Assessing the performance of different prediction market formats in forecasting tasks

John-William Awbrey

28529864

A research project submitted to the Gordon Institute of Business Science, University of Pretoria, in partial fulfilment of the requirements for the degree of Master of Business Administration.

9 November 2011

Abstract

Prediction markets have recently gained favour with the academic and business communities. Prediction markets have evolved a long way from their basic beginnings as friendly wagers among friends to become large scale markets connecting traders from around the world. They have been adopted into many large and dynamic corporations that require up to the minute information that can keep up with their business. Organisations like Google, HP, Yahoo! and Best Buy have been experimenting with prediction markets for demand forecasting tasks. Governments have also been using markets, although not always as successfully. The U.S. government looked at PAM which became the terrorist futures market in the post 9/11 world. This did not appeal to the American populous and it has since been withdrawn.

Through technological advancements the capabilities and availability of prediction markets has grown. With this the interest in how they work and what can be done to improve the accuracy of the markets. This research looked at the inclusion of a deliberative technique to the markets to improve that accuracy of the market. For this research, markets that made use of discussion boards were used. They were compared against traditional markets, which had no means of communication between traders.

The research took the form of a quantitative comparison between the two market types. Data was acquired from the Iowa Electronics Market (IEM) and Inkling Public Markets. The findings from this research indicate that there was a significant difference with α=0.012 for the markets at close. This indicated that there was a significant between the traditional (control) and non-traditional (experimental groups) markets from descriptive statistics it was indicated that the traditional markets performed better in the prediction tasks.

The conclusions of this research indicate that allowing traders to communicate and see the actions of others creates group biases which impacts on their independence when making trades and thus on the performance of the market.

Key Words

Prediction Markets, Delphi, Convergence,

I declare that this research project is my own work. It is submitted in partial fulfilment of the requirements for the degree of Master of Business Administration at the Gordon Institute of Business Science, University of Pretoria. It has not been submitted before for any degree or examination in any other University. I further declare that I have obtained the necessary authorisation and consent to carry out this research.

Signature: Date:

Name:

Acknowledgments

Thanks go to my supervisor, Linda Sing for the help with countless revisions and providing guidance on the completion of this research. A special word of appreciation goes to Craig Whitfield who provided assistance in the data collection, John Awbrey for editing and Tarryn Pheiffer for providing moral support throughout the research.

Table of Contents

1. Introduction

Is it possible that markets can foretell the future? Can markets rapidly incorporate new information to provide accurate predictions for future events? There have been several anecdotal events to support this theory. Firstly, the Challenger space shuttle disaster on 28th of January 1986. It took six months for the presidential commission to identify the faulty O-ring seals found to have caused the disaster. In comparison, within minutes of the announcement of the disaster the share prices of four major suppliers to NASA saw a drop in their share price. By the end of the day three, the suppliers' share prices started to creep up to within 3% of their original levels with the exception of one, Thiokol, who manufactured the seals saw their share price fall a massive 12%. Within hours of the disaster the markets had reacted to the information and identified the culprit (Surowiecki, 2004).

How were the markets able to rapidly gather all this information and discover the true cause of the disaster in such a short period of time? The ability of markets to rapidly gather information has been seen in many different settings. It forms the basis of prediction markets that have become very popular in organisations. A prediction market, (also known as prediction markets, decision markets, virtual markets and ideas markets) are speculative markets that have been created for the purpose of making predictions about future events. Equities are created in the market for the range of possible outcomes. Traders are then able to buy and sell the different equities. The market prices are then seen as the aggregated belief of the market on the probability of each particular outcome.

What is a key assumption for a prediction market is that they track definitive events or future events. That is there will be a correct answer that is currently unknown, but will be known within a given time frame (Kerr & Tindale, 2011). If this is not true for the market, then it cannot be a true prediction market. The outcome has to be definitive.

Prediction markets are distinctly different from traditional financial markets. In financial markets there is a concern with the diversification or pooling of risk. Traders are also attempting to make predictions about the values of commodities which will be linked to possible events that may occur over a limitless time frame. Prediction markets, by contrast have securities that are concerned with a limited number of uncertain events. These events are also definitive, in that they will occur, but the outcome is unknown (Snowburg, Wolfers, & Zitzewitz, 2009). In the same regard prediction markets are different from sports betting markets. While both have a similar structure it is the use of the markets that differentiates them. Prediction markets provide information on which policy or decision makers can act, which is not the case in betting markets (Berg, Forsythe, Nelson, & Rietz, 2001).

It is a commonly held belief that decisions are no better than the information that is behind them. We have come to accept that the majority of decision making in organisations is flawed due to poor information. Information is either not readily available or difficult to gather. The people who possess the information that would improve decision making are spread across the company or do not know that they have the information. Leaders very often find themselves in the situation of "ivory tower" decision making as they fail to consult sufficient sources when making decisions. In 2008, McKinsey Consulting hosted a roundtable discussion with some of the leading thinkers and implementers of prediction markets in the corporate space. The round table included Bo Cowgill, who manages Google's prediction markets; Todd Henderson, assistant professor from the University of Chicago; Jeff Severts, general manager of Geek Squad, the services arm of Best Buy; and Jeff Surowiecki, author of "The Wisdom of the Crowds" (Dye, 2008).

The round table discussions looked at the implementation of prediction markets at Google and Best Buy that have seen excellent results. Research conducted by Cowgill (2009) at Google is discussed in more detail later in this report. The round table also highlighted the need in business for a more efficient process of information aggregation. Through their discussions they highlighted how prediction markets are able to take large amounts of information from different people and provide an aggregated result that is on par with, if not better than, the top forecast analysts' predictions. Severts, from Best Buy, had seen some excellent results through his small scale markets, but has not published any of the findings. Severts' small scale markets looked at predicting greetings card sales in the United States. He invited everyone on the company's intranet to participate in forecasting what the greetings card sales would be in a given month. A relatively simple spread sheet was developed which allowed everyone from the shop floor up, to estimate what they thought the sales would be for the given month across Best Buy. Severts then aggregated all the estimates in the spread sheet to arrive at a forecast of greetings card sales for Best Buy.

The process was easier to operate than the centrally planned forecasting model which Best Buy had developed for greetings cards and proved to be 98% accurate, while the central forecasting model could only achieve 95%. While this study was never published in any academic sources it has been referenced in popular media and liturature such as “Think Twice: Harnessing the power of counter intuition” (Mauboussin, 2009) and “The Wisdom of the Crowds” (Surowiecki, 2004).

Prediction markets are not new to the business world, yet with the rise of technology they have seen a resurgence in popularity. An article in BusinessWeek written by King (2006) showed that several large corporates were experimenting with prediction markets. This included the likes of Google, Yahoo! Microsoft and Hewlett-Packard (King, 2006). Prediction markets work on the basis of aggregation of collective information through market forces. This allows the market to incorporate an infinite amount of information in the aggregation process. The final trading price that is shown in the market is a result of the accumulation of all the information present in the traders based on their imperfect information. The trading price of the share reflects the group’s belief in the likelihood of that event occurring (Borison & Hamm, 2010). Through the trading of these shares groups aggregate their specific knowledge to produce probabilities for the given outcomes.

Hewlett-Packard (HP) looked to internal prediction markets to determine the future price of DRAM chips. An article in Business 2.0 by Schonfeld (2006) commented on the success that the DRAM purchasing team had seen. The prediction market had outperformed the official HP forecast six out of eight times and drawn with it on the other two occasions. Bernardo Huberman, a senior fellow at HP Labs, was responsible for the creation of their prediction market and felt that this was a smarter way to make predictions. He also believed that the new tool worked better than the best person (Schonfeld, 2006). The best person in this case was not necessarily an expert in a particular field, but rather the person who performed the best in a given market. It was also expected and has been seen in the prediction markets that the best person in the particular market is always changing. This means that the best person is not the expert but rather who ever the top performer is at that point in time.

Collective judgment has been used in many different situations. An example of this would be committees, boards of directors and even, juries (Kerr & Tindale, 2011). We are relying more on groups to facilitate knowledge development in organisations, the rise in the number of meetings that are held and the amount of time that is consumed through meetings proves to be a common gripe among office workers who supports this contention. Yet even though meetings are inefficient and costly, there has not been an alternative solution suggested as to how information can be more effectively generated and dispersed through an organisation.

Prediction markets have also been used effectively in forecasting tasks (Van Bruggen, Spann, Lilien, & Skiera, 2010) and to predict rare events in the future (Goodwin & Wright, 2010). Recently prediction markets received a lot of media attention in the US after the Government launched Policy Analysis Market (PAM) in 2006. The intention of PAM was to forecast aggregated measures of political stability in various regions, particularly the Middle East. Traders would buy and sell futures in the outcome of U.S. policy choices for given regions eight times a year (Hanson, 2006). The failure of PAM was not due to it being ineffective at predicting future events, but rather due to the misunderstanding that it would be used to predict terrorist events, and in so doing allow people to profit from terrorism. Due to the political climate in America and the resulting outrage, the PAM project was cancelled. Hanson (2006) evaluated the social outcry about PAM and critically analysed the concerns that were held against prediction markets and predicting terrorist futures. His findings were that many of the concerns of the public and politicians could be overcome, and that prediction markets would be a very useful tool in determining possible effects of US policy decisions in the Middle East.

Prediction markets are a growing field of interest in the academic and business communities. Their ability to gather information and aggregate outcomes has shown to be remarkably efficient and accurate. The longest running prediction market has also been one of the most accurate. The Iowa Electronics Market was started in 1988 to teach finance students how markets work. Students would buy and sell wagers on the outcomes of presidential elections. Since the incorporation of the market, it has outperformed various exit polls for elections held not just in the U.S. but internationally. The Hollywood Stock Exchange has also been successful at picking the winners among new movie releases (Foutz & Jank, 2007)

There is a growing body of knowledge around the role of prediction markets and how they function. Yet prediction markets are still susceptible to many of the vices that befall traditional markets such as the formation of bubbles. While bubbles in real markets are seen relatively regularly this is due to the method of compensation in a traditional market, where traders can make money through buying low and selling high. Due to the compensation model used in many prediction markets the effects of bubbles are hardly seen. Prediction markets tend to operate on a winner takes all approach, as seen in the Iowa Electronics Market. As the markets are based on decisive events, the outcome of the event will be known and a fixed reward is paid given the outcome (Abramawicz, 2006).

Prediction markets represent another form of decision support, along with the Delphi method and to some extent scenario planning. Yet the efficiency of prediction markets and the improved accuracy over other prediction tools has created a high level of interest in them and the capabilities they may have. With this in mind, this research looks to build on the work of Van Bruggen, Spann, Lilien & Skiera (2010) who highlighted the need for integrating deliberative approaches into the prediction market framework looking to compare it with a repeated-Delphi type methodology.

Graefe, Luckner and Weinhardt (2010) showed similar findings from their study. This showed that while patterns of trading may suddenly increase, showing that the market is responding to new information, it is not always possible to know what that information is, as participants do not reveal their reasons or arguments for buy and/or sell decisions. Their findings highlighted the need for research to ascertain whether improving communication between market participants would improve the prediction accuracy of the market.

Prediction markets are being used in more business and organisational settings. They have found favour particularly in the forecasting divisions of many organisations. Forecasting has traditionally been an area of business that is more art than science. It has become characterised as being very cumbersome and slow when making forecasts and a very expensive process. The aim of this research is to build on the work on prediction markets, incorporating a method of in-market communication and comparing it with a traditional market. The role of prediction markets in demand forecasting needs further investigation as this is an area of business that would be interested in the findings of this and other work into the functioning of prediction markets.

One area in which prediction markets are finding significant application in many businesses is demand forecasting. Demand forecasting is a critical activity in many organisations. It is a lengthy, time consuming process that involves a small group of experts in the company. The process they follow is usually that of multiple face-to-face meetings, where information is presented and a group decision is reached as to what the demand for the products will be in the coming quarters or years. This output will then directly affect the budgeting of the organisation and production for the coming months. This one key decision would have a dramatic effect for the future of the business. Difficulty in reaching consensus is common in all forms of behavioural driven aggregation, even when a consensus is reached it is most likely based on individuals' personalities rather than through knowledge (Van Bruggen *et. al,* 2010). The consensus is very often driven by power relations, censorship and groupthink creating a difficult environment in which to accurately forecast future demand.

There are many key limitations to the process as outlined above, namely that it is very expensive to gather a group of experts in a room, as they usually charge by the hour, if consulting to the company, and very difficult to find a time in which all experts are available to meet. There is also a limitation on the amount of information that can be adequately discussed in a given period of time so that all members of the meeting have a similar understanding of the content. This leads to another problem, as there is a limit to the amount of information that can be included. There is therefore a large amount of information that cannot be included. There are also a large number of people who are not included in the process due to the nature of face-to-face meetings. It is an imperfect process.

Prediction markets provide an alternative to the traditional face-to-face meeting of a small group of experts. The process allows for much greater inclusivity as the market will allow a larger number of traders to incorporate a wider variety of information into the final outcome of the market. This can also be done in real-time, creating an environment that can adapt to real-time changes in market conditions thereby providing a more accurate prediction to that of a smaller group of experts.

Prediction markets have shown to be successful at predicting the outcome of future events and have been implemented at several companies for a wide variety of tasks, as discussed earlier. This research looks to determine how to make the markets more effective and efficient. The outcome can then be used as an input to the demand forecasting meetings.

This research is looking to expand on the current knowledge of prediction markets. The aim of this paper is to evaluate the contribution of deliberative methods in prediction markets. Deliberative methods are defined as any process that allows information to be shared among traders. In this paper, two market types are compared to determine if there is a significant difference between markets that contain a discussion / comment board, and markets that do not.

In this paper prediction markets will be discussed in more detail, looking at key requirements for prediction markets to be successful. Biases in individuals and groups will be explored as this has a fundamental impact on how we make decisions ad evaluate risk and reward. The history of the two markets will be looked at in more detail, with the key differences being highlighted. Finally the outcome of the markets is analysed with the results being presented and discussed further.

2. Literature Review

Prediction markets cover a variety of categories. In one form or another we have all at some point in time made use of a prediction market. We may have not been aware of the fact that it was a market as they often do not look like markets, but the end result is always the same. This comprises large amounts of information from independent individuals being aggregated into one outcome. The oldest example of this is a common fair ground game. "Guess the number of Smarties in a jar, balloons in a car, coins in a bottle and win a prize" in this game we are making a prediction about how many items are in a given object, to do this we make use of a variety of tools and previous experiences. We then write our estimate on a card and await the final outcome.

We have only ever thought of this game from one perspective, namely ours, but if we look at this game from the perspective of a market, each person who is taking part is essentially a trader. They buy a ticket with a specific outcome. As this is a basic market all outcomes cost the same, in more sophisticated markets the price of the popular outcomes will increase while the less popular outcomes will decrease. At the end of the day we have either correctly predicted the number of items in the object or not, that is we have either made a profit or a loss. This would be a single trader's view of this simple market, but if we look at the market as a whole we would see that, as traders buy outcomes, all the information they provide can be aggregated to determine what the actual outcome will be.

Financial markets have taken this principle and made it far more complex. Traders can buy or sell a variety of shares in many different companies. The idea though is that the money they are prepared to pay for the share is related to what they believe the share is worth. The worth of the share comes from the cash flows they anticipate the share will generate, so essentially they are predicting the future cash flows of the company to determine what the share should be worth today. In our simple fairground market there is a definitive outcome. That is they close the market and count how many beans were in the jar, unfortunately this never happens with a company. At the end of the day you will never see if you were correct with you valuation of its performance. This adds to the complexity of the financial market.

The basis of both of these markets is still the same. Through the aggregation of a group of people's individual opinions we are able to estimate with a reasonable amount of accuracy, the outcome for a specific event. Be that beans in a jar, or the winners of a sporting contest.

The fundamental principle behind prediction markets (in fact all markets) is to create a platform where individuals can buy and sell shares based on specific outcomes (Borison & Hamm, 2010). The activity of trading the shares sets a price for the given share that indicates the market's aggregated view of the likelihood of a given event. Prediction markets can take place in both the public and private domains, and as discussed, can range in complexity from simple spread sheet aggregations to the complexity seen in our financial markets.

In this section it is important to look at the various factors that make prediction markets work. As the mechanism is purely that of aggregation of individuals' information we need to look at a few of the factors that are unique to individuals that allow us to obtain the variety of information. To this end diversity, heuristics and biases will be discussed. We will then look at the role group's play and finally, some uses of prediction markets. As the focus of this research is on prediction markets and demand forecasting we will also look at other forecasting methods.

2.1.Examples of Prediction Markets

2.1.1. The Iowa Electronics Market (IEM)

The Iowa Electronics Market (IEM) has been predicting the results of presidential elections in the U.S. since 1988. This makes the IEM one of the longest running computer based real time prediction market. The IEM allows traders to invest between $5 and $500 of real money, creating the element of risk among traders, as there is real cash at stake. The securities that are used in the IEM have set payoffs. Any given contract pays out $1 if the contract occurs and $0 if the contract does not occur. This allows the current traded price to indicate the markets perceived belief in the given outcome.

The IEM began its life as an internet based teaching and research tool. Since its inception, over 100 universities, throughout the world, have enrolled in the IEM, including institutions like Harvard, MIT, Michigan and North Western. The IEM is used to teach students about business, economic and technological concepts through practical hands on experience. While the political markets are the most well known, other markets are run for different classes, such as box office markets. Each designed to allow students to learn about the complex relations in the business or political world.

While the IEM was developed for learning and research purposes its ability to forecast has not gone unnoticed. It too has been the basis of many published articles for its ability to outperform many political polls and expert panels.

The IEM is fundamentally different to election polls or expert panels. The traders in the IEM are self selected, not randomly selected as with election polls or specifically chosen as in an expert panel. This self selection ensures that only people who are interested in the markets and events will continue to take part. It also means that for specific topics, only traders that have an interest or understanding in that topic will be willing to buy securities.

The market does not weight traders' views. Every trade carries the same weight. This means that topic experts can not present any more influence than a regular trader. The market price is a metric, which is affected through market behaviours and the behaviours of traders. This builds in many hidden factors of traders behaviours, such as risk aversion, over confidence or timing.

The IEM allows the views of traders to be aggregated through the market price. It does not show the individual views of any particular trader, but rather the aggregated view of all traders. This is fundamentally different to expert panels, in which experts provide their individual views on any given subject.

2.1.2. Inkling Public Markets

While the IEM has the academic prestige, Inkling has capitalised on the business interest in prediction markets and set up a public market, along with custom markets for corporate clients. Inkling was founded in 2006 by Adam Siegel and Nathan Kontny to help organisations decrease operational and strategic risk. Inkling has also been named as one of America's most promising start ups by Business week (Tozzi, 2010). Inkling allows corporations access to prediction markets for staff for as little as \$10 - \$15 per employee per month.

Along with the corporate business Inkling also has public markets that anyone can access. Through creating a user account traders can set up their own questions and share them with the Inkling community. In this market traders are allocated a set amount of virtual currency with which to make their trades. As of October 13, 2011 Inkling markets had over 88% of registered users still active in the different markets (Inkling Public Markets, 2011).

There is a key difference between the Inkling markets and that of the IEM. That is, Inkling has allowed traders to communicate through comment and discussion boards available in all the different markets. This is a key difference and may impact how the markets function. No research has as yet looked at the inclusion of discussion and comment boards in prediction markets as to how they impact on the ability of the market. Currently no research has looked at the difference in performance of these two distinct markets, this research will compare the results of the two markets to determine if including a discussion or comment board has any impact on the performance of the markets.

As with the IEM, Inkling also has some political markets, but as users are able to create their own markets, there is a huge variety of different markets. These markets look at a wide range of topics, from politics and economics through to weather forecasting of tropical storms.

As with the IEM, traders self select into the different markets based on their interests. The markets aggregate the view of the traders into a price for each tradable option. This is done through a proprietary algorithm that has been developed by the Inkling founders. This algorithm allows the traders to influence the price of the different securities through buying and selling their options.

2.2. The Role of Diversity in Decision Making

Prediction markets require individuals to make them function. It is not possible for the markets to aggregate information if there is no input from individual traders. At the individual level, diversity in thinking is key to the success of prediction markets. We all differ in how we interpret the world. This is due to the different perspectives and mental models that we all possess. Our perspectives provide us with an internal problem solving language, as we code what we see in the world into words to explain it. Our mental models help us to form perspectives on the world based on experiences and assumption we make about the world. Mental models are images or assumptions of the external world that we carry with us. They influence how we see the world, acting like lenses or filters on the world (Senge, 1994). Mental models as with perspectives are tacit, existing in our subconscious. Our mental models will impact how we learn and how we approach different scenarios. An example of a mental model would be the assumption among teachers, that parents don't know what is best for their children. This assumption plays out as well intentioned reforms in schools that end up alienating parent groups. (Senge, 1994). Mental models help to form our perspectives. The experience of each individual allows them to form their own collection of mental models and perspectives.

This difference in perspectives allows us to see problems from a different angle, and thus find new and novel solutions. Examples of how these different perspectives play out in the real world have been extensively studied (Page, 2007). A common example is the game, Sum to Fifteen, which is discussed in Page (2007).

In this game the objective is to create three card combinations that sum to 15, while at the same time preventing other players from doing the same. Figure 1 below shows what cards would be required by a player. The table shows all the pairs of cards that each player holds and the cards needed to win the game.

From the given board it shows that the first player will win, as if player two takes the nine, player one can take the six and vice versa.

While this game looks nothing like Tic Tac Toe, simply re-arranging the cards into what Page (2007) refers to as the magic square shown in Figure 2, the game suddenly looks exactly like Tic Tac Toe. Perspectives are all about how we view the world. Initially this example started with what looked like two very different games, yet by changing our perspective on how the games are set up we have now been shown that actually they are the same game. The magic square has given us a new perspective on the Sum to Fifteen game and on the Tic Tac Toe board. This shows the power that perspectives have on how we view the world around us and solve problems that we face on a daily basis. Once we are aware that the problem is like Tic Tac Toe, we can apply a heuristic to the problem to solve it. The challenge is in how our perspectives interact with our heuristics to determine if this is a problem we have had to solve before, or if it is a completely new problem.

Cards	Sum	Needed Card
2,4	6	**9**
2,5	7	**8**
4.5	9	**6**

Figure 1 - First players cards and needs

8	3	4
1	5	9
6	7	2

Figure 2 - The Magic Square of Sum to 15

2.2.1. Heuristics

As individuals, we think in different ways, due to our unique perspectives. These perspectives develop through heuristics. Heuristics are rules that have been applied to existing solutions represented in the form of perspectives. The application of existing heuristics to new problems allows the generation of new solutions. We all possess a variety of different heuristics of which we are constantly making use. Common heuristics that are taught to small children is for example to "mind the pennies, and the pounds will take care of themselves." This heuristic indicates that by looking after the details, the bigger picture will look after itself. A conflicting heuristic would then be "don't sweat the little stuff."

The diversity of our individual collection of heuristics allows us to apply multiple solutions to a given problem, and to even view a problem from conflicting perspectives. The diverse combination of perspectives and heuristics in individuals allows for groups to be collectively intelligent. The odds of every individual in the group possessing all the same perspectives and heuristics is infinitesimally small. This means that a group of people collected together will inevitably allow for a problem to be viewed from multiple different angles and many more solutions offered, than if just one individual had to tackle the problem (Hong & Page, 2004).

Diversity comes in many forms. Above, we have discussed functional diversity, or how individuals represent problems and look for solutions. The diversity perspective that most of us share refers to identity diversity, referring to common demographic characteristics, cultures and identities (Robbins & Judge, 2007).

2.3.Biases in decision making

From the discussion above it can be seen that individuals take short cuts when making decisions. We rely on previous experiences to make decision making quicker. These short cuts create biases in our decision making process. There has been a lot of research into the problems of working in groups. Most of it looks at the biases that develop and the social structures which exist that shape the views of the group. The wisdom in crowds has seen extensive publication through scientific and general publications (Surowiecki, 2004)

In the Iowa Electronics Market, research has been conducted into the effect of overconfidence bias and the “long shot” bias. These biases are commonly seen in traders and punters in betting circles (Berg & Rietz, 2010). These biases look at individual’s beliefs in their own abilities and how they believe the market will play out over long and short periods of time. With the over confidence bias we see a majority of people rating their abilities above that of the average person. This indicates that they all believe they will do better than the average, which is not possible. The long and short bias shows how people will behave when they believe the outcome is in the short term versus the long term. This also shows that they are applying different heuristics and mental models to their decision making process based on the time frame they have been given.

Surowiecki (2004) found many examples of how prediction markets have been successful in the real world. His investigations were then taken a step further by Page (2007) in his book, "The Difference". Page looked at some of the conditions that need to be in place for a group to make a smart decision or to predict correctly. Much of his work was based on mathematical modelling. The role of diversity was key to the success of groups of individuals outperforming higher-performing individuals or experts (Hong & Page, 2004).

A further bias that may impact how prediction markets is known as the "Pollyanna Principal", this principal has been used to describe how people are inherently biased to a positive outcome (Matlin & Stang, 1987). A prerequisite to this principal was the optimism bias. If we do not have an optimistic outlook on the future we are unlikely to act (Tyebjee, 1987).

2.3.1. Groups

While we may develop as individuals, humans essentially tend towards forming groups, not only in or social lives, but in business. The formation of groups creates a complex system of behavioural interactions among individuals within the group, which impacts not only on our individual identities but how we form perspectives and heuristics. This will also impact on the diversity of individuals and perspectives with in a group. The importance of diversity (identity and functional) has been shown in many different situations (Robbins & Judge, 2007). Yet the ability of diverse groups to outperform groups of high-ability has been shown through mathematical modelling by Hong and Page (2004). In their research they create problem-solving agents of limited ability who differed in their heuristic and perspective pairing. They found that a random collection of agents would outperform a group of the best agents selected from the same group. Their experiments have interesting implications for how groups of people should work, yet never do. The agents in the experiments do not have to deal with intergroup dynamics that exist in the real world situation.

Groups are defined as two or more individuals that interact and are interdependent. They have come together to achieve a particular objective. The grouping of individuals brings with it perceptions and expectations of the individuals in the group. These revolve around the roles that each individual will have to fulfil, status of individuals within the group, group size, cohesiveness and established norms (Robbins & Judge, 2007). The interaction of these group dynamics can lead to problems at the group level, such as “group-think” and “group-shift”. Group-think shuts out idea generation and development of mutational ideas, thus disabling evolutionary learning (van der Heijen, 2004).

This process has been illustrated in a story, the trip to Abilene (Harvey, 1974). A group of travellers endured a difficult trip, wasting an entire afternoon, only to discover on their return than no one in the group actually wanted to make the trip. The group all believed that the other members of the group wanted to make the trip, thus creating an image of what the group wanted which did not reflect reality. Group-shift was first introduced to the academic world by Stoner (1961) he found that groups are more likely to make riskier decisions than any individual in the group would make on their own (Stoner, 1961). Many studies have looked at the causes and effects of group-shift, Rutledge (1993) has summarised many of these studies in his work into the effects of anchoring on group-shift (Rutledge, 1993)

A group phenomenon that was described before group shift was “Group Think” this was first described in the early 70’s and is the tendency of a group to strive for unanimity over the actual appraisal of reality. That is the group would rather reach a consensus about the outcome, than if the consensus makes sense (Janis, 1972). Janis (1982) indicated that for group think to be present there are pre-requisites that need to be in place. Firstly, the group needs to be cohesive, with elements of strong or respected leadership, time pressures and complex decisions that need to be made. While this phenomenon has not yet been described in prediction markets, it would be possible for the elements to come together under certain circumstances.

From the research into group think, it has also been observed that group polarisation can take place. Group polarisation occurs when the majority of a group holds a position and this position gets further intensified through in-group discussions (Jones & Roelofsma, 2000). Group polarisation differs from group think is the actual outcome that the group reaches. In a group think scenario, the outcome would be a unrealistic assessment, but one that everyone has now reached consensus on. In group polarisation, the group had already reach a consensus that was common to the in-group in terms of their beliefs, this consensus is then intensified through the group discussion, an example that has been used in the social behaviour research looks at the views of pro-choice groups (Yadi & Boyd, 2010).

The dynamics of working in groups, while fascinating, falls beyond the scope of this report, which looks to the outcomes of the process which allows for more accurate forecasting through aggregation of individual abilities.

In a traditional prediction market, such as the IEM, traders are unable to form groups, as the market keeps the independent of each other. They have no means of communicating or seeing who has done what. Yet in the Inkling markets this has been overcome. Through the creation of discussion boards, individual traders can communicate with one another and see reasons for specific trades.

When Google started experimenting with prediction markets they looked at how the physical proximity of different traders affected their performance in the market. Through the analysis it was seen that the closer the physical proximity of two traders the similar their performance and trading patterns (Cowgil, Wolfers, & Zitzewitz, 2009). This indicated that even though the market that

Google was using did not allow traders to communicate, the mere fact that they worked together enabled them to collaborate in the markets.

In a real world environment such as at Goole, this indicates that traders form groups and discuss the trades and markets that they are trading in. This creates a possibility of group based biases to slip into the markets and affect the outcome of the market. No research has as yet looked to see how the inclusion of discussion and comment boards in a market will affect the performance of the market in a virtual world, where traders have no physical proximity.

This highlights a potential problem for the Inkling market, how to ensure that the traders do not start to behave in similar ways, thus falling prey to group biases, such as group think. It has also been theorised that for prediction markets to function optimally traders need to maintain their independence. That is, it needs to be ensured that traders are not able to share insights or thoughts with one another as this will change the way in which they behave in the market and thus affect the overall ability of the market (Surowiecki, 2004). It has not been show empirically if groups form among traders on virtual markets, and if so how this affects the market's performance.

A further bias that is worth exploring is that of group polarisation. That is, when the views of a group become increasingly more polarised through a series of positive reinforcement signals from interactions with in the group (Jones & Roelofsma, 2000). This phenomenon has been described in economic and social settings. The most recent of which looked at how social media affects groups, and if groups with opposing view on a topic experience polarization even though they are forced to interact.

Yadi & Boyd (2010) looked at how conversations on Twitter over the shooting of an abortion doctor in the United States of America lead to heterogeneity and homophily among the pro choice and pro life groups, with each group more likely to interact with people who had shared views. The internet has had a huge impact not only on how we communicate with one another but how we source information on which we build our mental models of the world. While this is important when considering the group dynamics in social settings, it is also going to be an important consideration into the functioning of prediction markets. Traders will have access to more information than ever before, and can quickly learn about developments in the market, which are all reason why prediction markets are successful. But they may also look to gather views of people who have similar beliefs, thus reinforcing what they already know. This combined with the ability to rapidly share your point of view through a social media channel or even a comment board in the market may impact on the performance of the market. the direction of this impact is currently unknown. The increased speed and availability of information may improve the speed at which a market converges onto the correct outcome.

Convergence is a newly describes phenomenon in prediction markets. It has been explored in research on the performance of prediction markets that were looking at specific outcome for defence contracts (Davis, 2011). Davis (2011) describes convergence in a prediction market as the time, measures in days, for a market to pick the correct outcome and remain in that position till the market close. Convergence can only be measures at the close of a market, as the correct outcome is unknown prior to the markets close. Figure 3 - Convergence explainedshows a representation of when convergence is said to have taken place

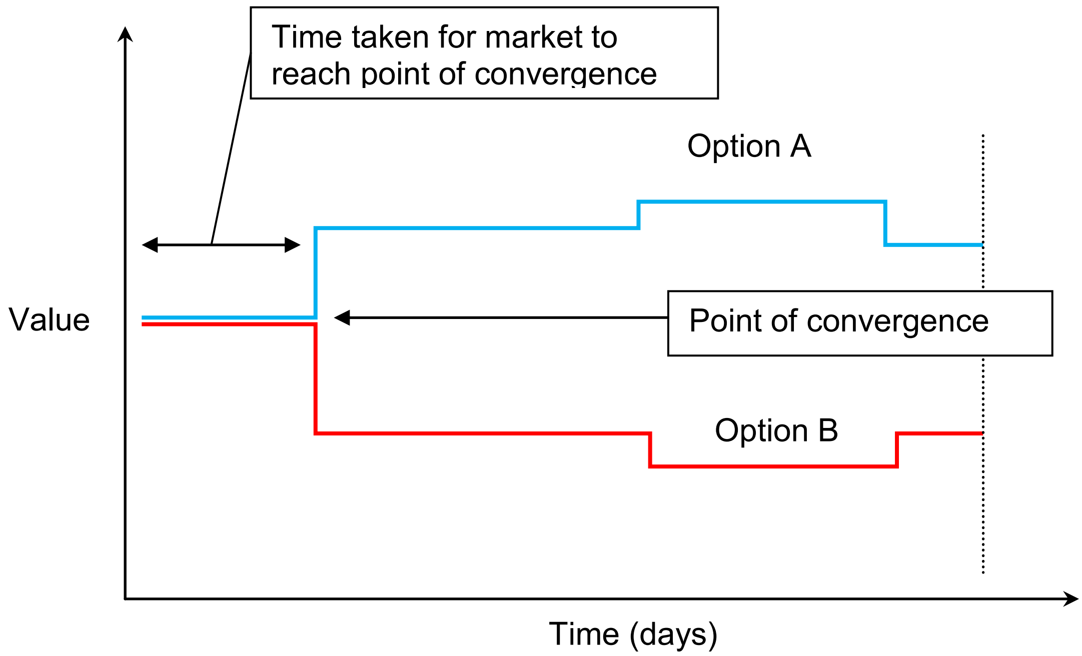


Figure 3 - Convergence explained

For convergence to have taken place the value of the correct outcome has to exceed that of the incorrect value and remain that way till the market closes. Convergence is still a new topic in the prediction market literature. While this term is new research has previously looked at the long run success of markets.

This has been measured previously using the IEM data for political election results. Berg *et. al* (2008) looked at the performance of the IEM versus traditional election polls over a time span of more than 100 days prior to the election. They examined the closing prices of election markets with more than 100 days to the election, from 66 - 100 days, 32 - 65 days, 6 - 31 days and the last 5 days prior to the close of the market on election eve. The findings supported their hypothesis that the prediction markets were superior to traditional election poll even more than 100 days prior to the event. While they did not measure how quickly the election markets converged on the correct answer, the long run accuracy of the markets is demonstrated.

2.4. Uses of Prediction Markets

Prediction markets are mainly seen as supplementary to other forecasting techniques, and can be applied in many different areas to improve the accuracy of forecasting (Graefe, Luckner, & Weinhardt, 2010). Such applications would be in environments that are constantly changing as the prediction market is able to rapidly incorporate the new information. The market can also be used for idea generation and to motivate creative thinking. The individual performance of members in the markets will also allow the organisation the ability to identify people who may possess unique knowledge in given fields to be used in other forecasting techniques (Graefe, *et. al,* 2010).

Newly launched products in the consumer goods space tend to have high failure rates. Companies can decrease the risks associated with product failure through the use of a lead user (Spann, Ernst, Skiera, & Soll, 2009). The lead user is a person who usually possesses expertise in the product field or realises the needs of customers earlier than the rest of the market. With the use of prediction markets lead users can be identified through the rankings of their performance in the given market (Spann, Ernst, Skiera, & Soll, 2009).

2.5.Forecasting Techniques

Prediction markets have seen a lot of research in the fields of forecasting and foresight. Studies have compared the performance of prediction markets to that of other forecasting approaches, such as Delphi, face-to-face meetings and statistical aggregation techniques (Graefe, Luckner, & Weinhardt, 2010; Van Bruggen, Spann, Lilien, & Skiera, 2010; Kerr & Tindale, 2011; Burt & Wright, 2006). This report looks to build on their work through combining the benefits of prediction markets with the sharing of information implicit in Delphi designs.

The Delphi forecasting technique is also a popular method for aggregating information from a small group to determine future events and outcomes. It is a relatively simple and efficient process that has diverse application in many business settings (Green, Armstrong, & Graefe, 2007). The Delphi process is designed to reveal the knowledge of panellists and allows for testing as they have to disclose the reasoning behind their decision making, yet this process can also lead to conformity of opinion due to group pressure (Green*, et. al,* 2007). The Delphi technique can be useful in many problems that involve forecasting or estimation, provided that complexity and ignorance do not preclude the use of expert judgement. This is an important consideration as if a group is misinformed about a topic or issue, the Delphi process will only add confidence to their ignorance unless disparity among opinions is uncovered, which may help to alert decision makers to the potential problem (Green, *et. al* 2007). One of the key aspects of the Delphi technique is the reasoning that is provided for decision making. This is only useful so long as the reasoning provided on each round provides new information, thus the experts can revise their opinion based on the new information.

The Delphi process is very time consuming, requiring panellists to dedicate time to producing their opinions and reviewing subsequent rounds of estimation, predictions and reasoning. This use of time can make the process very costly, as with the traditional face to face meetings (Rowe, 2007).

From the literature it can be seen that there are many different techniques to make forecast demands, each with its own advantages and disadvantages. As business is progressing, it is requiring more information and faster dissemination of information. Decision makers need information that is more accurate and they need it quickly. These requirements have made prediction markets look very appealing, as they are able to incorporate new information very rapidly and present possible outcomes very quickly.

A new area of research has been looking at convergence in the prediction markets. This is the speed at which a market converges on the correct outcome.

Research has found that in a series of markets looking at the awarding of defence contracts and forecasting the performance and demand for the different contracts showed that on average the markets had converged on the correct outcome in three days or so (Davis, 2011).

Once the outcome is known, the convergence of the market can be calculated. To calculate the convergence on a particular market, the first step is to look at the random probability of the possible outcome. So for a market with two possible outcomes, that would be equal to 50% for each outcome. Once the correct outcome probability goes above 50% and stays there till the close of the market it is said it has appropriately converged on the correct outcome (Davis, 2011).

Adding this new dimension to the prediction market body of knowledge improves the understanding of the markets, but also highlights why prediction markets are becoming so popular in the business world. Adding to this the possibility of improving the accuracy of prediction markets through incorporating a deliberative feedback loop through the use of discussion boards and comments may increase the popularity of the prediction markets in the business environment.

3. Research Question

Forecasting is a costly and time consuming process in many organisations. The process involves many people, but all too often never enough information is incorporated into the process due to constraints on time for those involved. Coordinating all the parties concerned becomes more difficult as the number of people involved increases, especially if the chosen method involves face-to-face meeting or interviews. Due to these constraints the process becomes very lengthy and seeks to forecast for a much longer time frame, so that new information can be introduced closer to the time. But with this said, many organisations will have budget implications linked to the outcomes of the forecasting process that are fixed and cannot change.

This increases the importance of having accurate processes that incorporate as much information as possible to provide accurate forecasts of demand in the forecast period. Prediction markets have shown their uses in many different demand forecasting tasks, such as user volumes at Google (Cowgil, *et. al,* 2009). This research seeks to improve on the accuracy of a prediction market by incorporating deliberative elements to the process. This would include discussion or comment boards in the actual market.

In a quest to improve the speed of forecast, prediction markets have been used and shown to be successful. Not only for their accuracy, but also for the speed in which large amounts of information can be aggregated and an outcome produced. With the growing use of prediction markets, interest has arisen around the use of more deliberative approaches in the prediction markets. Van Bruggen *et. al*. (2010) found that prediction markets were as effective, if not more effective, than many other forecasting techniques. The Delphi technique was also investigated in their research and seen to be an accurate, but more time consuming technique. With that in mind it was suggested that through including a deliberative approach into the prediction market format would improve the accuracy of the markets (Van Bruggen, Spann, Lilien, & Skiera, 2010)

This research looks to continue the work of Van Bruggen *et. al*. (2010) through investigating how including a deliberative feedback loop into the prediction market affects the performance of the prediction market

3.1.Hypothesis

The hypothesis that we are looking to test is that a prediction market that uses a more deliberative approach, that is a market that incorporates information sharing among participants, will produce more accurate forecasts than a prediction market that does not.

H_0 There is no difference between the traditional and non-traditional (markets that incorporate deliberative feedback) prediction markets at the close of the market.

This is further examined to look at work conducted on the IEM assessing the ability of prediction markets to arrive at the correct outcome long before the event. This phenomenon will also be considered in this research, looking at how the two market types perform at different time intervals prior to the close of the market.

H_0 There is no difference between the traditional and non-traditional (markets that incorporate deliberative feedback) prediction markets at measured time periods prior to the close of the market.

Sub hypotheses to be investigated:

H_0: The distribution of predicted winner (100 Days) is the same across both groups

H_0: The distribution of predicted winner (30 Days) is the same across both groups

H_0: The distribution of predicted winner (10 Days) is the same across both groups

H_0: The distribution of predicted winner (2 Days) is the same across both groups

The results from the test at the close will show how all the markets in the study performed overall. Looking at the different time periods has been done previously in studies on the Iowa Electronics Market (Berg & Rietz, 2010).

Looking for differences in the long run accuracy of the markets will allow for a better understanding of how comments and discussion held by traders on a particular market affects the outcome of the market.

4. Research Methodology

4.1.Research Design

As discussed in previous chapters the aim of this research is to improve the accuracy of prediction markets by making use of feedback loops and a Delphi-style process to provide participants with more information around decision making. Van Bruggen *et.al* (2010) suggested that markets would be better predictors if they made use of some form of deliberative information feedback process. The market structure also varies greatly from study to study with the number of participants varying from a small group of six as seen in Abramowicz (2006), to how participants interact with risk and reward in the market.

The Iowa Electronic Market allows traders to invest from $5 to $500 in the market. This is real money so the traders have to make decisions about risk and reward as they have something to gain. As the IEM is a teaching aid it has managed to receive exemption from many of the regulations that other financial and betting markets are subject to. The Inkling public markets on the other hand, do not use real money. Traders in this market are given a virtual balance with which to make trades. This virtual balance is not transferable. To help instil a sense that there is something to play for, Inkling uses scoreboards, ranking traders by their overall performance. This creates a sense of competition among the traders. With both market types the payouts are fixed. In the IEM traders can only receive a maximum of $1 for a correct outcome. In the Inkling markets this value is $100.

Yet the common thread through all these markets is that participants act like stock traders, buying and selling ideas, or predictions that they feel likely to happen in the scope of the market. Van Bruggen *et.al* (2010) provides the design that the market created for this research will follow.

The research will take the form of a quantitative design. The results of different formats of prediction markets will be compared, such as the IEM and the Inkling public markets. Previously, models have been created with agents programmed to behave in certain ways, (Hong & Page, 2004) yet more recently the development of actual markets with real players has provided a platform for real world testing (Van Bruggen, *et. al,* 2010). Mathematical models have been popular in many forecasting tasks but often fail to incorporate and accurately reflect the nature of complex systems (Orrell & McSharry, 2009). The growing popularity of prediction markets in organisations also provides evidence for the use of a real world market to test the hypothesis listed above.

The aim of this research is to look at how to make forecasting more efficient in an organisation with the use of a prediction market. To achieve this goal the research will compare results of different prediction market techniques. Software is currently available to companies that allow them to set up their own internal prediction markets. This software incorporates deliberative techniques such as comment boards and allows reasoning to be captured for given prediction tasks. This research seeks to determine whether this method is the most efficient method for forecasting tasks. Through comparing the accuracy of different market structures that are currently available, this report will ascertain which approach is more efficient.

1.1.Market Structure

The look and feel of the IEM and Inkling market are very similar to that of other online trading platforms such as traditional stock or futures exchange and functions in a similar manner. Previous studies (Abramawicz, 2006; Borison & Hamm, 2010; Cowgil, *et. al.*, 2009; Van Bruggen, *et.al,* 2010) have used this format and refined the design, allowing this research to confidently follow. The market format has also been successfully implemented at companies like Google and HP (Cowgil, *et. al* 2009; King, 2006) as discussed in previous chapters. The two market platforms have been widely recognised in both the business and academic circles and as such are appropriate to be used in this research. The IEM has been cited in over 15 published articles (The University of Iowa, Henry B. Tippie College of Business., 2010). While Inkling is seem as one of the U.S.'s most promising start up businesses and has been backed and supported by many published and respected authors in the prediction market field such as Bo Cowgill (Google Prediction Markets).

Both the IEM and Inkling allow users to bid on outcomes for definitive events. Possible outcomes are loaded into the market and given an initial trading value. The value of all possible outcomes must add up to 100. Shares are then traded and the prices are adjusted based on the individual's belief in what the outcome will be. At the close of the market, the correct prediction is paid out. In the case of the experimental public market, traders are able to discuss what is happening and post reasons for their prediction while the market is open. The control market does not allow for this, and only records the trades that have taken place.

The Iowa Electronic Market will act as the control market for this experiment. The IEM was created by the University of Iowa in 1988 to predict the outcome of presidential elections in the United States (Berg & Rietz, 2010). The IEM is a real time, real money market. The traders buy and sell assets for their own account. This means that they bear real money risk and reward for all tier trades (Berg & Rietz, 2010).

Figure 4 - The Market structure explained, shows how prediction markets function, and at what level this research will focus. While the interactions of the individual traders taking part in the market is an interesting subject and a topic of extensive research, the outcome of the market, namely the group’s beliefs expressed through the prices of the different shares they trade, is the focus of the study. Moreover, the study will seek to determine how providing feedback to traders improves their ability to predict outcomes for events.

As the hypothesis we are testing is, how providing traders in prediction markets with additional information, will affect the performance of the market. It is tempting to measure the performance of individuals in the market, but as has been shown in previous studies (Hong & Page, 2004) and commented on by

Page (2006) and Surowiecki (2004) that the performance of a given individual is not as consistent at prediction tasks as a group of diverse individuals.

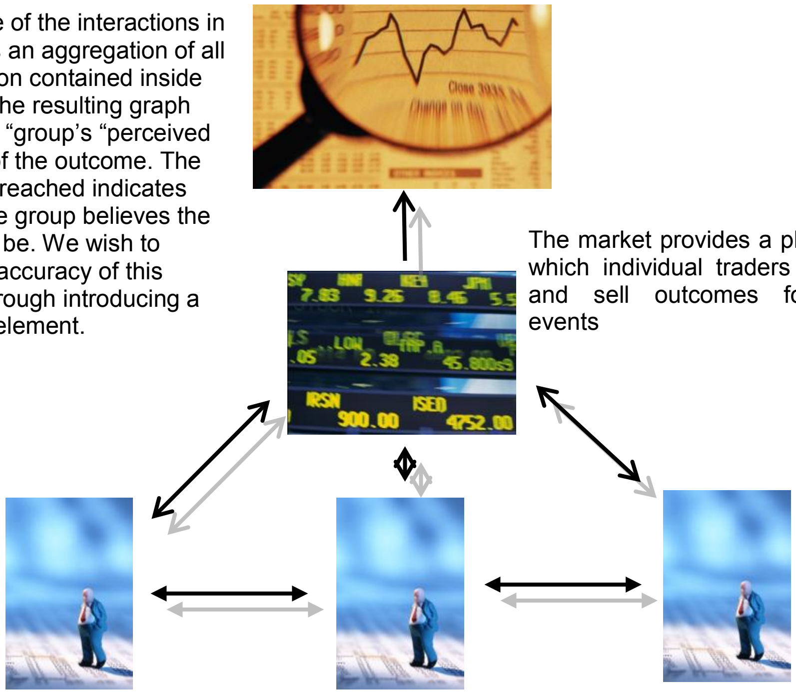


Figure 4 - The Market structure explained

To test the hypothesis that adding deliberative tools to the market increases its accuracy, the overall performance of the experimental market will be compared to that of the control market. To ensure that the additional deliberative techniques are the independent variable in the experimental design, the markets will be compared on similar topics, such as predicting election results, movie takings at box offices, closing prices of gold and stock exchanges.

The research is only interested in the outcome of the market and not the individual trades that take place or the traders who are participating in the markets. For this reason only trade prices are recorded and not the actions of individual traders. The price history graph for the different markets is based on the final trade price for each day for the given market and does not show the actions of individual traders.

To measure the performance of the markets in their actual ability to predict event outcomes, the predicted winner from the market will be correlated back to the actual outcome. This ensures that the markets are accurate in their prediction. Research has shown that prediction markets are an accurate indicator of future events. This is taken as a basis for this study. Research has looked at the ability of prediction markets to predict future outcomes in both the long and short term. The IEM has outperformed more traditional election eve polls and been more accurate than polls up to 100 days prior to the election result (Berg, *et. al*, 2008). The track record of the IEM allows for it to be a stable control market against which to test the effectiveness of the experimental market.

As the IEM has had long run studies looking at traditional election polls, we will also be looking at the long run performance of the IEM and Inkling markets. The analysis and results will compare the control markets and experimental markets prices at set intervals, from 100 days, 30 days, 10 days, 2 days and at close. In the analysis significant differences in performance will be looked for between the two samples.

Appendix 1 contains screenshots from the different markets.

4.2.Sampling/Population

A prediction market will track definitive events or future events. That is there will be a correct answer that is currently unknown, but will be known within a given time frame (Kerr & Tindale, 2011). The prediction market is a tool that has be widely used, it has been shown that individuals who have had previous trading experience performed better in markets, but that anyone who was interested could develop skills as a trader (Cowgil, *et. al*, 2009).

The sample will be randomly selected from recently closed (past five years) markets on the IEM and Inkling platforms. Data will be captured from these markets looking at the closing prices for the tradable securities at specific time intervals, namely 100 days, 30 days, 10 days, 2 days and at close of the markets. The markets have a database of traders who regularly participate in the markets. This research will not focus on the traders, but rather the performance of the market. The base of traders for each of the platforms is sufficiently larger as to not have an issue with liquidity in the market. There will not be an upper limit to market size yet from previous research the smallest markets have had only six members (Van Bruggen, *et. al*, 2010). For this reason we will look to only run markets that have more than six participants. The lower level for a prediction market to run on is of interest but falls beyond the scope of this study.

When registering for either the Inkling or IEM prediction market, traders are required to set up user accounts. Demographic and descriptive data is collected from participants and analysed for trends in terms of previous experience with markets and performance. Research conducted at Google found that participants who had previous experience in markets were more comfortable and initially performed better that those who had no experience in markets. This gap closed as their exposure to the markets increased. In their study they also initiated several "for fun" markets that helped participants build skill and comfort in the markets (Cowgil, *et. al*, 2009)

4.2.1. Acquiring the Data

Prediction market data and prices will be gathered from public sources. The trade price information is freely available to the public from both Inkling and the IEM. For a trader to participate in the prediction market they will have to agree to the terms of the market. Individuals participating in the prediction markets will not be identified in any of the research.

4.2.2. Analysis

The collected data will be analysed descriptive statistics to ascertain normality in the distribution. The Kolmogorov-Smirnov test for normality was selected to determine the distribution of the data. To ensure that the sample selected is also performing as expected, the individual markets for each group will be correlated against the actual outcome to ensure that there is a strong positive correlation. This would indicate that the higher the value of an item in a market the more likely it is that that event will occur. This will help to ensure that a representative sample has been selected.

After testing for normality the data will be analysed using the Mann-Whitney U test for independent samples. This is a non-parametric test designed to compare two independent samples. This test has been chosen in the case of a distribution that is not normally distributed. In the case of normality being proven the equivalent test would be the student's t-test for independent samples.

The hypothesis will be tested at a significance level of 0.05, with the null hypothesis, that there is no significant difference being either accepted or rejected.

4.3.Limitations

The markets that have been used for this study come from a variety of topic areas, thus allowing the results to be transferable to many different types of forecasts. The deliberative approach that was tested in this study was an unmonitored comment board made available to all traders in the Inkling prediction market. Traders also had the option of stating a reason for their selected trade which would be visible to other used on the market. This limits the findings of this study to other prediction markets that have an un-monitored comment or discussion board available to traders

The outcomes of this study will applicable to organisations that are considering using prediction markets for their business forecasting. The results will help identify which methods will provide the most accurate results. There is a selection bias in the participants of the market as they are self-selecting based on their experience and comfort with on-line markets. The participants in the market are also not necessarily experts in their particular field, and as such, will be affected by biases. The traders in the market are mistake prone and also tend to include biases in their individual decision making. Behavioural finance has shown how many traders will act and what biases tend to cloud their judgment (Oliven & Rietz, 2004).

Traders in both markets will be impacted on by external sources of information, such as the newspapers and television. Traders in the Inkling market may also be impacted on through the comments left by other traders. We will not be able to differentiate between impacts of external influence on the traders decisions, only note which markets is free of the influence of other traders beliefs directly. It is assumed that the external sources of information are present in both situations and will exert the same effects.

There is also a possibility that the risk and reward approach used in each market is not exactly the same. The IEM uses real money, while Inkling does not, but does allow users to see how they are ranking in the market, creating internal pressures to succeed. Research has been conducted that shows both methods are able to produce markets that are accurate, yet not if the two approached are quantitatively similar (Spann *et. al.* 2009; Van Bruggen,*et. al.* 2010).

5. Results

The results were firstly analysed to ensure that the prediction markets performed in a manner that is to be expected. Following this, descriptive statistics were run to determine the most appropriate test to run. The control and experimental groups were finally compared for differences using the Mann-Whitney U test. All tests were performed at the 0.05 significance level.

5.1.Descriptive Statistics

The initial phase in the analysis is looking at how the markets perform to actual events. Table 1 shows the accuracy of the markets in the long run. The predicted winner for the market is taken as the share with the highest value at the close of the market. It is taken as the last traded value at the close of the market. Looking at the accuracy of all 21 markets the total accuracy was 95% at the time of the markets closing. This does not indicate if the differences were significant or not, but rather in which direction the markets have moved. Table 2 show basic descriptive statistics from the markets. It also indicates the level of skewness of the data. It can be seen that the data is sufficiently skewed for non-parametric test to be an appropriate test statistic for this study.

5.2.Test for Normality

To test for normality, the Kolmogorov-Smirnov tests were conducted and p-plots were analysed. From the p-plots there was insufficient evidence to suggest that the data was normally distributed. For this reason non-parametric tests will be used to analyse the data. Table 3 shows the results of the test for normality, while the p-plots are show in

The only set of data that was normally distributed was for the 100 days set. As we are interested in all the different sets, non-parametric tests will be carried out. The Mann-Whitney U test for independent samples was used to look for differences between the experimental and control groups for the different time periods, and to compare the at close results for all markets. A significant result from this test indicates that there is a significant difference in the performance of the two markets. This would then indicate that either the presence or absence of the deliberative feedback loop has an effect on the performance of the market.

Appendix 2 contains the distributions and p-plots for the Kolmogorov-Smirnov test for normality.

5.3.Correlations

A simple correlation was run to determine if the selected sample was able to predict the correct winner and thus behave as defined for a prediction market. The outcomes for each market were correlated against the actual outcome. The positive correlation shown in Table 4 indicates that the sample behaves in the expected manner. This has been empirically proven previously, that prediction markets are good indicators of the outcome of unknown events (Berg, *et. al.*, 2008; Borison & Hamm, 2010; Van Bruggen, *et. al.* 2010). This has been shown in many studies, including work done on the IEM looking at the long run efficiency of the markets (Berg, *et. al*. 2008). To ensure that the markets which were used in this study conform to this proved hypothesis, the predicted outcome was correlated against the actual outcome (controlling for the effect of comments in the markets).

Table 4 shows the correlations for all the markets between the predicted outcomes from 100 days to the close of the market. What is of interest is to see how the predicted outcome correlates to the outcome that was the actual winner. There is a strong correlation for all the markets which improves and gets stronger as the market moves closer to the market close. Even 100 days prior to the closing of the markets, the predicted outcome has a stronger than 50% correlation, indicating that it is a better estimate of the winner than a coin flip.

From table 4 it can be seen that the markets performed as would be expected for a prediction market. The markets have correlated strongly with the actual outcome and in a positive manner. This shows that as the predicted outcome from the market increase in value it is getting closer to the actual outcome. Table 4 also shows how the long run accuracy of the two different markets perform. It is clear from this figure that the two markets have different levels of success when predicting events.

Table 1 – Accuracy of markets by group

		Control Group (No Comments)	Experimental group (Commented on)
At 100 days	Number of markets	6	3
	Market wins	4	1
	% Market wins	66.0%	33.3%
At 30 days	Number of markets	8	6
	Market wins	7	4
	% Market wins	87.5%	66%
At 10 days	Number of markets	10	8
	Market wins	8	4
	% Market wins	80%	50%
At 2 days	Number of markets	10	11
	Market wins	10	9
	% Market wins	100%	81%
At close	Number of markets	10	11
	Market wins	10	10
	% Market wins	100%	91%

Table 2 - Descriptive Statistics

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Predicted winner (100 days)	48	.0000	.8600	.188694	.2207118	1.248	.343	.585	.674
Predicted winner (30 days)	100	.0000	.9780	.140393	.2156575	2.580	.241	5.935	.478
Predicted winner (10 days)	175	.0000	.9990	.163976	.2456323	1.897	.184	2.607	.365
Predicted winner (2 days)	188	.0000	.9990	.168614	.3006422	1.783	.177	1.782	.353
Predicted winner (At close)	188	.0000	.9990	.171690	.3126696	1.767	.177	1.661	.353
Valid N (listwise)	47								

Table 3 - Test for normality

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Commented on	Statistic	df	Sig.	Statistic	df	Sig.
Predicted winner (100 days)	no comment	.172	18	.168	.909	18	.082
	Comment	.233	29	.000	.677	29	.000
Predicted winner (30 days)	no comment	.232	18	.011	.860	18	.012
	Comment	.346	29	.000	.480	29	.000
Predicted winner (10 days)	no comment	.289	18	.000	.807	18	.002
	Comment	.340	29	.000	.444	29	.000
Predicted winner (2 days)	no comment	.420	18	.000	.604	18	.000
	Comment	.497	29	.000	.422	29	.000
Predicted winner (At close)	no comment	.420	18	.000	.604	18	.000
	Comment	.498	29	.000	.425	29	.000

a. Lilliefors Significance Correction

Table 4 - Initial correlations

Correlations

Control Variables			Predicted winner (100 days)	Predicted winner (30 days)	Predicted winner (10 days)	Predicted winner (2 days)	Predicted winner (At close)	Actual Winner
Commented on	Predicted winner (100 days)	Correlation	1.000	.656	.617	.582	.581	.551
		Significance (2-tailed)	.	.000	.000	.000	.000	.000
		df	0	44	44	44	44	44
	Predicted winner (30 days)	Correlation	.656	1.000	.988	.912	.911	.871
		Significance (2-tailed)	.000	.	.000	.000	.000	.000
		df	44	0	44	44	44	44
	Predicted winner (10 days)	Correlation	.617	.988	1.000	.929	.928	.888
		Significance (2-tailed)	.000	.000	.	.000	.000	.000
		df	44	44	0	44	44	44
	Predicted winner (2 days)	Correlation	.582	.912	.929	1.000	.998	.952
		Significance (2-tailed)	.000	.000	.000	.	.000	.000
		df	44	44	44	0	44	44
	Predicted winner (At close)	Correlation	.581	.911	.928	.998	1.000	.968
		Significance (2-tailed)	.000	.000	.000	.000	.	.000
		df	44	44	44	44	0	44
	Actual Winner	Correlation	.551	.871	.888	.952	.968	1.000
		Significance (2-tailed)	.000	.000	.000	.000	.000	.
		df	44	44	44	44	44	0

5.4.Hypothesis testing

The hypotheses that were tested:

H_0: A prediction market that uses a more deliberative approach, that is a market that incorporates information sharing among participants, will produce more accurate forecasts than a prediction market that does not.

H_0 There is no difference between the traditional and non-traditional (markets that incorporate deliberative feedback) prediction markets at measured time periods prior to the close of the market.

To test for this hypothesis, sub hypothesises were looked at:

1. The distribution of predicted winner (100 Days) is the same across both groups
2. The distribution of predicted winner (30 Days) is the same across both groups
3. The distribution of predicted winner (10 Days) is the same across both groups
4. The distribution of predicted winner (2 Days) is the same across both groups
5. The distribution of predicted winner (At close) is the same across both groups
6. The distribution of actual winner is the same across both groups

Error! Reference source not found.6 shows the results of the hypothesis testing. From the hypothesis tests it can be seen that there are differences in the distribution of the predicted winners for the 100 day (α=0.007) 2 day (α=0.004) and at close of market (α=0.012). This indicates that the presence of comments in the experimental market significantly affected the markets'

Figure 5 - Results summary for Mann-Whitney U-tests

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Predicted winner (100 days) is the same across categories of Commented on.	Independent-Samples Mann-Whitney U Test	.007	Reject the null hypothesis.
2	The distribution of Predicted winner (30 days) is the same across categories of Commented on.	Independent-Samples Mann-Whitney U Test	.167	Retain the null hypothesis.
3	The distribution of Predicted winner (10 days) is the same across categories of Commented on.	Independent-Samples Mann-Whitney U Test	.086	Retain the null hypothesis.
4	The distribution of Predicted winner (2 days) is the same across categories of Commented on.	Independent-Samples Mann-Whitney U Test	.004	Reject the null hypothesis.
5	The distribution of Predicted winner (At close) is the same across categories of Commented on.	Independent-Samples Mann-Whitney U Test	.012	Reject the null hypothesis.
6	The distribution of Actual Winner is the same across categories of Commented on.	Independent-Samples Mann-Whitney U Test	.947	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

From the tests, it can be seen that there is a significant difference in the distribution for the 100 day (α=0.007) the 2 day (α=0.004) and the “at close” groups (α=0.012). these results are significant at the 0.05 level. This indicates that comments have had an effect on the markets. Due to the nature of non parametric tests it can de statistically shown the direction f the difference, yet from our earlier analysis looking at the accuracy of the markets, it can be assumed that the control group has outperformed the experimental group.

From the previous descriptive statistics in Table 1, it was shown that the control markets predicted more winners than the experimental markets for all groups. Given the significant difference between the groups for the 100 day, 2 day and “at close” groups it can be concluded that the presence of comments negatively affected the markets.

Further research would need to be conducted to discover why the comments negatively affected the outcome of the markets. This indicates that there is sufficient evidence to suggest that the null hypothesis of this research, can be rejected at significance level α=0.05

> *“H_0: A prediction market that uses a more deliberative approach, that is a market that incorporates information sharing among participants, will produce more accurate forecasts than a prediction market that does not.”*.

The results from this study indicate that there is a significant difference between markets that make use of deliberative approaches, such as un-monitored comments boards and markets that do not. The indication is that the markets that have no deliberative approach are more accurate.

5.5.Limitations

The statistics that were used in this research are non parametric in nature and thus not as strong as parametric results. The sample size for some of the groups was also small, which may have had an effect on the result for the smaller groups, such as the 100 day comparison. Due to the nature of the markets that were investigated, finding sufficient long run markets for each group proved to be challenging. For this reason it was decided to include the 100 day sub group even with the small sample size.

Due to the nature of non parametric tests the direction of the significant difference could not be calculated, yet from the descriptive statistics an assumption has been made that the accuracy of the control group was superior to that of the experimental group. Further investigation would be required in this regard.

The nature of the comments was not analysed as it fell beyond the scope of this research. Content analysis would need to be conducted on the comments of the different trades to determine how they may have affected the outcome of the markets.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

6. Discussion

Research has shown that prediction markets can effectively be used in forecasting for internal products at companies (Cowgil, *et. al,* 2009), and as such have an interesting application to business and how organisations make forecasts in the future. Prediction markets are also being proposed to help organisations learn. Through the aggregation of knowledge, a prediction market provides organisations with a large source of internal information, which can be used to develop strategies for the organisation moving forward (van der Heijen, 2004). With the promise and potential of prediction markets in the business world the effective use of them has been an interesting source of research.

A question would be how to optimise the function of the markets, determining exactly how many people are needed in a market to make it function optimally. Another question is concerned with how to improve the accuracy through incorporating a more deliberative approach (Van Bruggen, *et. al.* 2010). The aim of this research was to look at the use of a deliberative approach, such as a feedback loop or comment board, in the prediction markets and then determine how this affected the accuracy of the market. Looking to see if it improved the market's ability to forecast more accurately by being more open about what reasoning had been incorporated into the price of the market, or if it hindered traders' ability to value the information in the market.

This research forms part of the growing body of knowledge on prediction markets and their application to the business world. The experimental approach was used in this research comparing the results of completed prediction markets from the IEM and Inkling markets. Inkling has made use of comment boards in its markets and included an option allowing traders to state reasons for why they have made their trades. The IEM is a more traditional market and does not allow traders to communicate through such methods. The topic of the markets vary, with Inkling looking at social, political and technological topics, the IEM has a core focus on political markets, but also has markets that look at entertainment. The completed markets were used in this research. Each market was assigned to either the experimental (markets that contained comments) or the control (markets that had no comments) groups.

The initial analysis looked purely at the accuracy of the markets over time. Showing how the markets in each group would have fared had they been closed at 100 days, 30 days, 10 days and 2 days prior to the actual closure of the market. This compared favourably to work that had been conducted previously on the IEM looking at the long run accuracy of the IEM and traditional polls (Berg, *et. al,* 2008). The accuracy of the markets in this sample was correct 56% of the time. The accuracy did improve closer to the closure of the markets with markets at closure being 95% accurate. This was looking at the results across both groups and did not take account of the possible difference of having comments on markets. When looking at the two groups the markets that do not contain any comments performed better across all the time breaks.

The statistics showed that there was significant differences in the performance of the two groups at the 100 days (α=0.007) 2 day (α=0.004) and at close of market (α=0.012) at the 95% significance level. Of particular interest is that there was a significant difference between the two groups for the "at close" time break. This break takes all the markets into account and would as such be of interest when answering the question about the accuracy of the markets for the two groups.

From the initial analysis it shows that the group that did not have any comments performed better. This is contrary to the proposed hypothesis that adding a deliberative approach would increase the accuracy of the prediction markets as suggested in Van Bruggen, *et. al. (2010).* While this seems contrary to the power of prediction markets it does seem to follow findings that have been seen in Delphi studies.

As has been discussed previously, the Delphi technique has been used to aggregate the opinions of experts through successive rounds of prediction. Delphi has been successful as it makes use of anonymity of experts. It greatly improves the chances of forecasters obtaining unbiased forecast estimates. The key to the success of Delphi forecast is in the panel design. The panel needs to be made up of experts in particular fields. If any members of the group are ignorant about the issue that is being discussed the outcome suffers. Misinformed individuals who take part in the Delphi process derail the process (Green, *et. al*, 2007). From the findings of this study it can be suggested that a similar issue has been observed in the prediction markets that contained comment boards. The comments may have negatively affected the outcome of the market as other individuals who did not have sufficient knowledge followed the logic of a single trader's comment. This is speculative, but would be an interesting avenue for further research by looking at what comments were made and how this affected the traders' behaviour.

A second possible issue could be that discussions can take place among fellow traders on the comment boards available to traders in the Inkling markets prior to them making trades. This type of behaviour will lead to consensus seeking, as is seen in many meetings. This opens the deliberative element of the experimental group up to many of the problems associated with meeting and the biases that people have when communicating in this manner. The issue of group think also comes up again as the comments will cause people who hold similar views to reinforce their beliefs and place larger wagers than they probably would have without having the reinforcement of the misinformed commenter (Janis, Groupthink: Psychological Studies of policy decisions and Fiascos, 1982; Kerr & Tindale, 2011).

Cognitive biases and errors stem from the limited mental processing capacity the people have, and the simplification tools and techniques that they then use to overcome this limitation. The strategies that are used have been discussed previously, these heuristics are mental shortcuts that allow us to filter out some of the information that we are constantly being bombarded with on a daily basis

(Jones & Roelofsma, 2000). When looking at biases in decision making, having discussion boards on the markets appears to open them up to many of the biases that are seen in forecasting. The “Pollyanna principle” was described in 1978 by Matlin and Stang, this principle looked at how people had inherent biases towards favouring positive outcomes (Matlin & Stang, 1987). The optimism bias is a prerequisite to action, without us having an optimistic view of the future outcome of an event it is unlikely that we will act (Tyebjee, 1987). To back up this belief that there will be a positive outcome we often look for supporting evidence that corroborates our views. In the banking and trading circles this is often referred to as the herding instinct, there is a desire to conform to the behaviour of others (Roxburgh, 2003). Traditionally this can be seen in the markets through the formation of bubbles. Traders conform in their beliefs about an outcome and then look for evidence that reinforces their beliefs, causing the price to increase and as such further reinforces what they believe to be true, creating the market bubble, the “self-fulfilling prophesy”.

Looking at the prediction markets that were created there is a possibility of biases creeping in from interactions through the comment boards at the social level. These biases can result from social interactions, through discussions and viewing the trade patterns of more respected traders. In this situation members may be making assumptions about others’ ability. These assumptions often are anchored on their own abilities and referred to as social projection (Jones & Roelofsma, 2000). Other biases that make up the social or group biases are group think, group polarisation and group escalation of commitment.

As discussed previously, “group think” as a concept has been around since the early 1970s. It describes a tendency for members of a group to strive for unanimity over an actual appraisal of reality. Group pressures cause the group to look at achieving a consensus rather than whether the decision was a good one or not (Janis, Victims of Groupthink, 1972). As a theory group think is still being investigated and its application to different situations is being tested, yet the value of the concept has been widely accepted and investigated (Jones & Roelofsma, 2000). According to the original research for group think to take place there are certain pre-requisites. Groups need to be cohesive, primarily. There also need to be elements of strong leadership, time pressures and complex decisions that need to be made (Janis, Groupthink: Psychological Studies of policy decisions and Fiascos, 1982). Looking at the structure of the prediction markets, it is possible that there are cohesive groups that take part in the different markets. The ability of groups to form through online markets would have to be investigated further in future research.

Research shows that the proximity from co-workers affecting their predictions. In this study it was shown that traders would show similar trading patterns to fellow traders with whom they shared space (Cowgil*, et. al*, 2009). While this is not in a virtual world, but in an office, it does show that traders tend to form groups and discuss their trades. This behaviour would allow for the formation of group think in small trader groups. With the addition of open forums in online markets, traders that have no actual physical proximity can now communicate and increase the group, on-line access now provides traders with a virtual proximity.

A further bias which may have had an impact on the results of the experimental markets is a phenomenon called group polarisation. In group decision making, polarisation occurs when the majority of the group holds a position, and through discussion the position is intensified. This leads groups to further strengthen the views that are already held in the group (Jones & Roelofsma, 2000). The effects of group polarisation have been studied not only in economic circles, but also in the realms of social behaviour, as group polarisation is one of the first steps towards extremism and hate speech (Yadi & Boyd, 2010). The internet has fundamentally altered how people produce and consume information, yet the availability of a diversity of information does not cause people to actively look for contradicting views (Yadi & Boyd, 2010). Rather, they tend to find information that supports their current view point. In prediction markets the comment board would facilitate this polarisation. A study conducted by Yadi & Boyd (2010) looked at how conversations on Twitter over the shooting of an abortion doctor in the United States of America lead to heterogeneity, groups interacting with like minded individuals, and homophily, individuals sharing the same view points, among the pro-choice and pro-life groups, with each group more likely to interact with people who had shared views.

This bias would not be an issue in the control group markets as the trades are all anonymous. Groups would not be able to form views based on beliefs about the future, and then discuss these views openly. On the contrary this could be an issue in prediction markets with open discussion forums. The discussion would lead to strengthening views for opposing sides and affect the accuracy of the market in a non-rational manner. While debate around topics is necessary, it does lead to many issues that are seen in traditional meetings and many of the biases that prediction markets were developed to overcome.

In the traditional view of how markets work, it is believed that biases do not affect the outcome of the markets, yet with the recent events it can be seen that even the aggregated view of the market can produce an incorrect outcome. The rational arbitrage that overcomes irrational emotions should stabilise the markets and provide a rational outcome, yet this process alone is not able to eliminate all the effects of biases at an individual and group level in markets (Oechssler, Roider, & Schmitz, 2009). The results from this market show that there are issues which are present in the experimental group that are not seen in the control group.

While the research question has been answered this work does continue to add strength to the prediction market field. Through the use of the correlation matrix it was shown that the markets do give a clear indication of the outcome of an event prior to the actual event. Looking at markets 100 days prior to the event there was a positive correlation (r=0.551) this correlation grew in strength as the markets got closer to the close, with markets at the close having a very high correlation (r=0.968). As discussed previously, prediction markets have been used very effectively in demand forecasting tasks. For businesses like Best Buy, their markets were able to achieve 98% accuracy while their forecasting department only managed 95% accuracy (Dye, 2008).

From the correlation data it can be shown that even at a period of 3 months prior to the close of a market there is more than a 55% chance of the markets predicting the correct outcome, and a 96.8% chance at the close of the markets, for the combined group. This information is powerful for a forecaster, as it is able to be generated relatively quickly and easily in a very cost effective manner. Looking at the markets which ran for a short period of time, 30 days or less, the results showed that they were predicting the correct outcome 87% of the time (30 days). For many business questions they would want to run shorter time frames and have accesses to real time changes in the markets. This would require that they may have to use results of markets that have not closed yet. To do this with any sort of confidence they would need to know how quickly the market is able to predict the correct outcome prior to the close of the market. Convergence is a new term to do just that.

Convergence was defined as the ability of the market to move in the direction of the correct outcome and remain there till the end. An example of this would be in a choice between two possible outcomes, each has a 50% chance of occurring, priced at $50, thus as soon as one outcome moves above the $50 mark, it has moved in the correct direction. If the price never drops below the $50 mark for the rest of the life span of the market it is said to have converged on the correct outcome. The time taken to reach this convergence was measured in days from the opening of the market (Davis, 2011).

The convergence of markets on the correct answer is a new dimension to prediction markets and the body of knowledge that is developing, as described previously in Figure 3.. While it fell beyond the scope of this research, looking at how comments affect convergence would be an interesting topic for the future. The inclusion of comments on conversion study may help to indicate how rapidly the market responds to a new comment, and how quickly this comment may change the direction of the markets. This would help to identify the effects of some of the biases that have been previously discussed.

Looking at how the market reacts to the comments and how biases show up in the behaviour of traders around the comments would be an avenue for future research.

From the literature, Surowiecki (2004) indicated that there were some fundamental elements that are required for prediction markets to function effectively. These were having a group of diverse individuals and ensuring that the traders are independent of each other (Surowiecki, 2004). The diversity was crucial to the performance of the market, as the diverse heuristic set that the individuals had allowed the market to aggregate information for every possible outcome. Markets that were filled with a homogenous group would all share a similar view of the future and solve problems in a similar manner, thus not providing any useful information to the market. The outcome of this market would be the same as any individual's prediction.

The second criteria, that of independence, looks to avoid many of the issues associated with the biases that individuals and groups have and share. Compromising on the independence of the traders negatively impacted on the performance of the market (Surowiecki, 2004).

While trader independence is seen to be important to the functioning of the markets, deliberative approaches such as Delphi and other “con-joint decision making” (CJD) analyses make use of the opinions of groups of people to formulate a prediction. The hypothesis was that through allowing communication between traders, the knowledge transfer would increase and the performance of the markets would improve, that is the markets would predict the correct outcome more often.

It appears though that there is a fine line between allowing information sharing within a market and compromising the independence of traders in that market. This study has indicated that having un-moderated discussion between traders through the use of comment boards and discussion forums does not improve the accuracy of the markets, but rather decreases the accuracy. It could be hypothesised that through allowing traders to communicate freely on the site their independence becomes compromised and they start to fall prey to cognitive and behavioural biases in their decision making.

This is an avenue which would require further research. Being able to identify how much communication is possible within a market before independence becomes compromised would be of great interest and value. It would also be interesting to look at how a more structured approach to providing deliberative feedback into the markets affects not only the pricing but the speed of the response.

7. Conclusions

Prediction markets have been of interest to the academic community for a long time. The phenomenon was first noticed in farmers' markets where groups of individuals were able to accurately predict the weight of oxen during agricultural shows (Surowiecki, 2004). While the initial markets where not very sophisticated, taking the average value of large groups' predictions, as time progressed, the level of complexity increased. The application of prediction markets has also become more main stream, and new applications are been found regularly.

The growth in interest of prediction markets has come on the back of increasingly advanced technologies. Through powerful computer processing large amounts of data can be analysed very rapidly. This has made the analysis of prediction markets far easier and allowed users to reach a much larger audience through the markets. The larger participation pool for prediction markets has help to increase the accuracy of the markets and how quickly they are able to respond to changes in the environment.

Most recent applications of prediction markets have been in large corporations, such as Google, HP, Best Buy, Yahoo! and Microsoft (Cowgil, *et al*, 2009; King, 2006; Dye, 2008). They have also been used in the defence world, for demand forecasting and predicting terrorist events, with the controversial PAM market (Hanson, 2006). Markets have been used in these examples to predict usage of products, such as how many users Google will have in the next six months, the completion dates of new projects, looking specifically what the potential for a delay is or how likely the market believes a critical deadline will be missed. They have also found a place in predicting the unknown. While PAM was never fully operational, Inkling markets have been predicting the intensity of tropical storms in the Atlantic.

The exact mechanism of how the markets work is still being understood, and much work has been conducted on the role of diversity of participants, the sizes of markets, the role of experts in the groups and the role of individuals in the markets (Hong & Page, 2004; Page, 2007; Surowiecki, 2004). One thing that is becoming clearer as the prediction markets mature is the role of the individual traders within markets. Biases are starting to be uncovered in our decision making processes and these are being translated into the markets.

The traditional view has been that markets are able to overcome these biases, as the market aggregates information it should stabilise and provide a rational outcome (Oechssler, Roider, & Schmitz, 2009). There is a certain amount of evidence that is beginning to build that shows irrational patterns in the markets and the creation of bubbles is one such example (Kunz, 2008). This highlights the human element to markets. While the aggregated view of the market may be rational, the inputs from traders very often are not.

This highlights how important it is to understand the behaviour of traders and other participants in markets as this will affect their performance. Cowgil (2009) saw that traders who worked in close proximity tended to have similar trading patterns and made similar predictions. It was also shown that experience in trading on an online platform also affected the traders' ability and thus the ability of the markets. Further evidence of the different biases that appear in groups, such as groupthink and group polarisation, have also been shown to affect how people are affected by the interactions with others. Large online prediction markets have been able to overcome these biases to a large extent as traders' interactions are limited and generally there are sufficient traders to correct poor decisions of small groups.

In small markets however, with as few as six traders, there are other issues that need to be considered. Abramawicz (2006) found that markets with as few as six traders were still able to predict events successfully, based on the traditional approach to a market, that is, the trader's only interaction is through the market price and the information that they have (Abramawicz, 2006). The role of independence of traders has been explored in the book "The Wisdom of Crowds" and set up as a key requirement for a prediction market to work (Surowiecki, 2004). It is argued in this book that with independent traders many of the biases and issues of groups can be overcome and the market will then be able to stabilise around a rational outcome.

The level of independence and amount of interaction between traders has become an area of interest for some researchers, especially in the demand forecasting industry. There have been many tools available for demand forecasters, from unsophisticated, easy to implement, face-to-face meeting and expert groups, to more sophisticated and time consuming processes like Delphi and CJD analysis. (Green, *et. al*, 2007). These methods make use of rounds of interaction followed by re-forecasting based on the opinion of experts in given fields. They also require facilitation from skilled professionals, making the processes very time consuming and expensive. Yet they have shown to be an effective method to elicit predictions from people.

With this in mind, including more deliberative techniques in a prediction market should affect the market and make it more accurate. This question was posed in recent research by Van Bruggen *et.al*. (2010) through this research however, the evidence suggests that allowing traders a platform on which to discuss their views and indicate reasons for their trading behaviour, the market is negatively impacted. Biases may have started to creep into the markets and then become amplified by the actions of groups of traders, breaking down the effectiveness of the market.

From this research it can be argued that having unregulated open communication between traders affects the ability of the market to produce prediction. Yet there is still an opportunity to investigate what impact a more structured information feedback process may have on the outcome of the market. For this research the open forums and discussion boards were not regulated in any manner. This allowed traders to interact freely and could have voided the assumption of independence of traders that Surowiecki (2004) implied in his work. Finding the appropriate balance of information feedback while still maintaining the independence of traders would be a topic for future study.

Further areas for future investigation would be investigating how information feedback affects the convergence time of the markets. Convergence time in markets was found to be very rapid, averaging 3 to 13 days (Davis, 2011). This indicates that the markets are able to make rapid decisions. This is of great value to an organisation which is looking for information on new product development and demand forecasts. Yet it does not indicate how quickly the market may react to new information that is presented to it. Through the use of a deliberative feedback tool, this could be tested further to determine how quickly the market reacts to information that is presented to it in a controlled manner.

While outside of the scope of this study there is also an opportunity to analyse what discussions take place through the forums and comment boards, looking for what information is passed and how it affects trading. This can be correlated to the availability of news reports on the subject. Through this analysis it would be possible to uncover how biases creep into the market and if it is affected by group bias. Looking at this would also help identify how the market values information. This has been an issue in traditional markets, while it is evident that new information is being included it is not evident what this information is and how it is affecting the market.

Prediction markets are adding value to many organisations. Obtaining a better understanding of how they function and methods to improve their functioning will be of great value. There are still many different avenues to be followed in this field and many opportunities to increase the body of knowledge on this topic.

Technology is constantly improving and with it the ability of crowds to help answer many of the world's issues. Crowd sourcing as a concept has come of age and matured along with the internet, which acts as a tool for facilitating communication and allows groups of diverse individuals the opportunity to collaborate on a wide range of projects. This research builds on the theories postulated by Surowiecki (2004) indicating that traders in prediction markets need to maintain a lever of independence in order for the prediction market to function optimally. This goes against the postulations of Van Bruggen *et.al*. (2010) who believed that including more deliberative feedback would improve the accuracy of prediction markets.

The findings of this research indicate that the inclusion of discussion boards in a prediction market hamper its performance, decreasing the accuracy of the market when compared to a traditional market structure. Yet as seen in the literature there is reason to believe that through the inclusion of a more discreet method of information feedback, the efficiency of the markets can be improved.

Reference List

Abramawicz, M. (2006). Deliberative Information Markets for small Groups. In R. Hahn, & P Tetlock (Eds.), *Information Markets: A new Way of Making Decisions* (pp. 101-122). Washington: AEI-Brooking Press.

Berg, J, & Rietz, T (2010). Longshot, Overconfidence and Efficiency on the Iowa Electronic Market.

Berg, J, Forsythe, R., Nelson, F, & Rietz, T (2001). Results from a Dozen Years of Election Futures. In C. Plott, & V. Smith (Eds.), *Handbook of Experimental Economic Results.* Amsterdam: Elsevier.

Berg, J, Nelson, F, & Rietz, T (2008). Prediction market accuracy in the long run. pp. 285-300.

Borison, A, & Hamm, G (2010). Prediction Markets: A New Tool For Strategic Decision Making. *California Management Review , 52* (4), pp. 125-141.

Burt, G, & Wright, G (2006). "Seeing" for organisational foresight. *Futures , 38*, pp. 887-893.

Cowgil, B, Wolfers, J, & Zitzewitz, E (2009). Using Prediction Markets to track Information Flows: Evidence from Google. p. 3.

Davis, D (2011). Designing a viable prediction market to forecast defense acquisitions cost and schedule outcomes. pp. 351-366.

Dye, R (2008). The promise of prediction markets: A roundtable. *McKinsey Quarterly* (2), pp. 82-93.

Foutz, N, & Jank, W (2007, May 7). *The Wisdom of crowds: Pre-release Forecasting for New Products via Functional Shape Analysis of the Online*

Virtual Stock Market. Retrieved September 10, 2011, from Social Science Research Networks: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1432444

Goodwin, P, & Wright, G (2010). The limits of forecasting methods in anticipating rare events. pp. 355-368.

Graefe, A, Luckner, S, & Weinhardt, C (2010). Prediction markets for foresight. *Futures , 42*, pp. 394-404.

Green, K, Armstrong, J, & Graefe, A (2007). Methods to Elicite Forecasts from Groups: Delhpi and Prediction Markets Compared.

Hanson, R (2006). Designing real terrorism futures. pp. 257-274.

Harvey, J (1974). The Abilene Paradox: The Management of Agreement. *Organisational Dynamics* .

Hong, L, & Page, S (2004). Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers. *Proceedings of the National Acadamy of Science , 101* (46), pp. 16385-16389.

Inkling Public Markets. (2011, October 13). *Ratio of Registered Users vs. Inactive*. Retrieved October 13, 2011, from Inkling Markets: http://home.inklingmarkets.com/site/leaderboard#site_active_users_ratio

Janis, I (1982). *Groupthink: Psychological Studies of policy decisions and Fiascos.* Boston: Houghton Mifflin.

Janis, I (1972). *Victims of Groupthink.* Boston: Houghton Mifflin.

Jones, P, & Roelofsma, P (2000). Social biases in team decision-making. pp. 1129-1152.

Kerr, N, & Tindale, R (2011). Group-based forecasting? A social psychological analysis. *International Journal of Forecasting , 27*, pp. 14-40.

King, R (2006, August 3). Workers, Place Your Bets. *BusinessWeek* .

Kunz, B (2008, October 15). Prediction Markets Meet Wall Street. *BusinessWeek Online* , pp. 23-23.

Matlin, M, & Stang, D (1987). *The Pollyanna pricipal: selectivity in language, memory and thought.* Cambridge: Sckenkman.

Mauboussin, M (2009). *Thnik Twice: Harnessing the power of counterintuition.* Boston: Harvard Business Press.

Oechssler, J, Roider, A, & Schmitz, P (2009). Cognative abilities and behavioural biases. pp. 147-152.

Oliven, K, & Rietz, T (2004). Suckers are born but Marktes are Made: Individual Rationality, Arbitrage, and Market Efficiency on an Electroni Future Market. pp. 336-351.

Orrell, D, & McSharry, P (2009). System economics: Overcoming the pitfalls of foercastingmodelsvia a multidisciplinary approach. (25), pp. 734-743.

Page, S. E (2007). *The Difference.* New Jersey: Princeton University Press.

Robbins, S, & Judge, T (2007). *Organisational Behaviour* (12 ed.). New Jersey: Pearson Education.

Rowe, G (2007). A Guide to Delphi.

Roxburgh, C (2003). Hidden flaws in stratergy. pp. 27-39.

Rutledge, R (1993). The effects of Group decisions and group-shift on use of the anchoring and adjusting heuristic. pp. 215-226.

Schonfeld, E (2006). The Wisdon of the Corporate Crowd. *Business 2.0* , *7* (8), pp. 47-49.

Senge, P (1994). Mental Models. *The fifth Discipline Fieldbook* , 102-115.

Snowburg, E, Wolfers, J, & Zitzewitz, E (2009). Prediction Markets: From Politics to Business (and Back). In *Handbook of Sports and Lottery Markets* (pp. 385-402).

Spann, M, Ernst, H, Skiera, B, & Soll, J (2009). Identification of lead users for consumer products via virtual stock markets. *26*, pp. 322-335.

Stoner, J (1961). A Comparison of Individual and Group Decisions Involving Risk.

Surowiecki, J (2004). *The Wisdom of Crowds.* New York: Anchor Books.

The University of Iowa, Henry B. Tippie College of Business. (2010). *Research - Iowa Electronic Markets*. Retrieved October 13, 2011, from Iowa Electronic Markets: http://tippie.uiowa.edu/iem/research/

Tozzi, J (2010, March 30). *America's most promising start ups*. Retrieved October 12, 2011, from Bloomberg Businessweek: http://images.businessweek.com/slideshows/20110211/america-s-most-promising-startups/slides/17

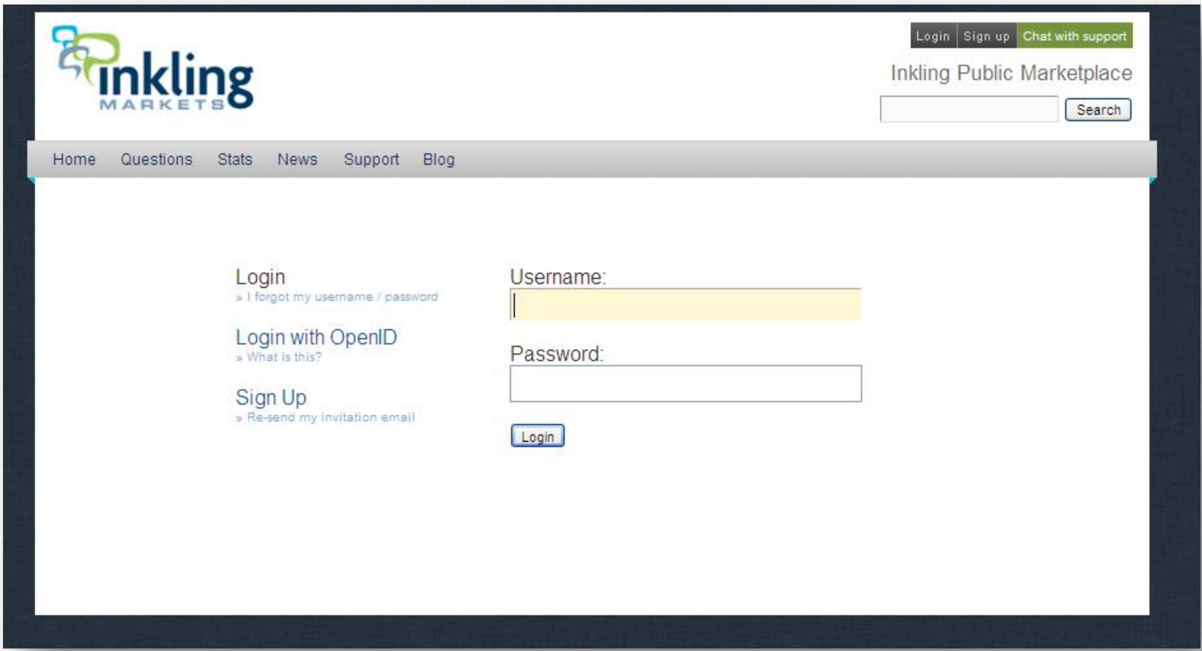
Tyebjee, T (1987). Behavioural biases in new product forecasting. pp. 393-404.

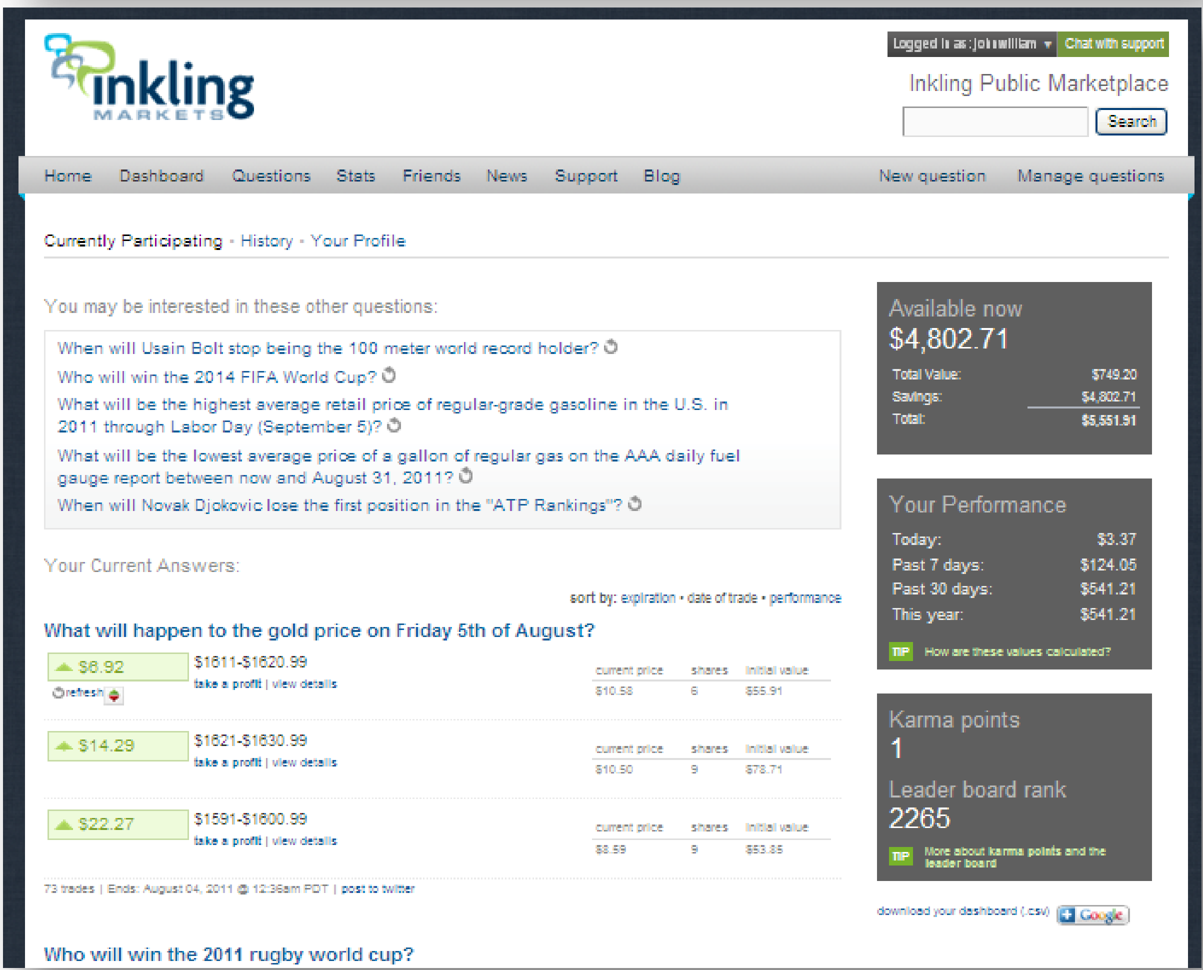
Van Bruggen, G, Spann, M, Lilien, G, & Skiera, B (2010). Prediction Markets as institutional forecasting support system. *Decision SUpport Systems* , *49*, pp. 404-416.

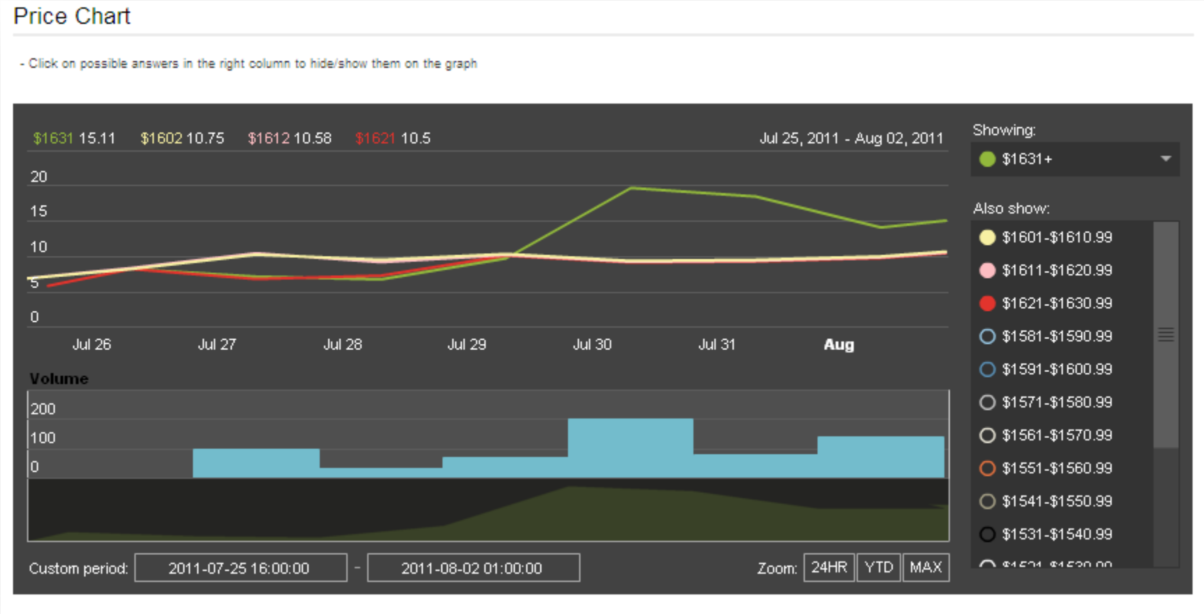
van der Heijen, K (2004). Can internally generated futures accelerate organizational learning? (36), pp. 145-159.

Yadi, S, & Boyd, D (2010). Dynamic Debates: A anlysis of group polarization over time on Twitter. pp. 316-327.

Appendix I – screens shots from the IEM and Inkling markets







Iowa Electronic Markets

Login as a Trader

July 25, 2008: New Survey Available

PLEASE COMPLETE OUR SURVEYS
Please complete our online surveys at: http://iemweb.biz.uiowa.edu/surveys. These surveys are a valuable part of our IEM research.
(newest survey dated Friday, July 25, 2008)

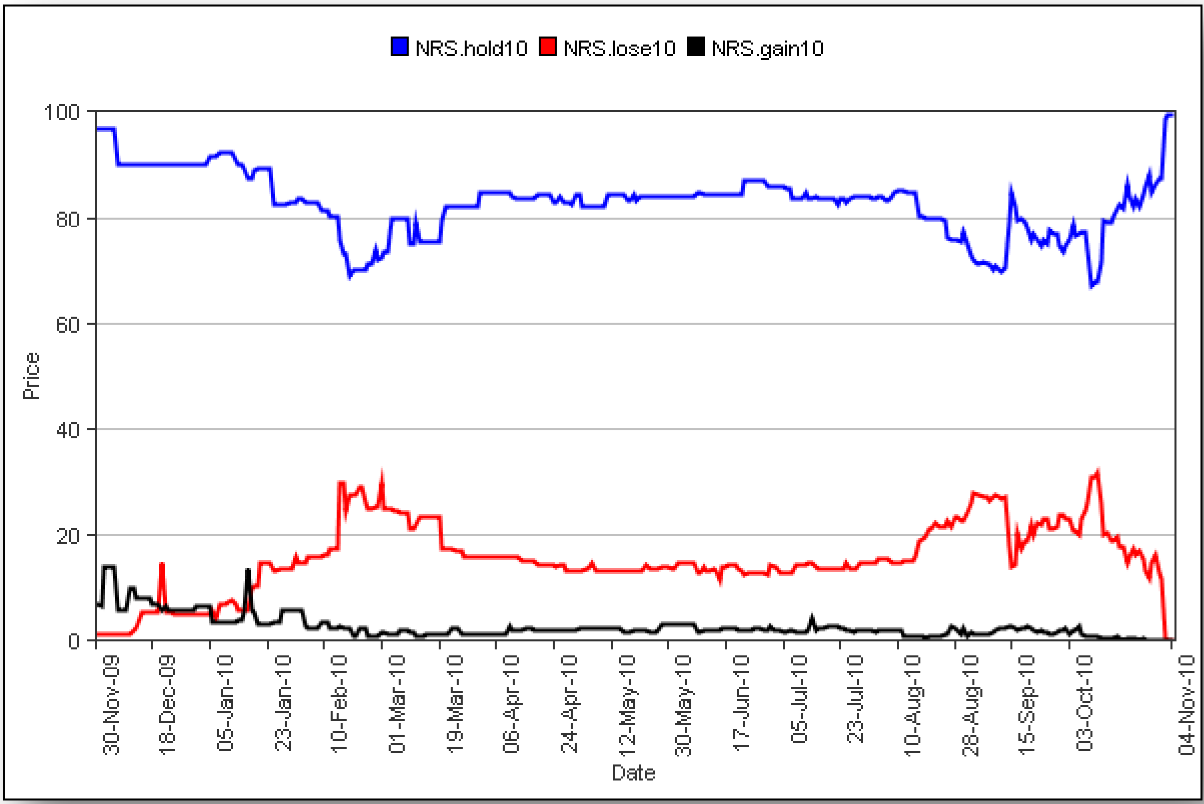
Login Name:
Password (case sensitive): (Forget your password?)
Language: english
IP Security: IP Security Enabled (Read about IP security)

Login

Guest Accounts: You can log in as a guest and trade in a practice market by using one of our practice account names and passwords. Practice account names are aaa, bbb, ccc, ..., zzz. Use the same three letter string as the password for these accounts.

| WebEx Manual | IEM Home Page |

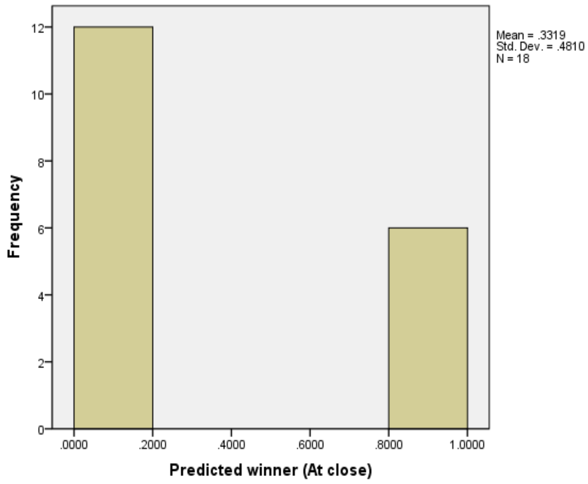
Site maintained by Iowa Electronic Markets .
WebEx Trading System © 1998-2010 copyright Joyce E. Berg and Forrest D. Nelson. All rights reserved.


NRS.hold10
NRS.lose10
NRS.gain10
100
80
60
40
20
0
Price
30-Nov-09
18-Dec-09
05-Jan-10
23-Jan-10
10-Feb-10
01-Mar-10
19-Mar-10
06-Apr-10
24-Apr-10
12-May-10
30-May-10
17-Jun-10
05-Jul-10
23-Jul-10
10-Aug-10
28-Aug-10
15-Sep-10
03-Oct-10
04-Nov-10
Date

Appendix II – Distributions and p-plots

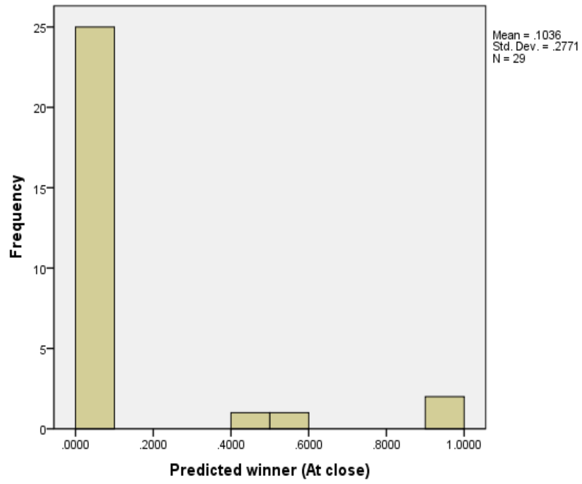
Histogram

for Commentedon= no comment



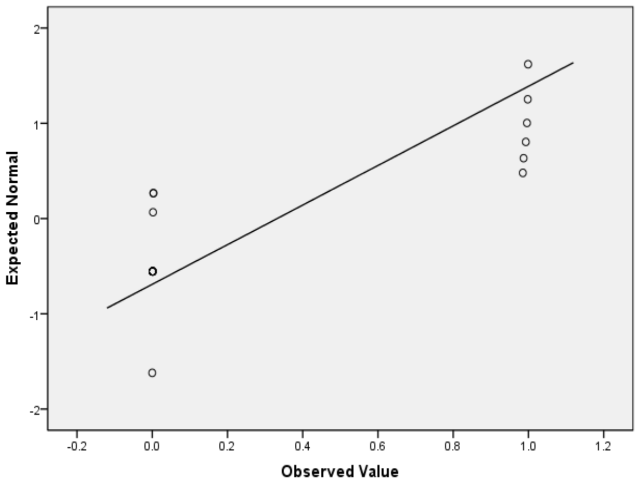
Histogram

for Commentedon= comment



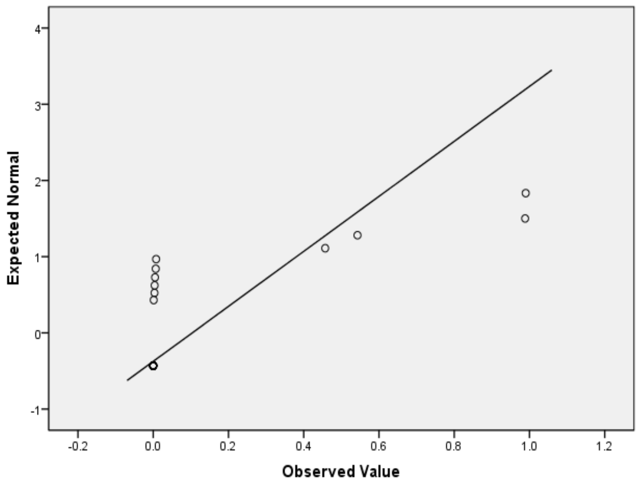
Normal Q-Q Plot of Predicted winner (At close)

for Commentedon= no comment



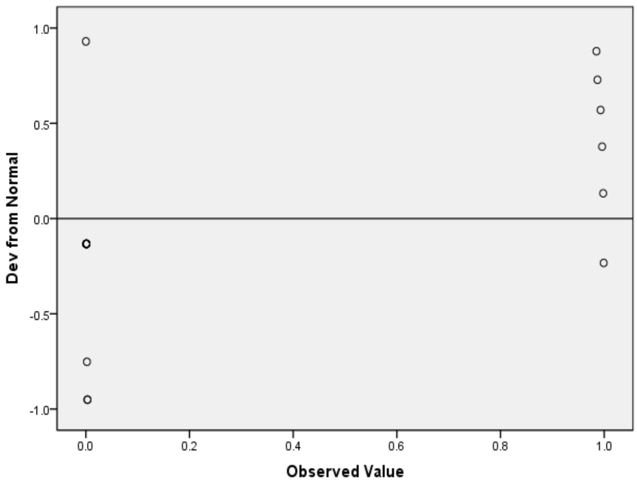
Normal Q-Q Plot of Predicted winner (At close)

for Commentedon= comment



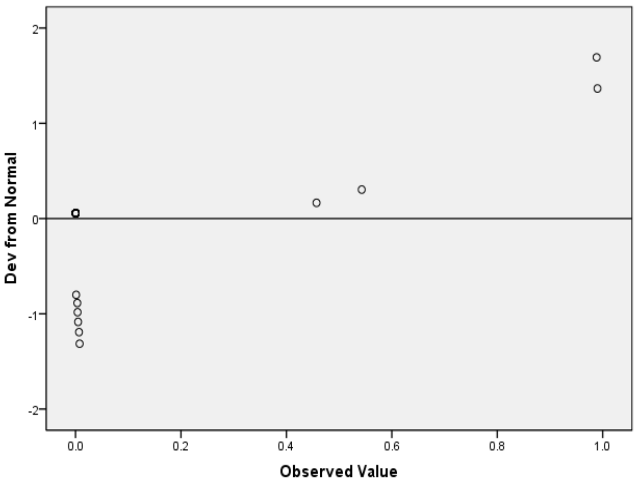
Detrended Normal Q-Q Plot of Predicted winner (At close)

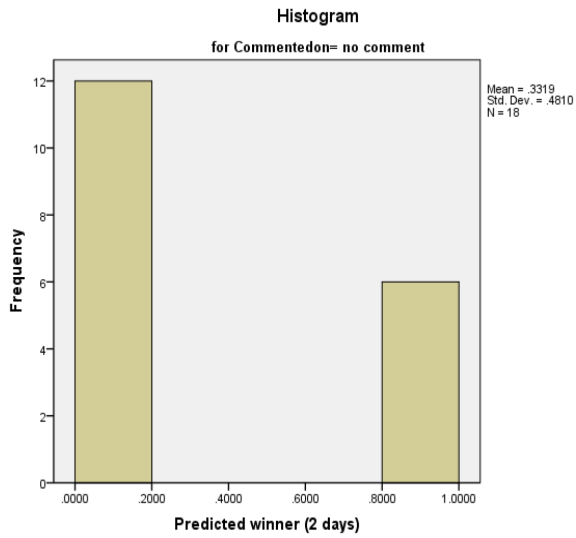
for Commentedon= no comment

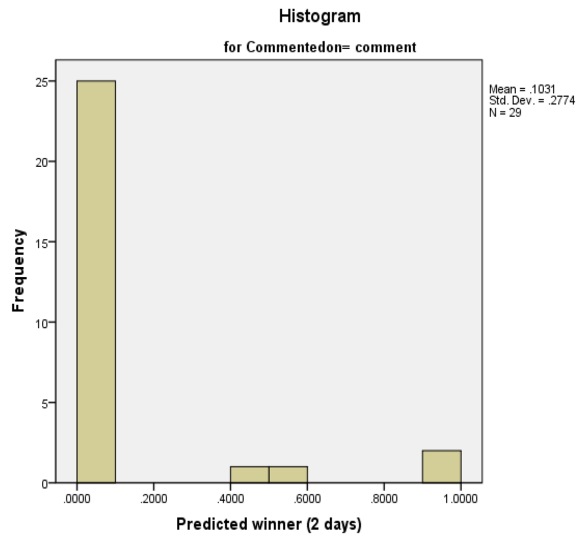


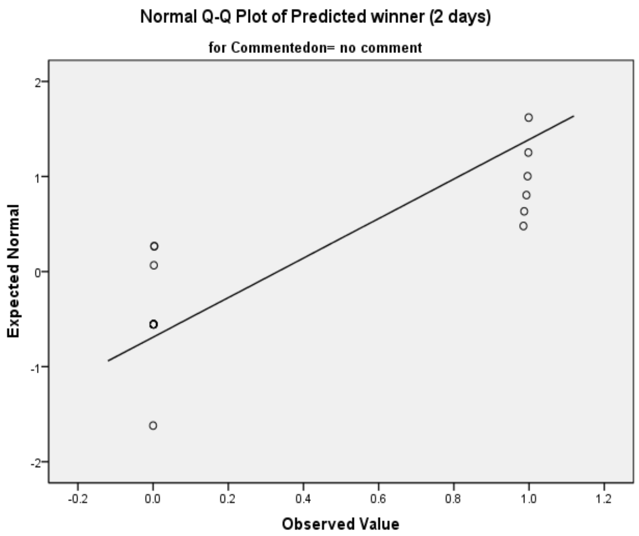
Detrended Normal Q-Q Plot of Predicted winner (At close)

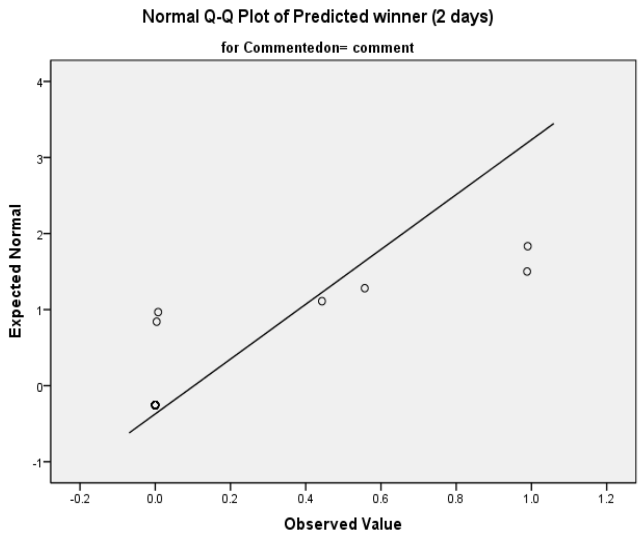
for Commentedon= comment

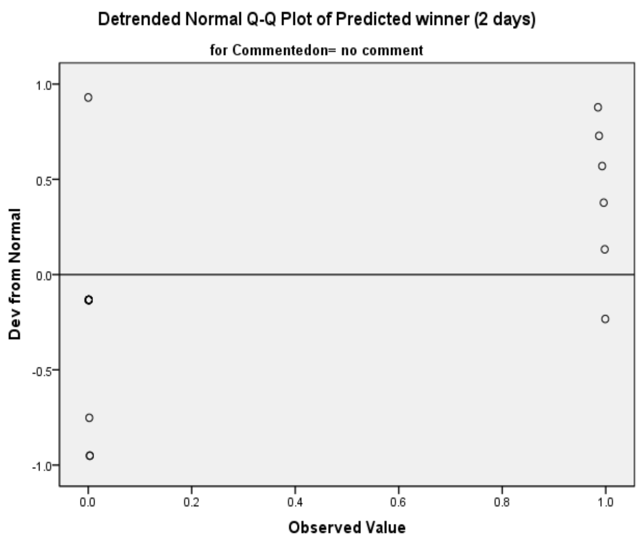


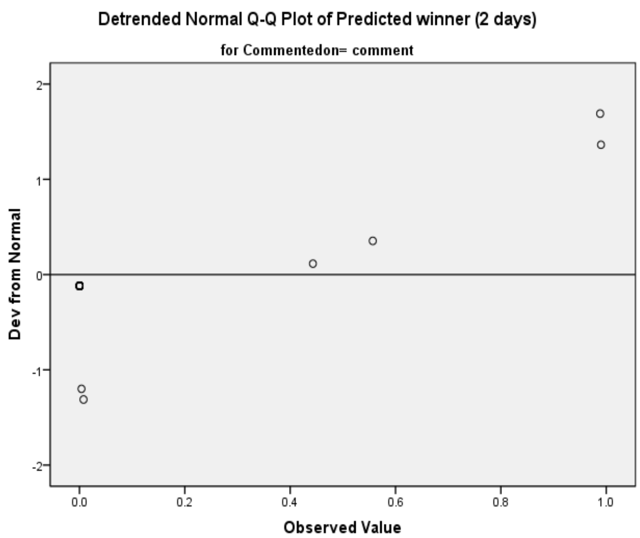

Histogram
for Commentedon= no comment
Mean = .3319
Std. Dev. = .4810
N = 18
Frequency
Predicted winner (2 days)

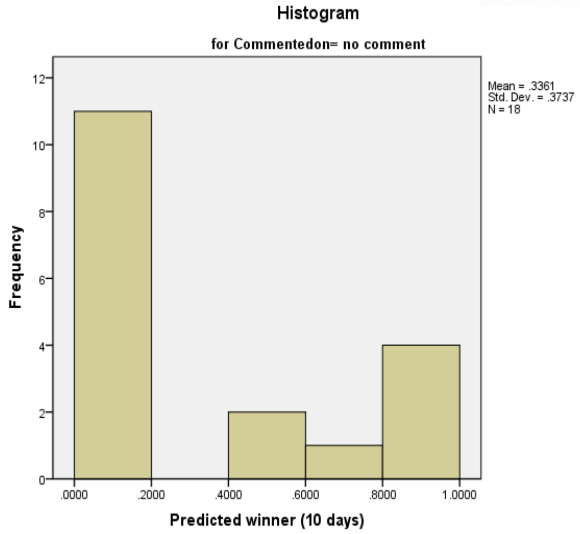

Histogram
for Commentedon= comment
Mean = .1031
Std. Dev. = .2774
N = 29
Frequency
Predicted winner (2 days)

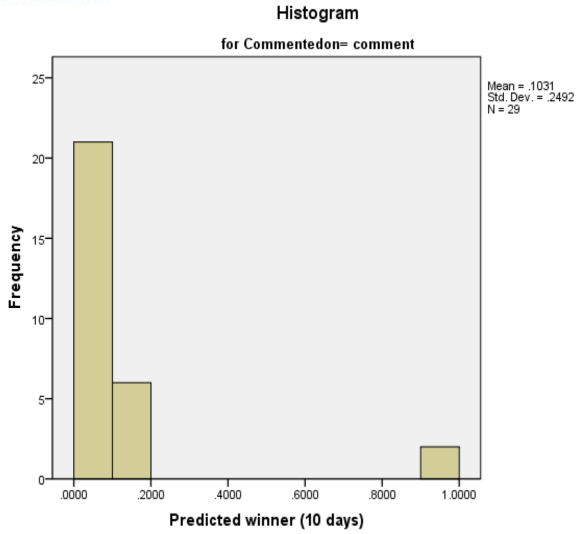

Normal Q-Q Plot of Predicted winner (2 days)
for Commentedon= no comment
Expected Normal
Observed Value

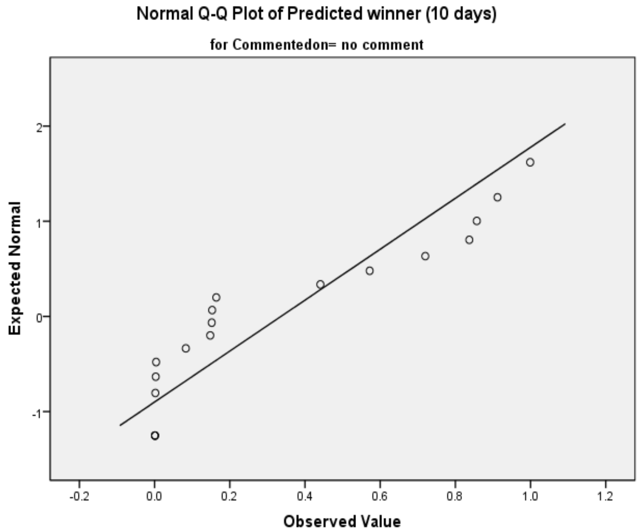

Normal Q-Q Plot of Predicted winner (2 days)
for Commentedon= comment
Expected Normal
Observed Value

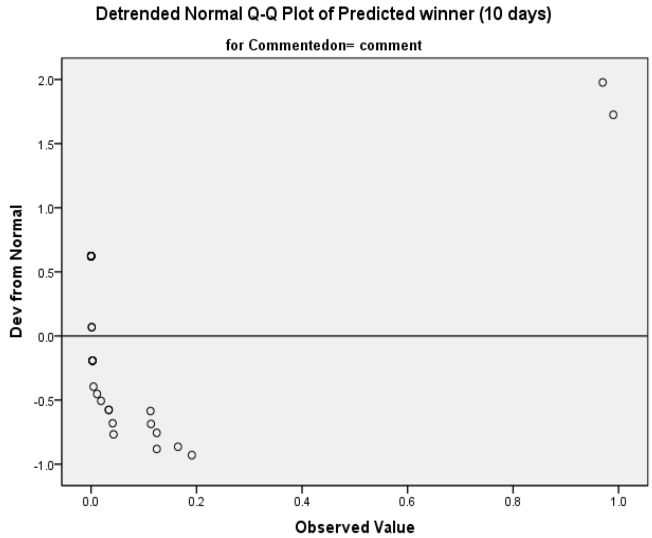

Detrended Normal Q-Q Plot of Predicted winner (2 days)
for Commentedon= no comment
Dev from Normal
Observed Value

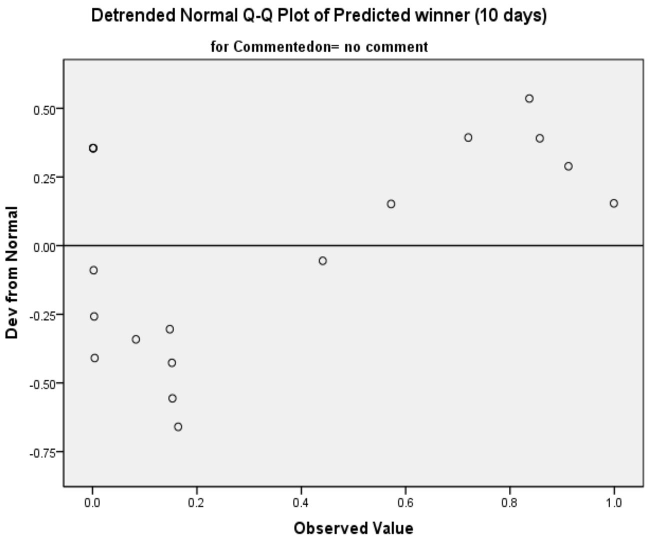

Detrended Normal Q-Q Plot of Predicted winner (2 days)
for Commentedon= comment
Dev from Normal
Observed Value

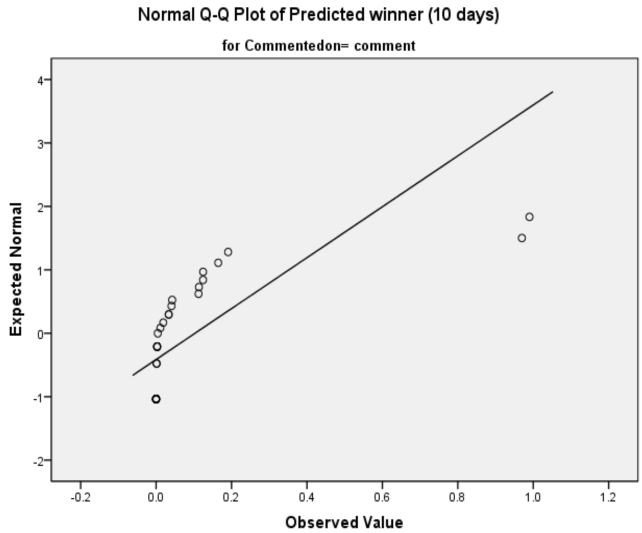

Histogram
for Commentedon= no comment
Mean = .3361
Std. Dev. = .3737
N = 18
Frequency
Predicted winner (10 days)

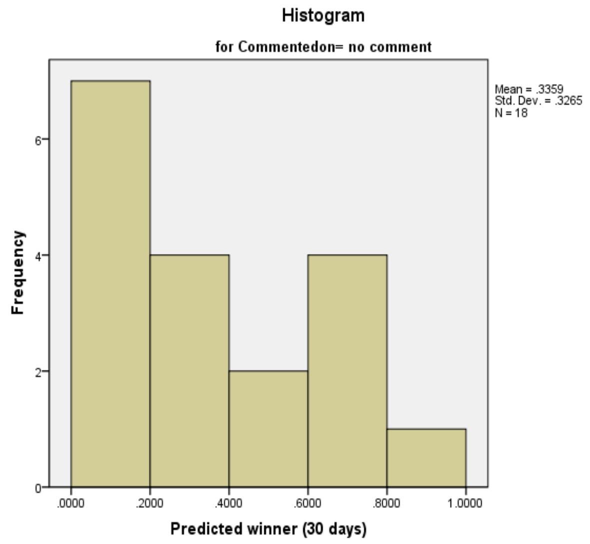

Histogram
for Commentedon= comment
Mean = .1031
Std. Dev. = .2492
N = 29
Frequency
Predicted winner (10 days)

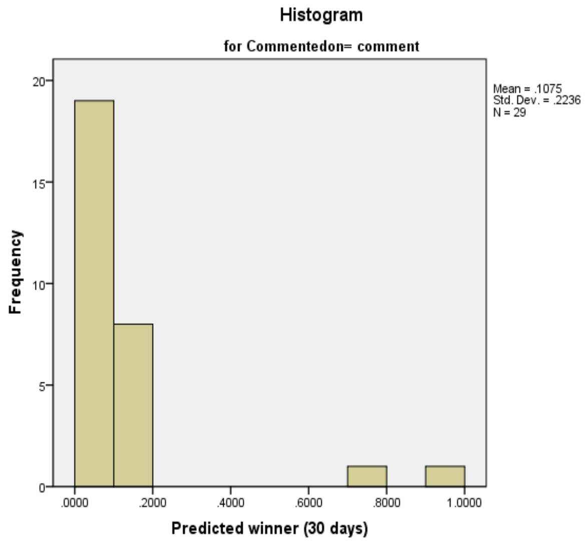

Normal Q-Q Plot of Predicted winner (10 days)
for Commentedon= no comment
Expected Normal
Observed Value

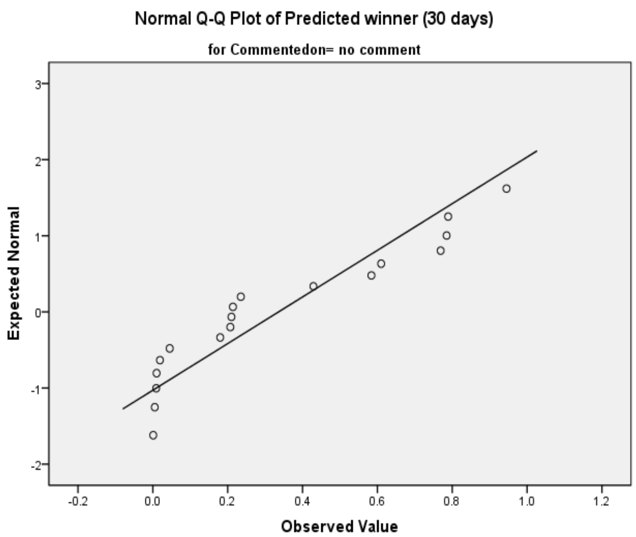

Detrended Normal Q-Q Plot of Predicted winner (10 days)
for Commentedon= comment
Dev from Normal
Observed Value

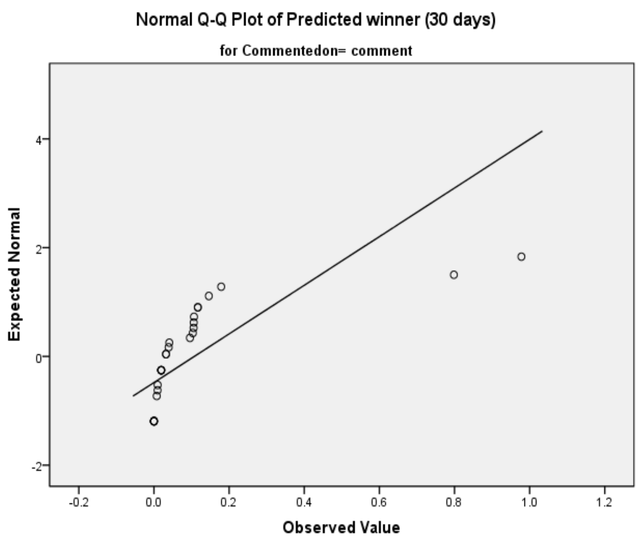

Detrended Normal Q-Q Plot of Predicted winner (10 days)
for Commentedon= no comment
Dev from Normal
Observed Value

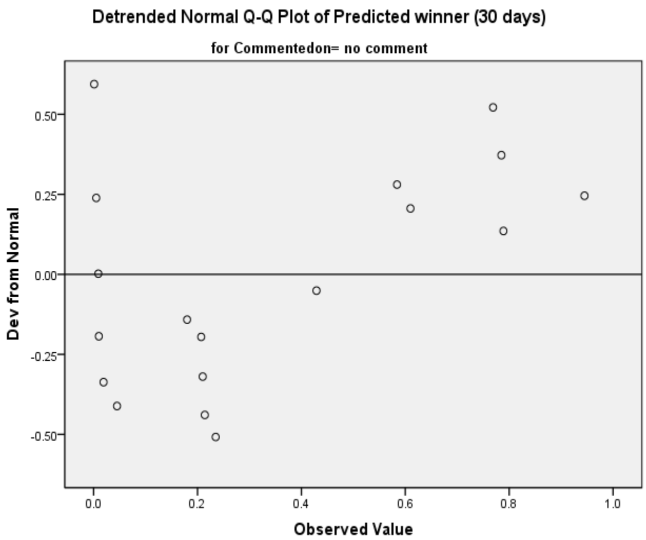

Normal Q-Q Plot of Predicted winner (10 days)
for Commentedon= comment
Expected Normal
Observed Value

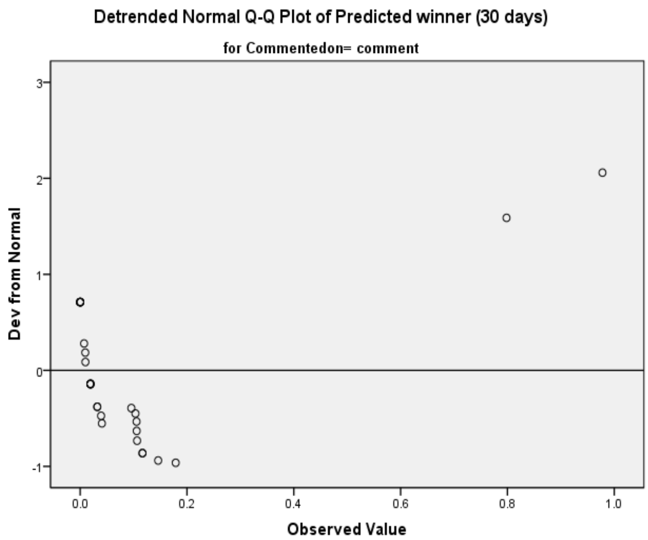

Histogram
for Commentedon= no comment
Mean = .3359
Std. Dev. = .3265
N = 18
Frequency
Predicted winner (30 days)

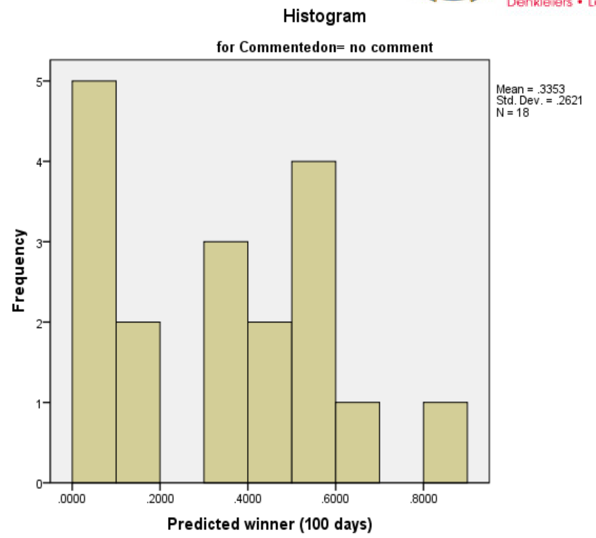

Histogram
for Commentedon= comment
Mean = .1075
Std. Dev. = .2236
N = 29
Frequency
Predicted winner (30 days)

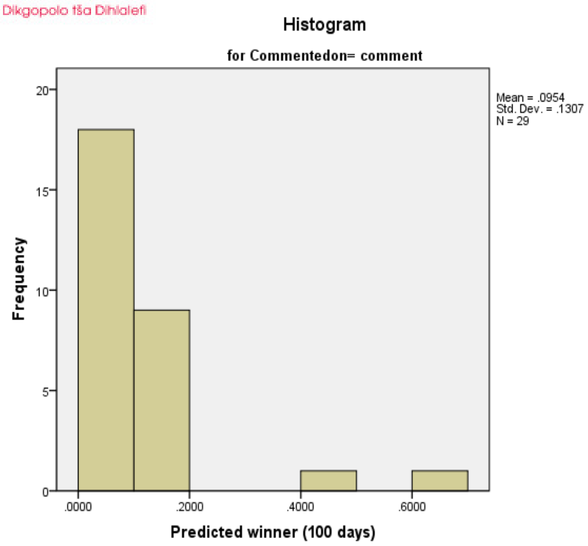

Normal Q-Q Plot of Predicted winner (30 days)
for Commentedon= no comment
Expected Normal
Observed Value

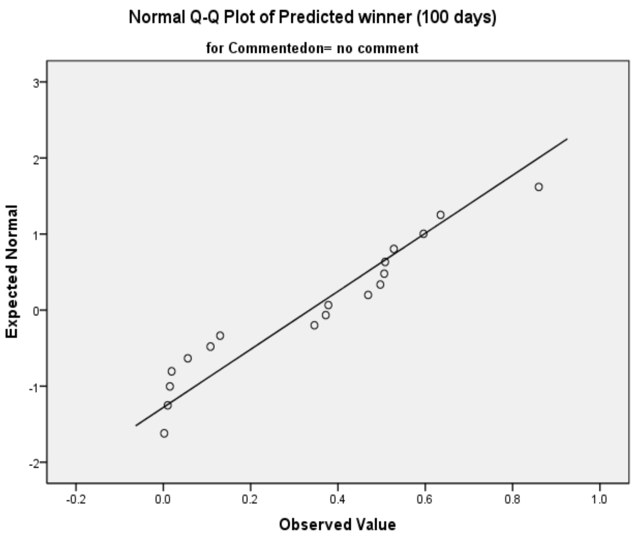

Normal Q-Q Plot of Predicted winner (30 days)
for Commentedon= comment
Expected Normal
Observed Value

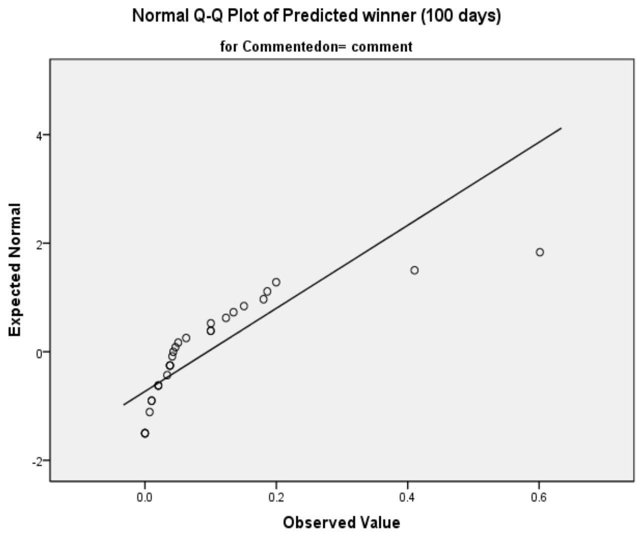

Detrended Normal Q-Q Plot of Predicted winner (30 days)
for Commentedon= no comment
Dev from Normal
Observed Value

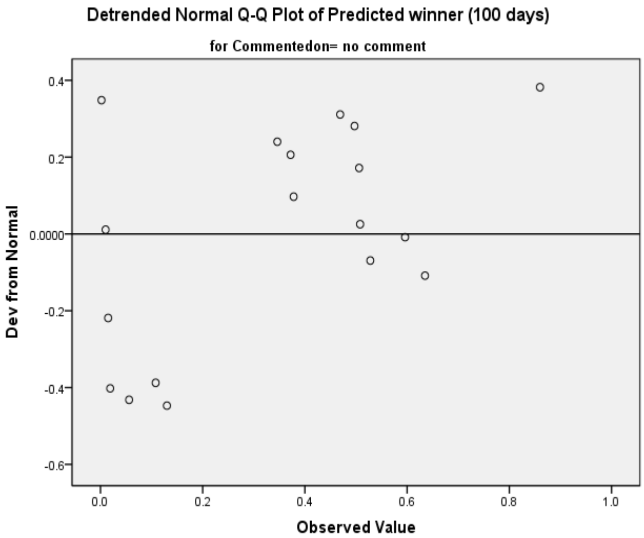

Detrended Normal Q-Q Plot of Predicted winner (30 days)
for Commentedon= comment
Dev from Normal
Observed Value

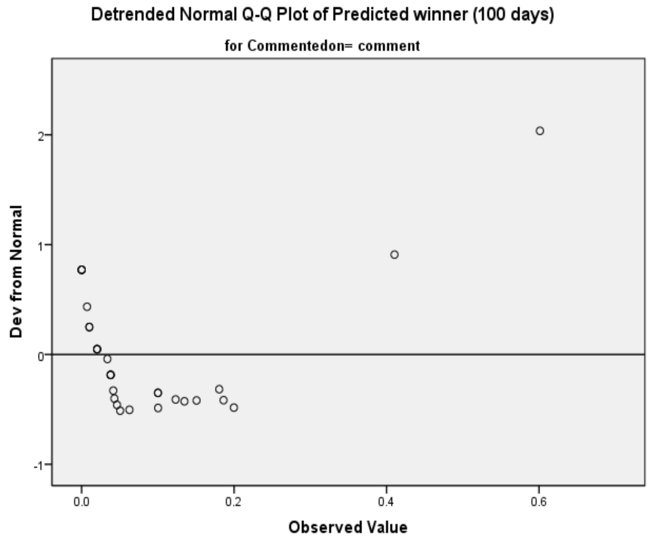

Histogram
for Commentedon= no comment
Mean = .3353
Std. Dev. = .2621
N = 18
Frequency
Predicted winner (100 days)


Histogram
for Commentedon= comment
Mean = .0954
Std. Dev. = .1307
N = 29
Frequency
Predicted winner (100 days)

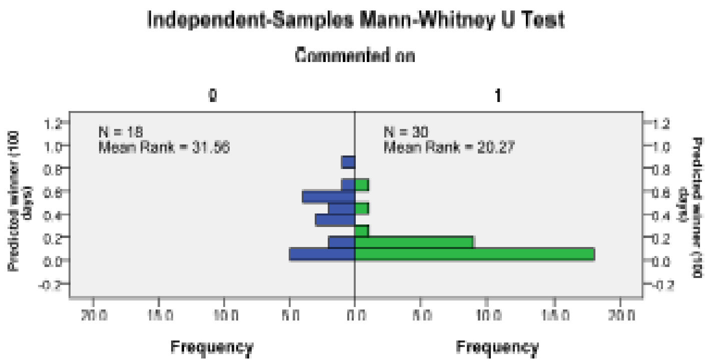

Normal Q-Q Plot of Predicted winner (100 days)
for Commentedon= no comment
Expected Normal
Observed Value


Normal Q-Q Plot of Predicted winner (100 days)
for Commentedon= comment
Expected Normal
Observed Value

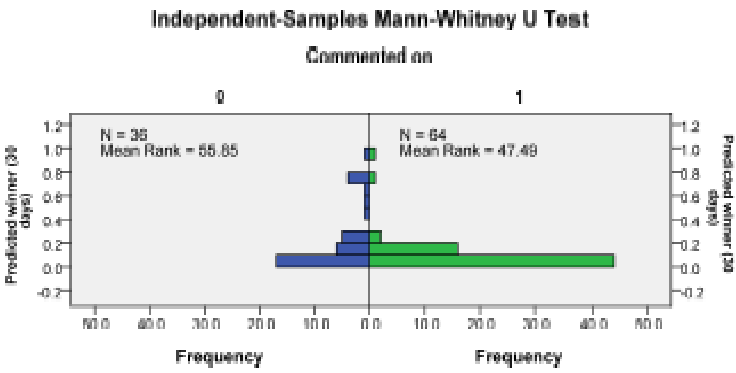

Detrended Normal Q-Q Plot of Predicted winner (100 days)
for Commentedon= no comment
Dev from Normal
Observed Value


Detrended Normal Q-Q Plot of Predicted winner (100 days)
for Commentedon= comment
Dev from Normal
Observed Value

Appendix III – Mann-Whitney U tests



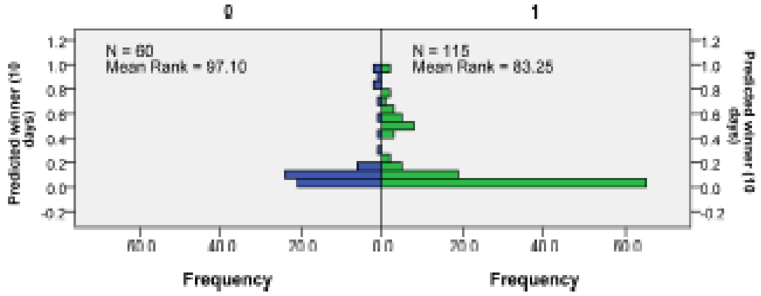
Total N	48
Mann-Whitney U	143.000
Wilcoxon W	608.000
Test Statistic	143.000
Standard Error	46.940
Standardized Test Statistic	-2.706
Asymptotic Sig. (2-sided test)	.007



Total N	100
Mann-Whitney U	959.500
Wilcoxon W	3,039.500
Test Statistic	959.500
Standard Error	139.208
Standardized Test Statistic	-1.383
Asymptotic Sig. (2-sided test)	.167

Independent-Samples Mann-Whitney...

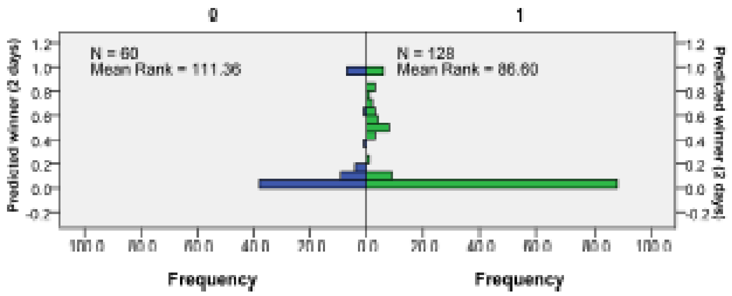
Commented on



Total N	175
Mann-Whitney U	2,904.000
Wilcoxon W	9,574.000
Test Statistic	2,904.000
Standard Error	318.047
Standardized Test Statistic	-1.717
Asymptotic Sig. (2-sided test)	.086

Independent-Samples Mann-Whitney...

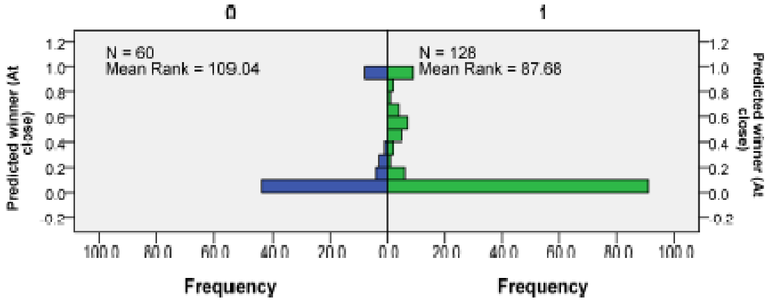
Commented on



Total N	188
Mann-Whitney U	2,828.500
Wilcoxon W	11,084.500
Test Statistic	2,828.500
Standard Error	346.819
Standardized Test Statistic	-2.917
Asymptotic Sig. (2-sided test)	.004

Independent-Samples Mann-Whitney...

Commented on



Total N	188
Mann-Whitney U	2,967.500
Wilcoxon W	11,223.500
Test Statistic	2,967.500
Standard Error	347.180
Standardized Test Statistic	-2.513
Asymptotic Sig. (2-sided test)	.012