Rational Market Making with Probabilistic Knowledge

Abraham Othman
Computer Science Department
Carnegie Mellon University
aothman@cs.cmu.edu

Tuomas Sandholm
Computer Science Department
Carnegie Mellon University
sandholm@cs.cmu.edu

ABSTRACT

A market maker sets prices over time for wagers that pay out contingent on the future state of the world. The market maker has knowledge of the probability of realizing each state of the world, and of how the price of a bet affects the probability that traders will accept it. We compare the optimal policy for risk-neutral (expected utility maximizing) and Kelly criterion (expected log-utility maximizing) market makers. Computing the optimal policy for a risk-neutral market maker is relatively simple, while computing the optimal policy for a Kelly criterion market maker is challenging, requiring advanced techniques adapted from the computational economics literature to run efficiently. We show that while a risk-neutral market maker has an optimal policy that does not depend on the market maker's state, a Kelly criterion market maker's optimal policy has an intricate dependence on both time and state. Counterintuitively, a Kelly criterion market maker may offer bets that are myopically irrational with respect to the market maker's beliefs for the entire trading period. In contrast, a risk-neutral market maker never offers a myopically irrational bet.

Categories and Subject Descriptors

J.4 [**Social and Behavioral Sciences**]: Economics; I.2.11 [**Distributed Artificial Intelligence**]: Multi-agent Systems

General Terms

Algorithms, Economics, Experimentation

Keywords

Agent Design, Computational Economics, Interpolation, Numerical Dynamic Programming, Kelly Criterion, Wagering

1. INTRODUCTION

Market makers are trading agents that set the prices for assets in exchanges. Market makers profit in two ways: first, by the *bid/ask spread* imposed when they buy a contract at a lower price than they sell it for, and second, by speculatively taking on positions and holding that *inventory* for a profit. Many realistic market-making settings, like Las Vegas sports betting or a proprietary trading desk at a bank, are characterized by a market maker that has a good prior on the future state of the world and on how traders will bet as prices change. A complication arises when traders and the market maker have substantially different beliefs. Then, the market maker must balance two competing factors: the desire to hedge bets for a certain profit, and the desire to profit in expectation from wagers made at favorable prices. For instance, a market maker could find itself in a situation where it could either increase its exposure to an event it thinks will probably occur at a bargain price, or hedge out its current risk on that event in order to guarantee a small but certain profit.

In this paper, we compute the policy of a *Kelly criterion* market maker over a series of interactions with traders. The Kelly criterion [Kelly Jr, 1956] is a way to make bets that mandates maximizing the expected log utility of a setting. While simple as a guiding precept, the Kelly criterion accomplishes a broad range of objectives: over a series of bets, it is the fastest way to double an initial investment, produces the highest median wealth, and produces the highest mode wealth. Poundstone [2006] provides a compelling introduction to the Kelly criterion and its use in practice, particularly by the mathematician and hedge fund manager Ed Thorp.

There are two prior literatures that deal with the sequential interaction of a market maker with traders: artificial intelligence, and finance. In the AI literature, there are *cost function based market makers* [Chen and Pennock, 2007, 2010]—agents which price bets so that they are neutral between them being accepted and them being not accepted [Ben-Tal and Teboulle, 2007, Agrawal et al., 2009]. Another closely-related branch of the AI literature involves *Bayesian market makers* which attempt to learn the correct value of a security by applying Bayes Rule to a series of interactions with traders [Das, 2008, Das and Magdon-Ismail, 2009, Chakraborty et al., 2011]. In contrast to these agents, the market makers we consider here behave rationally in the classical sense: they maximize utility given knowledge of the future and of how their prices affect trader actions.

In the finance literature, Glosten and Milgrom [1985] defined the basic framework used in many later works, including ours, for a market maker interacting with an anonymous pool of traders. Kyle [1985] considers a game-theoretic interaction between a profit-seeking monopolistic market maker and a mix of noise traders and informed traders. Our setting is similar, in that we consider utility-maximizing pricing by a monopolistic market maker, but we assume that no traders are privileged in their information. The most challenging part of our work involves computing the policy of a risk-averse market maker. The notion of a risk-averse, rather than risk-neutral, market maker was introduced in Rock [1996].

Our experimental results show that a Kelly criterion market maker follows a complex time-dependent strategy. In the early stages of wagering, the market maker will attempt to match orders to profit from the bid/ask spread. Towards the end of trading, the policy

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

gradually shifts to myopic optimization on the market maker's private beliefs. Perhaps surprisingly, we show that in the early stages of the market, profiting from the bid/ask spread dominates the desire to sell inventory at agreeable prices, that is, if it facilitates more trade, a Kelly criterion market maker should buy obligations at a price higher than, or sell obligations at a price lower than, its private beliefs. Moreover, because the inventory a risk-averse market maker accumulates affects the prices it offers, the market maker could offer bets that are myopically irrational for the entire trading period. This is in contrast to a risk-neutral market maker that would never offer a myopically irrational bet.

2. MODEL

Following Glosten and Milgrom [1985], the setting is a repeated sequential interaction between the market maker and a set of traders. In each period, the market maker sets prices for a finite set of bets, and then a trader is drawn randomly from a large pool of potential traders. That trader enters the market and selects one of the offered bets to make with the market maker (or none at all). After a finite number of periods the process halts, one of the n events is realized, and the bets are settled with the traders.

Our setting is closest to Das and Magdon-Ismail [2009], which also involved a dynamic stochastic optimization. In that work, however, the market maker was responding to a shock in the state of the world and was attempting to learn the new, correct values for contracts. In contrast, here, the market maker's beliefs over the probabilities of the future state of the world do not change, and the market maker seeks to maximize expected utility over the interaction with the traders.

2.1 Traders

The traders have the following features:

- Traders are *anonymous*, so there is no way for the market maker to distinguish between traders. Anonymity is a standard component of many models in the literature (for example, Feigenbaum et al. [2003] and Das [2008]), because it is natural for settings where prices are posted publicly, as is the standard in electronic markets.

- Traders are *myopic*, not strategic. They exist for only a single period: they enter the market, perceive the prices offered by the market making agent, select a bet to take (or no bet), and then exit. The traders do not learn from historical prices or strategize about their behavior. Myopic traders (also known as *noise traders*) are a feature of much of the literature [Glosten and Milgrom, 1985, Kyle, 1985, Othman and Sandholm, 2010]. Empirical studies of market microstructure have shown that the behavior of these agents is qualitatively very similar to behavior observed in real markets with human traders [Gode and Sunder, 1993, Othman, 2008]. However, in some settings the simple behavior of these agents may be an unrealistic model [Chen et al., 2007, Dimitrov and Sami, 2008, Chen et al., 2010].

- The number of trading periods is drawn independently of the market maker's policy. Since traders have the ability to decline to place a bet with the market maker if they do not find the offered bets agreeable, this condition means that the number of traders placing bets with the market maker is *not* a constant—instead, it will depend on the market maker's policy. We assume the market maker knows the true distribution of the number of trading periods.

2.2 Utility and the Bellman equation

The market maker's *state* can be represented by a tuple $(t, \mathbf{w})$ of the index of the participating agent $t \in \{1, 2, \ldots\}$, and the *wealth vector* $\mathbf{w}$, where w_i is the market maker's wealth (payoff) if state of the world $\omega_i \in \Omega$ is realized. (Since exactly one trader appears in each period, the variable t can be thought of as an index over discrete time.) There is a termination state $(\bar{t}, \mathbf{w})$, where the market maker gets an expected utility payout based on his subjective beliefs $\hat{\mathbf{p}}$, which he believes to be the correct distribution over the possible futures:

$$V(\bar{t}, \mathbf{w}) \equiv \sum_{i=1}^{n} \hat{p}_i u(w_i)$$

Without loss of generality, a risk-neutral market maker receives its expected linear utility on termination:

$$V(\bar{t}, \mathbf{w}) \equiv \sum_{i=1}^{n} \hat{p}_i w_i$$

A Kelly criterion market maker receives its expected $\log$ utility on termination:

$$V(\bar{t}, \mathbf{w}) \equiv \sum_{i=1}^{n} \hat{p}_i \log(w_i)$$

The *bets* a market maker offers can be expressed by vectors in payout space $\mathbf{x} \in \mathbb{R}^n$, so that x_i is the *trader's* payoff (that is, the market maker's loss) if ω_i is realized. For instance, imagine that the market maker is fielding bets on which of three horses will win a horse race. A bet that pays the trader 10 dollars if the first horse wins, 5 dollars if the second horse wins, and nothing if the third horse wins, is represented by the vector $(10, 5, 0)$.

The market maker's *policy* when interacting with trader t, $\pi(t, \cdot) : \mathbb{R}^n \mapsto \mathbb{R}$, maps these vectors to the amount the market maker would charge the agent for each bet. We denote by the zero-vector bet $\mathbf{0}$ an agent declining to make a bet with the market maker, and set $\pi(\mathbf{0}) = 0$. (This can be interpreted as the intersection of the individual rationality constraint of the traders (who would want $\pi(\mathbf{0}) \leq 0$) and of the market maker (who would want $\pi(\mathbf{0}) \geq 0$).) The market maker knows the probability that an agent will accept a bet given the prices. Because traders are anonymous, the market maker has no way to distinguish between traders and so these probabilities are the same for all traders.

In full generality, there is a chance $\delta(t)$ of the interaction terminating immediately before the t-th trader participates. Consequently, the value of being in state $(t, \mathbf{w})$ is

$$\begin{aligned} V(t, \mathbf{w}) = (1 - \delta(t)) \sum_{\mathbf{x}} & \mathbb{P}\left(\text{Trader takes bet } \mathbf{x} \text{ at price } \pi(\mathbf{x})\right) \\ & \cdot V(t+1, \mathbf{w} - \mathbf{x} + \pi(\mathbf{x})) \\ + \delta(t) V(\bar{t}, \mathbf{w}) & \end{aligned}$$

In every state $(t, \mathbf{w})$, a utility-maximizing market maker employs the optimal policy π^* defined by the Bellman equation

$$\begin{aligned} \pi^*(t, \mathbf{w}) = \arg\max_{\pi} (1 - \delta(t)) \sum_{\mathbf{x}} & \mathbb{P}\left(\text{Trader takes bet } \mathbf{x} \text{ at price } \pi(\mathbf{x})\right) \\ & \cdot V(t+1, \mathbf{w} - \mathbf{x} + \pi(\mathbf{x})) \\ + \delta(t) V(\bar{t}, \mathbf{w}) & \end{aligned}$$

with respective values V^* defined by

$$V^*(t,\mathbf{w}) = (1-\delta(t)) \sum_{\mathbf{x}} \mathbb{P}\left(\text{Trader takes bet } \mathbf{x} \text{ at price } \pi^*(\mathbf{x})\right) \cdot V(t+1, \mathbf{w} - \mathbf{x} + \pi^*(\mathbf{x})) + \delta(t) V(\bar{t}, \mathbf{w})$$

Solving these equations when the market maker has log utility is very challenging. We proceed to discuss how we solve for the optimal policy and values in this case.

3. COMPUTATION OF THE POLICY OF A KELLY CRITERION MARKET MAKER

When given a specification of the value function $V^*(t+1,\mathbf{w})$, it is simple to calculate the optimal value V^* and policy π^* of any state in the previous time step t. Thus, backward induction from the termination state is a straightforward way to solve for optimal values and policy across every time step. A complication arises from the difficulty in representing arbitrary $V^*(t+1,\mathbf{w})$. While the termination state is closed form, the previous time steps will generally not have closed form representations. In order to solve a Kelly criterion market maker's problem with backward induction, we must find a way to approximately represent the value function concisely.

3.1 Shape-preserving interpolation

While the value function for an arbitrary time step may have a complex, non-analytic form, we know a great deal about its *shape* from the properties it inherits from the log utility of the terminating state [Stokey et al., 1989]. In particular: (1) it is increasing in wealth, (2) it is concave, and (3) it goes to minus infinity as the wealth in any state goes to zero.

Since these properties are intrinsically linked to the logarithmic utility of the Kelly criterion market maker, we choose to adopt an approximation technique that preserves these properties, *shape-preserving interpolation*. Specifically, we employ the shape-preserving interpolation developed theoretically in Constantini and Fontanella [1990]. By shape-preserving, we mean that the technique retains the partial derivatives, concavity, and monotonicity of the original function, and by interpolation, we mean that the approximated function precisely matches the actual function at a set of interpolating points. While shape-preserving interpolation is well-known in the scientific computing literature [Judd, 1998], this specific technique has been featured rarely. Perhaps the most practical example is Wang and Judd [2000], who study a tax planning problem with stochastic stocks and bonds.

Because the theory of shape-preserving interpolation developed in Constantini and Fontanella [1990] is complete only for two dimensions, we focus only on settings with two events for the rest of the paper. While it does appear possible to extend the interpolation into n dimensions, it would suffer from the curse of dimensionality and take significantly longer to compute the approximate value function. The restriction to two events is not as limiting as it might first appear, because many realistic and popular settings involve wagers on binary events. An example from sports betting is whether the Red Sox or Yankees will win their upcoming match. An example from finance is credit-default swaps, where a bond either does or does not experience a default event.

In order to properly preserve the shape of the function, shape-preserving interpolation requires computing the partial derivatives with respect to the wealth in each state at the interpolating points. We compute these values by using the *envelope theorem*; since $V^*(t,\mathbf{w})$ is given by the maximizing policy π^*, we calculate the partial derivatives with respect to wealth by numerically differentiating the value function when the maximizing policy is followed [Mas-Colell et al., 1995, Wang and Judd, 2000].

We proceed to describe the interpolation procedure at a high

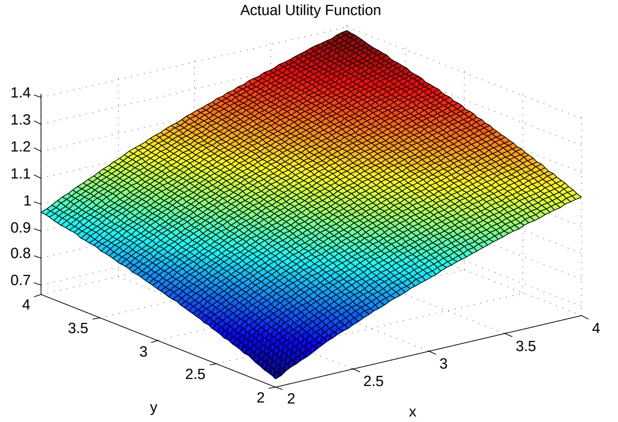

Figure 1: The utility function $u(x,y) = .6\log x + .4\log y$ on the rectangle $[2,4]^2$.

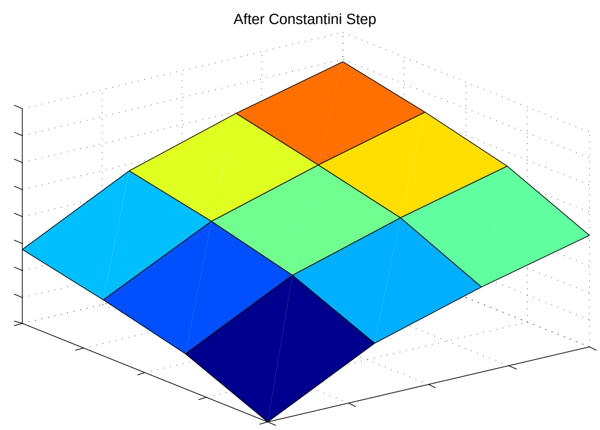

Figure 2: The quilt which matches the function values and partial derivatives.

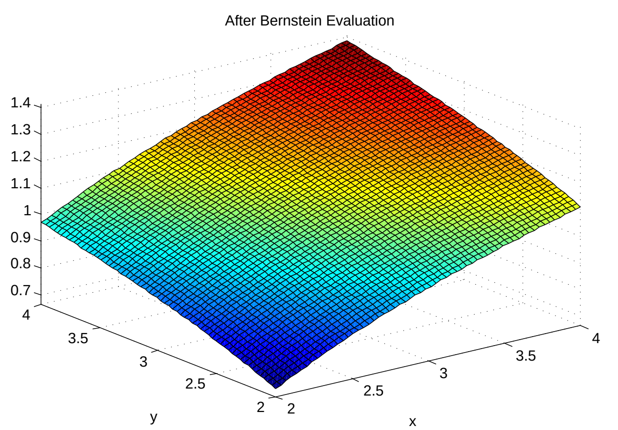

Figure 3: Evaluating the quilt using Bernstein bases produces a good approximation.

level, first on a single rectangle and then over the whole positive orthant. Figure 1 shows a sample expected utility function over a single rectangle.

The first step to creating an approximate interpolating function on this rectangle is to generate a three-by-three *quilt* (continuous, piecewise-linear approximation) of the function by matching the function values and partial derivatives at the vertices of the rectangle. Figure 2 shows the quilt that results from the utility function in Figure 1. This quilt retains the monotonicity, concavity, and partial derivatives at the vertices of the original function.

The final step is to evaluate the quilt using *bivariate Bernstein basis functions*. These are a variation-minimizing set of functions that retain the monotonicity, concavity, and partial derivatives at the vertices of the quilt. Of course, the quilt retained these properties from the original function itself, and so the interpolation is shape preserving. By variation minimizing, we mean that the bases are weighted to produce a polynomial that minimizes the sup ($\mathcal{L}_\infty$) norm error. It is therefore accurate to think of the Bernstein bases as smoothing the piecewise linear quilt [Judd, 1998]. Figure 3 shows the interpolated function that results from the process. Since the Bernstein evaluation step works directly on the quilt, the function is approximated concisely: for each interpolating rectangle we only need to store the sixteen values that create the quilt.

Computing the shape-preserving interpolated function is more involved than a simple linear interpolation (table lookup). However, the benefit of these extra steps is the dramatically improved accuracy of the evaluated function or, put another way, a substantial decrease in the degree of grid fineness required to compute the value function to the same level of accuracy. Table 1 compares the accuracy of the shape-preserving interpolation versus a simple linear interpolation at an arbitrary collection of wealth vectors for the representative utility function used in Figure 1.

Wealth vector	Shape-preserving error	Linear error	Ratio
$(2, 2)$	.0020	.14	72
$(5, 1.1)$	.0026	.36	135
$(20, 25)$	1.2×10^{-6}	.0011	895
$(50, 10)$	1.0×10^{-5}	.0021	210

Table 1: Relative errors for shape-preserving interpolation versus linear interpolation on identical rectangles. At each wealth vector, the interpolating rectangle is $(w_1 \pm 1, w_2 \pm 1)$, i.e., a square with side length 2 centered at the wealth vector.

The shape-preserving interpolation is between 72 and 895 times more accurate than a linear grid at the example points. Perhaps unsurprisingly, we found that the inverse of this relation also appeared to hold—to achieve the same level of accuracy as shape-preserving interpolation, the grid used in linear interpolation would need to be roughly one thousand times finer. We estimate that the running time of our experiments on a commodity PC using linear interpolation would take about a week; in contrast, solving the dynamic program took about ten minutes using shape-preserving interpolation.

3.2 Extending the technique

We have described how shape-preserving interpolation works on a single rectangle over which the function to be approximated is finite. It is straightforward to extend this technique from a single rectangle to a finite grid of rectangles over which the function to be approximated is finite. (In this case, care must to be taken to ensure that the function approximation is continuous at the boundaries of the individual interpolating rectangles, but this can be accommodated without too much additional complexity, see Constantini and Fontanella [1990] for details.)

However, the value function we are approximating is not just a finite function over a finite grid: it fails this in two separate ways. First, since $\lim_{x \downarrow 0} \log x = -\infty$, we have that at the lower boundary of the positive orthant (i.e., values close to zero along either dimension) the value function goes to $-\infty$. Second, the value function has no finite upper bound on its input—it is defined over the entire positive orthant. Consequently, we must extend the interpolation technique from the literature to accommodate the specific properties of a Kelly criterion market maker. Our solution is to have a large finite grid of interpolating rectangles on which we can apply the standard shape-preserving technique, and then to employ custom extensions to approximate below the lower boundary and above the upper boundary of the grid.

3.2.1 *Beyond the lower boundary of the grid*

We interpolate beyond the lower boundary of the grid as if the value were given by setting the value of a state equal to its termination value plus a constant that ensures continuity at the boundary of the grid. Formally, to approximate the value of state $\mathbf{w}$, with nearest point on the interpolating grid $\mathbf{w_g}$, we set

$$V(t, \mathbf{w}) \approx V(\bar{t}, \mathbf{w}) + (V(t, \mathbf{w_g}) - V(\bar{t}, \mathbf{w_g}))$$

(Observe that as $\mathbf{w} \to \mathbf{w_g}$, $V(t, \mathbf{w}) \to V(t, \mathbf{w_g})$). This approximation ensures the monotonicity of the value function and that it goes to negative infinity as the wealth of either state goes to zero, but, it is only an exact approximation for the termination function itself. To ensure that this extension does not change the overall value function substantially, in our experiments we start the interpolating grid at a small value, so the additional interpolation is only relevant over a small fraction of the state space. In our exploratory data analysis, we experimented with different lower bounds for the interpolating grid and found that different small values did not noticeably affect calculated optimal policies. We attribute this to states at the lower boundary of the grid having such low utility that they will be avoided, and are therefore largely irrelevant to the optimization problem as a whole.

3.2.2 *Beyond the upper boundary of the grid*

Consider the market maker's pricing problem at the upper boundary of the grid at time t. If the size of the trader's bet is bounded (say, to be no larger than c), then the market maker can approximately compute the optimal pricing policy by using an interpolating grid at time step $t + 1$ whose upper boundary is larger than the grid at time t by at least c. Using this insight, we eliminate the need to calculate a value beyond the upper boundary of the grid by increasing the upper boundary of the grid as time proceeds. (In fact, recalling that we solve the dynamic program through backward induction, from an algorithmic perspective we are actually reducing the upper boundary of the grid as we solve backwards through time.) In contrast to our extension to compute values below the lower boundary of the interpolating grid that we discussed above, this extension uses the same mechanics as the rest of the shape-preserving interpolation process and so suffers from no additional loss of accuracy.

3.3 Alternative approaches

As an alternative to the gridded approach here, we also considered but rejected a global shape-preserving approximation technique along the lines of De Farias and Van Roy [2003]. This would involve selecting basis functions ϕ_i that are each monotonic and concave, and representing the value function in each time step as a conical combination of these functions:

$$V^*(t, \mathbf{w}) \approx \sum_i \gamma_i^t \phi_i(\mathbf{w}), \quad \gamma_i^t \geq 0.$$

Such a representation retains the monotonicity and concavity properties of the value function and is concise. We rejected this approach for two reasons. First, the heuristic selection of the basis provides little guidance. Which set of functions is a good choice, and why? Observe that many standard basis function selections, such as radial basis functions, will not in general preserve the monotonicity or concavity of the value function and so could lead to nonsensical policies.

The second reason we chose to reject this technique is the difficult optimization to select the weights γ^t. In particular, a standard linear regression that maximizes the deviation from sum of squares at a set of relevant nodes can create aberrant behavior and an approximation that deviates significantly from the actual value function [Gordon, 1995, Guestrin et al., 2001, Stachurski, 2008]. The correct optimization to use to determine the weights is to minimize the sup norm (that is, $\mathcal{L}_\infty$, rather than $\mathcal{L}_2$), which is a significantly more challenging problem to solve numerically [Judd, 1998].

4. EXPERIMENTS

With only two possible events, it is possible to characterize bets in terms of a single event. In particular, setting p to be the probability that the first event occurs implies $1-p$ is the probability that the second event occurs. Applying this logic to the market maker's policy, we can without loss of generality have the market maker buy and sell contracts on the first event only, because buying (selling) a contract on the first event implicitly yields the sale (purchase) of a contract on the second.

The *ask* is the price at which the market maker will sell a contract, and the *bid* is the price at which the market maker will buy a contract. For non-degenerate settings, ask prices will always be higher than bid prices. In this section, we describe the optimal ask and bid prices for two different settings.

4.1 Parameterization

Following Das [2008], in our experiments, traders have a belief drawn from a Gaussian with a mean belief of $p = 0.5$ and standard deviation 0.05. The traders are zero-intelligence agents; a trader visits the market maker exactly once and behaves myopically. They purchase a unit contract if they see an ask price lower than their belief, sell a unit contract if they see a bid price higher than their belief, and do not transact with the market maker otherwise.

In our experiments, we set 50 trading periods (that is, $\delta(t = 51) = 1, \delta(t < 50) = 0$), although we found our results hold qualitatively for other distributions of traders. Recalling from Section 3.2 that the upper boundary of the interpolating grid increases in each trading period, we set the interpolating grid for trader t to $[1, 1.5, 2, 3, \ldots, 250, 250 + t]^2$.

In our experiments with Kelly criterion market makers, we consider only relatively small levels of wealth (alternatively, large bets relative to the amount of wealth). This is because for bets with large levels of wealth, a market maker maximizing the expected log of wealth can be well-approximated by a risk-neutral, linear utility agent. To see why, consider the Taylor expansion of log utility at wealth x:

$$\log(x + \epsilon) = \log(x) + \frac{\epsilon}{x} - \frac{\epsilon^2}{x^2} + \Theta\left(\epsilon^3\right)$$

If x is large enough that $x^2 \gg x$, then $1/x^2 \ll 1/x$. Consequently, at large wealths, the impact of small bets on the utility function can be well-approximated by the linear function $\log(x) + \left(\frac{1}{x}\right)\epsilon$, with negligible higher-order effects.

We now turn our attention to how to calculate the optimal policy for a risk-neutral market maker, and the qualitative properties of that policy.

4.2 Optimal risk-neutral policy

For this setting, a risk-neutral market maker's optimization problem is significantly simpler than the general case. Recall that in the two-event case, the market maker's knowledge of the future, the vector $\hat{\mathbf{p}}$, can be represented by a single scalar $\hat{p}$ (e.g., "Team A has a 50 percent chance of winning the game"). Then the termination state $V(\bar{t}, \mathbf{w})$ is

$$V(\bar{t}, \mathbf{w}) \equiv \hat{p} w_1 + (1 - \hat{p}) w_2$$

Let the agents have beliefs on the first event distributed according to the cumulative density function F with probability density function f. In the penultimate step $\bar{t} - 1$, a risk-neutral market maker sets their bid and ask price to maximize their utility in the termination state, conditioning on three cases: the bid being taken, the ask being taken, and neither offer being taken. Formally,

$$\begin{aligned} V^*(\bar{t}-1, \mathbf{w}) \quad = \quad & \max_{b,a} F(b) V(\bar{t}, (w_1 - b + 1, w_2 - b)) \\ & + (1 - F(a)) V(\bar{t}, (w_1 + a - 1, w_2 + a)) \\ & + (F(a) - F(b)) V(\bar{t}, \mathbf{w}) \end{aligned}$$

and since $V(\bar{t}, \mathbf{w}) = \hat{p} w_1 + (1 - \hat{p}) w_2$ the right-hand side optimization simplifies to

$$\begin{aligned} \max_{b,a} \quad & F(b)(\hat{p}(w_1 - b + 1) + (1 - \hat{p})(w_2 - b)) \\ & + (1 - F(a))(\hat{p}(w_1 + a - 1) + (1 - \hat{p})(w_2 + a)) \\ & + (F(a) - F(b))(\hat{p} w_1 + (1 - \hat{p}) w_2) \end{aligned}$$

which further simplifies to

$$\max_{b,a} V(\bar{t}, \mathbf{w}) + F(b)(\hat{p} - b) + (1 - F(a))(a - \hat{p}) \tag{1}$$

which implies

$$V^*(\bar{t} - 1, \mathbf{w}) + C = V(\bar{t}, \mathbf{w})$$

where C is a constant that does not depend on t or $\mathbf{w}$. Consequently, by inductive argument working back from the terminal state the optimal policy for a risk-neutral market maker does not depend on t or $\mathbf{w}$. Equation 1 also makes it easy to see that the optimal arguments (b^*, a^*) have $b^* \leq \hat{p} \leq a^*$, because if not changing to a policy satisfying that inequality would yield a higher value. Thus, a globally optimal risk-neutral market maker is always myopically rational.

This argument also applies to the general setting discussed in Section 2.2 with more than two bets and events. In that case, by similar reasoning, the result is that a risk-neutral market maker will always price a bet $\mathbf{x}$ such that $\pi(\mathbf{x}) \geq \hat{\mathbf{p}} \cdot \mathbf{x}$. In this more-advanced case, however, traders' demands could be a complex, combinatorial function of the price vector offered by the market maker. If so, computing the optimal policy could be infeasible.

We have shown that the optimal policy of a risk-neutral market maker is constant and invariant to time and wealth. To actually

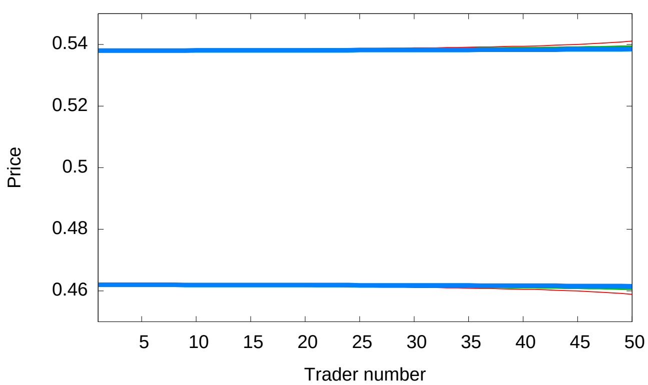

Figure 4: When the market maker's private beliefs align with those of the traders, the optimal ask prices (top lines) and bid prices (bottom lines) do not change significantly over the course of the interaction period.

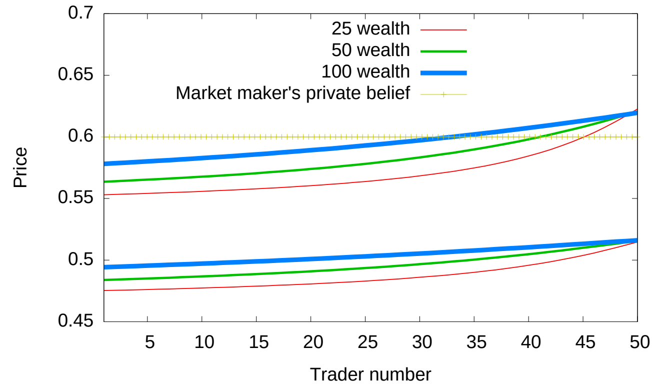

Figure 5: When the market maker's private beliefs do not align with those of the traders, the optimal policy is highly time and wealth dependent.

compute the optimal b^* and a^*, in the simple two-event, unit-bet case we can take the first-order condition of the optimization in Equation 1 to get

$$\begin{aligned} F(b^*)(\hat{p}-1) + f(b^*)(\hat{p}-b^*) &= 0 \\ (1-F(a^*))(1-\hat{p}) - f(a^*)(a^*-\hat{p}) &= 0 \end{aligned}$$

If $\hat{p} \in (0,1)$ and $f(x) > 0$ for all $x \in (0,1)$, the existence and uniqueness of optimal $b^* < \hat{p} < a^*$ are guaranteed. When F and f are well-behaved smooth functions (as is the case for our experiments where F is a normal distribution), the optimal values can be solved quickly by numerical root-finding techniques.

4.3 Optimal log-utility policy

Following the procedure outlined in Section 3, we computed the optimal value and policy functions for several different parameterizations of wealths and beliefs for both Kelly and risk-neutral market makers.

We begin by considering the case where the market maker's private belief aligns with the beliefs of the traders. Figure 4 shows the optimal bid and ask prices over the series of traders when the market maker has wealth $(100, 100)$ (thickest line), $(50, 50)$ (medium line), and $(25, 25)$ (thinnest line). That is, the plot shows $\pi(t, (w, w))$ for $t \in \{1, \ldots, 50\}$ and $w \in (25, 50, 100)$.

Here, prices throughout the interaction are very close to the myopic optimization for the last trader, and the prices are very similar for all of the sampled wealths. In this scenario, the prices are also essentially equivalent to the optimal policy of a risk-neutral market maker.

In contrast, Figure 5 shows the optimal policies when the market maker's belief is $p = 0.6$ (shown by the cross-hatched line). This value is two standard deviations higher than the mean of the traders' beliefs. The policies are calculated at the same wealths as in Figure 4, that is, the policy of a market maker with wealth of 25, 50, and 100 in both states at every time step.

Unlike in the previous figure, the optimal policies change over time and are wealth-dependent. In this scenario, the optimal risk-neutral policy is a bid of 0.52 and an ask of 0.62. Because with large wealth a logarithmic utility market maker making small bets can be approximated well by linear utility, we know that as wealth increases, the market maker's optimal policy throughout the trading period will converge to be the optimal linear utility policy. However, at the smaller levels of wealth in our experiments, for all except the last few traders, the asking price for a unit contract is below the market maker's belief that the event will occur. Thus, for much of the trading period, from a myopic perspective the Kelly criterion market maker offers irrational bets.[1]

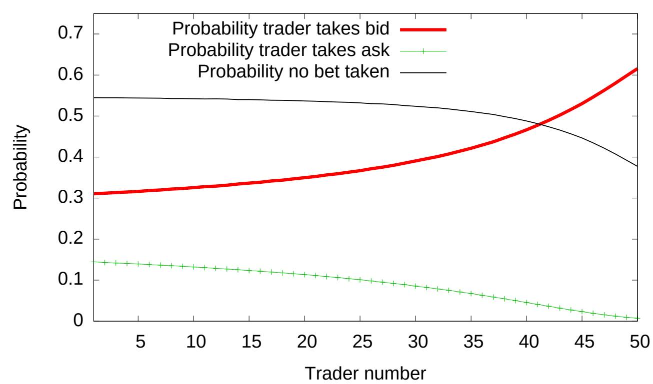

Figure 6: The probability of a trader taking each offered bet from a market maker with 25 wealth in each state over the entire trading period.

On the surface, this result seems confusing and even paradoxical. To see why it is the optimal policy, consider Figure 6, which displays the probability of each trader taking the bets offered by a market maker with a (constant) wealth of 25 in both states. It shows how the probability of a trader selling at the bid price rises over time, while the probability of a trader buying at the ask price falls. The first trader is about twice as likely to sell at the bid price than to buy at the ask price, while the last trader is about *87 times* more likely to sell than to buy.

For early traders, the market maker's bid and ask prices are roughly centered around 0.5, just like the distribution of agent beliefs. Con-

[1] One might think that this phenomenon could be explained by the market makers accumulating wealth from spread profits, and therefore becoming absolutely less risk-averse over time. However, even market makers with considerably larger endowments than could possibly be made through spread profits still display the same qualitative behavior.

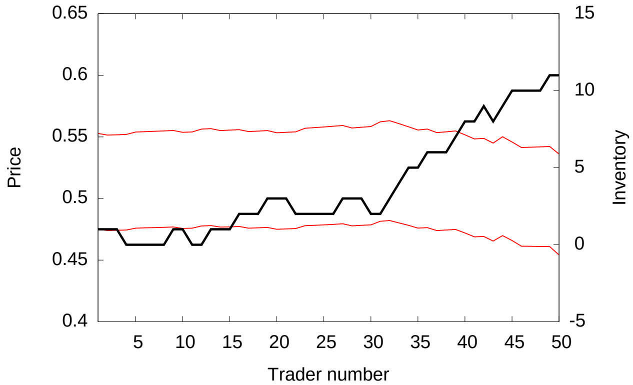

Figure 7: Simulating the prices (left axis; thin lines) and net inventory (right axis; thick black line) that result from the interaction of an optimal log**-utility market maker starting with 25 wealth in both states. In this figure, both the inventory and prices change over time.**

sequently, the market maker has a reasonable chance of matching traders' bids and asks and thus profiting off the bid/ask spread. A market maker that successfully matches the bids and asks of traders books a profit regardless of the personal beliefs of the market maker, even if those beliefs are, as in this scenario, very different from the prices in question. Of course, as fewer traders remain the setting more and more resembles a myopic optimization where, with equivalent wealths in both states, the market maker will employ a myopically rational strategy. This sophisticated policy emerges solely from the introduction of risk-aversion to the termination state, because a risk-neutral market maker in an identical setting displays none of this behavior.

Once the optimal policy is computed, we can simulate the behavior of the market maker against the pool of traders. Figure 7 shows the simulated prices of wealth 25 market maker in a sample interaction for the case where the the market maker has belief 0.6 and agents have belief mean of 0.5 (i.e., the setting for Figures 5 and 6). The thin lines, with values marked by the left axis, show the ask (upper line) and bid (lower line) prices of the market maker. The thick black line, with values market by the right axis, shows the market maker's net inventory, i.e., the market maker's payoff if the event occurs. The values on Figure 7 show the prices faced by trader i but the inventory *after* the participation of the trader. (The inventory line starts at 1 in this case because the first trader took the market maker's bid.) In this simulation, the market maker's expected utility from their wealth vector increases from 3.22 before any traders participate to 3.27 after all 50 traders participate.

Recall that in this setting the market maker has a significantly higher belief that the event will occur than does the pool of agents, so it is natural for the market maker to accumulate inventory. As the market maker accumulates inventory, its prices fall. This is because a risk-averse market maker prefers to take a small sure profit (the bid/ask spread from matching orders) over a somewhat larger speculative gain (from holding inventory). Consequently, the prices from the simulation are very different than the prices in Figure 5, because the prices in Figure 5 captured the prices of a market maker with constant wealth in both states over time. If the market maker were not taking on inventory, its prices would rise, as in Figure 5, but because the market maker in our simulation takes on inventory, that price rise is effectively dampened. Observe that in this simulation, because the price rise is dampened, the market maker's asking price is always less than 0.6 and therefore is myopically irrational for the entire trading period!

5. CONCLUSIONS

We initiated the study of rational market making where the market maker has knowledge about the probabilities of future events transpiring and of traders accepting bets that the market maker could offer them. We gave a general description of the optimization problem faced by a monopolistic market maker trying to maximize their expected utility. We investigated two cases in detail: a risk-neutral market maker (linear utility), that yields the highest expected wealth, and a Kelly criterion market maker (logarithmic utility) that yields the highest expected median and mode of wealth.

We showed that for a two-event setting, computing the optimal policy of a risk-neutral market maker is trivial, but computing the optimal policy of a log-utility market maker is not straightforward. Because there is no closed-form expression for the value function, we approximated it using Constantini shape-preserving interpolation. This interpolation technique preserves the concavity, monotonicity, and partial derivatives of the original function. Because it retains the shape of the approximated function, it is much more accurate than a simple grid interpolation. Since it is more accurate, to preserve the same level of accuracy we were able to solve the problem using a grid that is orders of magnitude coarser. Consequently, the time spent calculating the value function at each iteration of the dynamic program is orders of magnitude faster using the more-sophisticated shape-preserving interpolation technique than with a simple linear interpolation. Because our problem is defined over the whole positive orthant while the shape-preserving technique we used works only over a finite grid, we had to develop an extension of the technique at the lower and upper boundaries of the grid.

We showed that the optimal policy for a risk-neutral market maker is always myopically rational. In contrast, our experiments showed that the optimal policy for a Kelly criterion market maker is often myopically irrational, and that a Kelly criterion market maker could have a myopically irrational policy for the entire trading period. We showed evidence that a log-utility market maker would begin the trading period pricing in order to capture the bid/ask spread from agents. Recall that the traders' response to the market maker's prices is a random variable, so that whether or not a market maker acquires inventory is stochastic, not deterministic. If that pricing resulted in the market maker not taking on inventory, our results showed the market maker would gradually transition from pricing to capture the bid/ask spread to myopically pricing based on their beliefs. However, if the policy in early periods resulted in the market maker taking on inventory, we showed that the optimal pricing throughout the interaction need not deviate much from initial prices. Therefore, depending on how traders react to the market maker's policy, a Kelly criterion market maker could follow a myopically irrational policy for the entire trading period.

In sum, we can distill our results into three qualitative suggestions for monopolistic Kelly criterion market makers with good prior information facing a stream of anonymous traders, as in Internet sports betting: (1) change odds as bets are received and wealth changes, (2) begin pricing with the goal of matching orders and procuring a bid/ask spread, and (3) gradually transition into pricing to accumulate the market maker's desired inventory position.

There are several extensions to consider to our framework. Our model assumed the market maker was monopolistic, so that it could maximize profit without fear of competition. One extension could be to examine a competitive setting between several risk-averse market makers.

Another extension would be to incorporate informed traders into the pool of trading agents. These agents could have correct knowledge about the future, but, more importantly, the market maker could know of and react to their existence. The presence of informed traders that influence the market maker in this way would make our setting much closer to the Bayesian market maker setting explored in Das and Magdon-Ismail [2009], but would be even more complex because of the risk aversion of the market maker.

While our framework applies to any number of events and bets, our computational experiments focused on the binary case. An extension would be to develop algorithms to solve for optimal policy with multiple events. The Constantini shape-preserving technique used in this paper could presumably be applied to more than two events, although it will suffer from the curse of dimensionality. Perhaps an alternate approach to approximating the value function could be used in this case, such as a spline of radial basis functions, although this would be unlikely to preserve the monotonicity and concavity of the value function and so could lead to poor or unrealistic policies.

Acknowledgements

This work was supported by NSF grants CCF-1101668, IIS-0964579, and IIS-0905390, and by the Google PhD Fellowship in Market Algorithms.

References

S. Agrawal, E. Delage, M. Peters, Z. Wang, and Y. Ye. A unified framework for dynamic pari-mutuel information market design. In *ACM Conference on Electronic Commerce (EC)*, pages 255–264, 2009.

A. Ben-Tal and M. Teboulle. An old-new concept of convex risk measures: The optimized certainty equivalent. *Mathematical Finance*, 17(3):449–476, 2007.

M. Chakraborty, S. Das, A. Lavoie, M. Magdon-Ismail, and Y. Naamad. Instructor rating markets. In *Conference on Auctions, Market Mechanisms and Their Applications (AMMA)*, 2011.

Y. Chen and D. M. Pennock. A utility framework for bounded-loss market makers. In *Proceedings of the 23rd Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 49–56, 2007.

Y. Chen and D. M. Pennock. Designing markets for prediction. *AI Magazine*, 31(4), 2010.

Y. Chen, D. Reeves, D. Pennock, R. Hanson, L. Fortnow, and R. Gonen. Bluffing and Strategic Reticence in Prediction Markets. In *WINE*, 2007.

Y. Chen, S. Dimitrov, R. Sami, D. M. Reeves, D. M. Pennock, R. D. Hanson, L. Fortnow, and R. Gonen. Gaming prediction markets: Equilibrium strategies with a market maker. *Algorithmica*, 58: 930–969, December 2010.

P. Constantini and F. Fontanella. Shape-preserving bivariate interpolation. *SIAM J. Numer. Anal.*, 27:488–506, April 1990. ISSN 0036-1429.

S. Das. The effects of market-making on price dynamics. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 887–894, 2008.

S. Das and M. Magdon-Ismail. Adapting to a market shock: Optimal sequential market-making. In *Advances in Neural Information Processing Systems 21 (NIPS)*, pages 361–368, 2009.

D. De Farias and B. Van Roy. The linear programming approach to approximate dynamic programming. *Operations Research*, pages 850–865, 2003.

S. Dimitrov and R. Sami. Non-myopic strategies in prediction markets. In *EC '08: Proceedings of the 9th ACM conference on Electronic commerce*, pages 200–209, 2008.

J. Feigenbaum, L. Fortnow, D. Pennock, and R. Sami. Computation in a distributed information market. In *ACM Conference on Electronic Commerce (EC)*, pages 156–165, 2003.

L. R. Glosten and P. R. Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of financial economics*, 14(1):71–100, 1985.

D. Gode and S. Sunder. Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy*, pages 119–137, 1993.

G. J. Gordon. Stable function approximation in dynamic programming. In *International Conference on Machine Learning (ICML)*, pages 261–268, 1995.

C. Guestrin, D. Koller, and R. Parr. Max-norm projections for factored MDPs. In *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI)*, pages 673–682, 2001.

K. Judd. *Numerical methods in economics*. The MIT Press, 1998.

J. Kelly Jr. A new interpretation of information rate. *IRE Transactions on Information Theory*, 2(3):185–189, 1956.

A. S. Kyle. Continuous Auctions and Insider Trading. *Econometrica*, 53(6):1315–1335, 1985.

A. Mas-Colell, M. Whinston, and J. R. Green. *Microeconomic Theory*. Oxford University Press, New York, NY, 1995.

A. Othman. Zero-intelligence agents in prediction markets. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 879–886, 2008.

A. Othman and T. Sandholm. When Do Markets with Simple Agents Fail? In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 865–872, Toronto, Canada, 2010.

W. Poundstone. *Fortune's Formula: The Untold Story of the Scientific Betting System That Beat the Casinos and Wall Street*. Hill and Wang, 2006.

K. Rock. The specialist's order book and price anomalies. *Review of Financial Studies*, 1996.

J. Stachurski. Continuous state dynamic programming via nonexpansive approximation. *Computational Economics*, 31(2):141–160, 2008.

N. Stokey, R. Lucas, and E. Prescott. *Recursive methods in economic dynamics*. Harvard University Press, 1989.

S.-P. Wang and K. L. Judd. Solving a savings allocation problem by numerical dynamic programming with shape-preserving interpolation. *Comput. Oper. Res.*, 27:399–408, April 2000. ISSN 0305-0548.