants with relevant knowledge of the contract outcomes, even within the select group of invited traders. The traders with little knowledge may be discouraged from participating because there are others in the market with a significant inside information advantage, making it difficult for less informed traders to profit from trading. For example, one participant sent a weekly message updating the current result for one of the securities. Other traders likely felt a significant disadvantage in trading this security.

A second problem is that the traders may not have the time to participate in the pilot prediction market. In fact, some may think that those who participate frequently may not be busy enough with their regular work responsibilities. Time requirements involve both individual trades and the initial investment to understand the Inkling prediction market platform and trading rules. While the Inkling platform is relatively straightforward, some participants had a difficult time grasping the concept of short selling—that is to say, betting on the probability of an event (i.e., the price) going down. On Inkling.com, all traders need to do is indicate that they think the actual probability is lower than the current price, then they are asked how much lower they think the true probability is. Based on this information, inkling.com determines a number of shares to sell, or the trader can enter the number of shares to sell. Therefore, it is straightforward, yet it is difficult to grasp. Confusion over trading rules could limit trader participation.

In addition, incentives to play may be inadequate. The Navy currently precludes using real money. In conducting the pilot PM, we were not able to offer iPods or other material incentives. Still, other incentives could be used, such as a high-quality parking space. However, such a system of rewarding the top money earner or the top few earners would compromise anonymity and create adverse trading incentives that could skew the true probabilities. As described previously, rewards for top performers would encourage lower performing traders to make risky investments to capture the large potential pay-offs. Thus, people would be more likely to buy securities for low-probability events and sell short against high-probability events. This would inflate the probability of low-probability events and understate the chance of high-probability events occurring.

It seems essential to provide some participation incentives. Participation in the pilot prediction market was very low, even with fun, non-military securities. The pilot market was initially described as a student project, as opposed to a sanctioned Navy pilot, which may have further compromised participation.

Lastly, implementing a prediction market is a time-consuming, difficult task. As highlighted previously, the devil is in the details when managing prediction markets. It is critical to ensure that the securities are worded well without ambiguities. Ambiguities could arise from many sources, including simple differences in how an outcome is defined. For example, the pilot market considered a security involving the number of people waiting to join the Navy via the Delayed Entry Program (DEP). But, different Navy organizations used different methods to calculated DEP numbers. Thus, this question could have caused too much confusion.

There was a different problem in defining a security on reenlistment rates for nuclear-trained sailors. Unbeknownst to the pilot market designers, the Navy had suspended retention bonuses for nuclear-trained sailors for the rest of the fiscal year, but the bonus was to be reinstated in the next fiscal year (recall the pilot market was conducted between August and September 2009, the last two months of the fiscal year). Thus, no nuclear-trained sailors would reenlist until the new fiscal year. While still feasible, the prediction market question on nuclear-trained sailor retention became uninteresting.

These lessons learned highlight a few of the implementation issues involved in designing and running a prediction market. If the details are improperly addressed, the markets will be confusing and uninformative, at best. If early pilots are poorly designed, prediction markets may well be perceived as a flawed concept as opposed to a useful concept with a flawed implementation. It is critical that early demonstrations be well designed and appropriately analyzed to allow the Navy and the Department of Defense (DoD) to accurately assess their potential value.

THIS PAGE INTENTIONALLY LEFT BLANK

12. Conclusion: The "Known Unknowns" of Military Prediction Markets

Prediction markets are both powerful and complicated. They provide the best available opportunity to aggregate disparate, decentralized information yet they are complicated mechanisms posing complex implementation issues. There is substantial information regarding prediction markets in general, including information regarding claim definition, claim structure, participation incentives, market participants, and trading mechanisms. At the same time, there are several factors regarding prediction market applications in Navy and Defense Department applications that are still poorly understood. Critical issues involve market participation, contract definition, and performance measurement.

One issue concerns whether the military-oriented prediction markets can elicit adequate participation to generate accurate predictions. Having sufficient knowledgeable participation requires a combination of an adequate informed trader population and adequate participation rates among those potential traders. While there is likely a large knowledgeable population for most defense-oriented issues, it is uncertain whether incentives are sufficient to draw traders into the prediction market and to keep them active over time. Would defense markets be able use sufficient incentives, and, if so, what incentives could they use?

A connected issue is how participation would be affected for short-term versus long-term securities. In the pilot prediction market, all securities were relatively short-term (less than two months). As mentioned previously, there were trades in the first two weeks, but the number of trades dropped off dramatically. This may reflect participants' loss of interest, or it may reflect other factors, such as little new information on the contract outcomes. If a prediction were to start at the beginning of a fiscal year on what some outcome would be by the end of the fiscal year, it is uncertain how the participation would persist over the fiscal year. It is possible that emerging information bearing on the outcome over the course of the year would keep people interested in the outcome. However, the new information might come out too sporadically to generate sufficient interest for participants to continue participating. It is unclear whether government employees can be offered incentives that promote participation. Using taxpayer money to pay incentives could cause problems, but there may be ways to address this.

Another issue involves the outcomes that are good candidates for PM securities but are not self-defeating (endogenous). As discussed above, outcomes in which policymakers would likely affect the outcome based on the prediction market contract prices are self-defeating. This endogeneity likely encompasses a large share of potential defense outcomes (particularly in manpower applications). Still, there are likely good security candidate outcomes. It is important to define contract outcomes so that they avoid policy intervention (e.g., predict intermediary variables as opposed to policy-relevant final outcomes), so that they involve ranges of potential outcomes, or so that they include payoff criteria if policy interventions alter the underlying conditions.

Another uncertain issue is whether the success or usefulness of PM's can be measured. Previous attempts have not been satisfactory. Outcomes are probabilistic events. We previously outlined a theoretical method of measuring the success of markets, but this required a large number of securities and known information conditions. Performance measurement requires carefully controlled information and bidder interaction relationships.

References

Berg, J., Forsythe, R., Nelson, F., & Rietz, T. (2001). Results from a dozen years of election futures markets research. In C. Plott & V. Smith (Eds.), *Handbook of experimental economic results*, 742-751. Amsterdam: North-Holland.

Berg, J., Nelson, F., & Rietz, T. (2008). Prediction market accuracy in the long run. *International Journal of Forecasting*, *24*, 285–300.

Chen, K. & Plott, C. (2002). *Information aggregation mechanisms: Concept, design and implementation for a sales forecasting problem* (Social science working paper No. 1131). Pasadena, CA: California Institute of Technology.

Chinn, M., & Huffman, L. (2009). *Prediction markets: A review with an experimentally based recommendation for Navy force-shaping application* (NPS-AM-09-139). Monterey, CA: Naval Postgraduate School.

Corporate Executive Board. (2006, June). Using internal markets to improve corporate decision-making: Lessons learned by early adopters. *Corporate Strategy Board Primary Brief.*

Dye, R. (2008). The promise of prediction markets: A roundtable. *The McKinsey Quarterly*, *2*, 83–93.

Gürkaynak, R., & Wolfers, J. (2007). Macroeconomic derivatives: An initial analysis of market-based macro forecasts, uncertainty, and risk. In Frankel, J. & Pissarides, C. (eds.) NBER International Seminar on Macroeconomics 2005 (NBER International Seminar on Macroeconomics Series). Cambridge, MA: The MIT Press.

Kiviat, B. (2004, July 06). The end of management? *TIME Magazine.*

Malone, T. (2004, April). Bringing the market inside. *Harvard Business Review ,82*, 106-114.

Ortner, G. (1998). *Forecasting markets—An industrial application.* Mimeo, Austria: Technical University of Vienna1998.

Plott, C., & Sunder, S. (1982). Efficiency of Experimental Security Markets with Insider Information: An Application of Rational-Expectations Models. *Journal of Political Economy*, 90(4), 663-698.

Servan-Schreiber, E., Wolfers, J., Pennock, D. M., & Galebach, B. (2004). Prediction markets: Does money matter? *Electronic Markets, 14*(3), 243–251.

Tziralis, G., & Tatsiopoulos, I. (2007). Prediction markets: An information aggregation perspective to the forecasting problem. *World Review of Entrepreneurship, Management and Sustainable Development, 3*(3/4), 251–259.

Wolfers, J., & Zitzewitz, E. (2004). Prediction markets. *The Journal of Economic Perspectives, 18*(2), 107–126.

2003 - 2011 Sponsored Research Topics

Acquisition Management

- Acquiring Combat Capability via Public-Private Partnerships (PPPs)
- BCA: Contractor vs. Organic Growth
- Defense Industry Consolidation
- EU-US Defense Industrial Relationships
- Knowledge Value Added (KVA) + Real Options (RO) Applied to Shipyard Planning Processes
- Managing the Services Supply Chain
- MOSA Contracting Implications
- Portfolio Optimization via KVA + RO
- Private Military Sector
- Software Requirements for OA
- Spiral Development
- Strategy for Defense Acquisition Research
- The Software, Hardware Asset Reuse Enterprise (SHARE) repository

Contract Management

- Commodity Sourcing Strategies
- Contracting Government Procurement Functions
- Contractors in 21st-century Combat Zone
- Joint Contingency Contracting
- Model for Optimizing Contingency Contracting, Planning and Execution
- Navy Contract Writing Guide
- Past Performance in Source Selection
- Strategic Contingency Contracting
- Transforming DoD Contract Closeout
- USAF Energy Savings Performance Contracts
- USAF IT Commodity Council
- USMC Contingency Contracting

Financial Management

- Acquisitions via Leasing: MPS case
- Budget Scoring
- Budgeting for Capabilities-based Planning
- Capital Budgeting for the DoD
- Energy Saving Contracts/DoD Mobile Assets
- Financing DoD Budget via PPPs
- Lessons from Private Sector Capital Budgeting for DoD Acquisition Budgeting Reform
- PPPs and Government Financing
- ROI of Information Warfare Systems
- Special Termination Liability in MDAPs
- Strategic Sourcing
- Transaction Cost Economics (TCE) to Improve Cost Estimates

Human Resources

- Indefinite Reenlistment
- Individual Augmentation
- Learning Management Systems
- Moral Conduct Waivers and First-term Attrition
- Retention
- The Navy's Selective Reenlistment Bonus (SRB) Management System
- Tuition Assistance

Logistics Management

- Analysis of LAV Depot Maintenance
- Army LOG MOD
- ASDS Product Support Analysis
- Cold-chain Logistics
- Contractors Supporting Military Operations
- Diffusion/Variability on Vendor Performance Evaluation
- Evolutionary Acquisition
- Lean Six Sigma to Reduce Costs and Improve Readiness

- Naval Aviation Maintenance and Process Improvement (2)
- Optimizing CIWS Lifecycle Support (LCS)
- Outsourcing the Pearl Harbor MK-48 Intermediate Maintenance Activity
- Pallet Management System
- PBL (4)
- Privatization-NOSL/NAWCI
- RFID (6)
- Risk Analysis for Performance-based Logistics
- R-TOC AEGIS Microwave Power Tubes
- Sense-and-Respond Logistics Network
- Strategic Sourcing

Program Management

- Building Collaborative Capacity
- Business Process Reengineering (BPR) for LCS Mission Module Acquisition
- Collaborative IT Tools Leveraging Competence
- Contractor vs. Organic Support
- Knowledge, Responsibilities and Decision Rights in MDAPs
- KVA Applied to AEGIS and SSDS
- Managing the Service Supply Chain
- Measuring Uncertainty in Earned Value
- Organizational Modeling and Simulation
- Public-Private Partnership
- Terminating Your Own Program
- Utilizing Collaborative and Three-dimensional Imaging Technology

A complete listing and electronic copies of published research are available on our website: www.acquisitionresearch.org

THIS PAGE INTENTIONALLY LEFT BLANK

Initial Distribution List

1.	Defense Technical Information Center 8725 John J. Kingman Rd., STE 0944; Ft. Belvoir, VA 22060-6218	2
2.	Dudley Knox Library, Code 013 Naval Postgraduate School, Monterey, CA 93943-5100	2
3.	Research Office, Code 09 Naval Postgraduate School, Monterey, CA 93943-5138	1
4.	William R. Gates Dean, GSBPP Naval Postgraduate School, Monterey, CA 93943	1
5.	Stephen Mehay Associate Dean for Research, GB Naval Postgraduate School, Monterey, CA 93943	1
6.	Peter Coughlan Associate Professor, GB Naval Postgraduate School, Monterey, CA 93943	1
7.	William Gates Associate Professor & Dean, GB Naval Postgraduate School, Monterey, CA 93943	1
8.	Jeremy Arkes Professor, GB Naval Postgraduate School, Monterey, CA 93943	1

Copies of the Acquisition Sponsored Research Reports may be printed from our website: www.acquisitionresearch.org

THIS PAGE INTENTIONALLY LEFT BLANK

ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF BUSINESS & PUBLIC POLICY
NAVAL POSTGRADUATE SCHOOL
555 DYER ROAD, INGERSOLL HALL
MONTEREY, CALIFORNIA 93943

www.acquisitionresearch.org