A Bayesian Market Maker

ASEEM BRAHMA, Qualcomm Inc.
MITHUN CHAKRABORTY, Rensselaer Polytechnic Institute
SANMAY DAS, Rensselaer Polytechnic Institute
ALLEN LAVOIE, Rensselaer Polytechnic Institute
MALIK MAGDON-ISMAIL, Rensselaer Polytechnic Institute

Ensuring sufficient liquidity is one of the key challenges for designers of prediction markets. Variants of the logarithmic market scoring rule (LMSR) have emerged as the standard. LMSR market makers are loss-making in general and need to be subsidized. Proposed variants, including liquidity sensitive market makers, suffer from an inability to react rapidly to jumps in population beliefs. In this paper we propose a Bayesian Market Maker for binary outcome (or continuous 0-1) markets that learns from the informational content of trades. By sacrificing the guarantee of bounded loss, the Bayesian Market Maker can simultaneously offer: (1) significantly lower expected loss at the same level of liquidity, and, (2) rapid convergence when there is a jump in the underlying true value of the security. We present extensive evaluations of the algorithm in experiments with intelligent trading agents and in human subject experiments. Our investigation also elucidates some general properties of market makers in prediction markets. In particular, there is an inherent tradeoff between adaptability to market shocks and convergence during market equilibrium.

Categories and Subject Descriptors: J.4 [**Social and Behavioral Sciences**]: Economics

General Terms: Algorithms, Economics

Additional Key Words and Phrases: Liquidity, Market Making, Bayesian Learning

ACM Reference Format:
Brahma, A., Chakraborty, M., Das, S., Lavoie, A., and Magdon-Ismail, M., 2012. A Bayesian Market-Maker ACM Trans. V, N, Article A (January YYYY), 24 pages.
DOI = 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

1. INTRODUCTION

Interest in prediction markets has increased significantly in recent years across academia, policy makers, and the private sector [Wolfers and Zitzewitz 2004b; Berg and Rietz 2003; Servan-Schreiber et al. 2004; Arrow et al. 2007; Chen and Pennock 2007]. Wolfers and Zitzewitz [2004a] discuss how prediction markets have gone from minor novelties to serious platforms that can have substantial impact on policy and decision-making. Companies like Google, Microsoft, and HP have deployed prediction markets internally for forecasting product launch dates and gross sales. Prediction markets have often outperformed opinion polling: for example, the Iowa Electronic Markets have usually outperformed opinion polling in predicting US political races [Berg et al. 2001]. There is little doubt that prediction markets are valuable for information aggregation for two reasons: (1) They produce meaningful quantitative forecasts; (2) Those who possess information are incentivized and held accountable more than they are in alternative information-gathering methods like surveys or polls.

Brahma's contribution to this work took place while he was a student at Rensselaer Polytechnic Institute.

DOI 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

Wolfers and Zitzewitz [2004a] identify five key challenges to the success of prediction markets . First among these is liquidity provision – can prediction markets attract sufficient uninformed trading to be liquid and attractive to those with information? Liquidity is the classic chicken-and-egg problem, in which some liquidity begets more liquidity. Historically, financial markets have often used market makers to provide initial liquidity to get the ball rolling; financial exchanges often provide specific incentives for firms to become market makers. Prediction markets have adopted the same idea. Typically, in prediction markets, the market maker is allowed to take on a loss, subsidizing the market, to facilitate more liquidity and faster price discovery; this loss is taken as a cost of operation. Robin Hanson suggested a family of inventory based market makers based on market scoring rules. Of these, the one based on the logarithmic market scoring rule (LMSR) is now the *de facto* standard for subsidized prediction markets [Hanson 2007]. If the main purpose of the prediction market is not commercial, loss-making market makers can make a lot of sense, but as prediction markets become less experimental, subsidies become a real loss which must be minimized.

The LMSR market maker is appealing on several levels: (1) it has a deterministic guarantee on the amount of loss it can suffer; (2) since it is purely inventory based, it is difficult to manipulate in some settings [Chen et al. 2009]; (3) it can be shown that, under certain conditions, particularly that participants are rational and learn from prices, a market mediated by an LMSR market maker will converge to the rational expectations equilibrium price [Pennock and Sami 2007]. However, the LMSR also suffers from several drawbacks, some serious: (1) The market maker does typically run at a loss, which can be large, (2) A single parameter, b, controls many different aspects of the market maker's behavior, including the loss bound, the level of liquidity in the market, and the rate of adaptivity to market shocks; setting b to optimally manage these tradeoffs is considered something of a "black art" [Othman et al. 2010]; (3) when the posterior belief of the trading population does not converge (which is likely when people have independent information and valuations), the price does not converge to a well defined probability estimate, instead fluctuating about the equilibrium price; the fluctuations are asymmetric and again sensitive to the choice of b, making it difficult to extract a quantitative probability estimate; (4) The market maker provides only point probabilities over outcomes and cannot be easily coupled with a measure of uncertainty; (5) the market maker cannot easily be applied to unbounded markets.

There have been recent efforts to address some of these shortcomings. A particularly interesting approach is that of Othman et al. [2010], who propose a variant of LMSR that effectively adjusts the b parameter as a function of how much trading has occurred in the market. Unfortunately, as we demonstrate later in this paper, this liquidity-sensitive market maker can become very slow to adapt to jumps in the underlying value if this value changes after a fair amount of trading has already occurred. An alternative to inventory based market makers is to use the information inherent in trades in order to set prices. The seminal paper of Glosten and Milgrom [1985] introduced a model of market making under asymmetric information. Building upon this model, Das (2005; 2008) and Das and Magdon-Ismail [2008] have described efficient market making algorithms for zero-profit (competitive) and profit maximizing (monopolist) market makers. These market makers address some of the drawbacks of the LMSR market maker. Specifically, in stylized market models where a single shock to the value occurs, the price converges rapidly to an equilibrium price, without expected loss; further, the markets need not have bounded payoffs. The drawback of these Bayesian market makers has thus far been the same as that of the liquidity-sensitive market maker: after quick convergence following an initial market shock to the true value, the convergence after a subsequent market shock is slow, because the market maker gets "overconfident" after initial convergence.

Contributions. We introduce a new adaptive Bayesian Market Maker (BMM) that builds on [Das 2005; Das and Magdon-Ismail 2008]. BMM provides liquidity by adapting its spread based on its level of uncertainty about the true value. It uses two inference processes: the first is on the actual price; the second is on whether a jump has occurred. This allows it to achieve small spreads in equilibrium-like states, while remaining adaptive to market shocks.

Evaluation of market making algorithms is challenging. We use two different evaluation mechanisms. First, a novel experimental paradigm for comparing market microstructures in trading experiments with human subjects. Two challenges with live trading are: 1) the same group of traders cannot be used first in an experiment with one market maker, and then in a second identical experiment with a second market maker. This is because traders get primed, and even if the experiments are identical, the results are incomparable; 2) the same experiment cannot be run on two separate groups of traders with a different market maker in each group, because the high variability in human traders results in a very high variance due to inevitable small sample sizes in controlled human experiments. Our design is based on a graphical 2-dimensional random walk which simulates the classic *Gambler's Ruin* problem. This allows us to *symmetrically* compare different market structures.

Human subject experiments are somewhat restricted – one cannot run them too often, so it is hard to experiment with many different parameter settings of different algorithms, or capture the effects of rare events. Therefore, we design a trading game in which trading agents receive information equivalent to that received by humans in the gambler's ruin game, and allow agents with different trading strategies to participate. We describe some plausible designs for intelligent trading agents and evaluate market makers in terms of the quality of the market they provide when different combinations of the types of trading agents participate in the market.

In all our experiments, the results are consistent: BMM provides benefits compared with LMSR. In particular, BMM quickly adapts and generally does not lose money, while providing a liquid market. The caveat is that BMM is not loss bounded, and there is some risk of substantial loss. We also note a property of market makers that we conjecture is universal: there is an inherent tradeoff between how adaptive a market maker is to changes in market conditions, its ability to converge during equilibrium, the liquidity it provides, and its potential loss. Market designers need to keep these tradeoffs in mind in designing market making algorithms.

2. MARKET MAKING

The key challenge in most prediction markets is liquidity. How can one incentivize participants with good information to trade? Without uninformed traders to exploit, informed traders will not trade (the No-Trade theorem of Milgrom and Stokey [1982]). Automated market makers [Hanson 2007; Pennock and Sami 2007] are a means of creating "uninformed" (or less informed) trades that can provide liquidity in modern prediction markets. A market maker is willing to take either side of every trade, buying (resp. selling) when someone wants to sell (resp. buy); the market maker sets the prices, which will affect whether the trade will actually execute or not. We consider a pure dealership market, where a market maker takes one side of *every trade*. This model of the market allows us to compare market makers in a fair and precise manner, but in the future it will be important to consider integrating market makers with limit order books (which poses more of a challenge for evaluation than design).

2.1. LMSR

Hanson [2007] describes a market maker for combinatorial prediction markets, which we briefly review here in the context of a single market. Hanson's technique adapts the idea of a scoring rule to a prediction market setting. While many different scoring rules are possible, Pennock reports that in practice the logarithmic scoring rule is the most useful.[1] The market maker will take the opposite side of any order at a price specified by the market maker. This price depends on a parameter b and the market maker's current inventory q_t, where t indexes the arrival of trade requests; the inventory starts at zero, $q_0 = 0$, which corresponds to an initial price of 0.5. The market maker sets prices so as to guarantee bounded loss, no matter what the true liquidation value is.

Specifically, the spot price is given by $\rho(q_t) = \frac{e^{q_t/b}}{1+e^{q_t/b}}$. At time $t+1$, if a trade arrives for quantity Q, the cost of the trade (to the trader) is given by $C(Q; q_t) = \int_{q_t}^{q_t+Q} ds\, \rho(s) = b\ln(1 + e^{(q_t+Q)/b}) - b\ln(1 + e^{q_t/b})$. The volume weighted average price is $|C(Q; q_t)|/q$, and it corresponds to the trader accruing a position of size Q using infinitesimal increments, paying the prevailing spot price at each increment. If the trader accepts the trade at this average price, then the market maker updates its state to $q_{t+1} = q_t + Q$. Since the starting inventory $q_0 = 0$, it is easy to verify that the maximum loss incurred by the market maker is $\lim_{Q\to\infty} Q - C(Q; 0) = b\ln 2$.

The parameter b is the only free parameter in the LMSR market maker; not only does it bound the loss of the market maker, but it also controls how adaptive the market maker is. If b is small, the market maker is very adaptive, taking on small loss; b also controls liquidity in the market. An adaptive market maker leads to large bid-ask spreads, implying less liquidity.

It is known that a market mediated by the LMSR market maker can yield a rational expectations equilibrium if traders incorporate information from prices into their beliefs in a rational manner. However, as a thought experiment, consider what happens in a case where a large trading population continues to maintain somewhat different beliefs, and some traders regularly come in and trade some typical trade size Q.[2] The bid-ask spread $\delta(Q)$ for quantity Q, given the current inventory q_t, is the difference between the average price paid for buying Q shares versus selling Q shares;

$$\delta(Q) = \frac{b}{Q} \ln\left(\frac{\cosh q_t/b + \cosh Q/b}{2\cosh^2 q_t/2b}\right).$$

At market inception ($q_t = 0$), the spread is decreasing in b, so higher b means more liquidity (in general the relationship between liquidity and b is not monotonic). Suppose the equilibrium price corresponds to an inventory q_{eq}; if typical trade sizes are Q, then the spot price fluctuations around this equilibrium have magnitude $\frac{\sinh(Q/b)}{\cosh(q_{\text{eq}}/b)+\cosh(Q/b)}$. These fluctuations are asymmetric about the equilibrium and persist, making it hard to extract a quantitative probability estimate. The choice of b is an important open problem; smaller b guarantees smaller loss, but a less liquid market with higher fluctuations around the equilibrium.

[1] http://blog.oddhead.com/2006/10/30/implementing-hansons-market-maker/

[2] In standard rational expectations models, traders learn, and their beliefs converge, leading to convergence in the market maker's price. However, this requires hyper-rational, potentially computationally unbounded, traders and it is unclear how much time is needed for such convergence. As long as there is any variance in the posterior distribution of beliefs in the trading population, there is continued potential for the price to fluctuate with LMSR. This may not be a bad thing, because the fluctuations are critical to maintenance of the loss bound, and may be the "right" thing to do. However, these fluctuations can abstractly be measured by deviation of the price from the "true" value, and we argue that for a given level of loss, a market maker with lower deviation from the true value is preferable.

2.2. A Liquidity-Sensitive Variant of LMSR

A major shortcoming of LMSR is its inability to adapt to market liquidity levels: the price response for a fixed trade volume is the same regardless of liquidity [Othman and Sandholm 2010; Othman et al. 2010]. Formally, $p_i(\mathbf{q} + \alpha\mathbf{1}) = p_i(\mathbf{q})$ for any α and $\mathbf{q}$. As a solution, Othman et al. [2010] propose a modification of LMSR where the exogenous constant b in the cost function $C(\mathbf{q})$ is replaced with an increasing function of market volume $b(\mathbf{q}) = \alpha \sum_{j=1}^{n} q_j$, where $\alpha > 0$ is a constant and n is the number of securities. The liquidity-sensitive price function is $p_i(\mathbf{q}) = \alpha \log\left(\sum_{j=1}^{n} e^{q_j/b(\mathbf{q})}\right) + \frac{\sum_{j=1}^{n} q_j e^{q_i/b(\mathbf{q})} - \sum_{j=1}^{n} q_j e^{q_j/b(\mathbf{q})}}{\sum_{j=1}^{n} q_j \sum_{j=1}^{n} e^{q_j/b(\mathbf{q})}}$. However, now the sum of prices across all securities is always at least unity although it never exceeds $\alpha n \log n$. The market has a loss bound of $C(\mathbf{q}^0)$ where $\mathbf{q}^0$ is the initial quantity vector and so by setting this close to zero, the loss can be arbitrarily diminished. In fact, the worst-case revenue is $\mathcal{R}(\mathbf{q}) = C(\mathbf{q}) - \max_i q_i - C(\mathbf{q}^0)$ and there is a wide range of terminal market states with reasonable uncertainty levels in which $\mathcal{R}(\mathbf{q}) > 0$ so that the market maker actually books profits in those situations.

2.3. Information-Based Market Making

Das [2005] and Das and Magdon-Ismail [2008] describe an information-based market maker that starts from the canonical Glosten-Milgrom model of price-setting under asymmetric information [Glosten and Milgrom 1985]. At time t, the market maker has some belief (prior probability density) about the value of the security $p_t(v)$ An arriving trader gets a signal s; the variance of s measures the uncertainty in the trader's signal (or *information set*). The market maker's only information is its prior belief on V. Hence the information available to the market maker and trader are different, and this information asymmetry can be measured by the information disadvantage of the market maker, the ratio of the variance in the market maker's prior belief and the trader's uncertainty.

Given this initial setting, the market maker must set a bid and ask price, and the trader trades accordingly: if $s <$ bid, the trader sells, and if $s >$ ask, the trader buys. In a competitive setting, the market maker sets prices so as to receive zero expected profit. This is achieved by solving two non-linear fixed point equations,

$$\text{ask} = E_{p_t(v)}[v|s > \text{ask}]; \qquad \text{bid} = E_{p_t(v)}[v|s < \text{bid}]$$

This model has been extended to the sequential setting with a Bayesian market maker [Das 2005; Das 2008; Das and Magdon-Ismail 2008]. After setting prices, the market maker can now observe what the trader does (buy, sell or no trade). This gives the market maker information regarding the trader's signal s, and hence information regarding the realization V. Thus, the market maker can update its prior beliefs $p_t(v)$ to $p_{t+1}(v)$ to incorporate this new information. The market maker is now ready for the next trader.

The learning market maker in the sequential model is composed of two related parts. The first maintains the belief distribution on the value of the market, $p_t(v)$; the second sets prices to achieve some goal, for example zero expected profit. From the reinforcement learning perspective, bid and ask prices serve as actions, and agents' decisions to buy or sell at those prices provide observations that allow the market maker to update its beliefs. As in most reinforcement learning problems, the actions (prices) serve the dual role of 1) eliciting information (setting the bid-ask spread too high will lead to a lack of trading, yielding little information about the trading public's beliefs) and 2) generating reward.

Das and Magdon-Ismail [2008] present efficient approximate algorithms for performing these updates for zero profit (ZP) as well as profit maximizing monopolist market makers. In the specific algorithm considered, the trader signals are drawn from a Gaussian distribution, and the initial market maker belief is also Gaussian.

2.4. Other Alternatives

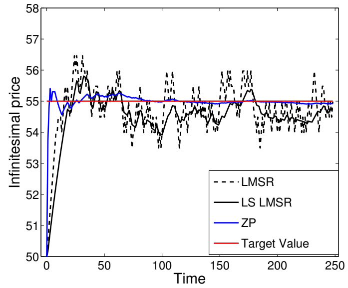
There have been other mechanisms proposed for market making in the literature. The dynamic pari-mutuel market (DPM) of Pennock [2004] is particularly interesting, since in practice it effectively offers a loss bound which determines initial liquidity in the market (theoretically there is no risk to the market institution, since it redistributes money from losers to winners like a standard pari-mutuel market, but in practice the initial subsidy is important to generating liquidity). However, one issue with DPM is that it does not offer liquidity for selling, only for buying, and selling has to occur through an alternative mechanism like a continuous double auction. Another issue is that when participants buy shares in particular outcomes, the payoff per share is not known at the time of purchase, since it depends on how many shares are outstanding at the time of payoff. Other mechanisms, such as the market maker used in the Hollywood Stock Exchange, are not well documented, although they appear to use many tuned heuristics.

2.5. Comparison

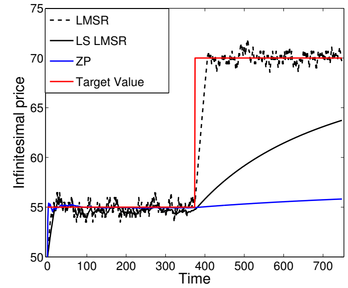
Since LMSR has become the *de facto* standard in information markets, we focus in this paper on LMSR, the recently proposed liquidity-sensitive variant, and Bayesian market-making algorithms. Figure 1 demonstrates some fundamental aspects of these market makers in a simple model – zero-intelligence traders arrive sequentially, with a belief that is distributed around the true value, which can jump (a "market shock" event). They are allowed to trade one share when they arrive – therefore they will buy when their belief exceeds the ask price and sell when their belief is below the bid price. While simple, this model illustrates some aspects of the different market makers (again the posterior distribution is non-convergent here – obviously reality lies somewhere between hyper-rational traders whose beliefs converge, and stubborn traders who never update their beliefs – this is investigated further in Section 6).

As we see, the LMSR market maker is adaptive but non-convergent; the information based market maker ZP [Das 2005; Das and Magdon-Ismail 2008] is convergent, but only slowly adaptive, potentially incurring large loss; liquidity-sensitive LMSR is also very slow to adapt, although it does not lose money – however, this is at the cost of less liquidity in the first phase of the market, because it has to maintain a significantly higher spread in order to recoup potential future losses because of such jumps.

Desiderata for a good market maker. We would like to design a market making algorithm that simultaneously satisfies several criteria. First, we would like the market maker to not make losses in expectation. It is unreasonable to expect real-money markets to be heavily subsidized in pursuit of liquidity. For now, we are willing to sacrifice loss-boundedness in pursuit of this goal. While LMSR is loss-bounded, it typically will substantially subsidize the market, taking on large loss. Alternative market making schemes necessarily incur more risk. In Hanson's words, "a computer program with less than human intelligence that attempts to make markets runs the risk of being out-smarted by human traders" [Hanson 2009]. This is because a market maker who makes offers to buy and sell any security runs the risk of losing out to either better informed or smarter traders. At the same time, "smart" market making algorithms may be able to exploit human trader errors or overconfidence. Thus, it might be possible to provide liquidity without substantial loss. Alternative risk-management strategies



(a) Behavior in a stable market.



(b) Adapting to a market shock.

Fig. 1: Behavior of LMSR, liquidity-sensitive (LS) LMSR and the Bayesian information based market maker ZP for an idealized setting. In both experiments, trades are a fixed size, and traders at every time step receive a Gaussian signal on the basis of which they trade. In (a), the trader value signals are distributed about a stable mean; in (b), there is a jump in the mean. In (a), we see that all market makers rapidly arrive at the target value (the mean), but LMSR fails to converge, whereas ZP converges. From (b), we see that LMSR is very adaptive to a market shock whereas its LS variant adapts much more slowly and ZP is the slowest of the three because it becomes overconfident.

can be considered in future work. For example Das [2005] proposes and evaluates one such scheme based on ruin models from finance theory [Amihud and Mendelson 1980].

Second, we would like the market maker to be *convergent in equilibrium*. This means that if the true value is not fluctuating, the market maker should maintain a reasonably stable price around that value. We measure this by root mean square deviation (RMSD) of the price from the true value (also see Footnote 1 for further discussion).

Third, we would like the market maker to be quickly adaptive to jumps in the underlying true value. This is measured by time of convergence to a region near the new true value if there is a shock to the underlying state of the world that changes the fundamental value of the security.

3. MARKET MICROSTRUCTURE

We consider a prediction market with a single binary outcome stock that trades between 0 and 100. Presumably, if the event occurs it pays off 100, and if not, it pays off 0. However, this can also be thought of as a stock with a liquidating dividend between 0 and 100. At any point in time, an arriving trader sees the history of trading in the stock, and the "current price," which can be thought of as either the infinitesimal price, the market maker's mean belief about the probability of the event occurring, or the middle of the bid-ask spread. The trader then chooses a quantity that she wants to buy or sell. The market maker observes the quantity demanded, and sets a price based on this quantity. The trader is informed of this price and can then choose whether or not to execute the trade at that price. Our market is structured as a pure dealer market, with the market maker as the only price setter. Only a single (infinitesimal) spot price is seen by arriving traders. However, they can query for the price of trading any quantity (for the trading agents described in Section 6, this becomes equivalent to knowing the entire order book, since it can query the market maker for the price of any quantity).

4. THE BMM ALGORITHM

BMM is based on the zero-profit market maker (ZP) mentioned in Section 2 and described in detail in [Das and Magdon-Ismail 2008], with two main innovations: 1) the ability to deal with trade sizes; and, most importantly, 2) the ability to adapt quickly to market shocks. A trader arrives, observes the spot price p_t and requests a trade for quantity Q in a direction $x_t = \pm 1$. $x_t = +1$ means the trader would like to buy. For concreteness, we will assume that $x_t = +1$; however, the process is completely symmetric. The market maker performs 3 tasks.

i. Provides a VWAP quote for Q shares;
ii. Updates its state depending on whether the trade is accepted or canceled;
iii. Maintains a validity measure for its current beliefs, which is crucial to being able to adapt to market shocks.

We briefly summarize ZP described in [Das and Magdon-Ismail 2008] first. The market maker's state is characterized by a Gaussianbelief for the value of the market V: $N(\mu_t, \sigma_t^2)$. The trader signal is assumed to be normally distributed around V, so $s \sim N(V, \sigma_\epsilon^2)$. The main relevant parameter (see [Das and Magdon-Ismail 2008]) is the information disadvantage of the market maker, $\rho_t = \sigma_t/\sigma_\epsilon$, the ratio of the uncertainties of the market maker and trader. A universal "Q-function", $Q(\rho)$ (see [Das and Magdon-Ismail 2008]) plays an important role in quoting prices. Specifically, the spot price is just the market maker's mean belief, $p_t = \mu_t$, and the ask price is

$$\text{ask} = \mu_t + \sigma_\epsilon Q(\rho_t)\sqrt{1 + \rho_t^2}.$$

This ask gives zero expected profit conditioned on the trade going through; this quoted price does not take quantities into account. Described in [Das and Magdon-Ismail 2008] is a range based update procedure for the market maker: if a trader's realized signal is known to lie in the range $z^- < s < z^+$, then the market maker updates its Gaussian belief to:

$$\begin{aligned} \mu_{t+1} &= \mu_t + \sigma_t . \frac{B}{A}, \\ \sigma_{t+1}^2 &= \sigma_t^2 \left(1 - \frac{AC + B^2}{A^2}\right), \end{aligned}$$

where A, B, C are functions of $z^-, z^+, \mu_t, \rho_t, \sigma_\epsilon$, the details of which are given in [Das and Magdon-Ismail 2008]. This range based update is used when the trader takes an action (accept or cancel the trade). So, for example, if the trader accepts a trade, then $s \in [\text{ask}, \infty)$, and so $z^- = \text{ask}$ and $z^+ = \infty$. If the trader cancels upon seeing the quoted price, $z^- = \mu_t$ and $z^+ = \text{ask}$.

4.1. Quoting a Price for Q Shares

ZP can only quote a price for a fixed trade size. To be practical, the algorithm needs to quote a price for an arbitrary number of shares. The spot price is $p_t = \mu_t$, and assume a trader wants to buy Q shares. We implement a heuristic of treating this order as independent orders of a fixed size α. There are thus $\lceil Q/\alpha \rceil$ independent orders; the sizes $\alpha_1, \alpha_2, \ldots, \alpha_k$ are all α, except possibly the last one.

The market maker starts in state $\mu_1 = \mu_t$, $\sigma_1^2 = \sigma_t^2$, and imagines the arrival of these k mini-orders in sequence; for each mini-order arrival, the market maker quotes the ZP price as in [Das and Magdon-Ismail 2008]; each mini-trade is accepted; the market maker then updates his belief and receives the next mini-trade. Specifically, consider mini-trade i, with market maker belief μ_i, σ_i^2. The price quoted is

$$p_i = \mu_i + \sigma_\epsilon Q(\rho_i)\sqrt{1 + \rho_i^2};$$

the trade is accepted, so the market maker updates his belief with $z^- = p_i;\ z^+ = \infty$:

$$\mu_{i+1} = \mu_i + \sigma_i \frac{B}{A}, \qquad \sigma_{i+1}^2 = \sigma_i^2 \left(1 - \frac{AC + B^2}{A^2}\right);$$

the market maker now processes the next mini-order in the sequence until all the mini-orders are processed. Note that these mini-orders are not real, they just describe the process going on in the market maker algorithm. Thus, $\alpha_1, \ldots, \alpha_k$ shares, with $\sum \alpha_i = Q$, get (fictitiously) executed at the prices $p_1, \ldots, p_k$. The price quoted to the trader for Q shares is the VWAP for this fictitious sequence of executions:

$$\text{ask} = p(Q) = \frac{1}{Q} \sum_{i=1}^{k} \alpha_i p_i.$$

Belief Update. Since the trader asked to buy, we know that $s \geq p_t$. The trader is quoted a price $p(Q)$, and so based on the trader's action, the market maker can update his beliefs to $\mu_{t+1}, \sigma_{t+1}^2$ using the range update:

$$\text{trade} \begin{cases} \text{accepted} & z^- = p(Q);\ z^+ = \infty; \\ \text{canceled} & z^- = p_t;\ z^+ = p(Q). \end{cases}$$

We described a buy order, but a sell is entirely symmetric.

4.2. Adapting to jumps

The original ZP algorithm leads to constantly decreasing variance of the market maker's belief. After a number of trades have been processed, the variance and therefore the spreads are significantly reduced While this increases liquidity and encourages further trading towards the true market valuation, it is also the root of the market maker's inability to adapt to multiple market shocks. In fact, the magnitude of each mean belief update is proportional to the variance of the market maker's belief. Large jumps in the true underlying value coupled with a small belief variance lead to very small relative update values. This causes the algorithm to be exponentially slow in adapting to a jump.

After a jump, the sequence of trades will be "one-sided", and hence inconsistent with a market maker's belief of the old valuation coupled with a highly confident low belief variance. The simple solution to this is to allow the market maker to become less confident as he sees a sequence of extremely one sided trades, i.e. an inconsistent sequence of trades. To accomplish this, we define a consistency index $C(\text{history})$, which measures exactly how likely the recent history of trades observed under the current uncertainty level is, as compared to a higher uncertainty. An intuitive solution is to increase the market maker's belief variance during periods of inconsistency.

Specifically, BMM keeps track of a fixed window of previous trades (including canceled trades), along with the z^+ and z^- values that are inferred from those trades. Then, at a particular time step, the probability of a sequence of trades over a window of size W, can be computed as:

$$L(\mu, \sigma) = \int_{-\infty}^{\infty} N\big(v, \mu, \sigma\big) . \prod_{i=1}^{s} \left(\Phi\big(z_i^+, v, \sigma_\epsilon\big) - \Phi\big(z_i^-, v, \sigma_\epsilon\big) \right) dv$$

The intuitive solution is to compare this probability against a fixed threshold; if the probability is too small, we are in an inconsistent regime, and so we increase the market maker's uncertainty level (variance). However, this solution is problematic because the threshold is highly sensitive to the choice of window size and particular features of the trade sequence. Instead, we make a relative comparison with the same probability computed at twice the uncertainty. We thus define our consistency index

$$C(\text{history}) = L(\mu_t, 2\sigma_t) - L(\mu_t, \sigma_t);$$

If $C > 0$, we increase σ_t, specifically $\sigma_{t+1} = 2\sigma_t$. The choice to double the variance is arbitrary, and any multiplier greater than 1 would do. Though we have only tried the multiplier 2, we expect that since this is a relative measure of consistency, the results would be robust to the choice of multiplier, unlike with the use of a fixed threshold.

This algorithm takes advantage of the fact that more "even" sequences of trades are more likely when the variance is lower, while sequences that are heavily biased in one direction or the other become more likely with higher variance. The key parameter for this algorithm is the window size W, which controls the balance between how stable the market maker is at equilibrium and how fast it can adapt to changes. The window size W also now becomes the dominant factor in measures like average spread, so that the particular value of σ_ϵ becomes unimportant.

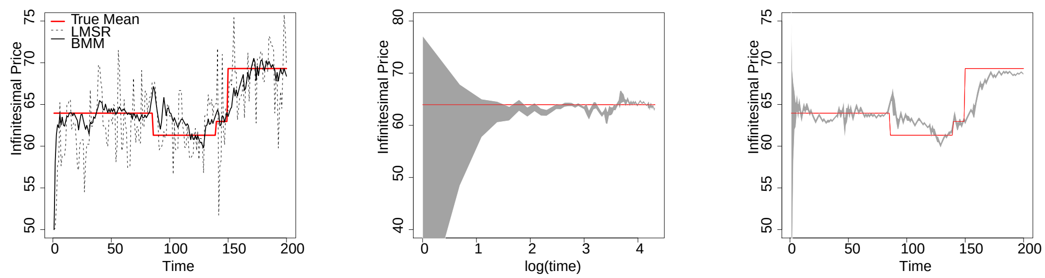
4.3. Preliminary Validation

In order to validate BMM, we conduct simulation experiments with the goals of (1) ensuring the adaptive capabilities of BMM (2) comparing BMM and LMSR on the basis of profit/loss, average spreads, and price discovery, and (3) calibrating parameters for further experiments with market makers that provide similar quality markets.

Each trading simulation consists of 200 discrete time steps. There is an underlying "true value" process. The initial true value is drawn from a Gaussian distribution with mean 50 and standard deviation 12 (in general, all values are truncated at 0 and 100 whenever that may be an issue). Then, at every time step, there is a probability p_j that the true value jumps. We consider two different types of jumps. In the first type, which is more realistic, the amount of the jump is drawn from a Gaussian distribution with mean 0 and variance σ_j^2. In the second type, which is meant to simulate a very problematic case for an information based market maker, the new value is itself drawn uniformly at random between 0 and 100. At any point in time, an arriving trader receives a valuation w_t drawn from a (truncated) Gaussian distribution with mean equal to the true value at that time, and variance σ_ϵ^2. If w_t exceeds the current infinitesimal price, the trader initiates a buy order, and if it is less the trader initiates a sell order. The quantity to be bought or sold is drawn at random from an exponential distribution with rate parameter α_q.[3] In our experiments, we set $p_j = 0.01$, $\sigma_j = 5$, and $\sigma_\epsilon = 5$. α_q is set to 0.05 so that the mean trade size is 20. The b parameter for the LMSR market maker was set to 125 and the window size parameter for BMM was set to 5. These choices of the MM parameters were in order to make the average spread approximately equal in the Gaussian jumps case, and were then used again for the human subject trading experiments described in the next section.

Figure 2 gives some intuition into the behavior of BMM as compared with LMSR. This is for a single experiment, and shows that BMM can adapt rapidly to changing valuations in the trading population, while at the same time settling into periods of low spreads and stable behavior at equilibrium. The typical behavior is to start off with a high variance (and hence high spread), and then quickly converge to a low variance regime. When a jump in the population belief occurs, the market maker can quickly pick up on that fact using the algorithm described previously, because the sequence of trades it sees is usually heavily biased in one direction, which would be more likely to occur if the market maker's beliefs had a higher variance (in contrast, series of

[3]This random quantity model is frequently used in models of zero-intelligence trading and models from the econophysics literature (e.g. [Farmer et al. 2005]).



Spot price versus true market value.

Initial convergence of MM's beliefs (spread) shown by the width of the gray region (log scale).

Spread convergence for window size 10.

Fig. 2: Behaviors of the market makers in an example simulation. The first two figures show BMM with window size 5, and the third shows window size 10. BMM is clearly capable of adapting rapidly to changing behavior in the trading population, but at the same time shows less jumpy and unstable behavior at equilibrium than the LMSR MM. This behavior is explained by the spread (which is a function only of the variance of the MM's belief). The spread starts off high, and increases around times of uncertainty, allowing the mean to move more quickly. At the same time, this can create occasional instabilities even when the underlying population mean has not jumped (note some of the periods of increasing spread when the true value is stable).

Table I: Performance of BMM and LMSR in simulated trading

	Gaussian Shocks		Uniform Shocks	
	BMM	LMSR	BMM	LMSR
Profit	2081.35	-2457.30	603.40	-1897.98
Max Loss	9479.82	8662.32	50183.77	8384.42
Spread	1.42	1.35	1.79	1.40
RMSD	2.92	5.38	8.78	10.79

trades that are more balanced are more likely to occur in a model with lower variance, since the probability mass is more concentrated in the "likely" region). Because of the adaptivity, in a long stable period there will be times when the variance (and spread) will increase even though no true change has occurred. This becomes more likely as the variance gets lower.

In simulations, for about the same average spread, BMM can in general achieve better market properties in terms of stability at equilibrium as well as profit. In this particular simulation, the average quoted (half) spread for BMM was 1.23 and its profit was 3080.29. The average quoted (half) spread for the LMSR MM was 1.91 and its profit was -956.27. Table I demonstrates this fact more generally by showing results from 1000 simulations. In addition to the average profit and spread, this table also reports the root mean square deviation of the infinitesimal price from the true value (population mean) at any given point in time (a measure of price discovery), and the single worst loss suffered by the market makers in 1000 simulations (in both cases the single worst loss suffered by the LMSR market maker is close to the theoretical bound of 8664.34. BMM performs better on average. However, it is worth noting that, as the probability of a jump goes up, especially in the case where new valuations are drawn uniformly at random, the loss suffered by BMM increases, so it may not be the best choice for highly unstable environments.

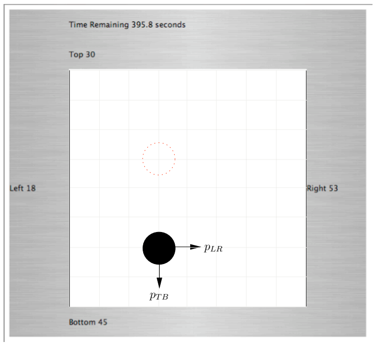


Fig. 3: 2-dim random walk

Table II: Random walk parameters for each of the 6 human subject experiments.

	p	S	z	V
Equilibrium	0.600	4	-1	0.7322
CommonInfoShock	0.600	2	-1	0.5846
jump to	0.600	2	+1	0.1231
LimitedInformation	0.764	4	-3	0.6912
Equilibrium(4)	0.533	4	-3	0.1897
Equilibrium(5)	0.866	4	-3	0.8453
IndivInfoShock	0.826	4	-3	0.7890
jump to	0.516	4	-2	0.2999

5. HUMAN SUBJECT EXPERIMENTS

Human subject experiments by their very nature use small samples; further, human subjects are diverse and very rapid learners, whose attention cannot reliably be maintained for extended time periods. This poses several challenges to live experimental comparison of market makers.

(i) Two comparably sized groups can display vastly different behaviors due to inherent diversity in backgrounds, skill sets and tendencies among human subjects.
(ii) Human subjects, being natural learners, build biases very quickly. So, for example, if you run an experiment for the first time with a market value of (say) 0.7, traders may take some time to become accustomed to the trading task. If you run exactly the same experiment again, it is possible that the second time around, the traders will display more intelligent behavior, with perhaps even a bias that the value is around 0.7, having "generalized" from the previous experiment.

In summary, the live trading experiment should use the same group of traders simultaneously to compare a pair of market makers. Further, the market makers should be compared in a completely symmetric way, using an intuitive interface.

We use a simple web-based trader interface (see Figure 5 in the Appendix). Traders can only place market orders, and in order to elicit information, only the spot price is displayed. A trader can request a trade (buy or sell) at a desired quantity, upon which the trader is quoted a (volume weighted average) price. The trader either accepts or cancels the trade.

5.1. Experimental Design

There are two markets, LR and TB, which are based on the 2 dimensional random walk illustrated in Figure 3. The 2 dimensional random walk is two independent 1 dimensional random walks: horizontal (LR) and vertical (TB). Each random walk is a classic Gambler's Ruin problem [Feller 1958]. The starting position (indicated by the dotted red circle) is (x_0, y_0), and there are two probabilities, p_{LR}, the probability of moving right in the horizontal dimension, and p_{TB}, the probability of moving down in the vertical dimension. The random walk (x, y) is bounded in the grid $[-S, S]^2$. So if $|x| \geq S$, the x-coordinate of the random walk is restarted at x_0 (the y-coordinate is left unchanged) and similarly if $|y| \geq S$, the y-coordinate of the random walk is restarted at y_0 (the x-coordinate is left unchanged).

The values of the markets LR and TB are defined before any particular experiment, based on how often the ball hits the right edge before the left edge, or the bottom edge before the top edge. The probability that the ball hits the right (resp. bottom) edge before the left (resp. top) edge can be computed analytically [Feller 1958]. In terms of $p_{\rm LR},\ p_{\rm TB},\ x_0,\ y_0,\ S$ these values are (For $p \neq \frac{1}{2}$)[4].

$$V_{\rm LR} = \frac{\lambda_{\rm LR}^{S-x_0} - \lambda_{\rm LR}^{2S}}{1 - \lambda_{\rm LR}^{2S}}, \qquad V_{\rm TB} = \frac{\lambda_{\rm TB}^{S-y_0} - \lambda_{\rm TB}^{2S}}{1 - \lambda_{\rm TB}^{2S}},$$

where $\lambda = p/(1-p)$. Traders can simultaneously trade both markets LR and TB. For the experiments, we set $p_{\rm LR} = p_{\rm TB} = p$ and $x_0 = y_0 = z$. Thus, modulo the (mild) vertical-horizontal asymmetry, the two markets are completely symmetric.

Trader Signals. Traders see a realization of the random walk unfolding over time. As shown in Figure 3, the number of times the walk has hit the left, right, top and bottom edges is shown, together with how much time is left. A trader can estimate $V_{\rm LR}$ and $V_{\rm TB}$ from these numbers; for example, from the figure, we can make out from this partial realization of the random walk that $V_{\rm LR} \approx \frac{53}{71} = 0.75$ and $V_{\rm TB} \approx \frac{45}{75} = 0.60$. Although these are realizations of the same random process, we immediately see that the trader is getting a noisy signal of the variable on the basis of which the market pays off (as $t \to \infty$, traders would have perfect information that determines the payoff). This signal improves with time as more information is revealed; in particular, in our example, the error in the traders signal decreases in proportion to $1/\sqrt{t}$. This gradual information revelation is akin to real markets, where traders get better informed over time.

Market Shocks. In a normal equilibrium setting the parameters p, S, z are fixed. We can institute a market shock during the random walk by changing one or more of these parameters. Changing these parameters can reflect different types of market shocks in the real world – for example, if p changes, there is no visible cue, and traders have to infer a change in the underlying dynamics from observables.

5.2. Description of Experiments

In our experiments, participants received an endowment of cash and shares. We ran several experiments summarized below (see the appendix for details).

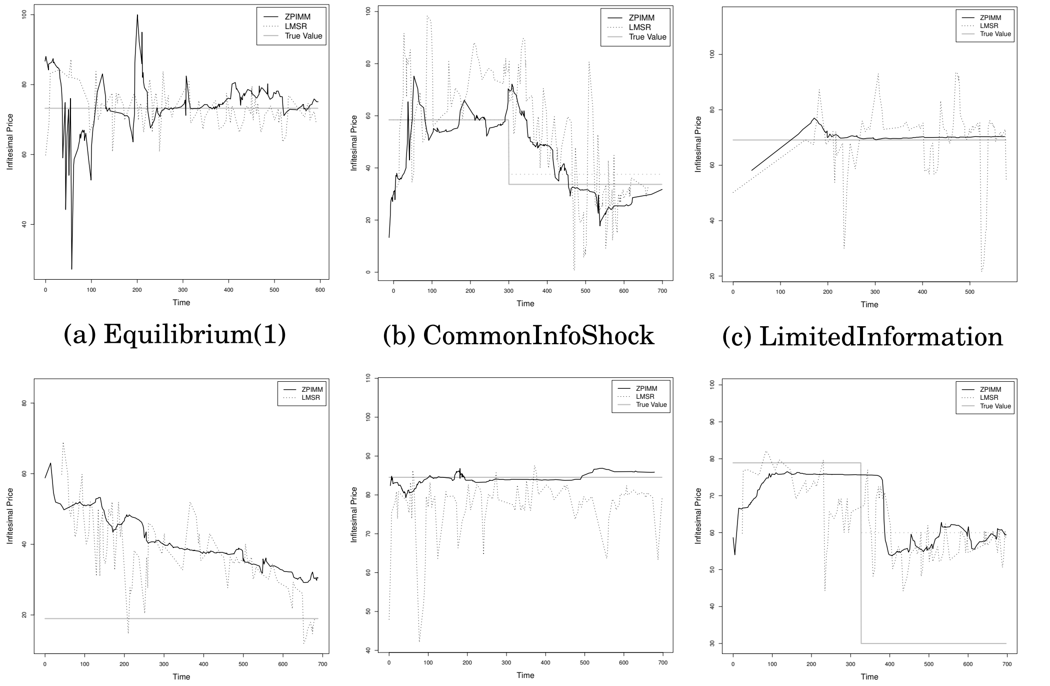
Experiment 1: Equilibrium Each trader viewed their own independent realization of the random walk for 10 minutes and had access to the market's price history (from which they may try to infer the information of other traders). We had 11 traders.

Experiment 2: Common Information Shock All traders viewed the same random walk realization. They were told that the payoff of the markets would be the actual realized ratios of the two random walks, rather than the analytically computed probabilities. The random walk parameters were "shocked" after 5 minutes. In this case, the traders' information gradually becomes completely correct, and the market maker is eventually trading against perfectly informed traders.

Experiment 3: Limited Information Equilibrium Similar to Experiment 1, except for the fact that traders only saw their personal realization of the random walk for 2 minutes. They were allowed to trade for 10 minutes. 17 students participated.

Experiments 4, 5, and 6: Equilibrium With Probabilistic Shocks Participants were told that they would be participating in 3 consecutive games. In each game, the random walk would start off with some combination of parameters. With a 50% chance, these parameters would change between minutes 3 and 7 of the random walk. Traders were not told whether or not there would be a jump in a particular experiment. There happened to be no shock (change in parameters) in Experiments 4 and 5, while there was a shock in Experiment 6 (therefore we call it IndivInfoShock below). Trading went on for 10 minutes. 17 students participated.

[4]For $p = \frac{1}{2}$ one has to take a limit (eg. $V_{\rm LR} = \frac{S+x_0}{2S}$)



(a) Equilibrium(1) (b) CommonInfoShock (c) LimitedInformation

(d) Equilibrium(4) (e) Equilibrium(5) (f) IndivInfoShock

Fig. 4: Summary behavior of LMSR and BMM in live trading experiments. Experiments (a) and (b) are from one deployment in which BMM used a window size of 5 and (c) through (f) from a later one in which BMM used a window size of 10. In general, BMM exhibits more stable behavior than LMSR. Additionally, In the "shock" experiments (b) and (f), BMM is now able to adapt as well as, if not better than, LMSR.

5.3. Results

Figure 4 shows the main results of the experiments, and Table III shows some statistics on the price processes. There are various interesting phenomena in the individual experiments, discussed below, but the big picture is relatively clear. BMM has a much higher profit than LMSR in five of the six live experiments, while at the same time producing a more stable price process, with better price discovery, as measured by distance from the "true" value (RMSD values in Table III). The behavior of BMM is improved in experiments 3 through 6, which were run with a longer adaptive window of 10, leading to more stability (and potentially slower adaptivity). While higher values of the b parameter for LMSR would lead to improved stability (and slower adaptivity), this would come at the cost of making even greater losses.

Experiments 1 and 2, and lessons learned. In the Equilibrium(1) experiment, there are some severe fluctuations at about the 75 sec and 200 sec marks. The fluctuations around the 75 sec mark are probably due to individuals who had outlier realizations early on. The fluctuations around the 200 sec mark are due to a single irrational "rogue" trader who was willing to buy at a price of 100. Unfortunately, since there is no penalty for random wild trading (unlike in real financial markets), such behavior can arise. Discounting these anomalous trades, BMM converges nicely to equilibrium, as does LMSR (except for its characteristic oscillations). Further, in the MarketShock experiment, BMM now adapts as fast if not faster than LMSR.

Table III: Summary Results of Human Subject Experiments

	Profit		Spread		RMSD		RMSDeq	
	LMSR	BMM	LMSR	BMM	LMSR	BMM	LMSR	BMM
Equilibrium(1)	-1350.12	47231.77	3.12	4.04	4.92	7.66	3.73	3.01
CommonInfoShock	-1510.89	8972.50	3.06	3.21	20.67	15.98	16.51	6.76
LimitedInformation	-1602.14	4083.95	1.61	0.49	14.43	2.15	14.56	0.93
Equilibrium(4)	-2619.07	-10588.86	1.81	0.95	21.67	23.13	14.82	17.05
Equilibrium(5)	-3168.55	9134.58	1.42	0.51	11.18	1.5	8.15	1
IndivInfoShock	-92.29	20226.44	1.89	0.89	8.87	6.47	6.88	6.6

Note: 'Spreads' for each trade are computed on an order size of 40 (the average trade size observed). RMSD is the root mean square deviation of the MM's belief from the true value. $RMSD_{eq}$ is the same metric evaluated after "convergence," defined for convenience as the time period halfway between the last change to the true value and the end of the trading period. The number of trades made with each market maker in each experiment was roughly comparable (see Appendix for further details).

The BMM profit in Equilibrium(1) is a little misleading because about 30,000 of it was due to the rogue trader; BMM does what it is supposed to do though, by adapting and making profit based on its Bayesian learned valuation. This wild trader also accounts for the increased RMSD of BMM in this experiment. After the market equilibrates and finds the true value, the RMSD of BMM is much better than that of LMSR (the RMSD_{eq} row in Table III). Similarly, when the market is close to equilibrium in the MarketShock experiment (after seven and a half minutes of trading time in total; since the jump occurs after five minutes, we give the market half of the remaining time to equilibrate) BMM outperforms LMSR in terms of RMSD by a significant margin.

These experiments reveal a couple of interesting facts. First, the behavior of some rogue traders can seriously impact outcomes. In this case, it seems that, when given large initial endowments and the ability to sell short, some traders use their market power to full effect without worrying about profit. So we decided to give people more "reasonable" endowments in the future, including an endowment of stock to start with, and prohibit short-selling. This likely leads to a psychologically more understandable scenario for participants, and less possibility for arbitrary manipulation by traders who are psychologically uninvested in the outcome.

Second, the spreads and behavior of BMM were somewhat less stable than we had expected based on simulation. BMM often increases the spread in response to market conditions, even though there are relatively few shocks in the system (see Appendix, Figure 6). While this still yields good behavior, we hypothesized that tweaking the window parameter would lead to more stable behavior without sacrificing adaptivity too much. Therefore, we changed the window size to 10 for the next set of experiments.

Experiments 3 through 6. Experiments 3 through 6 demonstrate the typical behaviors of BMM and LMSR clearly. There are a couple of interesting details that emerge from the experiments. First, in Experiment 4 (Equilibrium(4)), convergence to the true value is very slow for both LMSR and BMM. While LMSR comes close to the true value in the last few seconds before the end of trading, BMM fails to do so. We hypothesize that this is because this market was the only one in which the true value was below the starting value of 50, and thus necessitated people selling their initial endowment to get to the true value. In this case, BMM also takes a fairly substantial loss, because it was misled by the trading behavior.

Second, in Experiment 6 (IndivInfoShock), while participants were told that the liquidation value would only be the true value *after the shock*, later interviews revealed that they thought the true value would be the average of the two true values. Therefore, the stock ended up trading at around 60, instead of the final true value of 30.

In both these cases, it is nice that the symmetry of the experimental design enables fair comparison between LMSR and ZPIMM: trader behavior leads to anomalies for *both* market makers. Indeed, even the person running the experiment was unaware of which market maker was making which market (top-bottom vs left-right).

Summary. Our live trading experiments demonstrate several key facts. First, while LMSR has nice theoretical properties that suggest it will converge to rational expectations equilibria, in practice this is asking an awful lot of the participants in the market. As long as traders' posterior beliefs do not converge to a single point, there will remain trading incentives, and this is in evidence in all our experiments. LMSR suffers from characteristic fluctuations in the spot price even after it should have attained equilibrium. BMM, on the other hand, provides a tighter belief once it has converged, and has attractive potential to make markets without losing money, or even at a profit. It manages this while providing superior price discovery and spread properties in our live trading experiments.

Experiment 4 (Equilibrium(4)) provides evidence that BMM may sometimes suffer high losses, especially when the market behaves strangely. While occasional such instances are not a huge problem, it will be important to monitor and understand the circumstances that can lead to high losses so that we can ensure that they cannot be reproduced by manipulators intentionally deceiving the market maker.

6. TRADING AGENT EXPERIMENTS

We simulate realistic market conditions with complex trading agents. Automated agents are divided into technical and fundamentals traders, with the latter continuously receiving new information. By varying the proportions of these traders, we are able to simulate different market conditions.

6.1. Scenario

There are up to R rounds $1 \ldots R$. During round i, each fundamentals trader is told the outcome of one per-trader Bernoulli trial with success probability p_i. The traders can then buy and sell the security. Technical traders do not receive information about p_i directly, instead receiving a list of execution prices for trades of the security.

p_0 is chosen uniformly at random between 0 and 1. At the beginning of a round, p_i is calculated as follows: with probability $1/R$, $p_i \sim N(p_{i-1}, \sigma_{\text{jump}})$, and $p_i = p_{i-1}$ otherwise. If $p_i \leq 0$ or $p_i \geq 1$, the security liquidates prematurely at 0 or 100 respectively. Otherwise, the value of the security is $100p_R$ after round R. We generally use $R = 100$, with the exception of equilibrium results such as RMSDeq; here we use shorter $R = 50$ simulations. The shorter simulations do not have jumps, although traders are told there could be jumps. This allows us to measure equilibrium properties in a controlled setting. $\sigma_{\text{jump}} = 0.2$ in all trading agent experiments.

6.2. Agents

Fundamentals traders. We use two types of fundamentals traders, both based on maximizing the expectation of a linear utility function after round R, subject to a constraint on the variance of the final utility. One type of trader uses only the information it receives from its Bernoulli trials, while the other assumes rational expectations and also incorporates price history. The rational expectations trader uses an inference process very similar to BMM. See Appendix B for a derivation of the expectation and variance of the final utility for both traders.

Technical traders. Our technical trading agents are based on two simple stock market trading rules [Brock et al. 1992]. One agent trades based on two moving averages, and the other maintains a range based on price history; see Appendix B for details. These agents function primarily as noise traders.

Table IV: Results of trading agent experiments

	Average profit		Max loss		Spread		RMSD		RMSDeq	
	bmm	lmsr	bmm	lmsr	bmm	lmsr	bmm	lmsr	bmm	lmsr
10%	-823.74	-1915.51	-30471.51	-7435.33	2.38	2.35	16.09	19.27	5.97	6.63
40%	16630.89	-1496.90	-73382.41	-10396.75	1.24	1.94	12.19	12.95	3.58	6.30
60%	23630.75	-1097.00	-480742.72	-10397.06	1.06	1.88	10.81	14.05	3.10	6.15
100%	-295.61	-3055.04	-75582.42	-9679.35	0.94	1.95	9.28	8.42	3.04	4.87
RE40%	34494.88	-2008.72	-379003.83	-10396.80	1.62	2.02	13.32	14.61	4.87	4.59
RE60%	25223.28	-2312.65	-689535.67	-10397.03	1.28	1.99	11.60	12.05	3.62	4.81
RE100%	-738.83	-3077.43	-66105.85	-9807.95	1.03	1.98	9.67	9.10	3.15	4.56

Note: Only LMSR b=150 is shown; see Appendix B for additional results. Each experiment uses 10 traders: the left hand column indicates the %age of fundamentals traders in the experiment, and the remainder is split between each type of technical trader. RE denotes experiments with 2 rational expectations traders (which are counted as fundamentals traders); the other experiments include only pure fundamentals traders. The quoted spread is at 30 shares. Each row summarizes 1500 experiments per MM. Convergence for RMSDeq is defined from the first time the market price comes within 2 of the true underlying value.

6.3. Results

Table IV shows summary statistics for experiments with BMM and one parameter setting of LMSR. See Appendix B for more results and LMSR parameter settings. In all of the trading agent experiments, BMM uses an order size α of 3, and $\sigma_\epsilon = 5.0$.

BMM maintains a lower spread and RMSD from the true underlying market value, while at the same time losing less money on average than comparable LMSR b parameter settings. Increasing LMSR's b parameter reduces the average spread and RMSD in equilibrium, at the cost of an increase in the expected loss. Conversely, reducing b sacrifices RMSD and spread while still falling below BMM's expected profit for many combinations of technical and fundamentals traders. BMM's maximum loss across all 15000 experiments is 689535.67, while LMSR only achieves its theoretical loss bound in the worst case (10397.20 for b=150). BMM performs best in experiments with an approximately even mix of fundamentals and technical traders.

The addition of rational expectations traders, who perform the same inference process as BMM, significantly reduces LMSR's equilibrium RMSD in markets with technical traders. These rational expectations traders make more profit than the pure fundamentals traders, and LMSR loses significantly more in expectation with them (for b=150 with 60% fundamentals traders, a 95% confidence interval on LMSR's profit with rational expectations traders is [-2501.09, -2124.20], but is [-1434.49, -759.51] without). The rational expectations traders are taking the same risks that BMM does, and in doing so they provide more market stability (now LMSR is closer to convergence in equilibrium, as would be predicted by RE models) while extracting a profit.

7. CONCLUSION

We have presented an adaptive, information based Bayesian market maker BMM for binary (or continuous outcome) markets. In experiments with human subjects as well as with intelligent trading agents, when controlling for liquidity (as measured by spread), BMM demonstrates significantly better convergent behavior at equilibrium than LMSR, while being equally adaptive to changes in the market's valuation of the security. Further, in all our experimental settings, BMM on average loses much less money than LMSR, implying that it could provide substantially better liquidity at lower cost than LMSR. One caveat is that, unlike LMSR, BMM is not loss bounded. Another is that it is not as simple to extend BMM to combinatorial markets.

Future work includes understanding when BMM is likely to make losses and/or be manipulated (BMM has been successfully deployed in a longer-term prediction market with human subjects (the RPI Instructor Rating Markets), many of whom attempted to manipulate the markets, and did not succeed in exploiting BMM [Chakraborty et al. 2011]); extending BMM to combinatorial markets; and, further investigating the interplay between convergence, loss and adaptability for market makers.

REFERENCES

AMIHUD, Y. AND MENDELSON, H. 1980. Dealership market: Market-making with inventory. *J. Fin. Econ. 8*, 31–53.

ARROW, K. ET AL. 2007. Statement on prediction markets. AEI-Brookings Joint Center Related Publication No. 07-11.

BERG, J., FORSYTHE, R., NELSON, F., AND RIETZ, T. 2001. Results from a dozen years of election futures markets research. *Handbook of Experimental Economic Results*, 486–515.

BERG, J. AND RIETZ, T. 2003. Prediction markets as decision support systems. *Information Systems Frontiers 5,* 1, 79–93.

BROCK, W., LAKONISHOK, J., AND LEBARON, B. 1992. Simple technical trading rules and the stochastic properties of stock returns. *The Journal of Finance 47,* 5, pp. 1731–1764.

CHAKRABORTY, M., DAS, S., LAVOIE, A., MAGDON-ISMAIL, M., AND NAAMAD, Y. 2011. Instructor rating markets. In *Proc. of the Second Conference on Auctions, Market Mechanisms, and Their Applications*.

CHEN, Y., DIMITROV, S., SAMI, R., REEVES, D., PENNOCK, D., HANSON, R., FORTNOW, L., AND GONEN, R. 2009. Gaming prediction markets: Equilibrium strategies with a market maker. *Algorithmica*, 1–40.

CHEN, Y. AND PENNOCK, D. 2007. A utility framework for bounded-loss market makers. In *Proc. UAI*. 49–56.

DAS, S. 2005. A learning market-maker in the Glosten-Milgrom model. *Quantitative Finance 5,* 2, 169–180.

DAS, S. 2008. The effects of market-making on price dynamics. In *Proc. AAMAS*.

DAS, S. AND MAGDON-ISMAIL, M. 2008. Adapting to a market shock: Optimal sequential market-making. In *NIPS*. 361–368.

FARMER, D. J., PATELLI, P., AND ZOVKO, I. I. 2005. The predictive power of zero intelligence in financial markets. *PNAS 102,* 11, 2254–2259.

FELLER, W. 1958. *An Introduction to Probability Theory and its Applications, Vols. 1 and 2*. Wiley.

GLOSTEN, L. R. AND MILGROM, P. R. 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 14*, 71–100.

HANSON, R. 2007. Logarithmic market scoring rules for modular combinatorial information aggregation. *J. Prediction Markets 1,* 1, 3–15.

HANSON, R. 2009. On market maker functions. *Journal of Prediction Markets 3,* 1, 61–63.

MILGROM, P. AND STOKEY, N. 1982. Information, Trade and Common Knowledge. *J. Econ. Theory 26,* 1, 17–27.

OTHMAN, A. AND SANDHOLM, T. 2010. Automated market-making in the large: the gates hillman prediction market. In *ACM Conference on Electronic Commerce*. 367–376.

OTHMAN, A., SANDHOLM, T., PENNOCK, D. M., AND REEVES, D. M. 2010. A practical liquidity-sensitive automated market maker. In *ACM Conference on Electronic Commerce*. 377–386.

PENNOCK, D. 2004. A dynamic pari-mutuel market for hedging, wagering, and information aggregation. In *Proc. of the 5th ACM Conference on Electronic Commerce*. 170–179.

PENNOCK, D. AND SAMI, R. 2007. Computational aspects of prediction markets. In *Algorithmic Game Theory*, N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, Eds. Cambridge University Press.

SERVAN-SCHREIBER, E., WOLFERS, J., PENNOCK, D., AND GALEBACH, B. 2004. Prediction markets: does money matter? *Electronic Markets 14,* 3, 243–251.

WOLFERS, J. AND ZITZEWITZ, E. 2004a. Five open questions about prediction markets. Stanford GSB Working Paper.

WOLFERS, J. AND ZITZEWITZ, E. 2004b. Prediction markets. *J. Econ. Perspectives 18,* 2, 107–126.

APPENDIX

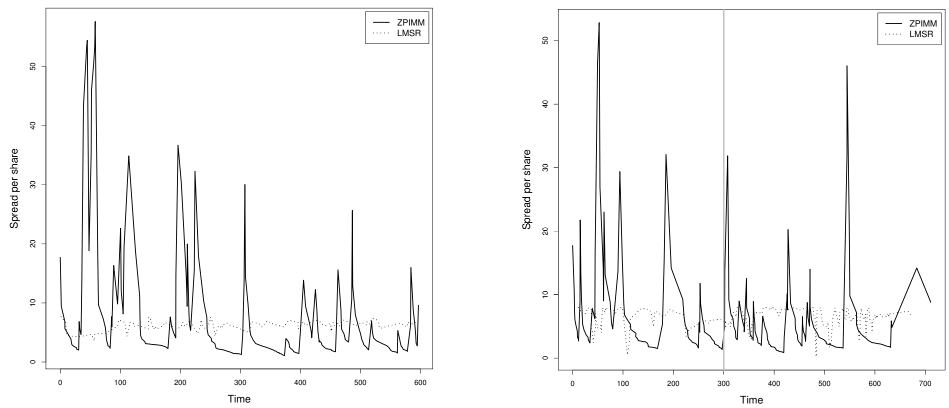
A. DETAILS OF HUMAN SUBJECT EXPERIMENTS

Our data were collected in three distinct trading sessions; we used results from the first two sessions in order to improve the design of the second session. All the traders in our experiments were relatively sophisticated; they all had prior experience with the trading interface and knowledge of prediction markets. In each case they sat in the same room and traded using the web-based interface on their personal laptops.

The first two sessions were individual experiments in which students from a graduate-level Computational Finance class were recruited to participate. 11 and 9 students respectively chose to participate in the two experiments. Participants were incentivized with gift certificates: the trader with largest return received $15; the second best received $10 and the three next best traders $5 each.

Fig. 5: Trading interface

The third session was an educational deployment of the market as part of a graduate / advanced undergraduate class on E-Commerce. Students were studying prediction



(a) Equilibrium(1) (b) CommonInfoShock

Fig. 6: Dynamic behavior of LMSR and BMM spread in Experiments 1 and 2. BMM often increases its spread in response to unlikely trade sequences, making it highly adaptive, while sacrificing some stability.

markets and participated in four trading games during the class. They were incentivized with the opportunity to earn extra credit in the class. 17 students participated. 10 points of total extra credit were allocated for the experiment, with the 10 points divided proportionally among all traders who overall made profit in the experiments.

In each case, the LMSR based market maker was configured with the loss parameter b set to 125. BMM was configured to begin with belief $\mu_t = 50$, $\sigma_t = 12$, and estimation of the trader noise given by $\sigma_\epsilon = 5$. The window of trade history for the adaptive mechanism was set to 5 for the first two experiments and to 10 for the last four. In the first two experiments traders started with $100,000$ units of currency and 0 shares, and were allowed to take both long and short positions in each market. In the four subsequent experiments, traders began with an initial endowment of $10,000$ units of currency, and 200 shares, but were not permitted to take short positions. We flipped the market makers being used in the TB and LR markets between experiments.

We now provide details of the experiments. For experiments 1 and 2, traders were told that there may or may not be a change in the underlying parameters governing the random walk. The conditions for the remaining experiments are described below. In all cases except for Experiment 2, final payoffs were based on the analytically computed probabilities described above. A summary of the parameters used in each experiment are shown in Table II.

Number of trades. The table below provides a summary of the comparison of the LMSR and BMM market makers in terms of the number of *confirmed* buy or sell trades executed in each experiment. The numbers are roughly comparable, although LMSR usually comes in a little bit higher, presumably because it provides profit opportunities due to price fluctuations when BMM has stabilized.

	# Buy/Sell LMSR	# Buy/Sell BMM	# Traders
Equilibrium	119	146	11
CommonInfoShock	213	131	9
LimitedInformation	116	63	17
Equilibrium(2)	109	113	17
Equilibrium(3)	118	85	17
IndivInfoShock	113	121	17

B. DETAILS OF TRADING AGENT EXPERIMENTS

B.1. Results summary

Table V: Market maker profits in trading agent experiments

	Average profit				Max loss			
	BMM	50	150	250	BMM	50	150	250
10%	-823.74	-510.01	-1915.51	-2577.45	-30471.51	-3465.73	-7435.33	-8903.50
20%	-565.58	-418.03	-1881.85	-3667.98	-159461.37	-3463.13	-10397.19	-13639.45
40%	16630.89	-5.44	-1496.90	-3518.06	-73382.41	-3465.74	-10396.75	-16977.65
60%	23630.75	-265.93	-1097.00	-2971.46	-480742.72	-3465.74	-10397.06	-17326.56
80%	10083.41	-924.73	-1661.31	-3169.62	-388195.68	-3439.30	-10396.75	-17327.48
100%	-295.61	-998.68	-3055.04	-4914.08	-75582.42	-3226.92	-9679.35	-16132.79
RE40%	34494.88	-762.72	-2008.72	-3844.91	-379003.83	-3379.01	-10396.80	-16476.47
RE60%	25223.28	-623.20	-2312.65	-3398.85	-689535.67	-3465.74	-10397.03	-17328.27
RE80%	7822.27	-953.80	-2318.88	-3340.74	-270165.67	-3232.11	-10397.13	-17297.80
RE100%	-738.83	-958.66	-3077.43	-5158.40	-66105.85	-3229.84	-9807.95	-16449.93

Table VI: Spread and RMSD in trading agent experiments

	Spread				RMSD				RMSDeq			
	BMM	50	150	250	BMM	50	150	250	BMM	50	150	250
10%	2.38	5.71	2.35	1.47	16.09	12.86	19.27	22.41	5.97	8.51	6.63	5.05
20%	1.70	5.75	2.15	1.40	12.85	12.16	14.70	18.42	4.48	7.51	6.76	5.57
40%	1.24	5.40	1.94	1.25	12.19	12.97	12.95	14.15	3.58	7.01	6.30	5.66
60%	1.06	5.45	1.88	1.18	10.81	12.40	14.05	12.78	3.10	6.64	6.15	5.73
80%	0.96	5.49	1.87	1.16	9.76	10.16	12.60	12.12	3.03	6.16	5.69	5.37
100%	0.94	5.66	1.95	1.21	9.28	8.60	8.42	8.89	3.04	5.71	4.87	4.71
RE40%	1.62	6.37	2.02	1.26	13.32	15.18	14.61	14.18	4.87	5.85	4.59	4.68
RE60%	1.28	5.87	1.99	1.21	11.60	12.30	12.05	12.60	3.62	6.05	4.81	4.22
RE80%	1.14	5.72	1.98	1.18	10.33	10.45	12.10	12.48	3.26	5.78	4.69	4.21
RE100%	1.03	5.83	1.98	1.20	9.67	8.97	9.10	9.52	3.15	5.63	4.56	4.05

Note: Results are shown for BMM, LMSR b=50, b=150, and b=250. Each experiment uses 10 traders, and the percentage of fundamentals traders is indicated in the first column; the remainder of the traders are technical traders, split evenly between range traders and moving average traders. RE denotes experiments with 2 rational expectations traders, which are counted as fundamentals traders. Statistics for each market maker and with each percentage of fundamentals traders are calculated from 1500 experiments. Convergence for RMSDeq is defined as the period after the market price comes within 2 of the true value for the first time.

B.2. Traders

B.2.1. Expected final utility. Given an assumption about p_i, a utility function $u(S, C, \hat{p}_R)$, a number of shares S of the security, and cash C, we calculate an expected final utility

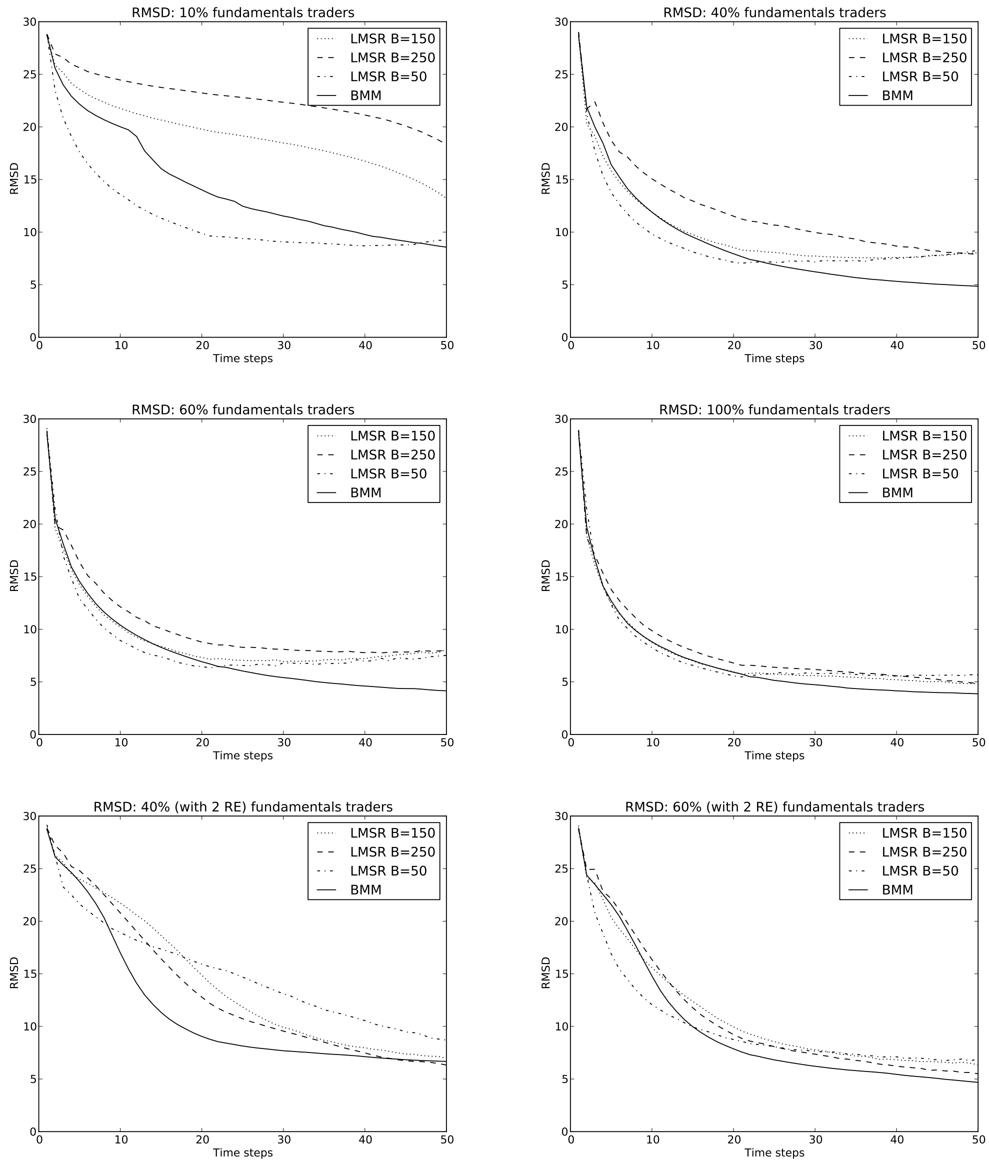


Fig. 7: Convergence of BMM and LMSR for 50-timestep trading agent experiments without jumps; traders still believe that jumps are possible. BMM generally reaches a lower equilibrium RMSD than the LMSR b parameters we tested. New liquidity is added to the market as traders become more confident that there will not be a future jump, lowering their estimate of the variance of their final utility and allowing them to borrow more while still satisfying their variance constraint. For markets without enough liquidity, such as the 10% fundamentals market above, this causes prices to become more accurate toward the end of the experiment. Discontinuities at low timesteps, such as in the 40% (no RE) plot for b=250, are due to technical traders beginning to trade. The top four plots show results for experiments with no rational expectations traders, while the bottom two plots show experiments with 2 rational expectations traders.

with the information available at round i as follows:

$$U_{\text{point}}(i, \hat{p}_i, S, C) = \sum_{k=0}^{R-i} \Bigg[B(k; R-i, 1/R) \left(\prod_{j=1}^{k-1} \Phi(0 < x < 1; \hat{\mu}_j, \hat{\sigma}_j) \int_0^1 d\hat{p}_R \phi(x = \hat{p}_R; \hat{\mu}_k, \hat{\sigma}_k) u(S, C, \hat{p}_R) + H(\hat{p}_i, k) u(S, C, 1) + L(\hat{p}_i, k) u(S, C, 0) \right) \Bigg] \tag{1}$$

$B(k; n, p)$ is the binomial PDF, $\phi(x; \mu, \sigma)$ is the normal PDF, and $\Phi(x; \mu, \sigma)$ is the normal CDF. The functions H and L are the probability that $p_R = 1.0$ (high) and $p_R = 0.0$ (low) respectively given that there are k jumps.

This is an approximation to the true distribution of p_R: each jump adds a normally distributed random variable, then truncates, leading to a non-Gaussian distribution after two jumps. However, the approximation is quite good for small numbers of jumps, and large numbers of jumps become increasingly improbable. Given this approximation, we calculate $\hat{\mu}_i$ and $\hat{\sigma}_i$ as the mean and standard deviation of a truncated Gaussian plus a normal with mean 0 and standard deviation σ_{jump}:

$$\hat{\mu}_1 = \hat{p}_i \tag{2}$$

$$\hat{\mu}_i = \hat{\mu}_{i-1} + \hat{\sigma}_{i-1} \frac{\phi(\frac{-\hat{\mu}_{i-1}}{\hat{\sigma}_{i-1}}) - \phi(\frac{1-\hat{\mu}_{i-1}}{\hat{\sigma}_{i-1}})}{\Phi(\frac{1-\hat{\mu}_{i-1}}{\hat{\sigma}_{i-1}}) - \Phi(\frac{-\hat{\mu}_{i-1}}{\hat{\sigma}_{i-1}})} \tag{3}$$

$$\hat{\sigma}_1 = \sigma_{\text{jump}} \tag{4}$$

$$\hat{\sigma}_i^2 = \sigma_{\text{jump}}^2 + \hat{\sigma}_{i-1}^2 \left[1 + \frac{\frac{-\hat{\mu}_{i-1}}{\hat{\sigma}_{i-1}} \phi(\frac{-\hat{\mu}_{i-1}}{\hat{\sigma}_{i-1}}) - \frac{1-\hat{\mu}_{i-1}}{\hat{\sigma}_{i-1}} \phi(\frac{1-\hat{\mu}_{i-1}}{\hat{\sigma}_{i-1}})}{\Phi(\frac{1-\hat{\mu}_{i-1}}{\hat{\sigma}_{i-1}}) - \Phi(\frac{-\hat{\mu}_{i-1}}{\hat{\sigma}_{i-1}})} - \left(\frac{\phi(\frac{-\hat{\mu}_{i-1}}{\hat{\sigma}_{i-1}}) - \phi(\frac{1-\hat{\mu}_{i-1}}{\hat{\sigma}_{i-1}})}{\Phi(\frac{1-\hat{\mu}_{i-1}}{\hat{\sigma}_{i-1}}) - \Phi(\frac{-\hat{\mu}_{i-1}}{\hat{\sigma}_{i-1}})} \right)^2 \right] \tag{5}$$

$H(\hat{p}_i, k)$ and $L(\hat{p}_i, k)$ are the probability of the stock liquidating at 1 and at 0 respectively, given that there are k jumps in the future; any of the jumps might cause the stock to liquidate, in which case the remaining jumps do not actually happen.

$$H(\hat{p}_i, k) = \sum_{t=1}^{k} \left[\Phi(x \geq 1; \hat{\mu}_t, \hat{\sigma}_t) \prod_{j=1}^{t-1} \Phi(0 < x < 1; \hat{\mu}_j, \hat{\sigma}_j) \right] \tag{6}$$

$$L(\hat{p}_i, k) = \sum_{t=1}^{k} \left[\Phi(x \leq 1; \hat{\mu}_t, \hat{\sigma}_t) \prod_{j=1}^{t-1} \Phi(0 < x < 1; \hat{\mu}_j, \hat{\sigma}_j) \right] \tag{7}$$

B.2.2. Beta trader. Assuming that an estimate $\hat{p}_i$ is the true p_i may be quite problematic, since the trader gets a very noisy signal. Instead, the beta trader assumes that there has not been a jump in the past W turns, maintaining a beta distribution over possible values of p_i. Even if this assumption is violated, the distribution shifts naturally as old information leaves the window and is replaced by post-jump samples. With k successes in the past $n \leq W$ trials (we can trade on less than W trials), we calculate expected final utility as follows:

$$U_{\mathrm{beta}}(i,n,k,S,C) = \int_0^1 d\hat{p}_i \mathrm{Beta}(\hat{p}_i; k+1, n-k+1) U_{\mathrm{point}}(i,\hat{p}_i,S,C) \tag{8}$$

$\mathrm{Beta}(x;\alpha,\beta)$ is the beta PDF. The beta trader attempts to maximize U_{beta} by picking the number of shares S which maximizes $U_{\mathrm{beta}}(i,n,k,S,C)$, taking transaction fees for exchanging cash and shares into account. In our experiments, $W = 20$.

Finally, we must pick a utility function $u(S,C,\hat{p}_R)$. One choice is log utility $u(S,C,\hat{p}_R) = \log(C + 100S\hat{p}_R)$. However, this requires that traders receive an initial allocation of cash and shares; tuning the allocation is problematic. Instead, we use a linear utility function with a variance constraint. The linear utility $u(S,C,\hat{p}_R) = C + 100S\hat{p}_R$, but we calculate both $\mathbb{E}_{\hat{p}_R}[u(S,C,\hat{p}_R)]$ and $\mathbb{E}_{\hat{p}_R}[(u(S,C,\hat{p}_R))^2]$: this allows us to calculate the variance of the final utility. In our experiments, the traders maximize the expectation of a linear utility function subject the standard deviation being less than 1000.

B.2.3. Rational expectations trader. The beta trader is a pure fundamentals trader. By incorporating price history, we can get a more accurate estimate of the true underlying market value than from private information alone.

The rational expectations trader is very similar to BMM, the main difference being that a trader does not need to quote prices, and does not receive information about cancels. Instead, we set the ask price in BMM's inference algorithm to an order's execution price. Belief updates and adaptivity to shocks then follow the same process as for BMM.

BMM's inference algorithm maintains a μ estimating the current market value, and a σ representing the uncertainty of that estimate. We combine this with the beta distribution from our private information as follows:

$$U_{\mathrm{RE}}(i,n,k,S,C) = \frac{1}{H}\int_0^1 d\hat{p}_i \mathrm{Beta}(\hat{p}_i; k+1, n-k+1)\phi(\hat{p}_i;\mu,\sigma) U_{\mathrm{point}}(i,\hat{p}_i,S,C)$$

$H = \int_0^1 dx \mathrm{Beta}(x; k+1, n-k+1)\phi(x;\mu,\sigma)$ normalizes the combined probability distribution. We use a lower bound of 2 for σ in the rational expectations traders, to avoid drowning out private information. One interpretation of this combined distribution is that we sample from each separately until drawing the same value from both. This favors values which are likely according to both our private information and the belief inferred from market prices.

B.2.4. Moving average trader. Keep track of a long and short execution price average, L_i and S_i respectively at time i. If we are in the "low" state and $S_i > (1+\alpha)L_i$, buy and move to the "high" state; if we are in the "high" state and $S_i < (1-\alpha)L_i$, sell and move to the "low" state. α determines a margin, which helps avoid responding to noise. α is 0.05 in our experiments, the short average is over 10 trades, and the long average is over 30 trades.

The amount to buy and sell is capped at 1000 shares, or up to a maximum market impact of 2 from the short average.

B.2.5. Range trader. In a fixed window, find the highest execution price $P_{\max}$ and the lowest $P_{\min}$. If the most recent execution price is more than $(1+\alpha)P_{\max}$, buy. If it is less than $(1-\alpha)P_{\min}$, sell. α again discourages responding to noise, and is 0.05 in our experiments. We use a window size of 20.

The amount to buy and sell is capped at 1000 shares, or up to a maximum market impact of 2 from $P_{\min}$ or $P_{\max}$.