An Explorative Study of Applying Prediction Markets in Course Evaluation

University of Oulu
Department of Information Processing
Science
Master’s Thesis
Kang Wang
Date 20.5.2013
Abstract

Prediction Markets is a new emerging efficient method for predicting the likelihood of uncertain future events, and also it has been used successfully in various fields. Compare to traditional forecasting methods, it can provide more accurate results and reflect continuous real-time information. In this research, we focus on its application in course evaluation.

This research includes a Prediction Markets-based systematic literature review, in which there are 120 articles included. They are analysed based on four main categories: description, theoretical work, organization management, and applications. The results demonstrate that few of research take Prediction Markets into consideration as an evaluation system and it has not been used for course evaluation purpose.

Course evaluation in universities is quite meaningful. Traditional course evaluation methods have been proved to be low response rate, time consuming, inaccurate data and lacking interactions between teachers and students. In this research, the Prediction Markets-based course evaluation system is designed as a new solution of evaluating the course quality. We present an explorative experiment which is conducted in a real course environment with 49 students involved. The students are divided into two groups: Prediction Markets-based group and Traditional group, with the purpose of comparing the difference between Prediction Markets-based course evaluation system and traditional in use course evaluation system.

In conclusion, the study confirms the feasibility of extending Prediction Markets in course evaluation use. Additionally, a possible solution of designing and implementing Prediction Markets-based course evaluation system is presented. Further studies should focus on the issues of optimizing the Prediction Markets-based course evaluation system.

Keywords
Prediction Markets, Course Evaluation, Students Feedback
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMM</td>
<td>Automated Market Maker</td>
</tr>
<tr>
<td>CA</td>
<td>Call auctions</td>
</tr>
<tr>
<td>CDA</td>
<td>Continuous double auctions</td>
</tr>
<tr>
<td>DA</td>
<td>Double auctions</td>
</tr>
<tr>
<td>DPM</td>
<td>Dynamic pari-mutuel market</td>
</tr>
<tr>
<td>GQM</td>
<td>Goal question metric method</td>
</tr>
<tr>
<td>HSX</td>
<td>Hollywood Stock Exchange</td>
</tr>
<tr>
<td>IEM</td>
<td>Iowa Electronic Markets</td>
</tr>
<tr>
<td>IT</td>
<td>Information technology</td>
</tr>
<tr>
<td>LMSR</td>
<td>Logarithmic market scoring rules</td>
</tr>
<tr>
<td>PMCES</td>
<td>Prediction Markets-based Course Evaluation System</td>
</tr>
<tr>
<td>UBC-ESM</td>
<td>UBC Election Stock Market</td>
</tr>
</tbody>
</table>
Foreword

During the thesis studying, I have learnt many things that they are not only the knowledge relating to the research topic, but also including many aspects of how to do a better research; for instance, I may rich the experience of holding a small experiment, managing a questionnaire survey and making research interviews. In addition, understanding a little bit more of the life living wisdom can be recognised as extra obtains.

I sincerely thank my supervisor Jouni Markkula and Muhammad Ovais Ahmad very much, and all the people who take part in my thesis work. Without their supports the thesis would not have been done. Especially, I have received much good guidance and significant helps from my supervisor. I owe many thanks to him.

Last but not least, I cherish the opportunity of study abroad and always holding gratitude for University of Oulu. Of course, my parents have been encouraging me and supporting my study, many thanks for them as well.

Kang Wang

Oulu, May 20, 2013
Contents

Abstract .................................................................................................................................................. 2
Abbreviations ......................................................................................................................................... 3
Foreword ................................................................................................................................................ 4
Contents ................................................................. 5
1. Introduction .............................................................. 7
2. Research problem and method ............................................................ 10
   2.1 Research problem .............................................................. 10
   2.2 Research strategy .............................................................. 10
      2.2.1 Systematic literature review .................................................. 13
      2.2.2 System development .......................................................... 13
      2.2.3 Experiment ................................................................. 14
      2.2.4 Interviews .................................................................. 14
3. Systematic literature review .......................................................... 15
   3.1 Background and related work .................................................. 15
   3.2 Classification of Prediction Markets ........................................ 15
   3.3 Review questions ................................................................. 17
   3.4 Methods ........................................................................ 17
      3.4.1 Inclusion and exclusion ..................................................... 18
      3.4.2 Search strategy .............................................................. 18
   3.5 Results ................................................................... 21
      3.5.1 Overview of results ......................................................... 21
      3.5.2 Analysis ................................................................. 21
   3.6 Relationship of results between systematic review and thesis work .... 25
4. Fundamentals of Prediction Markets .................................................. 26
   4.1 How does it work? ............................................................... 26
   4.2 Types of Prediction Markets ................................................ 27
   4.3 Trading mechanisms ......................................................... 27
   4.4 Does money matter? ........................................................... 29
   4.5 Why do Prediction Markets work so well? ............................... 29
   4.6 Previous usage of Prediction Markets ..................................... 30
      4.6.1 Iowa Electronic Markets .................................................. 30
      4.6.2 Prediction Markets for sports game ..................................... 30
      4.6.3 Hollywood Stock Exchange ............................................. 31
      4.6.4 Other applications ....................................................... 32
5 Course evaluation system in universities ........................................... 33
   5.1 Definition of course evaluation .............................................. 33
   5.2 Paper-based VS online-based course evaluation method ......... 34
   5.3 Current course evaluation in University of Oulu ............... 34
6 An explorative experiment of Prediction Markets-based course evaluation system ............................................ 36
   6.1 Design the Prediction Markets-based course evaluation system .... 36
      6.1.1 Contracts ................................................................. 36
      6.1.2 Trading mechanism ...................................................... 37
      6.1.3 Incentives ............................................................... 37
      6.1.4 Traders ................................................................. 38
      6.1.5 Trading platform ....................................................... 38
1. Introduction

Prediction Markets (also can be called as ‘Information Markets’ or ‘Decision Markets’ or ‘Virtual Markets’) is a new emerging forecasting mechanism which is able to collect various dispersed information efficiently (Tziralis & Tatsiopoulos, 2007). Since it was successfully implemented by University of Iowa as Iowa Electronic Market (IEM) in the year 1988, which is basically designed to forecast the results of US president elections (Boyle & Videbeck, 2005), many researchers were putting their focus on Prediction Markets area. For instance in the political election market (e.g. IEM), by transforming the results of future president to ‘contracts’ in the markets, people can bid (buy) or ask (sale) the contracts based on their expectations at a particular price. Thus the trading price can be reflected as the likelihood of becoming US president for each candidate. For example, in past 2012 US President Election between Barack Obama and Mitt Romney, the victory of Obama had been forecasted by Prediction Markets such as IEM (Electronic Market Predicts Presidential Election, 2012), and Intrade (Intrade Markets, n.d.). Besides of providing more accurate forecasting results, Prediction Markets has the features of reflecting real-time information (Berg, Nelson, & Rietz, 2003) and adopting the idea of “wisdom from crowds” (L.V. Williams & J.V. Williams, 2011). Actually, many existing running Prediction Markets have proved that it has a very powerful potential expansibility in wide range of areas, including politics (Forsythe, Rietz, & Ross, 1999), sports competitions (Debnath, Pennock, Giles, & Lawrence, 2003), entertainment (Gruca, Berg, & Cipriano, 2003), and also management purposes (Cowgill, Wolfers, & Zitzewitz, 2008).

To summarize the previous research on Prediction Markets, most of them studied the usability of Prediction Markets-based applications. They were trying to verify whether forecasted results are as accurate as when the events really happen, or testing the relevant theories which could be used to provide better predictions in Prediction Markets-based applications. For example, UBC Election Stock Market (UBC-ESM) which was created to forecast the results of the 1993 Canadian federal election (Forsythe, Frank, Krishnamurthy, & Ross, 1995), and Euro 2000 soccer championship experiment (Schmidt & Werwatz, 2002) verified that the market is more accurate than random predictor. Besides in Tziralis and Tatsiopoulos’s research (2006), they asserted that Prediction Markets could be recognized as an efficient information aggregation tool. Moreover, the rich literature and large achievements from Prediction Markets studies, they can be categorised into different topics. Some of them provide theoretical supports of markets modelling (Hanson, 2003; Luckner, 2008; Servan-Schreiber, Wolfers, Pennock, & Galebach, 2004). Some of them were trying to develop new Prediction Markets-based applications (Beckmann & Werding, 1996; Berlemann & Nelson, 2002; Gruca, 2000). For instance, in the research of Rajakovich and Vladimirov (2009), they extended the utilization of Prediction Markets into a healthcare environment in which a Prediction Markets was particularly built for forecasting the demands of people for the hospital, and Michael Berlemann built a prototype experiment to forecast inflations (Berlemann & Nelson, 2002). In addition, there is a published journal (The Journal of Prediction Markets) which is specially conducted with interests on relevant topics of Prediction Markets.

This research is put forward because of realizing the advantages of Prediction Markets. Many previous usages of Prediction Markets applications demonstrated that Prediction
Markets can be recognized as a good forecasting tool for predicting outcome of future events (Chen & Plott, 2002), or a management tool for projects (Ortner, 1998), or a part of decision making system (Berg & Rietz, 2003). People have captured its prediction function and developed well, but its dynamic nature has been somehow ignored. Tziralis and Tatsiopoulou (2006) indicated that Prediction Markets can continuously reflecting the possible results in high accuracy through its innate ability of continuously responding to new information. In this study, we are not only care about the forecasts from Prediction Markets, but also focus on its internal procedures. The whole procedures include all actions from traders who participate in the markets. Their buying/selling transactions, the changes of contract prices, discussions/comments and every specific changes should be recognized as data for analyzing.

Today, course evaluation is playing more and more important role in universities. Previous studies have confirmed that ratings from students are valid, reliable, and efficient way of evaluating teaching quality (Centra, 1977). The feedback from students is significant for improving teaching performance (Anderson, Cain, & Bird, 2005; Tiberius et al., 1989). Many universities have been asking students to take course evaluations during or after the course, in order to aggregate their opinions. As an information aggregation tool, Prediction Markets may have the possibility to be adopted in course evaluation.

Nowadays, almost all universities probably have implemented their own course evaluation system. Although, they may set up with different kinds of system, the main processes should be concluded in two ways: one is traditional paper based process, the other one is recognized as technology based online process (Anderson et al., 2005). Whether simply choose one of them or use a combination of both procedures, the course evaluation seems like cannot reach the original expectations. Typically, course evaluation was designed to get most valuable and accurate feelings from students in order to provide reliable data of each specific course to relevant teachers and macro point of view to school leaders. But low response rate, time consuming, inaccurate data and lacking interaction between teachers and students are summarised as obvious problems (Anderson et al., 2005; Tucker, Jones, Straker, & Cole, 2003). To build an efficient method for collecting students’ feedback is one of the most important goals that teachers and managers in universities are looking for. Thus, our hypothesis of Prediction Markets-based course evaluation system may become as the one remedy solution of doing course evaluation.

In addition, although many successful Prediction Markets applications have been known, such as Iowa Electronic Markets, Hollywood Stock Exchange, Tradesports and including lots of private markets which were adopted in organizations (Chen & Plott, 2002; Passmore, Cebeci, & Baker, 2005; Rajakovich & Vladimirov, 2009), but very few of them take Prediction Markets into consideration as an evaluation system and even none of them is used for course evaluation.

The purpose of this research is to extend the utilization of Prediction Markets as an evaluation tool, which may contributes to relative research of Prediction Markets-based applications. In the thesis, the idea of using Prediction Markets to develop a new method of organizing course evaluation has been put forward. The Prediction Markets-based course evaluation system can be a good way of revealing students’ feedback, and providing real-time reflections in order to let teachers to improve the teaching quality immediately. Meanwhile, students would also benefit from their own feedback.
Therefore, the objective of this thesis is to seek the possible solution for adopting Prediction Markets in course evaluation. In practice, we begin with a systematic literature review which analyzes the Prediction Markets studies based on four categories: description, theoretical work, organization management and applications. Then from the results of systematic literature review, the fundamentals of Prediction Markets have been studied. Based on the knowledge of key design elements of Prediction Markets (Luckner et al., 2011, p. 11), we design a Prediction Markets-based course evaluation system. Next, an explorative experiment in a real course environment is organized. Finally, we can study the feasibility and implementation issues of this new course evaluation system.
2. Research problem and method

The concept of Prediction Markets is recognized as a method that can help to forecast future events with lower prediction errors and higher accuracy results compared to the conventional forecasting methods (Arrow et al., 2008). It has been well known in economic research, especially its success in political election experiences (Berg, Forsythe, Nelson, & Rietz, 2001). The main goal of this research is to verify the possibilities of adopting Prediction Markets concept for the course evaluation purpose. Obviously, there is a mixed research area which combines with the pedagogy, economics, and information technology. In the following, the research question is defined and used method is presented.

2.1 Research problem

The research is not only to fix the problem from traditional course evaluation system, but also to extend the usage of Prediction Markets in order to contribute to the relevant research area. The main research topic is defined by answering the following questions:

*How to design and implement an efficient Prediction Markets-based course evaluation system?*

This main purpose can be divided into the following three sub-questions:

1. Can the Prediction Markets be adopted for course evaluation needs?
2. What are the requirements for developing a Prediction Markets-based application?
3. What are the potential improvements from Prediction Markets-based course evaluation system comparing to the traditional course evaluation method?

In order to answer the questions above, separately understanding course evaluation, Prediction Markets, Information System are not enough, they have to be linked up to consider. Hence, after making conceptual study and analysis of Prediction Markets and course evaluation, the first sub-question will have its answers. To answer the second sub-question, the Prediction Markets-centered systematic literature review has been conducted. At last, an explorative experiment is organized in order to observe the difference between Prediction Markets-based course evaluation system and traditional course evaluation method. The research method is described in detail in the following section.

2.2 Research strategy

This study generally follows the multimethodological approach method (see Figure 1) which developed by Nunamaker, Chen, and Purdin (1990). They argued that a research methodology is a combination of processes, methods, and tools. These are used to support the construction of research in a research domain. Four main stages are covered: *theory building, observation, experimentation, and system development*. The used research strategy is going to be described in detail based on these four stages.
Theory building represents the new ideas and concept (or the construction of the conceptual framework). Normally, the outcomes of theory building will display the limitation of relevant research domain and construct basic knowledge for coming research (Nunamaker et al., 1990). In the thesis, the literature analysis of both Prediction Markets and course evaluation is made. Particularly, a systematic literature review of Prediction Markets follows the guidelines of performing systematic literature review in Software Engineering (Kitchenham, 2007). The theory building stage of this research is corresponding with the research processes of constructing a conceptual framework and developing system architecture (see Table 1).

System Development is the hub of research which communicates with other research methodologies in order to integrate specific research program (Nunamaker et al., 1990). In this thesis, we build a Prediction Markets-course evaluation system based on the understandings of key design elements of Prediction Markets and course evaluation requirements. They are included in the research process “analyse and design the system” and “build the system” (see Table 1). In order to answer the research question, we present a possible guideline for designing and implementing a Prediction Markets-based course evaluation system.

Experimentation is depicted as the procedure which straddles the gulf between theory building and observation. Within the experiment, the concerned underlying theory (looking back to the research itself) or the issues of future development and technology

**Figure 1.** A multimethodological approach to IS research. (Nunamaker, Chen, & Purdin, 1990)
transfer (looking forward to the relevant research) will be validated. Results from the experiments may provide supports for refining the theories and improving the system (Nunamaker et al., 1990). In this study, the explorative experiment of adopting Prediction Markets into course evaluation system is defined. We test the feasibility of Prediction Markets-based course evaluation system in a real course.

Observation may in relation to case studies, field studies, and sample surveys to help to formulate particular conjunctions which are supposed to be tested by experiments (Nunamaker et al., 1990). Observations have been included many times during the whole research processing. Especially in the experiment, we arrange the tutoring presentation at the beginning, questionnaire survey in the middle, and afterwards interview session. These activities have provided much valuable information for optimizing the Prediction Markets-based course evaluation system and also have contributed potential questions for the future research.

According to the “A Process for System Development Research” (Nunamaker et al., 1990), authors depict five main stages: (1) construct a conceptual framework, (2) develop system architecture, (3) analyse and design the system, (4) build the system, (5) observe and evaluate the system. In this thesis, we have adopted the five main stages. The whole study processes are outlined in Table 1. Each stage has been allocated with sub-research issues which are listed in the right column.

**Table 1.** Research process adopted from the process for systems development research (Nunamaker et al., 1990).

<table>
<thead>
<tr>
<th>System Development Research Process</th>
<th>Research Issues</th>
</tr>
</thead>
</table>
| Construct a conceptual framework    | • Analyse the current situations of course evaluation system in the university.  
• Study the knowledge of Prediction Markets via making a systematic literature review.  
• Study course evaluation from literature as groundwork.  
• Adopt Prediction Markets as a solution of course evaluation. |
| Develop a system architecture      | • Study the literature and conclude key design elements of Prediction Markets.  
• Develop a conceptual Prediction Markets-based course evaluation system. |
| Analyse and design the system      | • Analyse the various existing Prediction Markets applications.  
• Design the system based on the course evaluation requirements. |
| Build the system                   | • Select Inking Markets (Inkling Incorporated, n.d.) as the platform.  
• Build a suitable Prediction Markets-based service for implementing in the department environment. |
| Observe and evaluate the system    | • Evaluate the system by a Prediction Markets explorative experiment.  
• Organize the questionnaire survey.  
• Interview participant. |

Note: Here, due to the specific research topic in the thesis, the research process (Nunamaker et al., 1990) is adopted in a general way.

By following the research processes and issues which are presented in the above table, this thesis is organized systematically and reported in systemic structure. At the first stage, we identify an idea of adopting Prediction Markets in course evaluation. After we collect
the knowledge of Prediction Markets and course evaluation as groundwork for thesis and understand the current situation of course evaluation system in the university, a conceptual Prediction Markets-based course evaluation system is developed. Later, in the *analyse & design the system* step, various existing applications and the key design elements of Prediction Markets have been analysed in order to build a suitable Prediction Markets-based service based on the requirements of course evaluation system. Finally, the feasibility of Prediction Markets-based course evaluation system is verified by a Prediction Markets-based explorative experiment. In addition, the questionnaire survey and interviews session are arranged at the end of the experiment.

### 2.2.1 Systematic literature review

Making a systematic literature review can be recognized as a combination procedure of summarising the existing evidences relating to specific research questions, identifying the gap from current research and providing the background/framework for supporting already processing research active. Compare to the literature review, systematic review is a more thorough and fair review which can provide higher scientific value not only for recent doing research but also benefit the potential future research. It must form up based on defined research strategy which consists with approval research question and included explicit inclusion and exclusion criteria (Kitchenham, 2007).

As a part of thesis, a systematic literature review for Prediction Markets has been put forward. Although, it does not follow the guidelines rigorously, but the outcomes are based on systematic searching strategy and systematic analyse. The three main procedures of doing systematic review are: *planning the review, conducting the review and reporting the review* (Kitchenham, 2007). In practice, Firstly, we identify the review questions by following the procedure of Goal Question Metric (GQM) method (Basili, Caldiera, & Rombach, 1994). Then, according to the inclusion and exclusion criteria, we select articles from predefined databases and journal. At last, a quantitative analysis based on the categories of Prediction Markets has been listed as the outcomes.

By making a systematic literature review based on Prediction Markets, it provides ample information of Prediction Markets including what are the key design elements of Prediction Markets, how is the previous experiments or applications (Berlemann & Nelson, 2002), and why Prediction Markets can make better forecasting (Gangur & Martinčík, 2011). In addition, previous experiences of Prediction Markets-based adoptions will contribute to the implementation of the explorative experiment of Prediction Markets-based course evaluation system.

### 2.2.2 System development

Based on the results from literature review, the relevant theories, practical ideas, methods as well as the various solutions of Prediction Markets are gathered for serving this research. The development of conceptual system is a process of constructing synthesis on Prediction Markets and course evaluation. By combining Prediction Markets with course evaluation, we design Prediction Markets-based course evaluation system in order to make students to evaluate the course quality in a new way. The system is realized by using an existing online service which is provided by Inkling Markets (Inkling Incorporated, n.d.).

Due to the thesis topic is to answer the question of how to design and implement a Prediction Markets-based course evaluation system, we present a possible guideline as
one of the outcomes in this thesis. It can provide supports for university stakeholders or teachers to adopt Prediction Markets-based course evaluation system in the future.

2.2.3 Experiment

The main task of this research is an explorative experiment of Prediction Markets-based course evaluation system. Considering, it is known that many comprehensive course evaluation systems have been adopted in different level of educational institutions. This research is restricted in department level, due to a Prediction Markets-based explorative experiment is held based on an existing real course, and all of the relevant information are collected and analysed in the environment of information processing science department in University of Oulu.

The purpose of the experiment is to test the feasibility of adopting Prediction Markets in course evaluation field. The Prediction Markets-based course evaluation system means to be a new course evaluation system, which will be tested on a real course. As a part of preparation work, the Prediction Markets-based course evaluation system is supported by a customized web based service. It is defined with proper settings to fit the experiment environment and research purpose. There are two groups of students included in the experiment. In order to reduce the effect of experiment bias, participant students are randomly allocated to Prediction Markets-based group and Traditional group. Each group is provided different beforehand training session. At the end of the experiment, we compare the performance and analyse the results from two groups. After analysing the experiment, the results will contribute to Prediction Markets research field with its extension adoption at course evaluation.

2.2.4 Interviews

There are two interview sessions during the research. They are arranged for different purposes by different questions. The interviews are conducted by face-to-face setting and recorded by voice recorder after asking the permission of interviewees. Time for each interview is around half an hour.

The first interview session is conducted to gain the situation of current course evaluation system in University of Oulu. We invite Chief of Academic Officer, recent and previous Deputy Director in department to join the interview, because they are the staffs in university who familiar with existing course evaluation system and respond for education management. After the interview, the realized situation of current running system from interviews is regarded as the motivation of selecting course evaluation area as the research goal. The interview question for staffs is attached in Appendix C.

Another round of interview is arranged for collecting the feedbacks of Prediction Markets-based course evaluation system. The interviewees are selected from two groups of students who participate in the experiment. They are consisted with the students who perform well in the markets from Prediction Markets-based group and the students who give interesting answers in the questionnaire survey from Traditional group. The results from the interviews have the functions on pointing out the problems of the Prediction Markets-based course evaluation system and producing interesting topics for further research. The interview question for experiment is attached in Appendix D.
3. Systematic literature review

In order to clearly understand the definition of Prediction Markets, the systematic literature review is conducted by following Kitchenham’s (2007) guidelines for performing systematic literature review. This systematic literature review does not rigorously follow the guidelines, but all the defined procedures are designed in systematic way as well as references are systematically searched.

The purpose of this systematic literature review was:

- By reviewing the existing literature relating to Prediction Markets concept, the knowledge of Prediction Markets can be gained systematically for clearly understanding. For example, the fundamentals of Prediction Markets including trading mechanism, incentives design, empirical evidence as well as its benefits and limitations.
- By summarizing the existing Prediction Markets-based applications and experiments including public- or private- markets, play-money or real-money based markets and also successful or failed markets, the experience of designing and implementing Prediction Markets-based applications can be studied.
- By completing the systematic literature review, it can be used for others and also give the opportunities for doing future relevant research.

3.1 Background and related work

Prediction Markets is an excited research area, partly because of it is a novel way of forecasting future events, and also adopted underlying the idea of “wisdom from crowds” (L.V. Williams & J.V. Williams, 2011). Due to there is no universally defined terminology for Prediction Markets, sometimes it referred to “Information Markets”, “Decision Markets”, “Electronic Markets”, “Virtual Markets” (Tziralis & Tatsiopoulos, 2007) or “Idea futures markets” (Passmore, Cebeci, & Baker, 2003).

Here, we prefer to use the definition of Prediction Markets from Arrow et al. (2008). They indicated Prediction Markets is used for yielding payments by trading contracts which is designed for forecasting the outcomes from uncertain future events. It is a new forecasting method with less error than traditional forecasting way (Arrow et al., 2008). In the markets, contracts will be paid off depend on the outcomes of uncertain future events. All transactions (buy or sell) from traders are based on their beliefs on the events. According to the efficient market theory (Malkiel & Fama, 2012), the prices of contracts reflect the aggregation of all available information and traders’ expectations on the coming uncertain events. It can somehow represent as the predictions for the future events.

3.2 Classification of Prediction Markets

Tziralis and Tatsiopoulos (2007) classified the topics belong to Prediction Markets into four basic categories; they are description, theoretical work, applications, and law and policy. In order to provide better support to this thesis topic, besides adopting the four categories, there was a new category added. It was named organization management (see Figure 2). This is because in fact there is a large part of Prediction Markets research
focusing on developing the applications which regard to organization management, and in addition course evaluation could be recognized as a part of management for university/school.

Figure 2. Classification of topics in PM literature.

1. **Description** – This category is built for collecting the papers which are regard as giving introduction for Prediction Markets research, and fundamentals description for beginners. It includes the subcategories of introduction, general description, and other descriptive issues. (Tziralis & Tatsiopoulos, 2007.)

2. **Theoretical Work** – The category is designed for gathering theoretical studies of Prediction Markets. They are summarised to subcategories of market modelling, information aggregation convergence and equilibrium, and other theoretical issues. (Tziralis & Tatsiopoulos, 2007.)

3. **Applications** – The literature in this category are the descriptions or comparisons of those Prediction Markets based applications or experiments. It contains experiments (academic or other environment), political stock markets, markets on sport events, and other applications. (Tziralis & Tatsiopoulos, 2007.)

4. **Law and Policy** – This category includes the issues law and policy research. They are legality and regulation of Prediction Markets, the analysis on government policy announcements, and other relevant issues. (Tziralis & Tatsiopoulos, 2007.)

5. **Organization Management** – This last category is designed for collecting the issues which are focusing on the benefits from Prediction Markets to improve the management in organizations.

The classifications of Prediction Markets topics were used to guide systematic literature review progress. Which means it was regarded as the key words for building the search strings in defined databases and the results were analysed by order of categories.
3.3 Review questions

Specifying the research questions is a most crucial part of doing systematic literature review (Kitchenham, 2007). In this systematic review, Goal Question Metric (GQM) method (Basili, Caldiera, & Rombach, 1994) had been applied. It includes five elements. They are *analyse* (the object under measurement), *purpose* (the reason why choose the object, e.g. to understand, to control, or to improve), *with respect to* (the quality focus of the object that the measurement focus on), *from the view point of* (the people’s view on the object measurement) and *context* (environment settings).

Table 2. First review questions underlying goal definition of GQM.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>For the purpose of (the reason why choose the object, e.g. to understand, to control, or to improve)</td>
<td>To understand</td>
</tr>
<tr>
<td>with respect to (the quality focus of the object that the measurement focus on)</td>
<td>Fundamentals of Prediction Markets, classification of Prediction Markets, as well as key design elements of Prediction Markets including contracts type, trading mechanism, incentive.</td>
</tr>
<tr>
<td>from the view point of (the people’s view on the object measurement)</td>
<td>System Designers, Economists</td>
</tr>
<tr>
<td>context (environment settings)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Second review question underlying goal definition of GQM.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>For the purpose of (the reason why choose the object, e.g. to understand, to control, or to improve)</td>
<td>To understand</td>
</tr>
<tr>
<td>with respect to (the quality focus of the object that the measurement focus on)</td>
<td>Previous Prediction Markets-based applications or experiments</td>
</tr>
<tr>
<td>from the view point of (the people’s view on the object measurement)</td>
<td>System analyser, users</td>
</tr>
<tr>
<td>context (environment settings)</td>
<td></td>
</tr>
</tbody>
</table>

Based on the above analysis which adopted Goal Question Metric method, the review questions would be defined as follow:

1. What are the key design elements for Prediction Markets based on the views from system designers? (see Tables 2)
2. What kinds of Prediction Markets-based applications have been adopted? (see Table 3)

3.4 Methods

This systematic literature review was totally adopted under predefined rules. They were database selection, inclusion and exclusion criteria setting, and scientific search strategy.
3.4.1 Inclusion and exclusion

Inclusion and exclusion criteria were predefined as follow which are based on the previous classification results. They would be used to identify the relevant papers which are relating to the review questions.

Inclusion criteria:

- Materials searching that based on at least one key word among prediction markets, information markets, decision markets, virtual markets, and electronic markets.
- Studies that could be recognized as any categories definitions which were showed in Figure 2. Such as descriptions on Prediction Markets and its related theories, Prediction Markets-based experiments and applications or information aggregation aspects.
- Studies that had any relations with the aspect of organization management.
- Studies were under covered from 1988 to 2012 and written in English.

Exclusion criteria:

- Studies were only focusing on the economic and finance aspects.
- Studies were just doing professional research for law and policy aspects.

3.4.2 Search strategy

Articles were selected in three phases, as it showed in Figure 3. Search strategy was comprised of manual scanning the Journal of Prediction Markets, searching electronic databases, and reviewing references lists of articles which were selected from previous two phases.

![Search strategy](image)

**Figure 3.** Three phases search strategy.
In the Phase 1, the articles in Journal of Prediction Markets from year 2007 to 2012 were completely reviewed. This was because the publications from the internet were available at the period. The Journal of Prediction Markets is an academic and peer reviewed journal which published by University of Buckingham, it is aimed at academics including business and economics and all those with an interest in the operation of markets and market efficiency (The Journal of Prediction Markets, n.d.). Furthermore, after making a completive review in phase 1, it was also help to identify the searching key words for forming up search strings of Phase 2.

In order to make an exhaustive systematic literature review, Phase 2 was designed for searching large amount of relevant literature. Three electronic databases were selected, because they are defined as key databases of computer science or business.

- Academic Search Premier (EBSCO)
- Computer and Information Systems Abstracts (ProQuest)
- Business databases (ProQuest)

Phase 2 was designed with four steps, and the number of selected articles for each step is presented in the right column boxes (Figure 4). In Step 1, the results from databases search were gathered as initial results. Then we removed the duplicates articles in Step 2. In Step 3, the articles were read through by titles, and those obviously inappropriate articles were excluded. In step 4, the rest of articles were scanned by abstracts. Each step followed the defined inclusion and exclusion criteria. If the studies whose titles and abstracts do not relate to any inclusion criteria (or belong to exclusion area), they were excluded. All the articles were searched at the period from 1988-2012, this was because Iowa Electronic Markets was launched at 1988 (Boyle & Videbeck, 2005), and then it provided big influence for the later research. In addition, the searches were scholarly limited to peer reviewed journals.
Based on the processes of forming up review questions by Goal Question Metric and studying classification of Prediction Markets, the key words relating to Prediction Markets were identified. Thus the search terms were defined as follow:

(1) prediction markets or information markets or decision markets or electronic markets or virtual markets or idea futures markets

(2) fundamental or description or functions or interpreting or performance

(3) theory or theories or requirement or model or modelling or aggregation or mechanism or contract or trader or platform or design

(4) applications or experiments or Iowa Electronic Markets or political or election or sports or business or entertainment or tool or implement

(5) organization or management or decision or manage or products or company

All these items for Prediction Markets were combined with the Boolean “AND” in order to link the major items to consistent with the defined review questions.

(1) AND (2) AND (3) AND (4) AND (5)

Phase 3 is designed to make systematic literature review more comprehensive. This was a process of doing backward searching, in order to ensure that all relevant important literature were included and also in case there was no big mistakes of missing significant
literature after getting the results from Phase 1 and 2. In the Phase 3, we reviewed on the reference lists of the selected articles from previous phases.

3.5 Results

The results are presented by overview section and quantitative analysis in accordance to the classification of Prediction Markets. From the results, it clearly shows that the situation of research based on Prediction Markets topics.

3.5.1 Overview of results

In total, 120 articles were collected as the result of systematic literature review. It included 52 articles from Journal of Prediction Markets (Phase 1), 62 articles from electronic databases (Phase 2), and 6 articles from reviewing references lists of selected articles (Phase 3). Figure 5 shows the number of articles according to the categories of Prediction Markets. There were 34 articles which focused on introducing the concept of Prediction Markets, and 33 articles belong to the description of theoretical work which includes markets supporting theories and markets designing issues. Besides the application category took almost one third of all articles (41 studies) which focused on Prediction Markets-based applications or experiments. For the organization management category, there were only 12 articles. All the articles were listed in Appendix A.

![Figure 5. Number of articles according to the classification of Prediction Markets.](image)

According to the Figure 5, the relevant research topics for Prediction Markets could be summarized that most of research focused on studying the aspects of intrinsic descriptions or theoretical work or external possible extension adoptions of Prediction Markets. But few of them connected with organization management aspect.

3.5.2 Analysis

In order to understand the situation of recent Prediction Markets research, and get the distribution of subjects, thus the classifications of Prediction markets were analysed by sub-categories.
Figure 6. Distribution of description category.

In the description category, the articles were divided into three subcategories. As the result is showed in the Figure 6, we had 24% articles which basically describe the concept of Prediction Markets. Half of articles made general description; they described various inherent aspects of Prediction Markets. And 26% articles contained with other aspects of Prediction Markets such as their possible development and potential usages. From the results, it clearly shows that the articles belong to description category introduce readers to understand the concepts and ideas of Prediction Markets, most of them toward to describe the inherent aspects of Prediction Markets. The reference literature for each subcategory is listed in the following Table 4.

Table 4. References lists of distribution of description category.

<table>
<thead>
<tr>
<th>Description</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>Hanson, 2002; Kambil, Van Heck, &amp; Van Heck, 2002; Spann &amp; Skiera, 2003a; Boyle &amp; Videbeck, 2005; R. Hahn &amp; Tetlock, 2006; Tziralis &amp; Tatsiopoulos, 2007; Williams &amp; Williams, 2011; Putintseva, 2011;</td>
</tr>
<tr>
<td>General description</td>
<td>Spann &amp; Skiera, 2003b; Schriber, 2004; Surowiecki, 2004; Williams, 2005; Mackie-Mason &amp; Wellman, 2006; Leigh &amp; Wolters, 2006; Tziralis &amp; Tatsiopoulos, 2006; Gruca &amp; Berg, 2007; Wolters &amp; Zitzewitz, 2006a; Arrow et al., 2008; Gruca, Berg, &amp; Cipriano, 2008; Hall, 2010; Rie &amp; Schoder, 2010; S. Luckner et al., 2011; Peeters &amp; Wolk, 2012; Berghford, Kildal, McPherson, Lothaas, &amp; Valvik, 2012; Antweiler, 2012;</td>
</tr>
<tr>
<td>Other descriptive issues</td>
<td>Lundhholm, 1991; Forsythe, Nelson, Neumann, &amp; Wright, 1992; Wolfers &amp; Zitzewitz, 2006b; Page, 2008; Ray, 2010; Slamka, Jank, &amp; Skiera, 2011; Teschner, Coblenz, &amp; Weinhardt, 2011; Saville, Stekler, &amp; Williams, 2011; Schnyter &amp; Schnyter, 2012;</td>
</tr>
</tbody>
</table>
Figure 7. Distribution of theoretical work category.

In the theoretical work category, we found 61% articles belong to the subcategory of market modeling; they consisted with the aspects of modeling Prediction Markets, framework design and supporting theories. And 27% focused on information aggregation convergence and equilibrium aspect. Rest 12% articles covered the theoretical aspects which could not be assigned to previous two subcategories, for instance the analysis on the Prediction Markets prices (see Figure 7). Theoretical work category provided much useful information for designing a Prediction Markets. We learnt and applied the key design elements of Prediction Markets from here. The reference literature for each subcategory is listed in the following Table 5.

Table 5. References lists of distribution of theoretical work category.

<table>
<thead>
<tr>
<th>Theoretical work</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market modeling</td>
<td>Pennock &amp; Wellman, 2001; Pennock, Debnath, Glover, &amp; Giles, 2002; Hanson, 2003; Simkins, 1999; Matthews, Rocchi, &amp; Gafni, 2002; Whinston, 2004; Chen, 2005; Hahn, Lien, &amp; Tetlock, 2005; McAdams &amp; Malone, 2005; Abramowicz, 2007; Hanson, 2007; Slamka, Soukhrooukova &amp; Spann, 2008; Ottaviani, 2009; Berg &amp; Proebsting, 2009; Hanson, 2009; Borghesi, 2009b; Borghesi, 2009a; Diemer &amp; Poblete, 2010; O. J. Bergfjord, 2011; McHugh &amp; Jackson, 2012; Teschner &amp; Rothschild, 2012;</td>
</tr>
<tr>
<td>Information aggregation convergence &amp; equilibrium</td>
<td>Pennock, 2004; Hanson &amp; Oprea, 2004; Koessler, Noussair, &amp; Ziegelmeyer, 2005; Ottaviani &amp; Sorensen, 2006; Hanson, Oprea, &amp; Porter, 2006; Sinha &amp; Bansal, 2008; Kalovicova &amp; Ortman, 2009; David, 2010; McKenzie &amp; Bullen, 2012;</td>
</tr>
<tr>
<td>Other theoretical issues</td>
<td>Deschamp &amp; Gergaud, 2007; S. Luckner &amp; Weinhardt, 2007; Graefe &amp; Weinhardt, 2008; Borghesi, 2012;</td>
</tr>
</tbody>
</table>
Figure 8. Distribution of application category.

In the application categories, we collected 41 articles which included empirical studies on Prediction Markets-based applications and various kinds of experiments. The result of subcategories is showed in Figure 8. We had 10% articles for describing Iowa Electronic Markets moreover 24% could be recognized as other political markets. Meanwhile, 24% articles were designed for predicting sports game results. At the subcategory of other applications, we received 27% articles which focused on extending the usage of Prediction Markets to other fields. In addition, 24% articles belong to the Prediction Markets-based experiments which were held in academic environment. According to the distribution of Application category, it clearly shows that Prediction Markets is popular in the political and sports fields. The reference literature for each subcategory is listed in the following Table 6.

Table 6. References lists of distribution of Application category.

<table>
<thead>
<tr>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other political markets</td>
<td>Forsythe, Frank, Krishnamurthy, &amp; Ross, 1995; Beckmann &amp; Werding, 1996; J. Wolfers &amp; Leigh, 2002; Brüggelambert &amp; Cruciger, 2002; Rhode &amp; Strumpf, 2006; O. Bergfjord, 2008; Tonkes &amp; Lesmono, 2010; Tung, Chou, Lin, &amp; Lin, 2011; Kros, Mai, &amp; Keller, 2012; Sinha, Sharma, &amp; Singh, 2012;</td>
</tr>
<tr>
<td>Markets on sport events</td>
<td>Schmidt &amp; Werwatz, 2002; Debnath, Pennock, Giles, &amp; Lawrence, 2003; Chen, Chu, Mullen, &amp; Pennock, 2005; Smith, Paton, &amp; Williams, 2006; Gil &amp; Levitt, 2007; O’Connor &amp; Zhou, 2008;</td>
</tr>
<tr>
<td>Other applications</td>
<td>Ortner, 1998; T. S. Gruca, 2000; Gruca, Berg, &amp; Cipriano, 2001; Gruca, Berg, &amp; Cipriano, 2003; Skiera &amp; Spann, 2004; Tetlock, 2004; Gruca, Berg, &amp; Cipriano, 2005; Passmore, Cebeci, &amp; Baker, 2005; Shilony, 2007; Ritterman, Osborne, &amp; Klein, 2009; Siegel, 2009;</td>
</tr>
</tbody>
</table>

In the organization management category, we only collected 12 articles which stepped into organization management aspect. Since it is known that Prediction Markets may contributes much helps for organization management, such as acting as decision making tool (Berg & Rietz, 2003; Hahn, Lien & Tetlock, 2005), but the result of this systematic
literature review shows less of articles focused on studying Prediction Markets usages in organization management and even few of them take it into consider as the evaluation system. The references literature for organization management category is listed in the following Table 7.

Table 7. References lists of distribution of organization management category.

| Organization management | Choo, 1996; Miller, Hickson, & Wilson, 1999; Lilien, Morrison, Searls, Sonnack, & Von Hippel, 2002; Chen & Plott, 2002; Abramowicz, 2004; Paul W Lail, 2005; R. Ray, 2008; Waitz & Mild, 2009; Spears, LaComb, Interrante, Barnett, & Senturk-Dogonaksoy, 2009; Rhode, 2009; Rajakovich & Vladimirov, 2009; Lavoie, 2009; |

The results of systematic literature review are listed as above based on the order of the categories and subcategories. The results were understandable and achieved the initial purposes of doing this systematic literature review.

3.6 Relationship of results between systematic review and thesis work

From the distribution analysis based on the categories of Prediction Markets, the results show that Prediction Markets has abundant literature resources. Most of them focus on studying the theories of Prediction Markets and extending its usage by various applications. But, less of them combine Prediction Markets with organization management aspect, even there is no literature concerning on course evaluation field.

As an important part of thesis work, the results of systematic literature review provided many valuable articles for constructing this thesis. At the beginning stage, the articles in description category were learnt to understand the concept and idea of Prediction Markets as well as its inherent aspects analysis. Moreover, the articles from theoretical and application category had been used for designing the Prediction Markets-based course evaluation system.

Actually, the review questions were designed to give background knowledge for the purpose of thesis. According to the review question 1, we could easily find out the useful articles from the description and theoretical category. The key design elements of Prediction Markets were learnt and summarized in the Chapter 4. In addition, review question 2 was designed to search for the studies of Prediction Markets-based applications. The relevant articles could be found in application and organization management category. The empirical knowledge of various applications would be used to organize the explorative experiment of Prediction Markets-based course evaluation system which will be described in Chapter 6.
4. Fundamentals of Prediction Markets

Since the definition of Prediction Markets was described in the previous chapter, in this chapter the focus is to study the main design issues of building a Prediction Markets and to describe previous usage of Prediction Markets. They can also be recognized as the answers for review question 1 and partly for 2. According to the systematic literature review, Prediction Markets is a popular research topic in recent years. In the Following, the useful literature including theories and empirical studies are described with corresponding to the research topic “Prediction Markets-based course evaluation system”.

4.1 How does it work?

The initial purpose of Prediction Markets concept is information aggregation. Later on its aggregated information is used for predicting, decision making and developing many extension applications in various fields. Give first impression of Prediction Markets, there could be confusions between Prediction Markets, stock markets and betting markets (or regards to Gambling). The analysis mentioned below makes them a clear boundary from the purpose perspective. It is known that stock markets are built depends on three main purposes of allocating resources, sharing risks, and raising capital. Undeniable it also provides the function of information aggregation, but it is only the pleasant byproducts from main purposes. And for betting markets, they never get away from the concept of entertainment, moreover, people recognize they are harmful and addictive. (Luckner, 2008)

The term Prediction Markets and its operational principles would be easily understand by a simple paradigm. Suppose there is a need from the board of directors in a company who want to know the sales forecasts on a specific product via Prediction Markets. The markets are open for all employees who have access to relevant information. Each of them is provided with a virtual sum of money and has the permission to access into the markets. Besides there is a contract question which was designed with connection of sales volume of defined product. For example, the question is “what will be the sales volume of product x in next year” and one of the optional answers is the products x will be sold out in the amount between 1000 and 2000. If the forecast result proved to be true then the contract would be paid off by €1, otherwise it is worth nothing.

Assume that at a certain point time the answer receive the price on p. This trading price indicate that the probability occurring of product x will be sold out between 1000 and 2000 is p%. If an employee who believes the likelihood of stock is at least p%, he/she probably buys the stock, while another one who expect that the answer is overvalued, definitely he/she won’t buy any stocks and even sale the stock holdings. With trading actions from employees, the price would be changed continuously. The transactions and prices are the useful indicators of revealing traders’ opinions on sales volume of product x. According to the efficient markets hypothesis (Fama, 2012) that prices of stock would fully reflect the relevant information and collect expectations of traders. Therefore, in this case the final value of the stock would be read as the probability of occurrence of product x will be sold out among 1000 and 2000 at p%. In addition, the markets should be the most efficient instrument to aggregate all distributed information of markets participants (Luckner et al., 2011, p 9).
4.2 Types of Prediction Markets

The types of Prediction Markets are providing a very useful way to explain the concept of prediction markets. The Prediction Markets can be classified by the payoffs which are tied closely to the outcomes of the future events. Regarding to the research formed by Wolfers and Zitzewitz (2004), there are three main types of Prediction Markets. Table 8 shows how different contracts have been run for different predictions.


<table>
<thead>
<tr>
<th>Contract</th>
<th>Example</th>
<th>Details</th>
<th>Reveals market expectation of</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winner-take-all</td>
<td>Event y: A1 Gore wins the popular vote</td>
<td>Contract costs €p. Pays €1 if and only if event y occurs. Bid according to value of €p.</td>
<td>Probability that event y occurs, p(y).</td>
</tr>
<tr>
<td>Index</td>
<td>Contract pays €1 for every percentage point of the popular vote won by A1 Gore.</td>
<td>Contract pays €y.</td>
<td>Mean value of outcome y: E[y].</td>
</tr>
<tr>
<td>Spread</td>
<td>Contract pays even money if Gore wins more than y*% of the popular vote.</td>
<td>Contract costs €1. Pays €2 if y&gt;y*. Pays €0 otherwise. Bid according to the value of y*.</td>
<td>Median value of y.</td>
</tr>
</tbody>
</table>

Wolfers and Zitzewitz (2004) summarized the types of Prediction Markets. According to the different contracts, the predictions also are revealed by three main ways. First, it is the Binary option contract but also going by name winner-take-all contract. The contract costs the price as €p and pays off depends on whether the specific event occurs. For instance, if a candidate wins an election at the end, the €1 (equals the event has happened in 100%) should be the last price of the contract. Otherwise the price is €0. The price which bidding by customers represents the markets’ expectation of event’s probability occurs in percentage.

Second, in an index futures contract, it is decided based on a number that act as a key index of the final result of unclear future events. The contract pays amount in continuous variable which the price following raise or fall. For example, the contract can be set as every percentage point of the popular vote for candidate is paid by €1, and the price from the markets represents the mean value that reveals the outcome of the event. (Wolfers & Zitzewitz, 2004.)

The third one is the spread betting. It is the combination of the above two. Traders bidding different prices on the cutoff of the event occur, like weather a candidate receive the exact percentage of the votes at the end is more than or less than the certain number percentage of the popular votes. The price reveals market expectation as median value of outcome. (Wolfers & Zitzewitz, 2004.)

4.3 Trading mechanisms

Trading Mechanism is another crucial element has to be considered, perhaps the most crucial parameter in market design (Slamka, Skiera, & Spann, 2012). Many kinds of trading mechanisms designed for different Prediction Markets applications exist. Simply, they can be classified into two categories to understand (see Figure 9). One is called Double Action (DAs) and its derivatives, another remains to automated market maker
(AMMs) which includes logarithmic market scoring rules (LMSR) (Hanson, 2007); dynamic pari-mutuel market (DPM) (Pennock, 2004), dynamic price adjustment (DPA) (Slamka et al., 2009).

![Diagram](image)

**Figure 9.** Functioning of a standard double auction and an automated market maker (Slamka et al., 2009).

In the DAs, when buyers submit bids (or sellers sell), if the order follows the specifications of market, it will be put into the markets to wait for trading. Then, if a matching order is found i.e. price bid from buyers is at least as high as the ask price in selling order, the order will be executed immediately, but if no matching orders exist, the order should keep stay in the market and remains until it expires, or wait for a relevant counteroffer, or be removed (Madhavan, 1992). DAs consisted with double auctions (DA), continuous double auctions (CDA) and call auctions (CA) which has been often used in Prediction Markets and financial markets. Both of them do not engage with transactions by themselves, but only depends on trades from others (Luckner et al., 2011, p. 14-18).

In contrast to the DAs, AMMs can provide traders opportunities to buy and sell shares at any point of time. The effective transactions do not depend on counteroffers from other market participants, but relies on a piece of software and its defined algorithms. Participants have no need to wait for matching orders anymore. If a matching order appears, their trades will be executed immediately. Otherwise, AMM will act as counterparty to execute the trade (Luckner et al., 2011, p. 14-18). For a successful Prediction Markets, market liquidity as a central problem cannot be ignored. In order to solve the “chicken-and-egg-problem”: traders are attracted to join the market with high trading frequency, but on other hand, liquid market needs more traders (Pennock, 2004). AMMs are applied as a possible solutions especially in small or internal Prediction Markets.
4.4 Does money matter?

Besides trading mechanism, it is also important to decide whether Prediction Markets will be operated based on real-money or play-money. Normally, people may believe that the markets where traders put into their own money should produce better forecasts than those markets where traders are risk-free (Servan-Schreiber et al., 2004). Such as the operations in other traditional markets (stock markets or betting markets), people are motivated by the return from their investments. But, in Prediction Markets the parameter which is decided on real-money based, or play-money based does not make that much sense on the final results. Previous research have shown the performance of using play-money is as well as real-money based Prediction Markets, in other words, the play-money based prediction Markets can just product the forecasts as accurate as those operated based on real-money (Servan-Schreiber et al., 2004).

So far, beside the money (stakes) issues, many other kinds of incentive schemes have been used in play-money based Prediction Markets in order to motivate people to take part in the markets and to reveal valuable information from them. Typical, there are three basic incentive schemes including to reward top performance participant, to set lotteries among traders, and to publish rankings (Luckner et al., 2011, p. 24). In addition, Robinson (2001) claimed that as the popularity of using play-money exchanges, such incentives are often enough to motivate trading actions (Robinson, 2001).

4.5 Why do Prediction Markets work so well?

First reason is from one of the significant intrinsic features of markets that can give continuously updated information via dynamic prices of markets which then can be conveyed as support information of a measurable future outcome (Berg et al., 2003).

The second reason is from the views of Gangur and Martinčík (n.d.) who thought that Prediction Markets are extremely efficient at aggregating information. This is because the trading process and prices from the Prediction Markets can reveal the information of people’s expectation on the future events. Based on the efficient market hypothesis, if all the relevant information is fully reflected by the markets prices, then the market is efficient (Fama, 2012). In addition, Hayek’s research shows the markets should be the most efficient instrument to aggregate all distributed information of markets participants (Luckner et al., 2011, p. 7).

Thirdly, Prediction Markets adopts the concept of wisdom from crowds. This is because groups are almost always smarter than individuals, even the smartest people (L.V. Williams & J.V. Williams, 2011). In addition, running a Prediction Markets is less expensive than hiring experts, especially, there is a series of complicated data needed to analysis, or has to conduct a large-scale poll (Ritterman, Osborne, & Klein, 2009).

Several studies show that Prediction Markets can result with high accuracy predictions compared to many traditional forecasting methods such as polls (Berg et al., 2001), expert predictions (Spann & Skiera, 2003), or official forecasts (Chen & Plott, 2002).

The last reason is from many successful experiments and applications, they give the evidences that Prediction Markets can overcome biases that possibly made by individuals and efficiently eliminating them from forecasts (Berg & Rietz, 2003).

In conclusion, an open market (especially for those which are under well designed) can produce better forecasting based on the given information. So why not try to adopt a
Prediction Markets–based system to aggregate the information from students instead of using traditional course evaluation method. An interesting and novel Prediction Markets-based course evaluation system is expected.

4.6 Previous usage of Prediction Markets

From the systematic literature review, the articles belong to the applications category clearly showed that Prediction Markets have been used in various fields, but most of them mainly focused on three distribution topics. They are political election, sports game, and academic experiment. In the following, Iowa Electronic Markets, Prediction Markets for sports game and Hollywood Stock Exchange are described. This is because they represent the three typical Prediction Markets applications. In addition, some other unfamiliar Prediction Markets applications are studied.

4.6.1 Iowa Electronic Markets

At the year of 1988, the University of Iowa got the permission from US Commodity Futures Trading Commission to operate the Iowa Electronic Markets which has become one of the most famous and successful cases that adopting the concept of Prediction Markets. It is also the first Prediction Markets employed by education institution (Boyle & Videbeck, 2005). The idea of IEM is simple: by posting claims into the market where the claims represent the contingent events and were traded continuously in crowds until did events happen. Within the fine procedures, the price should reflect from the expectations of pay offs to the claims. Depends on this relationship, it can be used to predict upcoming events (Berg et al., 2003).

Behind the Iowa Electronic Markets, it has been used for three aspects. One of the purposes is acting as a teaching tool. Students in University of Iowa can use IEM as their real study case for practicing and learning. In addition, it is also a good way to provide researchers first-hand data. Especially for those who are focusing on the prediction issues. The most useful part is the function of predicts contingencies named policy instrument such as president election, sales of new products (Boyle & Videbeck, 2005). The predictions derived from IEM have provided more accuracy results than natural benchmark, such as polls (Berg et al., 2001; Forsythe et al., 1999). As we know the accuracy of IEM is better than other survey companies. For instance, in 2008 US presidential elections, IEM has predicted the percentage of the voting result which has only average error of 1.5%, while the widely famous poll company Gallup has had the prediction error of 2.1% (Wolfers & Zitzewitz, 2004).

Berg et al. (2003) had given out the reasons why Prediction Markets (special for the market like IEM) should predict better than traditional forecasting methods. The first one is from the design aspect of markets itself. The market is totally designed for purpose of forcing traders focus on the specific event. Second, traders express their opinions by putting their money on. Presumably, the amounts of money reflect the confidence of event’s happening results. The last reason, the prediction markets are providing a dynamic, efficient and hopeful way to aggregate the huge information from crowds. (Berg et al., 2003.)

4.6.2 Prediction Markets for sports game

Later the Prediction Markets becomes mature to serve for predicting worldwide sports games, there are many companies doing such business, such as Betfair (Smith, Paton, & Williams, 2006), TradeSports (Chen, Chu, Mullen, & Pennock, 2005). In which, Betfair
is the world’s largest internet prediction exchange markets locate at London with over 4 million customers and £50 million turnover per week in year of 2011 (Sean, 2011).

These sports-based Prediction Markets focus on forecasting the results on kinds of sports game, sports tournaments. For instance Betfair including up to 30 categories of sports e.g. football, basketball, boxing, horse racing, and ice hockey. Previous studies on sports-based Prediction Markets showed that the results from these markets were just as accurate as predictions of experts (Chen et al., 2005; Servan-Schreiber et al., 2004) or even better (Spann & Skiera, 2009).

4.6.3 Hollywood Stock Exchange

Another famous application Prediction Markets is Hollywood Stock Exchange, which is speculate designed for serving movies, box-office returns, and stars. Visitors can predict their favor movies’ performance by using Hollywood Dollar at this world’s leading entertainment stock market (Gruca, Berg, & Cipriano, 2003). It seeks to aggregate information from traders that mainly formed up by fans.

Still Hollywood Stock Exchange is acting the most prominent Prediction Markets which organized by play-money. Players buy and sell shares in widely related issues of films e.g. box-office revenue, Oscars, or issues respect to actors and actresses. In HSX, there is a limitation of each trader would only hold 10000 shares per item, in order to avoid too much influence from a single player. It has also been proved that HSX offers good predictions of a film’s gross before release and even could provide better forecasts after the opening weekend (Levmore, 2003). After years developing, besides the “Movie Stocks” and “Star Bonds”, it includes “TV Stocks” and “Movie Funds” services (Hollywood Stock Exchange, n.d.).

Table 9. Summarize of prediction markets applications.

<table>
<thead>
<tr>
<th>Name of Market</th>
<th>Specific area</th>
<th>Markets services</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iowa Electronic Markets</td>
<td>Political and Economic Events Political election</td>
<td>Policy Market; Economic Indicator Market; Class Markets (Academic Traders Only)</td>
<td>Pedagogical tool, Research tool</td>
</tr>
<tr>
<td>BetFair</td>
<td>Most kinds of Sports and tournaments</td>
<td>Sports; In-play; Football; Hour Racing; Casino</td>
<td>Gambling</td>
</tr>
<tr>
<td>Hollywood Stock Exchange</td>
<td>Movies and Stars</td>
<td>Movie Stocks; TV Stocks; Star Bonds; Movie Funds; All Derivatives</td>
<td>By buying and selling shares of favorite and celebrity stocks to aggregate information from traders in order to forecast the revenue of movies, Oscar prize.</td>
</tr>
</tbody>
</table>

In above table, it lists the basic information of Iowa Electronic Markets, Betfair, and Hollywood Stock Exchange for comparison purpose (see Table 9). They are adopted in different areas and provide different services based on their purposes. Both of them could be recognized as excellent Prediction Markets-based applications in related field.
4.6.4 Other applications

Besides above typical Prediction Markets applications, there are amounts of studies which focused on making innovations on Prediction Markets. In the following, we briefly describe three of them.

In Berlemann and Nelson (2002) research, they argued that Prediction Markets could provide reasonable forecast of macroeconomic variables and built a prototype experiment of using Prediction Markets to forecast the future inflation. In their prototype experiment, the trading was based on real-money and most traders were students from Dresden University of Technology. The results from the prototype experiment demonstrated that well-defined Prediction Markets can forecast mean inflation as well as the different inflation scenarios at low costs. In addition, their research verified that Prediction Makers is a useful forecasting instrument. (Berlemann and Nelson, 2002)

Another application of Prediction Markets is used to forecast the demand for a hospital services. Rajakovich and Vladimirov (2009) conducted an experiment in the Royal Devon and Exeter Hospital, it continued for a week time and 65 participants were involved. They were the employees of the hospital. At that time, this research was the first trial of Prediction Markets in a healthcare environment. As a result, it confirmed the effectiveness of Prediction Markets. (Rajakovich and Vladimirov, 2009)

Third example is a trail which combines Prediction Markets with Twitter by Ritterman, Osborne, and Klein (2009). They explored the hypothesis that social medial encodes the belief of people about some future events of the world. The hypothesis was tested in the case of Swine Flu Pandemic in 2009. In the experiment, authors collected almost 50 million Twitter posts in a period of two month. They presented that a method of using the regression framework can reduce the error associated with modeling beliefs from Twitter data. (Ritterman, Osborne, & Klein, 2009)

The three applications of Prediction Markets are selected from the applications and organization management categories of Prediction Markets-based systematic literature review. This is because they did similar work as this research. Researchers explored the usages of Prediction Markets in three different areas. Their successful experiments provided useful experience for inspiring the idea of extending the usage of Prediction Markets in course evaluation and also providing empirical knowledge.
5 Course evaluation system in universities

As it is mentioned before, this thesis is proposed to extend the usability of Prediction Markets. To act as a new organization management tool, the course evaluation field has been set up as a focus. Thus, we should have a basic understanding of course evaluation. This chapter is consisted of two parts of work. One is the literature review which collects the knowledge of course evaluation. Another is the interviews for university staffs which provide the information of current existing course evaluation system in University of Oulu.

5.1 Definition of course evaluation

It is known that course evaluation is the most widely used measurement of evaluating the quality on a given course which is usually organized by a paper-based or electronic-based questionnaire (Anderson et al., 2005). Good-quality feedback from students that represent their real perceptions on a specific course are not easy to obtain. Incorrect information or lack of feedback may mislead teachers to evaluate the course quality on their own perceptions (Tucker, Jones, Straker, & Cole, 2003). Therefore, a cycle of evaluation based on collecting students’ feedback is essential to improve the course quality (Anderson et al., 2005; Tiberius et al., 1989). In addition, valid and reliable ratings from students could be acted as an efficient way of evaluating teaching quality (Centra, 1977). Typically, course evaluations are combined with multiple sources and types of data; they are students’ ratings, peer review and self-evaluation. The purpose of those data can either be used for forming summative evaluation or formative evaluation. (Paulsen, 2002)

In briefly, a summative evaluation is often organized at the end of the course or a semester. Students who participated in the course would be asked to give their feedback based on their feelings for a specific course. The process is completely confidential and anonymous in order to collect the real opinions (Mohanty, Grets, Flowers, Algozzine, & Spooner, 2005). The purpose of summative evaluation is to provide information for teachers, faculty committees or university leaders to evaluate the teaching quality of teachers and to make decisions for teaching management such as the issues of hiring, renewing, awarding, and promotion. (Paulsen, 2002). Another is called formative evaluation, it is typically designed for providing information to assist teachers or faculty to improve the effectiveness of teaching and it could take place during the current semester or period of a specific course. Generally, formative evaluation is performed by peer review from other experienced teachers but students can also be included if their information could directly help to enhance the teaching skills or teaching quality (Paulsen, 2002).

However, those student evaluations of teaching effectiveness have been widely criticized. Course evaluation often fails to capture the lecturer’s ability and could not be simply act as a tool to improve the teaching quality (Emery, Kramer, & Tian, 2003). The survey results from previous research pointed out that most of teachers believe that if a teacher raises the level of standards and/or content would also get worse results from evaluations, and the rating from students who participant in the evaluation session could be biased from their favour on certain teachers’ personalities, looks, gender, habits. (Gross & Small, 1979). Therefore, some teachers thought that using such method to evaluate the course
quality could not reflect accurate situation of the course. More studies demonstrated that the method of students evaluations which have been used by many universities are demeaning to instructors (Gray & Bergmann, 2003) and somehow misused results would bring a corrupting effect on students’ attitude (Platt, 1993). Due to the evaluations are usually organized at the end of the course or semester, there is typically no mechanism designed for responding the feedback from students and reporting the changes of the course arrangements. Students who contribute the opinions for improving course quality have no chance to get effects or see the immediate changes.

5.2 Paper-based VS online-based course evaluation method

Recently, most of universities probably have implemented their own course evaluation system in order to collect the students’ feedback for the purposes of enhancing the management and improving the teaching quality. Basically, most of them are realized by traditional or online evaluation system (Anderson et al., 2005). It is necessary to understand these two methods of organizing course evaluation, due to Prediction Markets-based solution could become as one possible new method to contribute in course evaluation field.

Traditional paper-based method could provide a quite good quantitative and qualitative data from students, but it often acts as big task for universities, teachers and students. For instance, it is time-consuming for teachers to collate and interpret large information. Especially, the teachers who are managing several courses at the same time, it is money-consuming for universities, a paper-based course evaluations can cost human resources, financial resources over years. In addition, the paper-based method can neither recognize nor evaluate the learning conditions of students. (Tucker et al., 2003)

Instead of paper-based method, online course evaluation system is supported by emerging technologies; for instance, web technologies, information and communication technologies. The advantages of online course evaluation system could be concluded with (1) students could provide rapid feedback; (2) less expensive than paper-based method; (3) require less time to arrange, or data analysis; (4) less vulnerable to professional influence; (5) provide students multiple methods to give feedback; (6) students have enough time to complete the evaluation. But the disadvantages are also significant and could not be ignored. Such as it requires the complex technology support as well as specific equipment (computers); its data would be considered inaccurate, and the situation of lower response rate is serious (Anderson et al., 2005). As it was mentioned that response rate for online course evaluation is lower than paper-based course evaluation and therefore it may bring the influence on the validity of the results (Anderson, Brown, & Spaeth, 2006). Especially for those teachers who prefer the traditional paper-based method may thought that online course evaluation could produce even worse data than traditional method.

5.3 Current course evaluation in University of Oulu

Before stepping into the research of extending Prediction Markets in use of course evaluation, the facts of current course evaluation in University of Oulu have to be understood. As a student, I observed that there are two ways for students to give their course feedback. One is an online course evaluation system which is consisted as one of information system for students, studies, and teaching administration and registration called Weboodi. Students could freely give their feedback by taking a questionnaire survey at a period of time. The system often informs students to participant the feedback at the end of the course. The other way often occurs at the last lecture of a specific course,
few of teachers would hand over a paper-based questionnaire to students. But there is no standard format; different teachers may use different questions.

To systematically understand the situation of course evaluation system in the university, two interviews with current and former deputy director of department were arranged. From the interviews, the course evaluation system in the university could be described that the course evaluation system in Weboodi is the official method which is organized at the university level. It is useful and valuable and the questions were predefined. Leaders of university do follow the data from the system and make decisions. However, the system is not satisfied, one of the big problem of current system is the low response rate, only small number of students really joined and given the feedback via Weboodi, the rate was around 10%-15% and the results from the system is somehow biased, such as mostly the students who are really satisfied the course would give the feedback or who are very negatively thinking on the course. In order to receive better feedback from students, additional methods are added in department level or organized by teachers who were responding on the specific course. In our department, multiple methods are included such as the feedback from students association (Blanko), actives on students’ feedback day and a teaching development work group which formed up by experienced teachers. For the teachers, they often organized own methods for receiving the course feedback. For instance, at the last lecture, a paper-based questionnaire survey which designed by course response teachers themselves would be handed to students or just arrange face-to-face interviews for some students. They thought these would provide more concrete ideas of what aspects need to be improved.

Obviously, the course evaluation system in University of Oulu has the same problems as learnt from the literature. The course evaluation seems cannot reach the original expectations well. Low response rate, time consuming, inaccuracy data and lacking interactive between teachers and students are summarised as common issues of course evaluation (Anderson et al., 2005; Tucker et al., 2003). The Prediction Markets-based course evaluation system which combined with Prediction Markets and course evaluation may become a solution for those issues.
6 An explorative experiment of Prediction Markets-based course evaluation system

From the previous chapters, theoretical backgrounds were summarized and ready to guide the following experiment. As the successes of Prediction Markets in various fields, we believe that it is possible to make an innovation on traditional course evaluation system by adopting Prediction Markets. The overall purpose of the experiment was to test the feasibility of adopting Prediction Markets in a real case of course evaluation. In order to achieve the purpose, each detail was taken into consideration as a whole course evaluation system. This research could be regarded as the first trail of applying Prediction Markets principles to the field of course evaluation.

In the explorative experiment, there were two parts which corresponding with two randomly allocated groups of participants (students). One was called “Prediction Markets-based group” which would be served by Prediction Markets-based course evaluation system; another was named “Traditional group” which was organized underlying the traditional method of course evaluation. The experiment supposed to make a course evaluation based on a real course environment which is named Research Method course. I assumed that the results are going to verify the extension use of Prediction Markets in course evaluation field. The experiment was running from 14th November to 23rd November, 2013 and supported by 49 participants in which 19 of them were served by a customized Prediction Markets-based site and the rest 30 students were asked to take traditional online-based course evaluation method. The trading questions in Prediction Markets-based Course Evaluation System and questionnaires in traditional method were predefined. In addition, the technical supports of this experiment were from Inkling Markets (Inkling Incorporated, n.d.), and an online-based survey service provider Surveygizmo (Surveygizmo, n.d.).

The explorative experiment of Prediction Markets-based course evaluation system included three steps. They were: design the system, establish the system, and defined the experiment. Besides during the experiments there consisted with training lecture session at the beginning and afterwards interview session. The details were described in the following.

6.1 Design the Prediction Markets-based course evaluation system

By following the conceptual Prediction Markets-based course evaluation system, it is clear that key design elements should be taken into consideration before running the experiment. They are the specification of contracts which are traded in the market, trading mechanism which would influence the liquidity of markets, incentives which are designed to motivate people to reveal information, as well as the selection of traders and the trading platform.

6.1.1 Contracts

Foremost, one of the crucial element should be defined is the contracts. The type of contracts is applied to decide the final value of the contracts. According to the classification of contracts types from Wolfers and Zitzewitz (2004), there are three basic
types: (1) Winner-takes-all contracts; (2) Index contracts; (3) Spread contracts. Each of them corresponds to different prediction requirements. Due to Index and Spread contracts should be adopted with strict requirements, one is referred to forecast “vote-share” in special context of presidential elections (Berg et al., 2003), while another is designed for predicting median value of outcomes. These two types are not satisfied the requirements of the experiment. In addition, the type of Winner-takes-all is common used in Prediction Markets applications and it is easier to follow than two others. Therefore, in this Prediction Markets-based experiment, all the stocks were underling Winner-take-all contracts. In this case, if the answer for specific course evaluation question fitted (or happens) the results from Traditional group students; its payoff value was €100, if it did not, then worth nothing. During the trading time, the price for each stock reflected the feelings of students for the corresponding aspects of course evaluation.

6.1.2 Trading mechanism

_Trading mechanism_ is another key element of Prediction Markets. Most existing mechanisms can be classified to two standards. They are Double Auction (DA) and Automated Market Maker (AMM), trading mechanism determines how to match buys and sales in the trading actions and influence the stock price after trading is completed (Luckner et al., 2011, p. 14). In consideration of Prediction Markets-based experiment, it was a small experiment. In order to ensure the liquidity of market, the experiment would adopt AMM mechanism. In this way, participants did not need to wait for their bids matching and the price of stock would be updated immediately. Relatively, AMM was easier for students to understand and make operations.

6.1.3 Incentives

_Incentive_ can regard as a key element to decide a successful experiment and also to attract traders and reveal their expectations. It is already known that play-money perform as well as real-money (Servan-Schreiber et al., 2004), and in university environment, it is not allowed to use real money even it is ridiculous for students to put their own money for giving course evaluation. Therefore in the case of Prediction Markets-based experiment, it would be operated under play-money by assisting with rewards and lotteries. The arrangements are described:

1. A ranking list of traders was shown in the stats page which ordered by their earned money.
2. A Dashboard of their account where they could see their account value and rank compare to others.
3. At the end of experiment, top three “richest” participants and one more activists would be rewarded by gifts. (see Table 10)
4. Students from Traditional group will join the lottery after they submit their surveys. Five of them would be rewarded with one coffee ticket.

**Table 10. Awards setting.**

<table>
<thead>
<tr>
<th>Award title</th>
<th>prize</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 1 “richest student”</td>
<td>3 coffee tickets</td>
</tr>
<tr>
<td>Top 2 “richest student”</td>
<td>2 coffee tickets</td>
</tr>
<tr>
<td>Top 3 “richest student”</td>
<td>1 coffee tickets</td>
</tr>
<tr>
<td>Best active award</td>
<td>2 coffee tickets</td>
</tr>
</tbody>
</table>
6.1.4 Traders

From the view of L.V. Williams and J.V. Williams (2011), the idea behind Prediction Markets can be easily defined as “The cleverness from crows”. In other words, the more traders are included in the Prediction Markets, the better results will be produced. Due to the limited resources, small experiment could be conducted. The traders were obviously from the students who supposed to take the course and volunteered to join the experiment. This was because in the existing course evaluation system, all the students were free to give course feedback after the corresponding course was finished. During the experiment, all the trading actions from traders were recorded and kept confidential. The information of trading activity including orders, trading time, quantity, price of stocks, comments, and account balance was available and could be exported at any time.

6.1.5 Trading platform

Trading platform could be recognized as a big issue for implementing Prediction Markets. From the view of Luckner et al. (2011, p. 43), a Prediction Markets is an online stock trading system; it is not only to make the abstract market system enough for properly trading, but also it should be implemented to satisfy the certain needs. Due to this was an explorative experiment of extending Prediction Markets usages, we decided to use existing Prediction Markets service from internet.

After assessed many existing Prediction Markets software sites, the Inkling markets (Inkling Incorporated, n.d.) can be a good solution for the experiment. Inkling Markets is a prediction markets-based platform which designed to offer intelligence solution to help organizations collect information efficiently and make better decision. Its products had been implemented in various fields including industries, government, small business and academic. The reasons to choose Inkling Markets service were because it is developed by following AMM trading mechanism and using play-money for trading, it can provide many incentives schemes settings, and easy to use. Besides, there was offering a free 14 days plan with unlimited full functions to customer to personalize their own Prediction Markets-based service. In the experiment, we applied this 14-day pilot in order to build up a customized service site for trading which is named Student Feedback Market.

6.2 Establish the Prediction Markets-based course evaluation system

Although Inkling Market can provide mature solutions for organizations and governments, it cannot suit for this specific experiment which was particularly designed for evaluating the course. In order to make a successful platform for experiment, the site should be customized. This site was named Students Feedback Market, and in the welcome page, the content included a description of the experiment, introduction of this market and a trading wizard to help participants to get started. As it is depicted in the Figure 10, it shows the front page of personalized Students Feedback Market. For the security settings, the market was not open to the public, only those people who received invitation could register into the market and traders were not allowed to ask questions i.e. they were not allowed to launch new contracts in the markets, due to the contracts were already prepared in correspondence with the research purpose. In addition, the logo had been changed to University of Oulu, the currency and date format and time zone were changed to fit Finnish usual practice. These attentions on details may provide better environment for participants. When they had the feeling of they are included in such an official experiment, they would put much concentrates on it.
Regarding to the questions setting in Weboodi course evaluation system, five questions were shown in the Students Feedback Market. Four of them were traded as contracts (see Table 11). They were designed to ask the traders’ opinions from four main aspects of course quality: lectures, course materials, exercises and general level of the course. Traders would buy (bid) or sell (ask) the possible answers of specific question based on their beliefs (expectation) for accessing how the “Traditional group” students would evaluate the course. The value of possible answers indicate the likelihood of the event happens and its equal to the value of the stock (e.g. 50% is the initial value for “yes” stock, that means the price of the stock is €50, trader can pay €50 to purchase one “yes” stock, or sell them with this price). On the last day of the markets, all the contracts were paid out according to the actual results, and the money will automatically be deposited or removed from traders’ accounts which have been involved in the relevant contracts.
Table 11. Questions list.

<table>
<thead>
<tr>
<th>Questions(contracts)</th>
<th>Possible Answers(stocks)</th>
<th>Initial value(predicted chance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Do the lectures support the learning and understanding of the contents of the course well?</td>
<td>Yes</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>50%</td>
</tr>
<tr>
<td>2. Do the course materials support the understanding and learning of the contents of the course well?</td>
<td>Yes</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>50%</td>
</tr>
<tr>
<td>3. Do the exercises support the understanding and learning of the contents of the course well?</td>
<td>Yes</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>50%</td>
</tr>
<tr>
<td>4. What is the most typical overall grading of the course by student?</td>
<td>Very Poor</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>Fair Well</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>Well</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>Very Well</td>
<td>25%</td>
</tr>
</tbody>
</table>

Besides, another question was designed to be a discussion area where traders can write their comments or improvements for the Student Feedback Market. It was “What is your opinion about express your feedback of the course (compare to the traditional survey)? How should it be improved?”

At the beginning of the Prediction Markets-based course evaluation experiment, all selected students data would be imported to the trading platform, and each of them would receive an invitation email, there was a link inside. By active their account via the link, they could login to the service with their user name and password which were defined by them, and also €5000 endowment (play-money) was composed in the account. The trading platform was open 24 hours a day and continues 10 days. People traded on above contracts based on their expectations relating to the likelihood of uncertain future. The goal of each trader was to increase the money in his/her account, by buying contracts when they were undervalued and selling those which were overvalued. In the trading platform, participants were provided two ways to expressing their opinions about which answers would happen and how confidences did they believe. One was express the opinions by trading. The common way to submit orders was directly click on the possible answer which you would like to trade. The trading screen is showed in Figure 11. Traders made a prediction just by click one of six buttons (color lump), the farther away from the center with deeper color the stronger traders believe the answer happen. Once, when clicking on the button, the amount of cost would be displayed.
Besides, another way was called “Advance model”. It provided more fine-tuned control and showed how much money you wanted to spend to make predictions and more information about how the money was being spent. It was featured with displaying amount of information (see Figure 12). They were the current price, the price after operating, shares currently held, shares want to buy, available cash, and available cash after the operation.

Let’s make an example, there was a trader, he was dealing with the question “do the exercises support the understanding and learning of the contents of the course well”. He thought the right answer is “YES”. As it is showing in Figure 12, a trader was going to buy 10 shares with the current price €50, and his available cash was €3305.18. After taking 10 shares, this trade would cost him with €505, and his rest deposited cash would becomes to €2800.18. Meanwhile, the current price of “YES” answer would be changed to €51 with €1 increase.

In order to provide clear historic transaction information to traders, in the page of each contract the price chart is available (see Figure 13). It was featured with the trading date, trading volume, and price and the data were updated automatically.
Market was implemented as a real-time system and continuously provides information for its users.

The other way for traders to express their opinions was by commenting. Along with each setting question was a discussion area. Traders could freely make comments for the specific question. In this way, it provided an opportunity of making interaction between traders; they could easily share the information. In addition, others might be persuaded by “you’re putting money where your mouth is”.

6.3 Define the experiment

The experiment comprised with two parts. The first one Prediction Markets-based experiment acted as main part which had been put into a lot of endeavors. The second part was designed to simulate currently using students’ feedback system which was acted as
the official means of course evaluation in the university. The results from “Traditional group” were mainly used as comparative data, and also can provide useful information in order to achieve the research purpose and optimize system later. The issues of the common course evaluation methods were summarized in previous chapter, so the different results in afterwards comparison are predictable.

At the beginning, it is necessary to understand the environment settings of the experiment which including the background course, group of participants. The situations are described in detail below.

The course Research Method was selected to serve the experiment. This is because the Research Method course is compulsory for students thus it may guarantee the number of potential participants to join the experiment, and make the experiment organized underly Research Method course environment may improve the experiment quality due to students were learning methods of doing high quality research. There were three modules in Research Method course. Finally, Design Science Research (DSR) module had been selected for experiment, just because its teaching period match the experiment plan and my supervisor was the module response teacher.

After deciding the course, we had to divide the students into two groups. The facts of Research Method course are it is a main course for the department students, and it includes the lecture courses and exercise groups. When students go and select the course at Weboodi, the system provides four exercise groups for the registration. To consider the schedule of students, different time and locations are arranged for different exercise groups. Each week has only one course lecture and one exercise. In consideration of the experiment needs two groups, four groups of students in the exercise arrangement would be divided into two groups for the experiment. The exercise groups’ time and students are described: group 1 was arranged in each Thursday afternoon with 26 students, group 2 was selected by 32 students with the time of Thursday afternoon, group 3 has 35 students and its time was arranged at Thursday morning, the last group was taking on Wednesday morning with 27 students. For the sake of balance time and number of participants, group 1 and 4 were decided as “Traditional group” which would be asked to take the online feedback questionnaire as traditional way as Weboodi to give course evaluation. The rest group 2 and 3 were formed up as “Prediction Markets-based group” to participant in Prediction Markets-based course evaluation system. In this way, the experiment would be organized in a relatively objective environment and to prevent much bias from participants’ aspect. Two groups of the experiment could be formed up with one morning exercise group and afternoon exercise group, and the number of participants was approximately balanced.

Later, in second part of experiment, in order to simulate the traditional course evaluation method, an online-based questionnaire survey was designed carefully. According to the Weboodi system, the questionnaire would be distributed to students after the DSR module and students can give their course evaluation on the internet. Its questions were comprised with a copy of existing questions in the Weboodi and several specific prepared questions for this research purpose (see appendix B). Surveys were available for every student from Traditional group via internet.

In addition, before opening the market, two separate groups have had training presentations. The purpose of the training was to briefly describe the experiment scenario including the goal of experiment and the rewards scheme arrangement. Especially, the training for Prediction Markets-based group students had a little difference, besides giving them the briefly descriptions for the experiment, it is more important to let them to
know the basic idea of Prediction Markets and teach them the operations in Prediction Markets-based course evaluation system.

At the end of experiment, afterwards work of interviewing students were arranged for the purpose of collecting the feedback based on their experiences in the experiment, and also looking for interesting further research topics. Students in different groups would be asked different questions. Two types of interview questions are listed in appendix D.

6.4 Results

During the experiment, all the collectable data was saved and backed up in order to support further research. In the following, the results of experiment are presented.

6.4.1 Participants

The first important data of experiment is number of participants. There were 49 students took part in the experiment, 30 out of them were belongs to Traditional group who were arranged to use existing traditional method to give course feedback, and rest 19 students were served by Prediction Markets-based course evaluation system. The number of participants was listed in the table 12.

Table 12. Number of participants.

<table>
<thead>
<tr>
<th>Group</th>
<th>Number of students in the group</th>
<th>Number of students participated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prediction Markets-based</td>
<td>31</td>
<td>19</td>
</tr>
<tr>
<td>Traditional group</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

The complete attendance in Traditional group is because their course feedback session setting was arranged in classroom. At the beginning 10 minutes of last exercise lecture, students were arranged to take an online-based questionnaire which was designed to simulate the service of students’ feedback in Weboodi system. In fact, the response rate in Weboodi system is very low. In Prediction Markets-based group, students was following voluntary principle, due to the experiments only continued for 10 days and it was initially designed for free use. There were 19 students included after sending 31 invitation letters; the participant rate is 61.3%.

6.4.2 Number of trades

Another significant reflection is the number of trades. In total, 19 participants made 149 trades on four contracts. In average, each student had made 7.8 transactions. Figure 14 presented the number of trades ordered by date.
The trading phenomena among students are interesting and fit the expectations of experiment. It can be easily explained. At the first day of the experiment, students were not familiar with the Prediction Markets-based course evaluation system, and markets had not received many activity traders. Then later two days, the transactions were increased by explosive growth. This is because the students had Research Method lectures and exercise on Tuesday, Wednesday and Thursday. It is understandable that students would give opinions and trading in markets after taking the lectures. The same situation happened at 20th and 21th of November again. It obviously reflected that most part of transactions were from the date when students have lectures. From 17th to 19th of November the amounts of trades were low. This is because it was a weekend, in addition on Monday there was no Research Method course in students’ schedule. In summary, students would like to trade the contracts after the lectures, due to the phenomena of most trades happened from Tuesday to Friday when there was related lectures or exercises.

6.4.3 Answers of the questions

In the Prediction Markets-based course evaluation system, four contracts were defined in the markets with corresponding to four aspects of evaluating course quality. Due to they were the key results of experiment and would be used for analyzing in next chapter, the data for each contract was summarized.

For the Question 1: Do the lectures support the learning and understanding of the contents of the course well? This question was designed to evaluate the lectures quality. There were 26 trades in total and most of trades happened in ‘Yes’ answer. The basic data is showed in Table 13.

<table>
<thead>
<tr>
<th>Possible answers (Stocks)</th>
<th>Initial value</th>
<th>Last value</th>
<th>Buy Trades</th>
<th>Sell Trades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.5</td>
<td>0.5755</td>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td>No</td>
<td>0.5</td>
<td>0.4245</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>1</td>
<td>18</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 13. Basic data of Question 1.
Figure 15 shows the price chart of Question 1 during the period of the experiment. The price of each answers moved smoothly. At the end of the experiment, students believed that the lectures support the learning and understanding of the contents of the course well in the chance of 57.55%; while there was 42.45% probability happen with adverse situation.

![Price chart of Question 1](image)

**Figure 15.** Price chart of Question 1.

For the Question 2: Do the course material support the understanding and learning of the contents of the course well? This question was defined to evaluate the students’ satisfaction on course material. There were 23 trades in total and most of them were traded with ‘No’ answer. This is because in research method course, there was no paper materials handed out to students. In the markets, one of students doubted the course material issue in the public comment area by asking question “what course material is? All we have is a copy of the lecture slides”, and then I gave the response that course materials should be the lecture slides or other recommend study materials from the teacher due to in the explorative experiment, there was no an separate authorized person set for answering students’ questions. Moreover there was another student wrote reasons of selecting on “no answer”. He thought the course materials were somewhat lacking. The basic data is showed in Table 14.

**Table 14.** Basic data of Question 2.

<table>
<thead>
<tr>
<th>Possible answers (Stocks)</th>
<th>Initial value</th>
<th>Last value</th>
<th>Buy Trades</th>
<th>Sell Trades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.5</td>
<td>0.4650</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>No</td>
<td>0.5</td>
<td>0.5350</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>1</td>
<td>16</td>
<td>7</td>
</tr>
</tbody>
</table>

The price chart of Question 2 (see Figure 16) shows that students had reflected negative opinions. At the end of experiment, students believe 46.50% probability of the course material support the understanding and learning of the contents of the course well, while there was 53.50% probability happens on adverse opinion.
For the Question 3: Do the exercises support the understanding and learning of the contents of the course well? This question was designed to evaluate the exercises quality of the course. There were 42 trades in total and students’ opinions were positive. The basic data is showed in Table 15.

Table 15. Basic data of Question 3.

<table>
<thead>
<tr>
<th>Possible answers (Stocks)</th>
<th>Initial value</th>
<th>Last value</th>
<th>Buy Trades</th>
<th>Sell Trades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.5</td>
<td>0.6638</td>
<td>34</td>
<td>6</td>
</tr>
<tr>
<td>No</td>
<td>0.5</td>
<td>0.3362</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>1</td>
<td>35</td>
<td>6</td>
</tr>
</tbody>
</table>

Figure 17 shows the price chart of Question 3 during the period of the experiment. The situation of Question 3 was not as stable as previous two contracts. This is because the number of trades was more than above two contracts. At the end of experiment, students believe 66.38% chance the exercises support the understanding and learning of the contents of the course well, while there was 33.62% chance paid on the disagreement.

Figure 17. Price chart of Question 3.
For the Question 4: What is the most typical overall grading of the course by students? This is a question which was designed to overall evaluate the course quality, thus received most trades from students. There were 58 trades in total and students’ opinions were positive. The basic data is showed in Table 16.

Table 16. Basic data of Question 4.

<table>
<thead>
<tr>
<th>Possible answers (Stocks)</th>
<th>Initial value</th>
<th>Last value</th>
<th>Buy Trades</th>
<th>Sell Trades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Poor</td>
<td>0.25</td>
<td>0.1190</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Poor</td>
<td>0.25</td>
<td>0.1189</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fair Well</td>
<td>0.25</td>
<td>0.2033</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Well</td>
<td>0.25</td>
<td>0.4348</td>
<td>34</td>
<td>2</td>
</tr>
<tr>
<td>Very Well</td>
<td>0.25</td>
<td>0.1238</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>1</td>
<td>49</td>
<td>9</td>
</tr>
</tbody>
</table>

Figure 18 shows the price chart of Question 4 during the period of the experiment. The situation of Question 4 was more complex than others. Due to it was designed with four possible answers and also it received most trades in the experiment. At the end of experiment, the course has 11.90% probability of becoming very poor course; 11.89% probability to be graded with poor level; 20.33% chance to be graded well; 43.48% probability to follow the students’ opinion on this is a very well course; and 12.38% chance to be graded very well.

Figure 18. Price chart of Question 4.

Deriving from the basic data of the experiment and reflection of contract price, it was reasonable to understand how Prediction Markets-based course evaluation system reflected students’ feelings on specific course and what were their reactions after the teaching actives. For instance, three of four questions had obtained more positive feelings from students than negative. Except for question 2, most transactions were traded on the answer ‘No’. This could be regarded as reflection of real fact. Besides, in the markets, students did follow “buy low and sell high rule”, it is obviously recognized from buy trades and sell trades actions and also the price chart reflection. Thus, there existed some interesting phenomena which caused by markets nature; for instance, I realized that a small amount of students had adopted speculative strategy in their trading actions. They
did not care to forecast the final results of the questions and got paid off by predicting which answer would happen or not. They increased their stakes by times of buying and selling actions. These could be clearly observed in each contact that few of sell actions always happened when the relevant got a high price. In generally, students prefer to express their expectations by buying contracts, and few would sell contacts only if they were senior player for the markets. These interesting situations of markets nature would be discussed later or acted as future research questions.

In Traditional group, students were able to give their course feedback via a questionnaire. Due to trading questions were designed in correspondence to four specific questions of questionnaire in Weboodi, therefore the answers of four questions were listed out in Table 17.

- Question 1: How well did the lectures support the learning and understanding of the contents of the course?
- Question 2: How well did the courses material support the learning and understanding of the contents of the course?
- Question 3: How well did the exercises support the understanding and learning of the contents of the course?
- Question 4: I would like to grade the Research Method course with.

**Table 17. Results from questionnaire.**

<table>
<thead>
<tr>
<th>Value</th>
<th>Question 1</th>
<th></th>
<th></th>
<th>Question 2</th>
<th></th>
<th></th>
<th>Question 3</th>
<th></th>
<th></th>
<th>Question 4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td>Percent</td>
<td>Count</td>
<td>Percent</td>
<td>Count</td>
<td>Percent</td>
<td>Count</td>
<td>Percent</td>
<td>Count</td>
<td>Percent</td>
<td>Count</td>
<td>Percent</td>
</tr>
<tr>
<td>very poorly</td>
<td>1</td>
<td>3.30%</td>
<td>0</td>
<td>0.00%</td>
<td>0</td>
<td>0.00%</td>
<td>1</td>
<td>3.30%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>poorly</td>
<td>1</td>
<td>3.30%</td>
<td>4</td>
<td>13.30%</td>
<td>1</td>
<td>3.30%</td>
<td>1</td>
<td>3.30%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fairly well</td>
<td>11</td>
<td>36.70%</td>
<td>10</td>
<td>33.30%</td>
<td>2</td>
<td>6.70%</td>
<td>5</td>
<td>16.70%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>well</td>
<td>14</td>
<td>46.70%</td>
<td>13</td>
<td>43.30%</td>
<td>17</td>
<td>56.70%</td>
<td>21</td>
<td>70.00%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>very well</td>
<td>3</td>
<td>10.00%</td>
<td>3</td>
<td>10.00%</td>
<td>10</td>
<td>33.30%</td>
<td>2</td>
<td>6.70%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>100%</td>
<td>30</td>
<td>100%</td>
<td>30</td>
<td>100%</td>
<td>30</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Each question was answered by 30 students in total. And in the table, it clearly shows what the result for each answer is.
7 Experiment analysis

At the beginning of building the experiment, its original purposes were to test the feasibility of adopting Prediction Markets in course evaluation and to compare the differences between Prediction Markets-based course evaluation system and traditional course evaluation methods. In the following, the data from the experiment were analyzed in two aspects. On the one hand, we compared the forecasts from Prediction Markets-based group and the results from Traditional group. On the other hand, the performance of Prediction Markets-based course evaluation system was analyzed based on four issues of course evaluation. In addition, the results from the interviews were included.

7.1 Defining the comparison method

Seemingly, the contracts questions in Prediction Markets-based course evaluation system were designed as similar as the questions in Weboodi. Both of them used to evaluate the course quality from four defined aspects including lectures, course materials, exercises and general feelings, but they could not be analyzed directly. This was because the different structures of the answers were applied in two groups. In Weboodi, the questions were matched with five answers consist of very poor, poor, fairly well, well, very well, which were designed by Likert scale types of measure. While in the Prediction Markets-based course evaluation system, the answers of contracts were just defined with ‘Yes’ and ‘No’. In order to make a relatively comparing on two different structured answers, it was necessary to define a method to transform the answers into the same scale type.

The data transforming method was decided with calculating the average value of the answers. At the beginning, the answers were valued by numbers. For the Likert scale answers of questionnaire, they were marked with number 1 to 5 (very poor = 1, poor = 2, fairly well = 3, well = 4, very well = 5), and then same to the answers of contracts (No = 1, Yes =2). By taking the number of people of specific answer, the average of each question would be calculated out.

After calculating the data, the question of Traditional group would be marked by $\overline{V_q}$, while $\overline{V_p}$ instead to the value of contracts question in Prediction Markets-based group. In order to avoid the inequality from answers value definition, the final value of contracts ($\overline{V_p}$) which were designed only with ‘Yes’ and ‘No’ answer would be taken into comparison after multiply by $\frac{5}{2}$. Thus, we would see the difference of those two methods. The results were listed and analyzed in the next section.

7.2 Comparison between two different group students

After transforming the answers into same scale, $\overline{V_p}$ and $\overline{V_q}$ are listed in the following table 18. Due to each of them represented the questions value in the different groups, the obvious difference indicated that the way of using Prediction Markets-based course evaluation system could produce totally different results for evaluating the course quality compare to traditional methods. If we assume that participants’ selections were rational
and the value of questions could fully reflect the course quality, then the different results showed differences between two course evaluations.

**Table 18. Value of questions in two different groups.**

<table>
<thead>
<tr>
<th>Question</th>
<th>$V_P$</th>
<th>$V_q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question1 (lecture aspect)</td>
<td>4.9</td>
<td>3.6</td>
</tr>
<tr>
<td>Question2 (course material aspect)</td>
<td>2.8</td>
<td>3.5</td>
</tr>
<tr>
<td>Question3 (exercise aspect)</td>
<td>4.9</td>
<td>4.2</td>
</tr>
<tr>
<td>Question4 (overall grades)</td>
<td>3.6</td>
<td>3.7</td>
</tr>
</tbody>
</table>

Obviously, in the Prediction Markets-based group, students gave higher grades for lectures teaching and exercises practicing aspects in which $V_P$ values 4.9 and $V_q$ values 3.6 at Question 1 (lecture aspect), and $V_P$ equals 4.9 and $V_q$ equals 4.2 at Question 3 (exercise aspect). Besides, Prediction Markets-based group gave 3.6 as overall grade for the whole course while in Traditional group received 3.7 grades. Especially, there exists difference between two course evaluations in course material aspect, while the value of $V_P$ was quite lower than $V_q$, this may because of in the Prediction Markets-based group, students not only realized there was no paper-based materials but also discussed the issue at the Question 2 website.

After comparing the results of four questions, it is clear that the results of Question 1, 2, and 3 had significant differences between two groups, except for the value of question 4 $V_P$ and $V_q$ were almost close to. This is because of the adoption of difference structure answers. In Question 1, 2, and 3, Prediction Markets-based course evaluation system was using ‘Yes’ and ‘No’ to act the possible answers, while in the Question 4, the answers were set by Likert scale method.

### 7.2.1 Response rate

It was already known that few students would like to give their course feedback via Weboodi. In the experiment, there was an extra question of asking students about their previous participation situation added in the questionnaire for Traditional group students. The result showed (see Table 19) that only 16.7% students would usually participate in the feedback system of Weboodi. This also echo the descriptions from the interviews on current and former deputy directors for teaching who indicated that the response rate in Weboodi is pretty low, the number of students who really participant to the course evaluation in Weboodi is less than 15%. While the situation of the experiment was totally different, the markets received 19 students participation after sending invitation letter to 31 students. The active rate was 61.3%, which was much higher than the response rate in Weboodi system. And in average each participant student had made 7.8 transactions in the markets, this means they gave their feedback many times for the course, compare to the feedback system in Weboodi, students only can give their feelings at the end of the course.
Table 19. Response rate VS Prediction Markets activity.

<table>
<thead>
<tr>
<th></th>
<th>YES (Active)</th>
<th>NO (Inactive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did you usually participate in the feedback system of Weboodi? (questionnaire)</td>
<td>16.7%</td>
<td>83.3%</td>
</tr>
<tr>
<td>Prediction Markets-based course evaluation system (actual number of participation in experiment)</td>
<td>61.3%</td>
<td>38.7%</td>
</tr>
</tbody>
</table>

In fact, this experiment was organized under a relatively good motivation pattern based on voluntary principle. The results of response rate showed here probably cannot give exact proof of that Prediction Markets-based course evaluation system would receive more students to join. However, in this specific experiment, the fact shows Prediction Markets-based course evaluation system brought positive influence on the response rate issue. And also we found an open question of whether Prediction Markets-based course evaluation system could receive higher response rate than traditional course evaluation system or not.

7.2.2 Time consuming

Here, we discussed time consuming issue on two aspects. One was from the perspective of teachers and the other one was under the consideration on students.

As it has been mentioned in Chapter 5 course evaluation, two types of course evaluation method. Paper-based course evaluation is time consuming due to teachers had to spent time to collate and interpret data. But for Prediction Markets-based course evaluation system, the data from students could be immediately displayed as same as online-based course evaluation. Both teachers and students could clearly understand the meaning of data. Especially for teachers they do not need to spend time to analyses the information from students; even they have many courses running at the same time.

For students, the situation is different. In the traditional course evaluation methods, the times of students participation are defined. Students are always asked to take a questionnaire at the end of the course or during the course. But in Prediction Markets-based course evaluation system, students evaluate the course continuously by making trades. They are not evaluating the course in a certain time by a predefined questionnaire but following the progresses of course instruction. Students have to spend time to trade contracts in order to increase their account value. Compare with the traditional methods, Prediction Markets-based course evaluation system is time consuming, especially for those students who have several courses running at the same time.

7.2.3 Accuracy of data

Not like political elections, or sports games, finally they will have objective results. In the course evaluation, it can be regard as a tool to collect students’ opinions. The opinion is always difficult to be identified by accurate. Thus we cannot make a conclusion of which course evaluation method has provided accurate data based on the experiment. But in this experiment, Prediction Markets-based course evaluation system indirect provided relatively closer results to the facts. It showed in the Question 2, students of Prediction Markets-based group argued the course materials issues in the markets, which was consistent with the fact that there was not provided any paper materials during the course.
While in the Traditional group, students were acting as nice roles to give all the questions positive answers, even ignored the materials issues. In addition, the continuously reflections and high degree of participations can somehow serve to get true feelings from students.

7.2.4 Interaction between students and teachers

Interactions could be a significant advantage for Prediction Markets-based course evaluation system which benefit by the dynamic nature of markets. Students could trade contracts based on their expectations at any time, and also write down the comments or describe the reasons of buying or selling contracts. From trading actions, students could express their feelings or feedback for the course. On the other hand, teachers could understand the situations of the course. The continuously reflection could not only reflect the course quality but also let teachers to track students actions in real time but also figure out the problems. When teachers realize the problems or issues from the Prediction Markets-based course evaluation system, they would be corrected immediately and avoided in the rest of lectures. And, students would also get benefits in the class. Thus the worry of feedback from current students for improving course would not make any sense for them is fixed.

Based on the comparison between two different groups, the difference between Prediction Markets-based course evaluation system and traditional system was clear. Prediction Markets-based course evaluation system was a totally different method of evaluating course quality. In this specific experiment, we may conclude that it performed better in above four common issues of course evaluation system.

7.3 Analysis of the results from interviews

After the experiment, students had their own opinions on the Prediction Markets-based course evaluation system. Some of them hold positive attitudes, they would like to participant in the markets to evaluate course quality. Some of them just gave negative views. They thought it would hardly be adopted in the reality.

One student from Traditional group mentioned that he only give feedback when teacher required and said “It is not so good to give feedback continuously, because students may give the feedback based on their emotion changes and the emotion is difficult to control and it cannot be trusted. Sometimes students’ feedback is strange to understand.”

There was another student from Traditional group. From him “in traditional way, feedback is given after the course, so it does not help me anymore at that point.” So he put much interest on the system and indicated that it is fun and easy to express the students’ opinions during the course. In addition, he really would like to use the system if it is available in the future.

In the Prediction Markets-based group, one student was interviewed. He was the “richest” trader in the markets who gained 20% benefits based on his skillful trading actions. This was because he had already knew the prediction markets before the experiment. He realized that by using Prediction Markets-based course evaluation system, his own opinions for the course were somehow ignored; instead of fully expressing own feelings, he was trying to guess what the common answer from students was. Then he gave a possible solution that in the markets “there should not have the surprise ending, so I should not be trying to guess what were someone else’s questionnaires doing”. In addition, same as other students, he did not like feedback system in Weboodi either, as he
said: “I actually had some bad experiences with Weboodi, I never did the digital feedback from Weboodi, I usually only do them when a professor hands me a paper copy.” However, he showed his interests for the future developments, “If it just only for one course, I would probably skip it. If it is a one market for all the courses and I can keep it up over the courses for my studies, I think that would be fun and interesting.”

The afterwards interview from students did help the research. From their opinions, many issues were recognized which could be used for further research and system optimization. They were discussed later.
8 Discussion

From the previous literature, the key design elements of Prediction Markets were understood. It seems reasonable to make a discussion of designing and implementing Prediction Markets-based course evaluation system based on description of the explorative experiment and its basic results.

The first issue is relating to the contracts design, we already knew the importance of contracts for the Prediction Markets. It was not only just to decide which type of contracts would be used, but also to deeply consider the structures, contents of questions especially for course evaluation system, because unlike the existing Prediction Markets-based applications, the purposes of adopting Prediction Markets were to evaluate the course quality by aggregating the perception of students and build an easy way to connect students and teachers. In addition, the data would also be helpful for university management. So, the contracts should be carefully designed before published in the markets. In the experiment, the contract questions had not been fully thought over. This was because they were set up in correspondence with the feedback system in Weboodi. Thus, we could compare and analyze the difference between Prediction Markets-based course evaluation system and Weboodi. It may hardly be agreed that the contracts questions which derived from Weboodi were perfectly fit the situation of Prediction Markets. In order to make better contract questions for Prediction Markets-based course evaluation system, there may require professional pedagogic knowledge and to redesign the questions underlying the consideration of different courses and objects. For the future system development, the issues including what the optimal number of questions will be, when to publish them into markets and how about opening the rights of lunching new questions for students would be took into consideration, and also these can become interesting topics for the further research.

Trading mechanism is another design issue of Prediction Markets. Continuous Double Auction and Automated Market Maker are the two of most common used. In the experiment, the Prediction Markets-based course evaluation system was equipped by Automated Market Maker, it performed very well. Thus the Automated Market Marker is a suitable trading mechanism of designing a Prediction Market for course evaluation purpose. Because other trading mechanism could be very complicated and based on the facts of Automated Market Maker tend to be very illiquid (Pennock, 2004), it can execute the bids and asks immediately, unlike the standard Continuous Double Action, that the orders have to wait until there is a match. As Prediction Markets-based course evaluation system was supposed to become as a small and private system, the limited number of transactions might not immediately match up. In addition, most of students were not experts for Prediction Markets, other trading mechanisms would let them confused and they might feel it is too complicated to trade the contracts. Therefore, to build a Prediction Markets-based course evaluation system, Automated Market Maker trading mechanism should be highly recommend. But it still needs to be reminded that different situations may have different considerations.

In a decision of incentives, it plays very important role in Prediction Markets-based course evaluation system. A well designed incentive scheme will not only motivate traders to express their opinions sufficiently via trading on contracts, but also attract more and more people to take part in the markets. In the experiment, the markets were operated
by play-money and there included internal and external intensives. Internal intensives were ranking list and dashboard function from Inkling markets, while external incentive included small gifts and lottery setting. After the experiment, the fact that 19 out of 31 students joined the markets proved those intensives worked well. But it was implemented in a specific course and few students were involved. If we consider adopting Prediction Markets-based course evaluation system in many courses or even whole university level, the existing incentives might be improved except that play-money was definitely decided. For instance, combine the system with social media concept in order to ensure students visiting times, or define course evaluation system as a part of students study duties. To find out a better incentive for the system, more experiments are needed.

To build a well Prediction Markets-based course evaluation system, just to consider on the common key elements of Prediction Markets are not enough. Due to this is an interdisciplinary topic, the fields of Prediction Markets and course evaluation should be linked up to consider. When adopting Prediction Markets in course evaluation, besides to consider the issue of how to efficiently evaluate the course quality by designing suitable contracts questions, it also require to analyze on the environmental issues such as schedule of course arrangement, types of course, psychology of students and their behaviors. If the Prediction Markets-based course evaluation system will be adopted in reality of the university level in the future, this experiment which implemented based on one specific course could contribute less. Therefore, in order to make a successful Prediction Markets-based course evaluation system, the objects for the markets should be decided carefully. The relationship between courses, schedule, and students has to be clearly defined. In addition, the technical support which concerns to the trading platform has to be improved. Due to this is an explorative experiment, it was just run by existing service. As we know, the trading platform which provided by Inkling company is not designed for course evaluation purpose, after customization settings it could just fit the situation of the experiment. During the experiment, its limitations were recognized. From a long-term and realistic consideration, the particular trading platform should be fully redesigned, redeveloped, and implemented in order to fit the requirements of complex environment. The issues of friendly-user interface, system security, and risk management should be taken into consideration.

From the observation of speculative actions in the experiment, the other considerations should be concerned to the nature issues of markets itself. The phenomenon of few students pursuing the speculative profits via frequently buying and selling on a specific contract was obvious. It seems like they did not care about the “true” answer of the contract questions. Since, at the beginning of the experiment design phase, the different types of trading strategies did not be studied. After the arbitrage phenomena were observed, we learnt the knowledge of trading strategies. Basically, there are four types of possible trading strategies existing in the markets, and somehow they will be combined to operate. They are the arbitrage strategy, the expectations strategy, the risk-adjusted expectations strategy and the speculative strategy (Berlemann & Nelson 2002). We thought the issue of trading strategy is significant for developing a Prediction Markets-based course evaluation system and derived many questions. For discussion, what will be the influence of the final results from those speculative actions? If those actions are helpful for revealing the students’ opinions of the course, definitely when designing a Prediction Markets course evaluation system, we have to make the markets more open and encourage the speculative actions. On the other hand, if the actions will affect the accuracy of data which represent students’ opinions, then we have to consider forbidding those speculative actions happen. The possible methods are to set up some restrict functions or make some changes for the structure of the Prediction Markets-based course evaluation system. In the future, the issues relating to the trading strategies should
be studied. In a summary, for building a Prediction Markets-based course evaluation, different trading strategies have to be taken into consideration. In addition, other nature issues of markets should not be ignored including trading period, price limits, and transaction quantity limitation.

Another issue towards to the limitation of Prediction Markets itself. Normally, it is known from the definition of Prediction Markets that it is designed for forecasting the results of the future events, the relevant contracts would be paid off depending on the actual outcome of the events. In this experiment, those contracts in Prediction Markets-based course evaluation system were paid off relying on the results of questionnaires from Traditional group. The bias of Prediction Markets-based group students was obvious. Students were expressing their own opinions underlying the course quality somewhat mixed with guessing what others opinions are. This issue was mentioned in the interview of one student also. There was a contradiction of setting Prediction Markets-based course evaluation system and traditional questionnaire method at the same time. At this point, the research seems like did not so helpful for course evaluation purpose, instead of choosing a more complex method. To solve the contradiction there should have no need for Prediction Markets to get supports from questionnaires for determining the payoff value of relevant contracts. Recently, the major shortcoming of Prediction Markets was solved by a rising interest named as Second-Generation Prediction Markets which referred as preference markets and idea market (Slamka, Jank, & Skiera, 2011). For the Second-Generation Prediction Markets, the payoff of contracts do not rely on the actual outcomes of events, in other words the underlying situation of the specific objects which were used to form up contract questions in the markets would never be known. The payoff values are determined by following three alternative mechanisms. They are based on volume-weighted average trading price (Lacomb, Barnett, & Pan, 2007), the last fixed price (Chen & Plott, 2002) or the last fixed price which is decided on random point of closing the markets (Dahan, Soukhoroukova, & Spann, 2007). These possible payoff mechanisms could be sufficient for optimizing the Prediction Markets-based course evaluation system in the future.

In conclusion, we provide a possible solution of adopting Prediction Markets in course evaluation. Basically, there are six steps to design and implement a Prediction Markets-based course evaluation system.

1. Identify the objects which would be forecasted in the markets.
2. Understand the relevant environment including internal environment such as the course arrangement, software, hardware and external environment such as school culture, students’ preference and teachers’ requirements.
3. Define the key design elements of Prediction Markets carefully with respect to the environment and objects.
4. Develop or select a suitable trading platform.
5. Organize preparation work such as operate training.
6. Implement and run the Prediction Markets-based course evaluation system.

The guideline of building a Prediction Markets-based course evaluation system is easy to understand. But the other issues which were observed in this thesis are more important. Theoretically, a well-designed Prediction Markets can provide valuable information of revealing students opinions. After the conceptual system is ready, the issues and relevant suggestions which were discussed above should be taken into consideration.
9 Conclusion

As a summary, the basic idea of this research was to extend the potential usage of Prediction Markets into course evaluation area. Firstly, multidiscipline had been learnt in order to build a Prediction Markets-based course evaluation system. Then the explorative experiment based on the system was organized for the purposes of verifying the feasibility of adopting Prediction Markets-based course evaluation system in a real course case and comparing the difference with the traditional in use system (Weboodi). The system was designed as a new evaluation tool for continuously revealing students’ opinions. After the experiment, the results were analyzed and they could be used as answers for the initial defined research questions. The answers were described as follow.

- Can the Prediction Markets be adopted in course evaluation?

Both from theoretical aspects and practical aspects, this research demonstrated that Prediction Markets do have potential usability of being adopted in course evaluation. At the beginning, the inspiration was from the idea of Prediction Markets itself including advantages on prediction and continuous information reflection. We studied the relevant knowledge and combined Prediction Markets and course evaluation in order to build the Prediction Markets-based course evaluation system. The already made explorative experiment and its promising results just enough to make us believe that Prediction Markets-based course evaluation system can not only be adopted but also acted as a useful method for course evaluation purpose.

- What are the requirements for developing Prediction Markets-based applications?

The purpose of this question was to summarize the foundations for designing a Prediction Markets-based course evaluation system. The answers were from previous research and relevant Prediction Markets-based applications and experiments. Since different kinds of Prediction Markets have different concerns that would relate to the different objectives, the focuses of this research were to study the general key design elements for developing Prediction Markets including contracts, trading mechanisms, incentives, and traders (Luckner et al., 2011, p.11). Both of them were studied and described in Chapter 4. To build a Prediction Markets is a holistic work; the elements could not be separated from each other.

- What are the potential improvements of Prediction Markets-based course evaluation system comparing to traditional course evaluation method?

Essentially, the Prediction Markets-based course evaluation system is different to traditional course evaluation method. The former is acting as efficient information aggregation method that based on the idea of wisdom from crowds, while the traditional method is based on the formative questionnaires that will be used to serve students after the courses. Therefore, in this research the differences between Prediction Markets-based course evaluation system and traditional method were only analyzed by four well known aspects of course evaluation system. They are response rate, time issue, accuracy of data and interactions (Anderson et al., 2005; Tucker et al., 2003). According to the comparison of two different student groups in the experiment, it could be concluded that Prediction Markets-based course evaluation system is an efficient system for continuously revealing
students’ opinions and it could provide rich data for teachers and school managers as well as the interactive feedback is the highlight point. In addition, it may also attract students’ interests better.

- How to design and implement an efficient Prediction Markets-based course evaluation system?

This is the main question for the research. To explore the feasibility of Prediction Markets-based course evaluation system, it is adopted by an explorative experiment. The experiment showed clearly that to build a Prediction Markets-based course evaluation system can be a challenge. Because the experiment is the first trial of combining Prediction Markets and course evaluation, and it is difficult because the concepts of Prediction Markets and course evaluation are intrinsic complex.

When building a Prediction Markets-based course evaluation system, we followed the key design elements of Prediction Markets. They are types of contracts (Wolfers & Zitzewitz 2004), trading mechanisms (Slamka, Skiera, & Spann, 2012), incentives (Luckner et al., 2011, p. 24), selection of traders (Luckner et al., 2011, p.35) and trading platform (Luckner et al., 2011, p. 43). After the experiment, we discussed many issues of designing and implementing Prediction Markets based on the experiment and presented a possible guideline in the chapter 8. The answer of this question can be briefly described as when to design and implement an efficient Prediction Markets-based course evaluation system, it is not only to take Prediction Markets key design elements into consideration, but also have to care on many aspects including the relations between courses, teachers and students, and trading strategies, especially, the junctions between Prediction Markets and course evaluation. The successful design and implementation of Prediction Markets-based course evaluation system depend on the careful planning and overall consideration.

From the course evaluation view, the experiment could be recognized as a successful course evaluation system which reaches the purposes both on summative and formative evaluations. Due to this is just an explorative research on combining Prediction Markets and course evaluation, many interesting issues have been recognized. For instance the contracts design issues, the speculative phenomena, and the incentive design. It is imaginable that Prediction Markets-based course evaluation system would show its impacts in the future. However, it is not a mature system yet. In order to optimize the Prediction Markets-based course evaluation system and realize its value, more studies are required.
References


Appendix A. Primary studies


Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*, Little, Brown.


Appendix B. Questionnaire of course evaluation

Course Feedback

Here is the Questionnaire which is totally designed for collecting course feedback. We need to hear from students to help improve the teaching quality, and also to understand how your study is. Please take a few minutes to answer the following questions. Your answer is anonymous and the data is under protected. Thank you for your participation. Your contribution makes course better.

1) How well did the lectures support the learning and understanding of the contents of the course?*
   ( ) very poorly
   ( ) poorly
   ( ) fairly well
   ( ) well
   ( ) very well

2) How well did the teacher present the contents of the course?*
   ( ) very poorly
   ( ) poorly
   ( ) fairly well
   ( ) well
   ( ) very well

3) How well were the contents of the course planned and presented?*
   ( ) very poorly
   ( ) poorly
   ( ) fairly well
   ( ) well
4) How well did the amount of the work correspond with the requirements of the course and the study points?*

( ) very poorly
( ) poorly
( ) fairly well
( ) well
( ) very well

5) How well did the course material correspond with the topics and contents of the lecture?*

( ) very poorly
( ) poorly
( ) fairly well
( ) well
( ) very well

6) How well did the courses material support the learning and understanding of the contents of the course?*

( ) very poorly
( ) poorly
( ) fairly well
( ) well
( ) very well

7) How well did the course support your previous studies?*

( ) very poorly
( ) poorly
( ) fairly well
8) How well did the arrangements and the schedule of the course work?*

( ) very poorly
( ) poorly
( ) fairly well
( ) well
( ) very well

9) How well did the course meet your expectations?*

( ) very poorly
( ) poorly
( ) fairly well
( ) well
( ) very well

10) How well did the exercises support the understanding and learning of the contents of the course?*

( ) very poorly
( ) poorly
( ) fairly well
( ) well
( ) very well

11) How well did you get assistance in learning and/or doing the exercises? *

( ) very poorly
( ) poorly
( ) fairly well
12) I would like to grade the Research Method course with*

( ) very poor
( ) poor
( ) fairly well
( ) well
( ) very well

13) I attended the lecture approximately (%)*

( ) 0-10
( ) 10-20
( ) 20-30
( ) 30-40
( ) 40-50
( ) 50-60
( ) 60-70
( ) 70-80
( ) 80-90
( ) 90-100

14) I attended the exercises approximately (%)*

( ) 0-10
( ) 10-20
( ) 20-30
( ) 30-40
( ) 40-50
The questionnaire above is the typically traditional way for collecting course feedback which has been used in Weboodi.

15) How well did the traditional way of giving feedback support delivering your opinion?*

   ( ) very poorly
   ( ) poorly
   ( ) fairly well
   ( ) well
   ( ) very well

16) Did you usually participate in the feedback system of Weboodi?*

   ( ) yes
   ( ) no

17) Please explain your previous selection that why you usually/ didn't usually participate in the feedback system of Weboodi? *

   _________________________________________________________

18) Any ideas for improving recently course feedback system?*

19) Four students will be randomly selected and rewarded with one coffee ticket. If you want to join the lottery, Please leave your name and email address.

   First Name: ____________________________________________
   Last Name: ____________________________________________
Thank You!

Thank you for taking our survey. Your response is very important to us.
Appendix C. Questionnaire of interview for university staffs

1. Will you go and check course feedback in Weboodi when the course ends?
2. Do you think course feedback is useful? How do the feedback results influence your future teaching plan?
3. How important is the feedback for the teaching quality improving?
4. Do you think that feedback from students represents how their learning was?
5. What are the problems for you (teachers) to receive students’ real voice (feelings)?
6. Do you think giving course feedback often be set after all the teaching activities done is a good way?
7. Some students argued that the feedbacks from them are useless for the current student. How do you think about it?
8. Have you used any other method for collecting feedback?
9. Would you like to adopt a new way for collecting students? (Prediction Market-based Course Evaluation)
Appendix D. Questionnaire of interview for students in the experiment

For Prediction Markets-based group

1. Did you know Prediction Markets before the experiment?
2. What is your first impression when you get in touch with Student Feedback Markets?
3. What is your feeling during using Student Feedback Markets? (Do you think the past two weeks make you good experience of giving course feedback?)
4. Did you know you were playing well with Student Feedback Markets? During the experiment, did you feel any trouble to use it?
5. After the experiment do you feel you understand Prediction Markets? What is Prediction Markets?
6. Did you think the first training lecture and the introductions on the front page are useful for you? If not, how should they be improved?
7. What are the differences between Student Feedback Markets and Traditional feedback upon the purpose of delivering your feedback (Weboodi)?
8. Would you use this kind of new service if it is available in the future?
9. Any improvements for the Student Feedback Markets?

For Traditional group

1. Do you think your feedback will make any influence for course improving?
2. Has your feeling always been same for each teaching activities? If it is changed, how you express it?
3. Do you think the traditional way of giving feedback should be changed? Would you like to use Prediction Market-based Course Evaluation?