The 54^{th} Annual Conference of the New Zealand Association of Economists

Wellington, July 3-5, 2013

ARBITRAGE IN PREDICTION MARKETS

Tram Cao

Victoria University of Wellington

PhD candidate

May 2013

I gratefully acknowledge the support of my PhD suppervisors Prof. Dean Hyslop, Dr. John Randal, Prof. Lew Evans and Prof. Neil Quigley (in 2010-2012). I am grateful to Matt Burgess for providing me with data of iPredict. Thanks to Dr. Andrew Fung for valuable feedback.

Abstract

A prediction market is a relatively new form of financial market whose ultimate purpose is to predict the outcome of an uncertain future event. This paper examines the extent of arbitrage in prediction markets and its implications for market efficiency. The most straightforward arbitrage opportunity in prediction markets is to exploit any divergence of the aggregate price of binary contracts predicting the same event from the contract payout. The existence of arbitrage is found in OCR prediction markets; however, its extent is small and negligible. The results found in this study provide supporting evidence for the efficiency of prediction markets.

1 Introduction

Prediction markets are a new form of financial market which was first developed in 1988 by three economists in Iowa College of Business. It served as a common place where people could trade on contracts whose price and payout depended on the United State presidential election in that year. The ultimate purpose of prediction market is to offer a vehicle to aggregate dispersed information about an uncertain future event. Unlike traditional financial markets, there are no real commodity or assets in prediction markets. Market participants trade contracts based on future events. Traders are rewarded for correct prediction with either real money or virtual money. For example, in order to predict the result of the General Election in New Zealand in 2011, iPredict[1] launched three contracts; each pays out $1 per share held if there will be a National or Labour or neither-National-nor-Labour Prime Minister after the 2011 Election and nothing otherwise. Those who believe in the National Party's victory will buy as many shares of National contract as he can.

Prediction markets are most valuable in situations where no futures market in the relevant commodity actually exists, for instance political events, and markets in which trading in the underlying commodity is so thin that it is not possible to discern an accurate market price. In these situations, prediction markets can provide a vehicle for people to trade on relevant information and valuable means of informing market participants.

Prediction market is also known as idea market, information market, decision market, forecasting market, artificial market, electronic market and virtual stock market. Berg and Rietz (2003) defined prediction markets as markets that run for "the primary purpose of aggregating information so that market prices forecast future events".

[1] iPredict is the first real-money prediction market in New Zealand
www.ipredict.co.nz

As a forecast tool, it is crucial for a prediction market to be efficient so that prices can serve as an indicator of the likelihood of a future outcome. Despite an increase in the number of studies claiming the efficiency of various prediction markets and the accuracy of prediction market prices, there has been no study investigating the efficiency of prediction markets in terms of the extent of arbitrage in those markets. Arbitrage should be the first element to be examined because the no-arbitrage condition is a necessary condition for market efficiency.

The theoretical foundation of the efficacy of prediction markets is based on the Efficient Market Hypothesis Hayek (1945). Information is never either in integrated form or possessed by an individual, instead it is dispersed among many individuals and in many cases bits of information owned by individuals are contradictory. A market serves as a common place where individuals possessing different information can trade on the basis of the information they have and as trading they reveal their information. In an ideal world, information is communicated to all and market prices will be the best indicator of the value of the commodity.

There are many types of contracts in prediction markets, each is designed to aggregate different forecast of future events (Wolfers and Zitzewitz, 2004). The most common is binary contracts which pays out only when the underlying event occurs and nothing otherwise. These contracts are designed so that they have a direct interpretation as the market's expectation of the likelihood of the underlying event. Wolfers and Zitzewitz (2005) provided sufficient conditions under which prediction market prices can be interpreted as an exact probability that an event will occur. For example, the National contract mentioned above pays out $1 if New Zealand has National Prime Minister after the 2011 Election and $0 otherwise. The contract price at a point in time, for example, is 70 cents; this implies that the market aggregately expect 70% likely that the Prime Minister is a National.

Under those conditions, prices are unbiased and accurate predictors of probabilities. The unbiasedness of prediction market prices is an essential condition of the Efficient Market Hypothesis: prices reflect fully information available of the underlying event. In generalizing their model, Wolfers and Zitzewitz (2005) first relax the assumption that traders' budgets are orthogonal with beliefs and replace it with that the prediction market prices are a wealth-weighted average of beliefs among market traders. Second, they relax the assumption of log utility and calibrate alternative utility functions. They show with both theoretical and empirical evidences that the prediction market prices can deviate from the mean belief but that this deviation is typically small.

2 Theoretical Background

The definition of market efficiency in literature has two main features: (1) prices fully reflect all available information, and (2) there are no trading strategies that produce positive, expected, risk-adjusted excess returns. The Efficient Market Hypothesis asserts that in an efficient market prices reflect information as soon as the information arrives. The news spreads very quickly and is incorporated into the prices without delay (Fama, 1969). An implication is that no trading rules would enable an investor to obtain returns that are greater than those that could be obtained from holding a random portfolio with comparably equal risk.

In other words, if prices fully reflect all available information, then it is generally true that there exists no trading strategy that produces risk-free positive expected (risk-adjusted) excess returns. Any advantage to trading on new information will be eliminated as soon as they appear.

In principle, arbitrage is defined as "the simultaneous purchase and sale of the same , or essential similar, security in two different markets for advantageously different prices (Sharpe and Alexander, 1990). Arbitrage involves no negative cash flow at any probabilistic state and a positive cashflow in at least one state. In simple terms, it generates a non-zero probability of a risk-free profit. Under the assumptions of a perfectly efficient market, market participants have access to all related information. Traders detect an arbitrage opportunity as soon as it arises and take advantage of it by bidding prices up and down until the mispricing is eliminated. If an arbitrage opportunity is considered public information then an efficient market (at least in the semi-strong state) should expect arbitrage to be eliminated very quickly because a sufficient number of traders will trade in order to take advantage of it and eventually bring price toward its arbitrage-free level. In other words, the existence of arbitrage itself might not a claim of market inefficiency if it is quickly driven away by the pursuit of profits. However a persistent arbitrage opportunity may indicate the inability of a market to eliminate the mispricing.

There exists no benchmark of determining how fast arbitrage opportunity is to be eliminated so that it does not indicate market inefficiency. The speed of elimination of arbitrage varies and determined by the characteristics of individual markets itself, and of traders trading in the market. Those factors include transaction costs, market liquidity, trader's attitude toward risk and their preference to trading strategy, trader's price assessment, trader's search strategy, the allocation of relevant information among market participants and trader's budget constraints. Transaction costs are added to the costs of buying and selling assets and consequently affect the profit optimizing behavior of traders, thus they must be taken into account. Naturally traders only engages into trading at a market price where it promises a positive return which is sufficient to cover all transaction costs. The presence of transaction costs in disequilibrium markets, thus, is expected to discourage trading by reducing expected returns and consequently restrain the price movement toward its equilibrium level. Therefore, transaction costs reduce the speed of price convergence to equilibrium. A liquid market should expect that arbitrage is detected and driven away very quickly because it has a large number of active traders watching the market closely. Another element constituting to market liquidity is the bid and ask spread; a liquid market which has a small bid and ask spread provides market participants with greater monetary incentives to trade. A large bid and ask spread, in the other hand, reduces expected trading profits as it plays a role of transaction costs and as a result, discourages trading which eventually reduces the chance mispricing is driven away. The speed at which a mispricing is corrected is also affected by trader's characteristics. Trader's attitude toward risk, preference toward trading on certain commodities, strategy of price search and rationality have impact on trader's utility optimizing behavior and how he trades and influence how prices are set. Budget constraints also play a role in incomplete markets; traders who detect arbitrage opportunities but are restrained by budgets might not be able to trade effectively to drive the market price to the point at which they think it should be. Last, information is dispersed among traders and hardly possessed entirely by one individual. The allocation of relevant information among groups of market participants may determine how fast information is incorporated into market prices and eventually affect how quick mispricing is detected. If relevant information is held mostly by a group with tight budget constraints, traders in that group may not be able to take advantage of their information to trade against the market in order to remove the mispricing.

An alternative approach to study the role of arbitrage in the context of market efficiency is indirectly via out-of-equilibrium. Market efficiency requires prices to reflect all relevant information. This only can happen in frictionless markets where there exists no constraint to arbitrage and information is held by all interested parties. As soon as new information arrives, the market will move instantaneously from the current equilibrium position to a new one. From the view of Rational Expectations theory, any disequilibrium will disappear very quickly if not instantaneously and competitive equilibria will be reached quickly and maintained thereafter until the arrival of new information. The First Theorem of Welfare Economics states that equilibrium in competitive markets without externalities and frictions are Pareto efficient. It means that a market in equilibrium is also trading at its efficient allocation. Generally, every market has certain frictions that decelerate the speed of market prices reaching a new equilibrium. There will always be trading out of equilibrium required to move to a new equilibrium. Therefore, during the convergence process, there are trades taking places at out of equilibrium prices and this is where arbitrage opportunities arise. However, these arbitrage opportunities are a result of information and price changes, and in fact, reflect the process of new information incorporated into prices. Thus arbitrage arising from the process of moving to a new equilibrium is temporary and irrelevant to market inefficiency. Plus, the speed of convergence to equilibrium reflects the speed of new information being incorporated in market prices and is a measure of market efficiency.

Once a market is trading at equilibrium, it is Pareto efficient. However, what is more interesting is that out-of-equilibrium trading does converge to an efficient allocation under specific conditions. There have been intensive studies on this subject, including Goldman and Starr (1982) in which they characterize the necessary and sufficient conditions for a disequilibrium to converge to a Pareto optimal allocation in an exchange market where traders involve in multiple transactions in small groups and they trade directly with each other instead of using an intermediary.

Fisher (1981) develops a model of equilibrium stability which gives insight into how the economy with different set of allocations and prices is forced to converge to a new equilibrium after a temporary shock. In the model set up, he allows the awareness of disequilibrium: Traders are aware that they are trading at out-of-equilibrium prices and prices are expected to change constantly. They are also aware of the risk that they may not be able to complete their transactions at the desirable prices due to the constant price changes. This is an advance to previous studies in which traders are assumed to trade naively. The new awareness will consequently change the trading behavior, as now they optimize their utility by taking into account the price changes, the disequilibrium status of market prices and the risk of transaction failure. The model shows that as soon as information arrives, the equilibrium allocation will move to a new level. Simultaneously at the current price arise arbitrage opportunities. Those opportunities will be arbitraged away quickly by the pursuit of profits and the economy will eventually converge to the new equilibrium. This finding at first conflicts with the Rational Expectations theory which argues that there exists no arbitrage at equilibrium. However, in a dynamic economy in which new information arrives continuously, equilibrium allocations are not unique and fixed. Arbitrage will cease as soon as the equilibrium allocation is reached. According to this, arbitrage plays an essential role in the convergence to equilibrium - it is the mechanism of driving mispricing away and moving the economy to new equilibrium corresponding with new information.

Ghosal and Porter (2010)'s study shows similar findings to Fisher (1981)'s. They study the convergence of an equilibrium to an efficient allocation in a pure exchange economy where there are only two traders (pairwise) and orders are matched randomly. They also allow disequilibrium awareness. In their model's set up, trader's preference is presented by Cobb-Douglass utility function. Traders trade cautiously acknowledging that their prediction may turn wrong and their knowledge about the preference of their trading partner is limited. They only involve in trading if it increases their utility. As a result, the trading process is path dependent. Ghosal and Porter (2010) confirms Fisher (1981)'s findings about the sources of instability, the certain convergence to optimal allocation and how it is achieved. The authors agree that if there is no arrival of new information or no new perception of opportunities, the economy will stop moving and stay at its equilibrium (i.e. being stable). The mechanism of how the economy reaches new equilibrium after a shock explained in their study is similar to that of Fisher (1981). The authors also recognize that arbitrage constitutes to the convergence to a new equilibrium by driving away old profitable trading opportunities. Their study provides numerical evidences that the trading process in their model converges with probability 1 to a pairwise optimal allocations. These allocations are Pareto efficient subject to specific conditions.

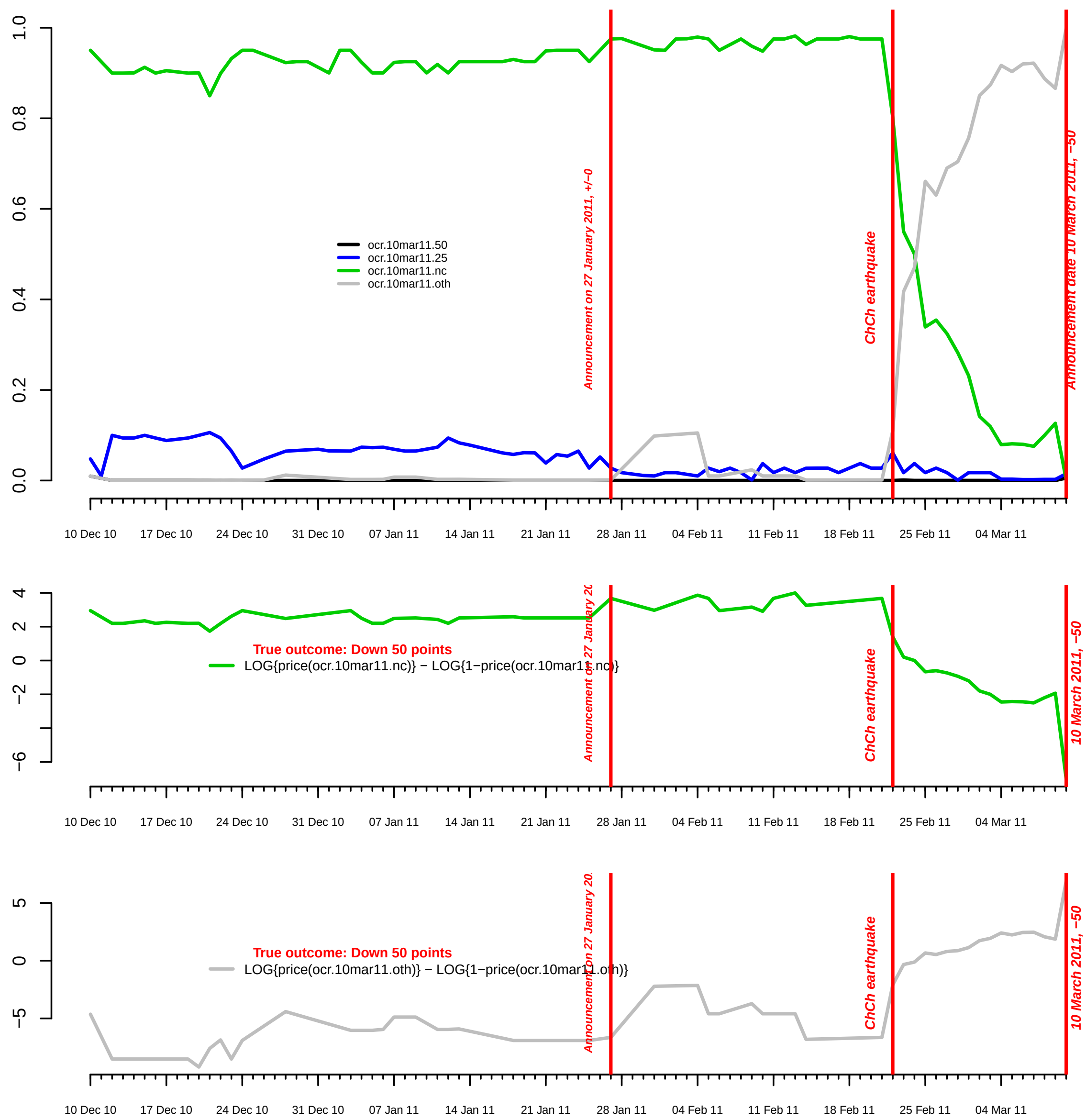
In summary, the convergence of out-of-equilibrium to efficiency provides further insight into the role of arbitrage. Most trading in a market is out of equilibrium and out-of-equilibrium trading converges to efficient allocations under some specific assumptions. Arbitrage arises as the market is trading out of equilibrium, or in other words, arbitrage is a result of price changes. In frictionless markets, as soon as those opportunities arise, traders will arbitrage them away. Arbitrage constitutes to the convergence process by removing mispricing in the market and driving market prices to their equilibrium level.

The speed at which it does so is a measure of market efficiency.

Therefore, the question is no longer whether arbitrage exits but how fast it is eliminated so that the market converges to equilibrium. As mentioned earlier , it is very difficult to specify the exact dynamic process of the adjustment to equilibrium and its speed because it depends on the properties of individual markets. The most relevant finding is in the study of Ghosal and Porter (2010) in which they use a numerical approach to study the average speed of convergence for Cobb Douglas utility function. They look mainly at the estimated convergence in average global utility and assess the performance of cautious trading. Their study shows that the speed of convergence is exponential with the utility function for a range of sizes of economy, both in terms of number of goods and number of agents.

As studying how information is incorporated into market prices and the role of arbitrage in the convergence to equilibrium in iPredict, I have found some supportive evidences. Arbitrage plays the same role in prediction markets as in any other financial markets. The speed of adjustment to new information will also depend on market friction, transaction costs and trader's trading preferences and budget constraints. Trading costs do not include only direct costs incurring with trading but also motivation costs (e.g. the bid and ask spread). In iPredict's markets predicting OCR announcements, the manifestation of new information and the speed of translating to contract prices are given by the odd ratio. The odd ratio of an OCR announcement reflects the market's aggregate expectation of any change in OCR. A positive (or negative) odd ratio implies that market predicts that Reserve Bank of New Zealand (RBNZ) will maintain (or change) the OCR. Figure 1 shows the odd ratio of the contracts predicting OCR announcement on March 10, 2011. The change of the odd ratio from positivity to negativity after the earthquake in Christchurch on February 22, 2011 reflects the shift in market expectation. The market had predicted a no-change outcome since the contract launching and as soon as the earthquake occurred, the market reversed its prediction to that RBNZ would change the OCR as a response to the earthquake event. In fact, in the RBNZ's announcement on March 10, 2011, OCR was to reduced by 50 basic points which was intended to leverage the negative effect of the earthquake to New Zealand economy. This can be seen more clearly if we look at the time series of prices of all contracts in this market. Price of contract ocr.10mar11.nc which bets on the no-change outcome fell sharply from almost $1 to less than 10 cents as soon as the earthquake occurred and offset by a strong rise in price of contract ocr.10mar11.oth which bets on an increase by more than 50 points or a decrease. In this instance, the market responded very quickly to the earthquake event in both contracts (ocr.10mar11.nc and ocr.10mar11.oth) reflecting a shift in market expectation to favor the outcome of a decrease in OCR.

Figure 1: OCR announcement on 10 March 2011



In summary, the relationship between arbitrage and efficiency is not a black and white story; their mutual interaction is not simply a causation or a reference. As investigating market efficiency with regards to arbitrage, the question is not about the existence of arbitrage; instead, it is rather about its persistence and significance and more importantly, how long it takes for the market to eliminate those opportunities. Arbitrage only indicates market inefficiency when it yields significant profits compared with alternative risk-free investments and persistent over a long period of time. Otherwise, the emergence of arbitrage in the market might be a result of price changes which results from market instability or an arrival of new information. Arbitrage, by its nature as a mechanism to remove the price discrimination, constitutes into the process of converging out-of-equilibrium to equilibrium and eventually efficiency.

Arbitrage in prediction markets is free from risks that arbitrage in financial markets is exposed to (Shlerifer and Vishny, 1997). A transaction in prediction markets has a human trader in one side and the other side can be either another human trader or the market maker. Trading with the market maker has no counterparty risk - the risk that one side fails to complete the transaction. The market maker never fails to fulfill its obligation of the trade because it can create as many shares as it needs to (when it is on the short position) and it is unlikely for a market maker not to have sufficient funds to complete the trade (when it is on the long position). Trading against a human trader is also free from this risk. The market operator ensures that the long side of the trade has sufficient funds to execute the purchase. Traders are required to make deposit before they enter the market and transaction will not be executed if they do not have sufficient fund in their deposit to fulfill their purchase. The short side has to pay upfront to insure for the maximum loss before a sale is executed, for example a trader short-selling a stock at 60 cents does not receive 60 cents immediately once the sale takes place. Instead he has to pay upfront 40 cents to the market maker. This is because in the worst case that the contract closes at \$1, he has to buy the stock at \$1 to cover for his precedent short position. The maximum loss of this shott sell is 40 cents. Also, while arbitrage across financial markets may have the risk that commodities are not entirely identical, this is not a problem in prediction markets because stocks are identical across contracts within the same market.

Arbitrage in prediction markets is exposed to only execution risk. It may be the case that having closed one side of the deal, there is a shift in price at the time of closing of the other side. However, this risk is not a serious issue in most prediction markets where traders can observe a schedule of prices to buy and to sell beforehand. Therefore, a shift in market prices is unavoidable but it will not be a surprise to traders and the new level of market prices, if there is a shift in price, is observable from the price schedule.

I do not consider arbitrage across events or across prediction markets in this paper. Arbitrage across events is not feasible because it is unlikely to have two (or more) different events that are entirely related and more importantly, share the same attributes such as the identical set of possible outcomes and their associated likelihood. Plus, arbitrage across prediction markets are possible and actually has been observed in practice. It is quite common that various prediction markets offer contracts predicting exactly the same event and it is quite possible that contracts offered by those prediction markets are of the same type (binary, index or spreading contracts) and have the same payout structure. However, conducting this trading practice requires sophisticated technique and a great amount of time and effort on watching the markets. The opportunity cost of arbitrage across prediction markets, therefore, will be much higher and consequently make it less profitable for arbitrageurs.

The remainder of this chapter is structured as follow. In Section 3, I analyze the arbitrage opportunity arising as using historical transaction prices. I construct a theoretical framework for a typical arbitrage strategy. Although the analysis simplifies the real world, it provides an initial insight into arbitrage in prediction markets and the role of transaction costs in this respect. In Section 4, I then expand the investigation by taking into account the spread between bid and ask offers. I modify the initial framework in order to capture the effect of the bid and ask spread to arbitrage. I use data of markets predicting OCR announcements for empirical analysis. The last section concludes.

3 Arbitrage: Without Bid-Ask Spread

3.1 Methodology

The most straightforward arbitrage opportunity in prediction markets is to exploit any divergence of the aggregate price of contracts in the same bundle from the contract's payout. A contract bundle consists of contracts that cover every possible outcome of an event. The bundle price is the aggregate price of all contracts in the same bundle at a single point in time. All binary contracts within a prediction market have the same payout rule: pay a certain amount of money if the underlying outcome of designated event occurs and nothing otherwise.

A prediction market seeking to predict an event i with N_i possible outcomes will launch N_i contracts, each pays k if the underlying outcome occurs and nothing otherwise. Often the possible outcomes are mutually exclusive so are contracts on these outcomes, i.e. only one contract in the bundle closes at price k and the rest at \$0. Some events have a discrete and finite number of possible outcomes while others have continuous, infinite or a large number of possible outcomes. For example, the forecast of the change in OCR in the next quarter has an infinite number of outcomes consisting of both negative and positive changes. In the case of continuous variable, the prediction market does not launch a large number of contracts to predict every possible outcome. Instead it groups all possibilities into a few sub-ranges depending on the information of interest. In the example of predicting OCR change, the prediction market may launch five binary contracts, each of which pays \$1 if the GDP change lies within the possibility such as up by 25 points, down by 25 points, no change and outside these ranges.

Because a binary contract pays k if the underlying outcome occurs and \$0 otherwise, the price of individual contract varies within \$0 and k. As the price is assumed to coincide with the probability of the likelihood of the underlying outcome (Wolfers and Zitzewitz, 2005), the prices of all contracts predicting the same event must sum up to k. The aggregate price of \$1 is a necessary but not sufficient condition of the unbiasedness of market prices. This will be discussed at the end of this section.

First, I assume that transactions are cost-free. Let $p_{ij}(t)$ be the price of contract j ($j = 1, ..., N_i$) at time t predicting event i whose maturity time is T_i. At time T_i, the outcome of event i is realized and all contracts predicting outcomes of event i are closed and pay out. All contracts j in the event i are launched at the same time and closed at the same T_i.

Let

$$p_i(t) = \sum_{j=1}^{N_i} p_{ij}(t) \tag{1}$$

be the aggregate price of all contracts j in event i at time t.

At maturity T_i, only one contract in the event i pays out k for each share being held and others pay nothing so at time T_i, the aggregate price of contracts in event i is exactly k.

$$p_i(T_i) = \sum_{j=1}^{N_i} p_{ij}(T_i) = k \tag{2}$$

The payout amount k is predetermined before the launching of contracts so during the lifetime of all contracts j and within event i, k is fixed and considered as a constant. k is also often fixed across event i because the payout structure is consistent within an individual prediction market.

At any time t, the aggregate price $p_i(t)$ should be k otherwise arbitrage will arise. Arbitrage exploiting the mispricing of the bundle price arises in two scenarios, each requiring a different strategy but both bear no risk.

- If $p_i(t) < k$, there is at least one underpriced contract in event i. Arbitrage requires purchasing the entire bundle (i.e. buying one share of each contract) and holding it until the close time T_i. At maturity T_i, the bundle will pay out k. Arbitrage profit is $k - p_i(t)$.

- If $p_i(t) > k$, there is at least one overpriced contract in event i. Arbitrage requires short-selling the entire bundle (i.e. sell one share of each contract) and wait it until time T_i. At time T_i, the bundle is worth k and the arbitrageur will buy the bundle which costs k to cover his precedent short position. Arbitrage profit is $p_i(t) - k$.

Let $y_i(t)$ be the difference between $p_i(t)$ and k. Arbitrage arises whenever

$$y_i(t) = p_i(t) - k \neq 0 \tag{3}$$

Accordingly, arbitrage profit is

$$\pi_i(t) = |y_i| = |p_i(t) - k| \tag{4}$$

Next, relaxing the assumption of no transaction costs, the condition for an arbitrage to be profitable in Equation (3) no longer holds. Instead, arbitrage is only profitable if the gain is sufficient to cover the transaction costs, otherwise it is better to make no attempt to arbitrage.

Transaction cost varies between prediction markets. In iPredict, there are three types of cost but only the trading fee of $0.0035 per share traded should be included as transaction cost because it is incurred as soon as the transaction occurs.

I assume that every transaction incurs a transaction fee of γ_i per share traded. This transaction fee does not vary within either event i or prediction markets. However, to generalize the framework I assume that γ_i is constant within event i and allow it to vary between events.

In the presence of transaction costs, arbitrage also arises in two scenarios as mentioned above.

- If $p_i(t) < k - cost$, arbitrage requires purchasing the entire bundle and holding it until the close time T_i. There are N_i transactions needed to be made. The purchase costs $p_i(t) + N_i\gamma_i$. At maturity T_i, the bundle will pay out k. The arbitrage profit is thus $k - p_i - N_i\gamma_i$.

- If $p_i(t) > k + cost$, an arbitrager will short-sell the entire bundle and receive $p_i(t) - N_i\gamma_i$ from the sale. At time T_i, he will have to buy the bundle to cover his precedent short position. At T_i, the bundle price is k. The purchase costs him $k + N_i\gamma_i$. Arbitrage profit is $p_i(t) - k - 2N_i\gamma_i$. This implies that the short-sell-and-cover strategy would cost more than the buy-now-and-hold strategy.

  However, iPredict offers a feature to eliminate this cost disadvantage of short-selling. It offers a feature called Buy-a-Bundle which allows purchasing a bundle at \$1 ($k = \$1$ in iPredict); this means that purchasing an entire bundle is not subject to the trading fee. So whenever $p_i(t) > k + cost$, arbitrage strategy is as follow: Use Buy-a-Bundle feature to buy the entire bundle at \$1 and right after the purchase sell it for $p_i(t)$, the purchase will incur transaction cost $N_i\gamma_i$. Arbitrage profit will be $p_i(t) - k - N_i\gamma_i$.

In summary, arbitrage profit in the presence of transaction costs equals:

$$\pi_i(t) = max(0, |p_i(t) - k| - N_i\gamma_i)$$

$$\text{or equivalently} \tag{5}$$

$$\pi_i(t) = max(0, |y_i| - N_i\gamma_i)$$

Within event i, the parameters k, N_i, T_i and γ_i do not vary so the source of randomness in arbitrage comes from $|y_i|$. The distribution of π_i is determined by the distribution of $|y_i|$.

There are a few issues with this framework which need to be addressed. First, the derivation of arbitrage profit in Equation (4) and (5) requires an underlying assumption that at any time t, trader can buy and sell shares at the same price, i.e the buy offer coincides with the sell offer. In order words, there is no price discrimination between sell and buy orders. In practice, the offer prices to buy and sell never coincide. From the perspective of traders, buy offer is always higher. This assumption will be relaxed in Section 4.

The second issue is that the framework ignores the fact that arbitrage at any time t has to be executed at the current offers in the market at time t, not the last traded price $p_{ij}(t)$. This price is historical and may be no longer available for trading at the current time. Arbitrage based on this last traded price may mislead the possibility and significance of arbitrage in the market. For instance, at time t, a transaction which is a purchase occurs at the price p_{ij}. If the volume of the this trade is sufficient to move up the market price above $p_{ij}(t)$ then any traders coming to the market wish to buy the stock will have to pay a higher price than $p_{ij}(t)$ (assume other things remain constant). This issue will be addressed in Section 4.

Third, if more than one unit of a bundle is traded at a time, it is not guaranteed that the second unit will be traded at the same price as the first unit or not all units may be acquirable or saleable at the same price. In order to keep this simple, I examine arbitrage on the basis of one share traded at a time. This means that every transaction is either to buy or sell a unit of share.

The fourth issue comes from the fact that the framework does not capture a cost associated with the timing of transaction. As capital is invested to buy shares, not until the close day of the contract, capital is tied in the form of shares. The foregone returns from reinvesting the capital is considered an opportunity cost. Of the two arbitrage strategies described above, only the buy-now-and-hold strategy bears this cost. This strategy requires investing capital to buy shares now and hold them until the contract is closed. The opportunity cost gets larger as the purchase of shares occurs at earlier time to the maturity T_i. For example, arbitrage occurring one week before the close day of contract bears less cost than that occurring one month before.

The fifth issue is that the derivation of arbitrage profit ignores the discount factor. The payoff of an arbitrage which is to be received at maturity should be discounted at an appropriate discounting rate. Accordingly, the present value of arbitrage profit will be the absolute value of the difference between the present value of the payoff ($1) and $y_i(t)$ instead. As a result, arbitrage profit calculated by this framework is over-estimated. However, this discrepancy is not significant because most contracts have a relatively short lifetime. Therefore the discounted value of $1 received in a very near future should not be significantly smaller than $1.

The last issue involves the relationship of the unbiasedness of prediction market prices. Assume that prediction market prices coincide exactly with the market's aggregate belief. i.e. the market prices are unbiased and accurate predictors of probability that the event will occur. Then the condition for an arbitrage to arise holds: the aggregate price of contracts predicting the same event diverges from $1. However, the framework is also able to detect arbitrage even when prediction market prices are not unbiased. Consider an event with two possible and mutually exclusive outcomes whose market prices are called p_1 and p_2 ($j = 1, 2$). Arbitrage arises whenever $p_1 + p_2 \neq \$1$ (assume no transaction cost). This condition does not require the unbiasedness of market prices. For instance, the market believes that the probabilities that the outcome 1 and 2 occurs are 30% and 70% but their current prices are 50 cents and 60 cents respectively. This means that the contract 1 is overpriced and contract 2 is underpriced. The market prices are biased but their aggregate price is greater than \$1 (\$1.1). According to the framework, arbitrage is detected which requires short selling a bundle consisting of one share of each contract for \$1.1. When the final outcomes are realized and contracts are closed, arbitrageur pays \$1 to cover their precedent short position and realizes the profit of 10 cents. Besides, there exists another strategy to exploit the mispricing in contract 1 and contract 2: short sell the overpriced p_1 and buy the underpriced p_2. However, this trading practice is not considered arbitrage because it is nor risk-less. In order to take advantage of this mispricing, a trader is required to acquire the information of which contract is underprice and overpriced. This sort of information can never be obtained with 100% of certainty until the outcomes are realized. This practice fits in better the definition of speculation than arbitrage and thus is not considered in this chapter.

If a market is efficient then the market price must be unbiased. The foundation is that if prices fully reflect all available information, they must be the most accurate predictor of the probability of an event. In other words, if we have market efficiency, we must have unbiasedness of market prices. Accordingly, if prices are unbiased, there will be no arbitrage in the market. However, the argument in the opposite direction is not sufficient. The existence of no arbitrage does not guarantee that the market prices are unbiased. Consider the above example, the market prices p_1 and p_2 are 40 cents and 60 cents instead. According to the framework, no arbitrage opportunity is detected but the market obviously misprices the likelihood of the two outcomes. In short, the existence of arbitrage is an indicator of market inefficiency but the existence of no-arbitrage is a necessary but not sufficient condition of market efficiency.

Further, the deviation of market prices from the mean of market expectation may be a result of out-of-equilibrium trading. When the unbiasedness of prediction market prices creates arbitrage opportunities as in one of the examples above, it does not imply market inefficiency as long as traders, in the pursuit of profits, drive away those opportunities quickly. Arbitrage eventually constitutes to the process of bring the unbiased market prices to their fundamental level and simultaneously contributes to the process of converging out-of-equilibrium trading to efficient allocations.

3.2 Empirical Results

I analyze arbitrage opportunities using iPredict data for markets predicting OCR announcements over the period of October 23, 2008 to September 15, 2011. iPredict launches a market for each announcement consisting of a number of contracts which cover all possible outcomes of the OCR announcement. Between 23 October 2008 to 15 September 2011, there are 24 announcements so my data has 24 events ($i = 1, ..., 24$). The number of contracts in an event (i.e. N_i) varies. Out of 24 events, there are two events consisting of 6 contracts, six have 5 contracts, eleven have 4 contracts, and five have 3 contracts (for more detail see Appendix **??** and Appendix **??**). There are in total 101 contracts across all 24 events. The length of an event's lifetime (i.e. T_i) varies. Six markets were launched more than 3 months before the announcement day, fifteen launched within 3 to 1 months before, and three less than one month before.

For each contract j in event i, $p_{ij}(t)$ is the price of the transaction executed at time t. This transaction price can be the price of either a sale or a purchase from the perspective of the human trader who trades against available offers in the market. As soon as a transaction occurs, $p_{ij}(t)$ is recorded. Across all OCR events, the number of observations of $p_{ij}(t)$ is approximately 175,000.

All binary contracts in iPredict have the same payout structure: pays \$1 for each share being held if the underlying event occurs, and nothing otherwise. So k is \$1 in iPredict.

Ignoring the transaction costs, the gross arbitrage divergence is given by $y_i(t) = p_i(t) - \$1$. There are more than 45,000 observations of y_i across all events. The distribution of y_i directly relates to the distribution of $|y_i|$. As the absolute value of y_i measures the divergence of the aggregate price from its efficient level, \$1, $|y_i|$ determines the distribution of arbitrage profit net of transaction costs, π_i. Across 24 events, the average of y_i is 2 cents and of $|y_i|$ is 3.8 cents.

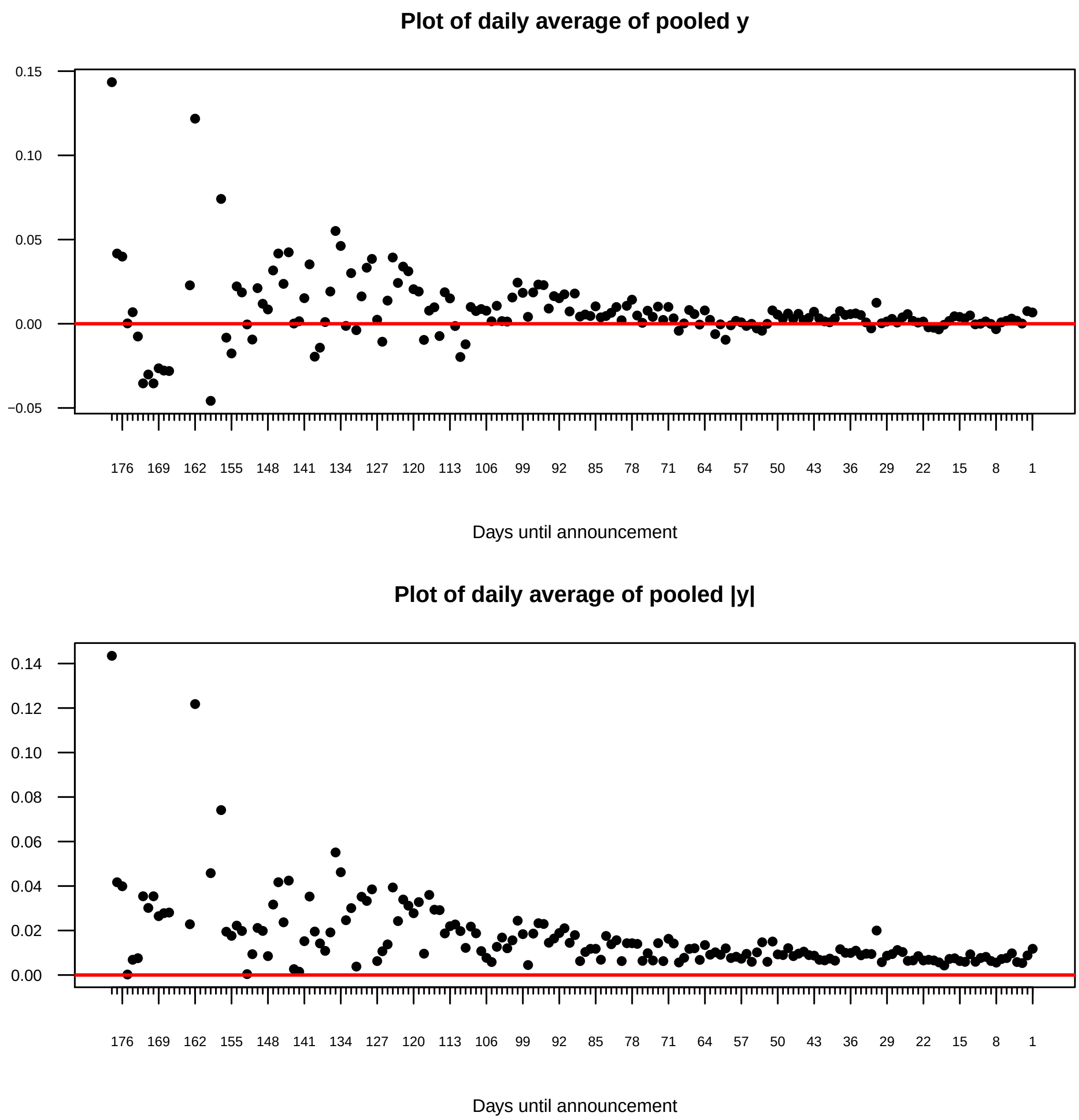
I define the gross arbitrage divergence as y_i. Thus $|y_i|$ is effectively the gross arbitrage profit, as arbitrageurs would be able to generate a profit, regardless of whether the divergence is positive or negative. For each event i, there are a number of readily observable features:

1. The gross arbitrage profit appears to be distributed around 0.
2. While a number of events show a degree of convergence in y_i toward 0, this is not universally observed. Indeed for some events, no pattern is discernible under visual inspection.
3. The number of events with a positive mean for y_i is dominating, i.e. the bundle tends to be underpriced more often.
4. In most events (17 out of 24) the distribution of y_i is unimodal
5. The bulk of the observed y_i appear to be relatively tightly distributed. Nevertheless, there are outliers which appear to be a large distance from the mean.

Figure 2 shows the time series of the average of y and average of $|y|$ as y is pooled from all events i. Pooled y is grouped on the basis of how far the transaction occurs from the announcement day (e.g. one day, two days,... until the announcement) and then averaged on the daily basis. One should expect a smaller divergence of y from 0 as the settlement date draws closer: more information revealed about the event helps to remove the mispricing. This expectation is met as looking at the time series of pooled y and

$|y|$: they both converge to 0 as the transaction occurs closer to the announcement day. The top plot of Figure 2 shows that y converges toward 0 from either below or above 0 as transactions occur closer to the announcement day. The bottom plots give a better insight to this convergence, $|y|$ shrinks toward 0 as it gets closer to the maturity time. Thus, graphically a positive correlation between time to maturity and $|y|$ is observed. The correlation between days to maturity and $|y|$ is 0.53 confirming what have been observed graphically.

Figure 2: Time series of y and $|y|$ as y is pooled from all events



However, as I repeat the test in individual events, I find that the positive correlation between $|y|$ and the time to maturity is not universally observed and where it exists, it is very weak. This is a finding supporting the efficiency of iPredict's OCR markets: the particular pattern in pooled y which represents arbitrage possibility is not universally observed in individual OCR events. In one hand, this means that arbitrage is independent to the time to maturity; an efficient market should not expect to have any predictable price patterns. In the other hand, the positive correlation of arbitrage possibility and the time to maturity gives some signals into how prices are adjusted over time in response to the revelation of new information.

In practice, there may exist situations of measured $y_i \neq 0$ that arise because the price level is in the process of change (there is trading out of equilibrium). This might show up as serial correlation in y_i; arbitrage profit is higher when prices change in response to the arrival of new information. The efficient market hypothesis says that in a complete or fully efficient market y_i should not be serially correlated. If serial correlation appears because of the adjustment in price level changes to the arrival of new information, it is unlikely to be persistent. This is due to trading out of equilibrium and these market level price changes will presumably happen quickly.

An efficient market should not have any serial correlation in market prices because it will enable to create trading technique based on historical prices in order to obtain abnormal returns. Within each event i, I test for the serial autocorrelation of $y_i(t)$ on $y_i(t-1)$ and $y_i(t-2)$, i.e. with lag of 1 and 2. Note that lag of 1 does not imply lag of one day. Since $y_i(t)$ is calculated based on $p_{ij}(t)$ which is recorded only when transaction occurs. Lag of 1 simply means y_i(s) of two adjacent transactions. The subscript of t indicates the order in which the transaction takes place. Also, it should be noticed that because the time interval between transactions is not constant.

For each event i, I regress:

$$y_i(t) = \beta_{1i} y_i(t-1) + \beta_{2i} y_i(t-2) + \varepsilon_{it} \tag{6}$$

Across 24 events the coefficient β_{1i} is significant and positive in all events and β_{2i} is significant in 17 events and positive in 17 events (at 5% error). It is shown that y_i in individual events exhibits significant and positive autocorrelation, especially at the first lag.

The most important notice is that the serial correlation in y_i might be the result of how y_i is constructed. Consider an event i has two contracts (say, contract 1 and contract 2). At a particular time $(t-1)$, both contracts were traded and at time t, only one contract, say contract 1, was traded. So $y_i(t-1)$ reflects contemporaneous trades on both contracts. However at time t, because contract 2 was not traded, I do not have $p_2(t)$. In order to construct $y_(t)$, $p_2(t)$ is required. Because there have been no actual trades on contract 2 at time t, the prevailing price must be the same as the price at time $t-1$. Effectively, there is autocorrelation in p_2. This may contribute into the autocorrelation in y_i. Another source of serial correlation in y_i may be the result of out-of-equilibrium trading. Price p_{ij} is in the process of change so there exists an inherent correlation in p_{ij} that is consequently carried forward to y_i.

In order to generalize this issue, I decompose $y_i(t)$ into two distinct parts:

- $p_{ij}(t)$ which captures the prices of contracts j which have traded at time t.
- $p_{ij*}(t)$ which captures the prices of contracts $j*$ which have not traded since some time s where $s < t$; i.e. $p_{ij*}(t) = p_{ij*}(s)$

Thus,

$$y_i(t) = \sum_{j \in J} p_{ij}(t) + \sum_{j* \notin J} p_{ij*}(t) - \$1$$

where J is the set of contracts traded at time order t. Here t is the ordinal index. For instance, a transaction occurring at time $(t-1)$ is the transaction immediately preceding that which occurs at time t, regardless of the actual time elapsed between the transactions. It should be noted that $p_{ij*}(t)$ may be the transaction price carried forward from $(t-1)$ or $(t-2)$ or even further back. I also note that most $y_i(t)$ have only one contract traded at each time order t, i.e. the set J consists of one element only.

From the decomposition of $y_i(t)$, it can be seen that there is a number of factors affecting the autocorrelation in y_i:

- There is an inherent autocorrelation derived from the construction of y_i as y_i comprises p_{ij*} carried forward from the last transaction.
- Time to the last trade: This is the time interval between transaction occurring at time s and t. The larger the difference between s and t, the longer the lag of the autocorrelation in y_i.
- Number of contracts in an event (N_i): The more contracts not in the set J, the more likely there will be autocorrelation in y_i.

If the market is efficient, the average of arbitrage profit π_i should not be significantly positive and the positive arbitrage profit should not be persistent over time.

The trading fee was first introduced on 20 June 2011. Therefore , any transactions occurring after August 1, 2011 would bear this trading fee. The formula calculating $\pi_i(t)$ (5) has an issue: it ignores the cap of fees. The trading fee is capped at 5% of any trade's gross value, and capped at $5 per month per user. The cap at 5% implies that buying a share at price less than 7 cents or short-selling a share at a price greater than 93 cent will be charged less than the standard fee of 0.35 cents per share. In order to take this fee cap into analysis, one would need to separate prices p_{ij} which lie outside the range $(0.07, 0.93)$ and apply a different transaction rate to them. However, in order to keep this simple, I apply the flat fee of 0.35 cents per share traded to every transaction.

The average conditional and unconditional mean of π_i across event i is 1.65 cents and 1.46 cents respectively while the possibility of arbitrage is 90%. In order to determine whether the arbitrage profit in OCR markets is significant, I compare its returns with the contemporaneous risk-free rate. The arbitrage return equals the ratio of profit $\pi_i(t)$ to the cost which is the sum of the bundle price and any trading fee.

$$ROR_t = \frac{\pi_t}{p_t + N_i\gamma}$$

The average arbitrage returns across events of 1.49% will be compared with the one-year secondary market government bond yield. The average risk-free rate during the same period of time with OCR markets is 3.2%. Because arbitrage in OCR markets yielding returns less than the risk-free rate, it should be neglected because its payoff is not sufficient to cover the opportunity cost. This is a supportive finding to the efficiency of OCR prediction markets.

One of the biggest issue in this framework is the fact that it detects arbitrage from historical transaction prices instead of currently available market prices and ignores the price discrimination of buy and sell offers. This issue will be dealt with in the next section where I expand the methodological framework in order to take the bid and ask spread into account.

4 Arbitrage: With Bid-Ask Spread

4.1 Methodology

The analysis in the previous section investigates the possibility of arbitrage based on prices at which transaction occurs. This implicitly ignores the fact that as transaction prices are historical and anyone entering trade in the market has to accept the currently available offers. If he wants to buy (sell) stocks, he will have to trade at the lowest (highest) available ask (bid) offer in the market. The previous framework in Section 3.1 ignored the effect of bid-ask spread on arbitrage in prediction markets. At any single point in time the bid offer has to be smaller than ask offer.

Let $p_{ij}^b(t)$ and $p_{ij}^a(t)$ be the bid and ask offers of contract j in event i at time t respectively. Any trader who wants to sell (or buy) shares of contract j in event i at time t has to accept $p_{ij}^b(t)$ (or $p_{ij}^a(t)$).

Let $p_i^b(t)$ and $p_i^a(t)$ be the aggregate bid and ask offer of all contracts j in event i at time t respectively.

$$
\begin{aligned}
p_i^b(t) &= \sum_{j=1}^{N_i} p_{ij}^b(t) \\
p_i^a(t) &= \sum_{j=1}^{N_i} p_{ij}^a(t)
\end{aligned}
\tag{7}
$$

Let $y_i^b(t)$ and $y_i^a(t)$ be the difference of the aggregate bid and ask offer of all contracts j in event i and k (which is \$1 in iPredict) at time t respectively.

$$
\begin{aligned}
y_i^b(t) &= p_i^b(t) - \$1 \\
y_i^a(t) &= p_i^a(t) - \$1
\end{aligned}
\tag{8}
$$

First I consider markets in which transactions do not incur costs. In the presence of bid and ask spread, arbitrage is no longer profitable whenever the aggregate price of the last transaction diverges from its efficient level (i.e. \$1 in iPredict). Instead, arbitrage only arises whenever one of the following conditions is satisfied.

- The aggregate ask offer of all contracts in the same market at any time t before the announcement is less than \$1. Arbitrage in this case requires buying a unit of contract bundle, paying $\$p_i^a(t)$ in total. The portfolio is held until the OCR announcement day when at least one contract pays out \$1 and the rest nothing. The payoff is \$1. Arbitrage profit is thus the difference between \$1 and $p_i^a(t)$.

$$\text{Condition: } p_i^a(t) < \$1 \leftrightarrow y_i^a(t) < 0$$

$$\text{Arbitrage profit: } \pi_i(t)(\text{buy and hold}) = \$1 - p_i^a(t) = -y_i^a(t)$$

- The aggregate bid offer of all contracts in the same market is greater than \$1. Arbitrage opportunity requires short-selling a unit of contract bundle for $p_i^b(t)$. On the maturity date, the bundle costs \$1 and trader will buy the bundle to cover his precedent short position.

$$\text{Condition: } p_i^b(t) > \$1 \leftrightarrow y_i^b(t) > 0$$

$$\text{Arbitrage profit: } \pi_i(t)(\text{short and cover}) = \$1 - p_i^b(t) = y_i^b(t)$$

In sum, in the absence of transaction costs, the arbitrage profit at any time t is:

$$\pi_i(t) = \begin{cases} -y_i^a(t) & \text{if } y_i^a(t) < 0 \\ y_i^b(t) & \text{if } y_i^b(t) > 0 \\ 0 & \text{otherwise} \end{cases}$$

The conditions of $y_i^a(t) < 0$ and $y_i^b(t) > 0$ are mutually exclusive because $y_i^b(t) < y_i^a(t)$ is always true. Thus $\pi_i(t)$ can be written as:

$$\pi_i(t) = max[-y_i^a(t), y_i^b(t), 0] \tag{9}$$

In the presence of transaction costs, buying and selling stocks incur costs. As mentioned earlier in Section 3.2, the only relevant cost associated with trading in iPredict is the transaction fee of \$0.0035 per share traded (denoted as γ). Taking this transaction fee into the analysis, arbitrage arises whenever:

$$-y_i^a(t) - N_i\gamma_i > 0 \text{ (buy and hold) , or}$$
$$y_i^b(t) - N_i\gamma_i > 0 \text{ (short sell and cover)}$$

Arbitrage profit in the presence of transaction costs thus equals:

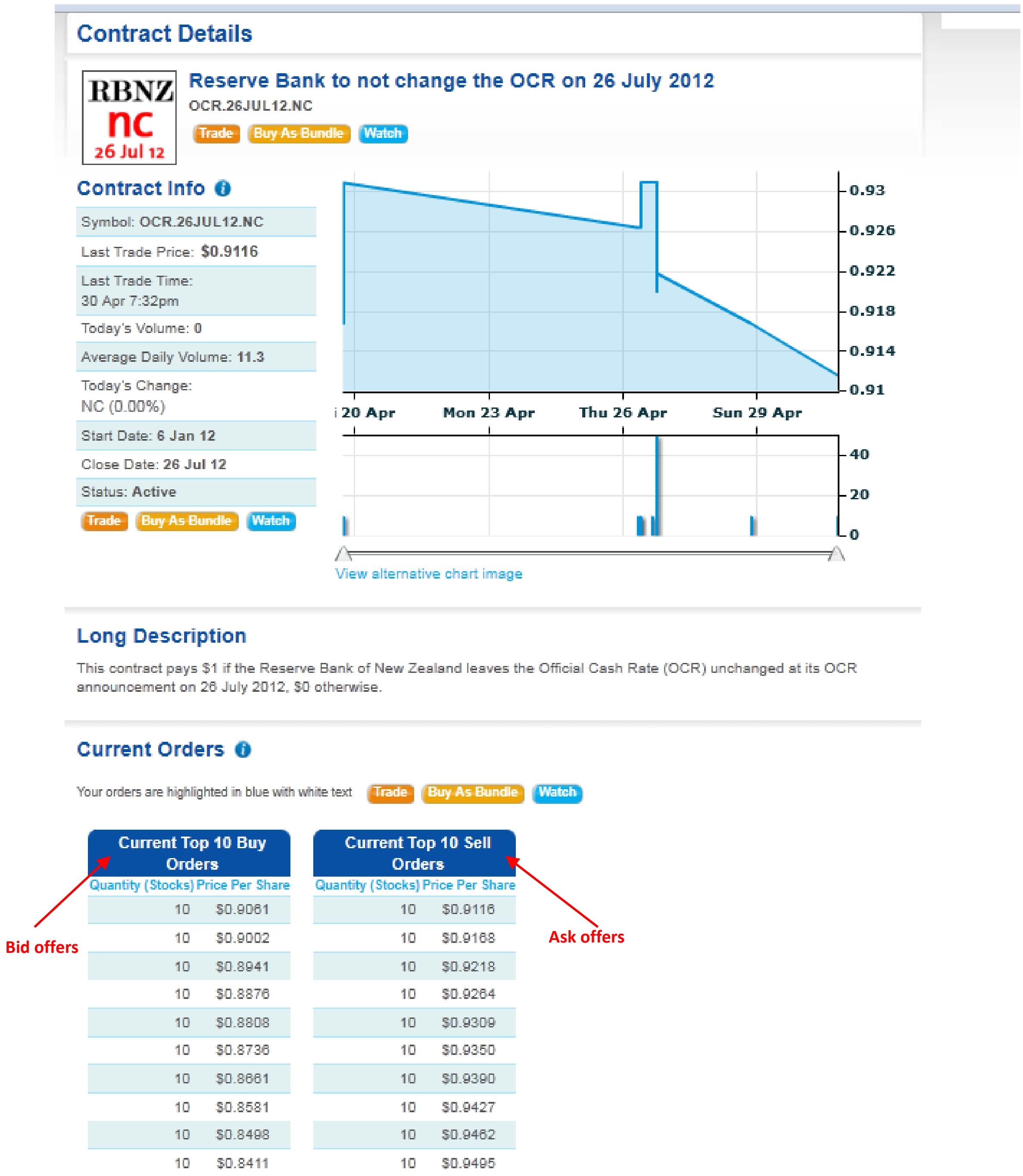
$$\pi_i(t) = \begin{cases} -y_i^a(t) - N_i\gamma_i & \text{if } -y_i^a(t) > N_i\gamma_i \\ y_i^b(t) - N_i\gamma_i & \text{if } y_i^b(t) > N_i\gamma_i \\ 0 & \text{otherwise} \end{cases}$$

or equivalently,

$$\pi_i(t) = max(-y_i^a(t) - N_i\gamma_i, y_i^b(t) - N_i\gamma_i, 0) \tag{10}$$

Empirical analysis requires a scheme to derive bid and ask offers at a single point in time. The system of iPredict only records prices and other information relevant to a transaction when a trade takes place. In other words, historical bid and ask offers are not available. Figure 3 shows the interface for the contract OCR.16JUN12.NC at 2.00pm on May 1, 2012, showing the contracts' current bid and ask offers. The contract pays \$1 if Reserve Bank of New Zealand leaves the OCR unchanged, \$0 otherwise. The contract's last traded price is \$0.9116. Any trader who wants to buy stocks knows that he would have to pay \$0.9116 per share for the first 10 shares and \$0.9168 for the next 10 shares and so on. Similarly, whoever wants to sell stocks knows that he would be able to sell the first 10 shares at \$0.9061 per share and the next 10 shares at for \$0.9116 per share. According to the notation of the framework, at time t which is 2.00pm on 1 May 2012, the last traded price of this contract $p_{ij}(t)$ is \$0.9116, the bid offer $p^b_{ij}(t)$ is \$0.9061 and the ask offer $p^a_{ij}(t)$ is \$0.9116 and so on. The last traded price and the ask offer happen to coincide in this case but this is merely coincidence and not necessarily true all the time.

Figure 3: Snapshot of contract OCR.26JUN12.NC at 2.00pm, 1 May 2012



While in practice the schedule of bid and ask offers has multiple units at each price level, for simplicity I assume that only 1 unit is offered for purchase or sale at each offer price. Thus, if a purchase (or sale) of a share of a stock occurs at price $p^a_{ij}(t)$ (or $p^b_{ij}(t)$, it will remove the ask (or bid) offer on the top of the price schedule and the market maker will fill a new buy (or sell) offer at the bottom of the price schedule. Any trader entering the market after a successful transaction has to trade at the new ask (or bid) offer.

Given this assumption, the derivation of bid and ask offer after a successful transaction relies on the S-curve price setting. The mechanism of deriving bid and ask offers are manually constructed by imitating the mechanism adopted by the market maker, assuming that the sensitivity of the S-curve is fixed/ unchanged within an individual event during its lifetime. In practice, there occurs cases in which the market maker adjusts the sensitivity of the S-curve and this accordingly affects the bid and ask spread and the liquidity of the market. The market maker tends to adjust the bid and ask spread in order to ensure its loss from subsidizing the market to stay within the allowed level.

4.2 Empirical Results

I repeat the same analysis that I have conducted on pooled y_i in Section 3.2 on pooled y_i^b and y_i^a. There are few observable features:

1. Both y_i^b and y_i^a exhibit a tendency to diverge from 0 as the transaction occurs further from the announcement.

2. The distributions of both y_i^b and y_i^a diverge from normality.

3. Observations of both are more dispersed at further from the contract's close day.

A special event occurred during the lifetime of OCR markets: the Christchurch earthquake on February 22, 2011. This unexpected event significantly changed the market's expectation of OCR announcement on March 10, 2011. The change in the market's expectation is reflected in the odd ratio. The ratio reflects the market expectation of the likelihood that the prospective OCR announcements is unchanged. The line is calculated as below:

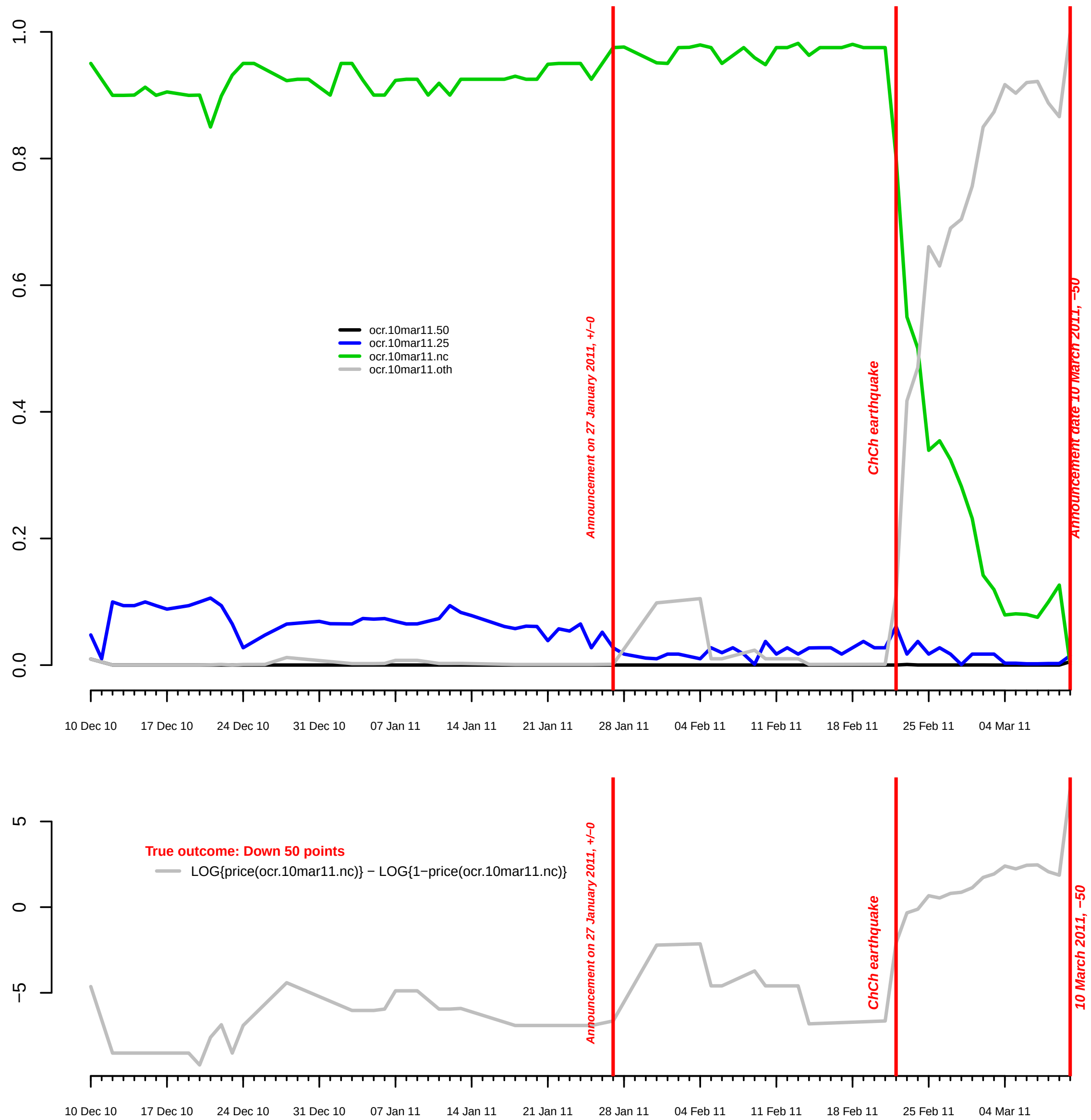
$$log(odd) = log\frac{\bar{price}}{1-\bar{price}}$$

where $\bar{price}$ is the last traded price of contract ocr.10mar11.nc which pays out if Reserve Bank leaves OCR unchanged.

A positive $log(odd)$ ratio implies that the market expects OCR not to change from its value at the previous announcement and vice versa. From the middle plot in Figure 1 before the earthquake the odd ratio is positive which indicates that the market in general expected that Reserve Bank was not to change OCR. However, the market revised its expectation and bid down the price of no-change contract right after the earthquake and remained low to 0 until the announcement was released. The market turned to be right about the no-change outcome as OCR was officially reduced by 50 basis points.

The top plot in Figure 4 gives more insight into what was going on in the market: price of no-change contract (ocr.10mar11.nc) which had been high to \$1 started dropping down after the earthquake and simultaneously the price of the contract predicting a decrease in OCR (ocr.10mar11.oth) rose up and remained high (close to \$0 till the announcement's release). The drop in price of contract ocr.10mar11.nc was almost perfectly offset by the rise in price of contract ocr.10mar11.oth. The bottom plot shows the $log(odd)$ ratio of the price of contract ocr.10mar11.oth which correctly predicted the true outcome of the OCR announcement. After the earthquake, the $log(odd)$ ratio jumped from below to above 0 which indicates that the market changed its belief from not expect to expect that Reserve Bank would increase OCR by more than 50 points or decrease OCR as a response to the earthquake.

Figure 4: Time series of prices of contracts in OCR announcement on 10 March 2011



Last, taking the bid-ask spread into the arbitrage analysis significantly reduces arbitrage in terms of the profitability of arbitrage and the chance it can arise. This is true within individual event i as comparing the conditional mean, unconditional mean of arbitrage profit and the probability of arbitrage with those in the absence of bid and ask spread. It shows that as the bid and ask spread is considered in arbitrage practice, arbitrage arises with smaller probability and when it does, it produces smaller arbitrage profit. With the bid-ask spread, on average the magnitude of arbitrage profit reduces from 1.41 cents per bundle unit to 0.6 cents and the probability of a profitable arbitrage opportunity reduces from 90% to 24.5%. The result is unsurprising. The introduction of bid and ask spread into the analysis of arbitrage is expected to limit the expected profitability of arbitrage in terms of both the magnitude and its associated probability.

The average rate of returns of arbitrage reduces from 1.49% to 0.6% as the bid and ask spread is taken into account. This returns is very insignificant as compared to the contemporaneous risk-free rate of 3.2%. Again, this finding confirms that arbitrage in OCR markets is very insignificant and negligible.

5 Chapter Conclusion

The empirical results has found the existence of arbitrage in iPredict's OCR prediction market. However, the more important question is that whether traders find it worthwhile to spend effort on. In the presence of bid and ask spread, the chance of a profitable arbitrage is 24.5%. This means that unless one can build a robot detecting arbitrage possibility 24/7, he should expect to capture an arbitrage opportunity with less than a quarter of his time spending on watching all transactions in the market. Plus he average return of arbitrage is 0.7% which is considerably less than the contemporaneous risk-free rate of 3.2%. In conclusion, arbitrage in OCR prediction market is not significant because it is outperformed by alternative risk free investment in the market.

Last, the existence of arbitrage itself is not a signal of market efficiency unless it is significant and persistent over time. Arbitrage in OCR prediction market is temporary. It arises as the market is adjusting its expectation due to the arrival of new information. Also, the arbitrage profit is too small compared with alternative risk-free investments. This concludes that the existence of arbitrage in prediction market does not violate the market efficiency.

Bibliography

Berg, J. E. and Rietz, T. A. (2003). Prediction markets as decision support systems. *Information Systems Frontiers*, 5(1):79–93.

Fama, E. F. (1969). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2):383–417.

Fisher, F. M. (1981). Stability, disequilibrium awareness, and the perception of new opportunities. *Econometrica*, 49(2):279–317.

Ghosal, S. and Porter, J. (2010). Out of equilibrium dynamics with decentralized exchange cautious trading and convergence to efficiency.

Goldman, S. M. and Starr, R. M. (1982). Pairwise, t-wise, and pareto optimalities. *Econometrica*, 50(3):593–606.

Hayek, F. A. (1945). The use of knowledge in society. *The American Economic Review*, 35(4):519–530.

Sharpe, W. and Alexander, G. (1990). *Investments*. Prentice Hall, 4th edition.

Shlerifer, A. and Vishny, R. (1997). The limits of arbitrage. *Journal of Finance*, 52:35–55.

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

Wolfers, J. and Zitzewitz, E. (2005). Interpretating prediction market prices as probabilities. Draft.