INTERTEMPORAL PRICE DISCOVERY BY MARKET MAKERS: ACTIVE VERSUS PASSIVE LEARNING

by

J. Chris Leach
Ananth N. Madhavan

15-90

RODNEY L. WHITE CENTER FOR FINANCIAL RESEARCH
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6367

Intertemporal Price Discovery By Market Makers: Active versus Passive Learning

J. Chris Leach
and
Ananth N. Madhavan [1]

The Wharton School
University of Pennsylvania
Philadelphia, PA 19104

Original Version, November 1988,
This Version, February 1990

[1]We would like to thank to Michael Brennan, David Easley, Richard Kihlstrom, Paul Laux, and an anonymous referee for their helpful comments. Any errors are entirely our own. Financial assistance from the Geewax-Terker Research Fund is gratefully acknowledged.

Abstract

This paper examines the role of the market maker in intertemporal price formation in securities markets. We argue that the market maker, in performing the critical function of price discovery, may set prices in a dynamic context that would be suboptimal in a single period context in order to learn more from the resulting order flow. Such actions constitute an investment in the production of information. Necessary and sufficient conditions for the existence of such price strategies are developed and several examples are presented.

Contents

1 Introduction

Most securities markets include professional intermediaries known as market makers or dealers whose central role in price formation and trading has been the subject of numerous studies. In a seminal paper, Demsetz (1968) modeled the market maker as a supplier of liquidity, permitting continuous trading. Others have viewed the market maker's primary function as that of providing price stability, accommodating transitory order imbalances by 'leaning against the wind.' Intertemporal price smoothing may reflect the market maker's inventory control policy which may be designed to minimize carrying costs as in Ho and Stoll (1983) or to maximize expected utility, where the utility function exhibits risk aversion, as in O'Hara and Oldfield (1987).

Models of price formation by Glosten and Milgrom (1985), Kyle (1985), Easley and O'Hara (1987), and Admati and Pfleiderer (1989), among others, have focused on the impact of asymmetric information. In these models, a portion of order flow originates from traders with private information, and rational market makers set prices on the basis of the information contained in the order flow and revise their beliefs using Bayes' rule. Since market makers in these models do not pursue dynamic strategies, learning takes on a passive character; competition dictates the prices in each period.

More recently, attention has shifted to *price discovery*, i.e., the process of finding market clearing prices.[1] In this paper, we show that market makers have the incentive and ability to perform the function of dynamic price discovery. This role for the market maker emerges as a by-product of intertemporal profit maximization. The market maker invests in information production through a process of price experimentation to induce statistically revealing order flow. The investment cost is recovered in the future since better information enables the

[1] Schreiber and Schwartz (1986), who argue that price discovery is the critical function of a market mechanism. Leach and Madhavan (1989) model the process of price discovery across different security market structures. See also Bronfman and Schwartz (1990) and Handa and Schwartz (1989) for alternative approaches.

market maker to increase his intertemporal expected utility, provided he has a monopolistic position in trading.[2]

Apart from Leach and Madhavan (1989), the issue of price experimentation by market makers has not been addressed in the microstructure literature. However, related work exists in consumption theory and models of monopolistic pricing under uncertainty. The optimality of price experimentation by a market maker is closely related to the consumption experimentation models of Grossman, Kihlstrom and Mirman (1977), (hereafter GKM), Kihlstrom, Mirman and Postelwaite (1984) and Fusselman and Mirman (1987). In these models, different levels of consumption of a good providing unknown marginal utility represent different experiments. The choice of levels of consumption higher than the myopically optimal level is deemed experimental consumption and is interpreted as the consumer's attempt to improve the information available for making future decisions, thereby increasing total utility from the stream of consumption choices over time.

The model of GKM is, perhaps, the most direct ancestor of our model: both models show the optimality of experimentation by an agent whose decisions have an unknown effect on his reward. Models of a monopolist facing an unknown demand curve like those of McLennan (1984) and Easley and Kiefer (1988a, 1988b) are also similar to our model. In these models, however, the focus is on whether the agent ever finds it optimal to learn the true state of the world.[3] The focus in this paper is on demonstrating the optimality, and not the effectiveness, of experimental pricing. In the market maker's world, the recurrence of information shocks may obviate the need for long run convergence.[4]

The contribution of this paper is to demonstrate that experimental behavior is a rational

[2]The argument is similar to Schumpeter's (1942) analysis of industrial organization in which he provides a rationale for the possible superiority of monopoly.

[3]A different strand of the literature examines whether agents in an economy can converge to rational expectations using least-squares criteria. See, e.g., Jordan (1982) and Marcet and Sargent (1989). See also Kiefer and Nyarko (1989) for a discussion of optimal control in the linear case.

[4]The analysis is also related to bandit problems (see Rothschild (1974) and Berry and Fristedt (1985)). In bandit problems actions are discrete; here, the actions (prices) are selected from a continuum.

component of price formation in the securities market. While the problem we analyze is similar to the experimental consumption literature discussed above, we cannot apply these results directly to a security market because of a crucial difference. The experimental consumption literature typically assumes that rewards (the discounted sum of which is to be maximized) are observed in each period after the disturbance term is realized. [5] In contrast, the rewards to a market maker are not a sequence of observable cash flows but are changes in expected terminal wealth due to current actions while the consumer (product monopolist) knows what he has consumed (made in profits). The market maker receives a reward that is still random (since it depends on a future price) after the current period's disturbance is realized. We also state our results for more general functional forms than is common in the literature in which, generally, linearity has been invoked.

The remainder of the paper proceeds as follows. Section 2 lays out a simple model in which an agent, the market maker, sets the price of a security in a dynamic context without knowing the parameters of the demand function for the security. We provide necessary and sufficient conditions for the existence of opportunities for active experimentation. Section 3 provides some examples to illustrate the results. The examples show that experimentation can have important effects on asset prices that are quite different from the predictions made by single-period models. Finally, Section 4 concludes the paper.

2 The Model

We consider the intertemporal pricing decision following an information event which induces a shift in the demand or supply of a risky asset. There is a single market maker who does not know the new parameters of the (excess) demand curve for the risky asset. We assume that it is common knowledge that an event has occurred and the functional form, but not the

[5] Kiefer and Nyarko (1988) have a periodic reward function that involves expectation over the parameter in a manner analogous to our problem. However, they consider only linear processes, a smaller class than we consider.

parameters, of the excess demand curve are known to the market maker. Trading occurs over a finite number of trading rounds, indexed by $t = 1, \ldots, T$. Implicit in this formulation is a known future event or liquidation date after the final trading round. The trading protocol is as follows: In trading round t the market maker quotes a single price, denoted p_t, and then observes the excess demand at that price, denoted $z(p_t)$. After observing an excess demand realization, he updates his beliefs about the form of excess demand and sets a price for the next round.[6] What distinguishes this mechanism from a Walrasian auction is that trading occurs at each round. We begin by considering a very general specification for demand.[7] Excess demand in round t takes the following form:

$$z_t = f(p_t, \beta) + \epsilon_t \tag{1}$$

where $f(p_t, \beta)$ is the price-sensitive portion of excess demand which is assumed to be differentiable and strictly decreasing in p_t. The function f also depends upon β, the unknown parameter, which is drawn randomly from $B = \{\beta^1, \ldots, \beta^n\} \subset \Re^n$.[8] Finally, ϵ_t is the stochastic disturbance term which we assume to be normally distributed with $\mathrm{E}[\epsilon_t] = 0$, $\mathrm{E}[\epsilon_t \epsilon_t'] = 0$ if $t \neq t'$ and σ^2 if $t = t'$. Let $h(.; \mu)$ denote the normal density function with mean μ and variance σ^2. Let $\theta \in \Re_+$ denote the market clearing price, where $\theta \equiv F^{-1}(0)$ and $F(\cdot) \equiv f(\cdot; \beta)$ is assumed to be invertible. The market maker knows this structure, but not the realization of β, and hence the market clearing price, θ. Each period, the market maker receives a noisy information signal about β in the form of the observed excess demand at the price set. The market maker's prior probability that $\beta = \beta^i$ in period t is denoted by μ_t^i, where $0 \leq \mu_t^i \leq 1$, for $t = 0, \ldots, T$. It is convenient to define $\mu_t = \{\mu_t^i\}_{i=1}^n$. An arrow over the variables will be used to indicate the time series vector of the variable from 0 to the subscript.

[6]The assumption of a single price greatly simplifies the exposition. In section 3.3 we provide an example that incorporates bid and ask prices.

[7]In section 3.1 below we provide an example where we explicitly describe the optimization problem that gives rise to such a function.

[8]Here β can be regarded as a vector of coefficients, where one element of the vector is unknown.

For technical reasons, we need to bound the feasible price space. We assume that the exchange requires that $p_t \in [0, \bar{p}]$, where $0 < \bar{p} < \infty$. Define $I \equiv [0, \bar{p}]$, the domain of p_t. We will refer to the set I as the action space. Let $u(p, \beta^i, z)$ be the utility from a trade volume of z at price p when $\beta = \beta^i$. We assume that u is continuous in p.[9] In what follows, we will show that the market maker has the ability and incentive to discover prices by setting prices in early periods to induce statistically informative order flow. This is analogous to experimenting with new wines, different medication, and so on. Such experimentation typically involves a departure from the short-run optimum in order to obtain information to increase the utility from future consumption. We turn now to a formal statement of the problem.

2.1 Myopic Optimization and Passive Learning

The single period expected utility in period t for a price choice of p_t and beliefs μ_t is:

$$U(p_t, \mu_t) = \sum_i \mu_t^i \int_{\Re} u(p_t, \beta^i, f(p_t, \beta^i) + \epsilon_t) h(\epsilon_t; 0) d\epsilon_t \tag{2}$$

which can be written (by change of variable) as:

$$U(p_t, \mu_t) = \sum_i \mu_t^i \int_{\Re} u(p_t, \beta^i, z_t) h(z_t; f(p_t, \beta^i)) dz_t. \tag{3}$$

A myopically optimal price in period t, given priors μ_t, is:

$$p_t^M \in P_t^M \equiv \text{argmax}_{p_t} \{U(p_t, \mu_t)\} \tag{4}$$

We will refer to a $p_t \notin P_t^M$ as *active control* of the learning process or *price experimentation.* For a rational market maker, a departure from P_t^M can only be explained by investment in the production of information.

[9]The utility function is quite general. For example, the market maker could maximize ex post profits or the utility of trading income. The reward function may also implicitly incorporate a brokerage fee which the market maker charges for the net order imbalance he accommodates. Without such compensation, the market maker would suffer expected losses, because traders would on average sell when prices were set too high and buy when prices were too low.

2.2 Intertemporal Optimization

To understand the intertemporal decision process in the dynamic learning environment more clearly, we need to specify the informational link between trading rounds. We adopt the notation and approach similar to that of GKM, and let $w_t = (z_t, p_t)$ for $t = 0, \ldots, T$, and $\vec{w}_t = (\vec{z}_t, \vec{p}_t) = ((z_0, .., z_t), (p_0, .., p_t)) = (w_0, .., w_t)$. Here, $\vec{w}_t$ represents the history of trading through period t. Let W_t designate the set of all possible $\vec{w}_t$, i.e., the set of all possible histories. We can write Bayes' rule for this problem in the form $\mu_t = g_t(\mu_0, \vec{w}_{t-1})$. Recall from our earlier definitions that μ_t is a $n \times 1$ vector. From the theory of sequential Bayesian updating, we can write $\mu_t = g_1(\mu_{t-1}, w_{t-1})$.

Let $G_t \equiv \{\mu_t : \mu_t = g_t(\mu_0, \vec{w}_{t-1}), \vec{w}_t \in W_t\}$, i.e., the set of all possible posteriors that could occur at time t. Finally, let $\delta \leq 1$ be the market maker's discount factor. Now define:

$$V_T(\mu_0) \equiv \max_{p_0, \{\vec{p}_t\}_{t=1}^{T}} \left\{ E\left[U(p_0, \mu_0) + \sum_{t=1}^{T} \delta^t U(p_t, g_t(\mu_0, \vec{w}_{t-1})) \right] \right\} \tag{5}$$

where $p_0 \in I$, and $p_t : G_t \rightarrow I$. In (6), $V_T(\mu_0)$ denotes the discounted sum of expected utility with T periods remaining starting from a prior of μ_0 when Bayes' rule is used to update prior beliefs. By Bellman's optimality principle for $p_t = p_t(g_t(\mu_0, w_{t-1}))$ we have:

$$V_T(\mu_0) = \max_{p_0} \{U(p_0, \mu_0) + \delta E[V_{T-1}(g_1(\mu_0, w_0))]\} \tag{6}$$

where

$$E\{V_{T-1}[g_1(\mu_0, w_0)]\} = \int_{-\infty}^{+\infty} V_{T-1}[g_1(m_0, z_0, p_0)] \sum_j \mu_0^j h(z_0; f(\beta^j, p_0)] dz_0 \tag{7}$$

An optimal price in period 0, given a prior of μ_0 is:[10]

$$p_0^* \in P_0^* \equiv \text{argmax}_{p_0} V_T(\mu_0). \tag{8}$$

[10]Technically, to ensure V_T is well-defined and P^* is non-empty, we need to show that the value function is a continuous function of p, for $p \in I$. This involves showing that $u(\cdot)$ and $f(\cdot, \beta)$ are continuous in p, and $V(\cdot)$ is convex and therefore continuous in μ (on open convex sets). For a prototype of the proof technique, see e.g., GKM or Prescott(1972).

Equations (6) and (7) are a formal representation of the decision problem facing the market maker. Our model differs from other models of intertemporal price formation because the market maker attempts to change the statistical properties of order flow (i.e., the signal-to-noise ratio) to induce the revelation of information. Before considering such policies, we provide a result which characterizes situations in which departures from the single-period optimal actions are never optimal in a dynamic sense, i.e., $P_t^M = P_t^*$.

2.2.1 Intertemporally Optimal Myopic Pricing

Suppose that the realization of β is β^i and $f(p_t, \beta^i)$ is such that the path of posterior beliefs regarding $\tilde{\beta}$ is invariant to the price choice. In this case, the market maker has no ability to control the learning process, i.e., to affect the informativeness of the signals conveyed by order flow. For our problem with additive disturbances, invariance of the posterior path in current price choice is equivalent to invariance of next period's posterior beliefs in the current price choice. Formally,

Lemma 1 *If $p_t = p(\mu_t, t)$ for all t, then $\vec{\mu}_s(\vec{z}_{s-1}(\vec{p}_{s-1}), \vec{p}_{s-1}, \mu_0)$ is independent of p_1 for all $s > 0$ if and only if $\mu_1(z_0(p_0), p_0, \mu_0)$ is independent of p_0.*

Proof: See Appendix

A market maker who is uncertain about the true β will know that attempting to control the learning process is futile if and only if f induces invariance of the next period's posterior in the current price choice for all $\beta \in B$. When this invariance holds it is not necessary to look more than one period into the future, as long as actions depend only upon current beliefs and the current time period, the 'state variables' for this problem.

Theorem 1 *The optimal dynamic actions coincide with the myopic actions (i.e., $P_t^* = P_t^M$) if for all $p \in I$, $f(p, \beta)$ satisfies:*

$$\frac{\partial f(p, \beta)}{\partial p} = \psi(p)$$

where $\psi(p)$ is a real-valued function that is independent of β.[11]

Proof: See Appendix.

The theorem implies that if there is no interaction between price and the parameter β, the market maker knows that his actions are not going to alter the statistical information received in the next period. Accordingly, if f has this property, then the price choice in any period is a single period or myopic utility maximizing choice given current information. Once we have described the market maker's objectives, myopic single period optimization will be optimal in the dynamic problem. Costs incurred in departing from myopic optimization are not investments in information refinement, since the latter occurs exogenously. Such costs would not be incurred by a optimizing market maker. For functions which do not satisfy this property, controlled learning opportunities arise at some value of the unknown parameter β. Since all β's carry positive weight, the market maker will believe that there is a chance that his prices affect future beliefs. However, this does not imply that the market maker will depart from myopic pricing. To address the optimality and form of price experimentation is our model, we restate the problem to make use of some well known results of statistical decision theory.

2.2.2 Active Control of the Learning Process

Lemma 2 *There exist strategies, $\xi = (p_0, \gamma)$, where $p_0 \in I$, and $\gamma_t : G_{t-1} \rightarrow I$ and functions $\eta_T(\xi, \beta^i)$ such that $V_T(\mu_0)$ can be written in the following form:*

$$V_T(\mu_0) = \sup_{\xi} \sum_i \eta_T(\xi, \beta^i)\mu_0^i. \tag{9}$$

Proof: See Appendix.

Recall that $h(z; f(p, \beta))$ represents a normal density function of z with mean $f(p, \beta)$ and variance σ^2, which we will write as $h(z; p, \beta)$. As the action variable p varies, we obtain

[11]The proof of the result is easily generalized to apply to functions which are not differentiable.

a family of distributions, $H = \{h(.;p,.), p \in I\}$, termed a family of experiments. Distinct elements of H correspond to different experiments. Our definition of a more informative experiment is the sufficiency condition of Blackwell (1953), which is also discussed by DeGroot (1970).

Definition 1 *An experiment $h(z;p,\beta) \in H$ is more informative than (or is sufficient for) an experiment $h'(z';p',\beta) \in H$ if there exists a real-valued function $\nu(z';z) \geq 0$ such that for all $z' \in \Re$ and all $\beta \in B$.*

$$h(z';p',\beta) = \int \nu(z';z)h(z;p,\beta)dz \tag{10}$$

where $\int \nu(z';z)dz = 1$.

The basic idea of sufficiency is that the results of the experiment h' can be replicated (in a probabilistic sense) given the results of the experiment h, so that h is more informative than h'. We will sometimes refer to the experiment p' for the experiment $h'(z';p',\beta)$ when there is no potential for confusion. We now state an important result for analyzing the possible optimality of controlled learning, Blackwell's Theorem on the comparison of experiments.

Lemma 3 *(Blackwell) Consider functions of the form:*

$$V_*(\mu) = \sup_{\xi \in \nabla} \sum_i \eta_T(\xi, \beta^i)\mu^i \tag{11}$$

where the set ∇ represents the set of possible decisions, an element of which is designated ξ, and $\eta : \nabla \times B \to \Re$. Then, if the experiment with p is more informative than the experiment with p' then the expected value of a function in this V_* class is higher under p than under p':

$$E[V_*(g_1(\mu,p,z))] \geq E[V_*(g_1(\mu,p',z'))] \tag{12}$$

where:

$$E[V_*(g_1(\mu,p,z))] \equiv \int_{\Re} V_*(g_1(\mu,p,z)) \sum_i h(z;\beta^i,p)\mu^i dz \tag{13}$$

and $E[V_*(g_1(\mu,p',z'))]$ is similarly defined.

Proof: See DeGroot (1970), page 436. By Lemma 2, we know that $V_{T-1}(\mu_1)$ is in the class of $V_*(\mu)$ functions where:

$$E[V_{T-1}(g_1(\mu_0, p, z))] = \int_{\Re} V_{T-1}(g_1(\mu_0, p, z)) \sum_i h(z; \beta^i, p)\mu_0^i dz \tag{14}$$

Using Lemma 3, we obtain:

$$E[V_{T-1}(g_1(\mu_0, p, z))] \geq E[V_{T-1}(g_1(\mu_0, p', z'))] \tag{15}$$

given that an experiment with p is more informative that an experiment with p'. Therefore, more informative experiments give rise to higher future expected profits. This is true for all functional forms for excess demand. Note that this result is independent of the discount factor, δ.

Theorem 1 showed that pricing strategies for functions where p and β are additively separable do not permit controlled learning. For functions that do not have this property, opportunities for controlled learning through price experimentation arise. When will such active learning strategies be optimal and what predictions can we make about price behavior under such strategies? The answer to this question will typically depend on the functional form of f, which in turn depends upon the behavior of traders in the market. Having eliminated additively separable functions as candidates for active control, we now consider functions that are multiplicatively separable in p and β.

Let $f(p, \beta) = \psi_0 + \psi_1(\beta)\psi_2(p)$ with the weak condition that $\psi_2 : I \rightarrow \Re_+$ is continuous and takes on at least two values.[12] Then, excess demand is:

$$z_t = \psi_0 + \psi_1(\beta)\psi_2(p) + \epsilon_t. \tag{16}$$

[12]The restriction of the range to $\Re_+$ is without additional loss of generality over the non-zero real line since the sign of ψ_1 is unrestricted. Zero is excluded because our proofs have division by ψ_2. The nondegeneracy of ψ_2 excludes the functional forms of excess demand characterized in the proof of Theorem 1 since

$$f(p, \beta^i) - f(p, \beta^j) = \psi_2(p)\left[\psi_1(\beta^i) - \psi_1(\beta^j)\right]$$

is equidistant in β if and only if $\psi_2(p)$ is a constant, i.e., excess demand is insensitive to price.

Theorem 2 *For excess demand functions of the form (16), the experiment conducted using price p is more informative than the experiment conducted using price p′ if and only if* $\psi_2(p) \geq \psi_2(p')$.

Proof: See Appendix.

The intuition for the result is easily couched in terms of signal-to-noise ratios. Consider the experiment with p. Suppose ψ_0, $\psi_1(\beta)$ and $\psi_2(p) \neq 0$ are known constants at the time of the experiment, observing z is equivalent to observing $r = \frac{z-\psi_0}{\psi_2(p)} = \psi_1(\beta) + \frac{\epsilon}{\psi_2(p)}$, which is distributed normally with unknown mean $\psi_1(\beta)$ and variance $\frac{1}{\psi_2(p)}^2\sigma^2$, which is decreasing in $\psi_2(p)$. So the higher the $\psi_2(p)$, the more informative the experiment.

Theorem 3 *Define* $\psi_2^* = \min\{\psi_2(p) : p \in P_0^*\}$ *and* $\psi_2^M = \max\left\{\psi_2(p) : p \in P_0^M\right\}$. *Then* $\psi_2^* \geq \psi_2^M$.

So, optimal prices induce higher values of ψ_2 than myopic prices.

Proof: See Appendix.

2.2.3 The Effect of Active Learning

Active control of the learning process involves price experimentation at cost to the market maker. In this section, we examine how price experimentation affects asset prices and volume. The following result shows that investment in the production of information implies above normal trading activity (relative to the single-period choice) for the market maker.

Theorem 4 *If z satisfies equation (16), then in the initial trading round:*

(i) The variability of trading volume is strictly higher under p^* *than under* p^M. *The expected magnitude of price sensitive trading is larger under* p^* *than under* p^M.

(ii) If the market maker is risk-neutral, his expected profits are lower under p^* *than under single-period choice* p^M.

Proof: See Appendix.

Theorem 4 has some important implications for observed returns and volume around information events. Suppose that an econometrician uses a model of the type $p_t^M(g(\mu_0, \vec{w}_{t-1})$ to predict prices in round t, where $g(\cdot, \cdot)$ is the Bayes operator and $\vec{w}_{t-1}$ is the history of trading. Such a model will work very well if information is homogeneous and the optimal and myopic policies coincide. Following an information event, however, the market maker's search for a new settling price will induce a higher level of price sensitive trade and volatility, accompanied by 'abnormal' price deviations relative to the predictive model. An econometric study that attempts to infer positive or negative stock price reactions using a predictive model based on historical data will also capture the effects of price discovery. In addition, if the market maker is risk neutral, the gross returns of traders in the early rounds will be biased. The existence of such 'anomalies' is not evidence that the market is inefficient in any sense.

3 Applications of the Model

In this section, we develop a series of examples to illustrate the theorems of the previous sections. Our first example shows that active learning can occur even if order flow contains no information about the liquidation value of the security. Price discovery takes the form of learning the market settling price, which depends on investors' time preferences. In the second example, the market maker trades with privately informed agents, but the demand function is such that the market maker has no opportunity to actively control the learning process. We then extend our results to the case of a two-sided market with an endogenously determined bid-ask spread. Finally, we provide a numerical example to illustrate that active control strategies can strictly dominate myopic strategies. The example suggests that price discovery can have a non-trivial effect on asset returns.

3.1 Example 1: Learning Unrelated to Liquidation Value

Consider a multi-period model where $N \times T$ investors trade claims to a risky security. Calendar time is divided into two 'days.' On the first day, day 0, there are T trading rounds. Following trading, the security pays a liquidating dividend at day 1 in the future. The dividend is distributed normally with mean v_0 and variance σ_v^2. We will assume that on each of the T pre-announcement trading rounds, N traders arrive at the market through some exogenous arrival process. We assume investors trade only once. An investor $i = 1, \ldots, N$ who trades at round $t = 1, \ldots, T$, has an endowment of cash, $W_{it,0}$, and risky securities, x_{it}. The endowment vector $\omega_{it} = (W_{it,0}, x_{it})$ is private information. Across all investors, x_{it} is drawn from a normal distribution with mean zero and constant variance. Let $C_{i,1}$ represent the cash flow to investor i in the final period, i.e., day 1. We assume that a trader's subjective future value of income, denoted by C_i, is given by:

$$\tilde{C}_i = \rho_i C_{i,0} + \tilde{C}_{i,1} \tag{17}$$

Investors maximize the expected utility of $\tilde{C}_i$, where for simplicity we assume a mean-variance utility function.

$$u_i(\tilde{C}_i) = E[\tilde{C}_i] - \left(\frac{r}{2}\right) Var[\tilde{C}_i] \tag{18}$$

where $r > 0$ is the coefficient of absolute risk aversion.[13] The parameter ρ_i is investor i's rate of time preference, where $\rho_i \geq 1$. Future income is a random variable, since it depends on the risky asset's total payoff. For investor i who trades in round t, let $q_{it} \in \Re$ denote the trade size, with the convention that q_{it} is positive for purchases and negative for sales. The investor's budget constraint implies that we have:

$$C_{i,0} = W_{it,0} - p_t q_{it} \tag{19}$$

$$\tilde{C}_{i,1} = \tilde{v}(q_{it} + x_{it}). \tag{20}$$

[13]Unlike the usual constant absolute risk aversion utility function, this is a two-parameter utility function.

Then, solving the utility maximization problem, we obtain:

$$q_{it} = \frac{E_i[\tilde{v}] - \rho_i p_t}{r\sigma_v^2} - x_{it}. \tag{21}$$

In what follows, we will assume that $\rho_i = \rho \in \{\rho_1, \ldots, \rho_n\}$, where $\rho > 0$. Then, in round t, we can write:

$$z_t = \sum_{i=1}^{N} q_{it} = N\left(\frac{v_0 - \rho p_t}{r\sigma_v^2}\right) + \epsilon_t \tag{22}$$

where $\epsilon_t \equiv -\sum x_{it}$ is distributed normally with mean zero and constant variance. We assume that the market maker does not know ρ. Putting $\psi_0 = \frac{Nv_0}{r\sigma_v^2}, \beta = \frac{N\rho}{r\sigma_v^2}, \psi_1(\beta) = -\beta$, and $\psi_2(p_t) = p_t$, this can be written as:

$$z_t = \psi_0 + \psi_1(\beta)\psi_2(p_t) + \epsilon_t \tag{23}$$

where $\tilde{\beta} \in \{\beta^1, \ldots, \beta^n\}$ is a random variable from the market maker's perspective. This model illustrates the difference between market clearing prices and liquidation values. The market clearing price is $\theta = \psi_0/\beta$ while the fundamental value of the asset is $\tilde{v}$. If we assume the market maker maximizes expected profits, myopic utility is given by:

$$U(p_t, \mu_t) = \sum_i \mu_t^j \int_{\Re}\int_{\Re} (p_t - \tilde{v}) z_t h(z_t; f(p_t, \beta^j)) h_v(v)\, dz_t dv \tag{24}$$

where h_v is the density of the unknown future liquidation value. This can be written as:

$$U(p_t, \mu_t) = (p_t - v_0)(\psi_0 - p_t E[\tilde{\beta}]) \tag{25}$$

where v_0 is the expectation of $\tilde{v}$. From the definition of ψ_0 and β, we see that the myopic action is:

$$p_t^M = \frac{v_0}{2}\left(1 + \frac{1}{E[\tilde{\rho}]}\right) \tag{26}$$

which is less than v_0. Contrast this with the expected price, $E[\theta] = E[v_0/\tilde{\rho}]$. Applying the previous theorem directly, we see that active learning leads to 'overpricing' relative to the passive or myopic learning, despite the fact that trading reveals no information about the value of the security, $\tilde{v}$. Why is this? The answer is that the optimal single-period price depends upon how investors view the trade-off between current consumption and future consumption. A market maker who knows the rate of intertemporal price substitution is better able to extract rents from traders. This effect leads to experimentation in the case where ρ is unknown.

3.2 Example 2: No Active Control Under Imperfect Information

Consider a modification of the model above where each investor receives a private information signal $y_i \in Y$, where Y is a finite set. Suppose y_i and $\tilde{v}$ have known non-zero correlation and $\rho_i = 1$. The market maker does not obtain a private information signal and his information is limited to the publicly available prior signal. In this case, the excess demand in round t is:

$$z_t = \sum_{i}^{N} \left(\frac{(E[\tilde{v}|y_i] - p_t)}{r\sigma_v^2} \right) + \epsilon_t \tag{27}$$

where $\epsilon_t = -\sum x_{i,t}$, as before. Uncertainty in this model concerns the future value of the asset, $\tilde{v}$. From (27), we can write f as:

$$f(p;\beta) = \phi\beta - \phi p_t \tag{28}$$

where we define $\phi \equiv \frac{N}{r\sigma_v^2}$ and $\beta \equiv \sum_i E_i[\tilde{v}|y_i]/N$ is the mean of the conditional expectations of the security's value.[14] The function $f(p_t;\beta)$ satisfies:

$$\frac{\partial f}{\partial p_t} = -\phi \tag{29}$$

and applying Theorem 1, no experimentation will occur because the price set does not affect the posterior distribution of $\tilde{v}$. The optimal dynamic action at any round t is the myopic action. If the market maker is risk neutral, $p_t = E_t[\tilde{v}]$, where the expectation is taken on public information at time t.

[14] Since the set Y is finite, β lies in a finite set given by the combinations of all possible conditional expectations.

3.3 Example 3: Learning and the Bid-Ask Spread

The introduction of a two-sided market complicates our model, but is desirable if we wish to discuss the effects of active learning on the bid-ask spread. Suppose the market maker now sets two prices in trading round t, a bid price, denoted b_t and an ask price, a_t. All buy orders are executed at the ask price and all sell orders are executed at the bid price. We assume that the supply of securities to the market maker, q_t^s is given by:

$$q_t^s = q^s(b_t; \beta_s) + \epsilon_t^s \tag{30}$$

where $q^s(\cdot; \beta_s)$ is increasing in b_t and $\beta_s \in \{\beta_s^1, \ldots, \beta_s^n\}$ is a parameter that is unknown to the market maker. The error term ϵ_t^s is assumed to be normally independently and identically distributed for all periods.[15] Similarly, the quantity of securities demanded from the market maker, q_t^d, is assumed to be a decreasing function of the ask price, a_t, and we write:

$$q_t^d = q^d(a_t; \beta_d) + \epsilon_t^d \tag{31}$$

where again, the error term ϵ_t^d is assumed to be distributed normally and independently of ϵ_t^s. The $q^d(.; \beta_d)$ function is assumed to be differentiable and strictly decreasing in a_t. We assume that there exists a unique $\theta > 0$ such that:

$$q^d(\theta; \beta_d) = q^s(\theta; \beta_s) \tag{32}$$

i.e., that at a price θ the price-dependent component of demand equals the price-dependent component of supply. The expected profit function in period t is:

$$\pi(a_t, b_t) = \sum_i \{ \int_{\Re} \int_{\Re} (a_t - \theta) q_t^d + (\theta - b_t) q_t^s h(\epsilon_s) h(\epsilon_d) \, d\epsilon_t^s \, d\epsilon_t^d \} \mu_t^i. \tag{33}$$

[15]Unfortunately, this implies that supply could be negative with positive probability. We assume the variance of ϵ_t^s is sufficiently small relative to supply that the probability of this event is arbitrarily small. Similar comments apply to the demand function.

where $\mu_t^i = \Pr[\theta = \theta_i]$ given information at time t. What is the impact of the possibility of experimentation on the bid-ask spread? The answer depends upon the specific functional forms taken by the demand and supply curves. Suppose that the demand curve is:

$$q^d(a_t; \beta_d) = \psi_0^d + \psi_1^d(\beta_d)\psi_2^d(a_t) \tag{34}$$

and the supply curve is:

$$q^s(a_t; \beta_s) = \psi_0^s + \psi_1^s(\beta_s)\psi_2^s(b_t) \tag{35}$$

where $\psi_2^j(\cdot) > 0$ and $j = \{s, d\}$ with the additional requirement that ψ_2^j is differentiable. Theorem 1 implies that active control opportunities exist if the ψ_2 functions are not single valued. Further, Theorem 3 implies that the experiment b is more informative than b' if:

$$\psi_2^s(b) \geq \psi_2^s(b') \tag{36}$$

Similarly, the experiment a is more informative than a' if:

$$\psi_2^d(a) \geq \psi_2^d(a'). \tag{37}$$

Suppose $\psi_1^s(\cdot), \psi_1^d(\cdot) > 0$. Then to ensure upward sloping supply we require $\psi_2'^s(\cdot) > 0$ and for the demand curve to be downward sloping we require that $\psi_2'^d(\cdot) < 0$. With this framework, the bid-ask spread is narrower under the active learning strategy. To see this, suppose the single period maximizing prices are $a_0^M \geq b_0^M$. Our assumptions concerning the independence of ϵ_t^s and ϵ_t^d imply that we can separate the two sides of the problem when considering learning. For any fixed bid price, say $\bar{b}$, suppose $a_0^* > a_0^M$. Since $\psi_2'^d(\cdot) < 0$, a_0^M is more informative than a_0^*. Therefore, a_0^M is preferred to a_0^*, which contradicts the assumption that a_0^* is the optimal decision. So, $a_0^* \leq a_0^M$.

A similar argument shows that $b_0^* \geq b_0^M$. So, if $a_0^* > a_0^M$ or $b_0^* < b_0^M$, the market maker can improve the current period expected profit by setting the inequality in question to an equality. Further, since this action is more informative, the value function would increase with this action. Therefore, to avoid contradiction, $a_0^* \leq a_0^M$ and $b_0^* \geq b_0^M$. This example shows that experimentation can induce a narrower bid-ask spread than the single-period optimal choice. This suggests that empirical studies that examine the size of the bid-ask spread for evidence of information-based trading around information events may be misleading if market makers adopt active learning strategies. This result is, in part, a direct consequence of a transfer of trading profit from the market maker to early traders in return for more accurate information regarding the fundamental value of the security.

3.4 Example 4: A Numerical Example

In order to illustrate that the inequality in Theorem 3 can be strict, we used numerical methods to compute the optimal and myopic strategies for a linear model with unknown slope coefficient. This example has the added virtue that it shows that price discovery has non-trivial effects on returns around an information event. Specifically, we have $\psi_1(\beta) = -\beta$, $\psi_2(p) = p$, and $\beta \in \{\beta^1, \beta^2\}$. The initial parameter values are as follows: $\mu_0 = 0.5$, $T = 2$, and $\{\beta^1, \beta^2\} = \{20, 30\}$. The quantities are interpreted as round lots of 100 shares. We allowed the intercept of the excess demand function, ψ_0 to vary between 100 and 1,000, in increments of 20. The full information market clearing price in this model is $\theta = -\psi_0/\psi_1$, so that the parameter values implied prices ranging from 3.33 to 50. In order to assess the impact of noise, we also varied the standard deviation of the error term, σ, from 20 to 1,000 in increments of 20.

For computational reasons, we chose the price set P to be discrete in increments of one cent.

The myopic and optimal prices were solved using Gauss.[16] Figure 1 displays the absolute deviation from the myopic price and the ratio of the optimal price to the myopic price. By Theorem 3, the ratio of the two prices is at least 1. The optimal price reached a maximum

[16]Computing the optimal two-period strategy required over 135,000 separate numerical integrations.

of 1.375 percent above the single-period price at $(\psi_0, \sigma) = (200, 40)$ and and the largest absolute deviation was 53 cents at $(\psi_0, \sigma) = (1,000, 180)$ and at $(1000, 200)$. Extending the time horizon should increase the benefit from experimentation (See, e.g., Prescott (1972)), and hence the 'abnormal' return following an information event.

Examining the surface plotted in figure 1 shows that larger intercepts generate greater price deviations from myopia. Intuitively, for a given standard deviation, the spread between the two possible clearing prices is $\psi_0(\frac{1}{\beta^1} - \frac{1}{\beta^2})$, which is linear in ψ_0. So, larger intercepts imply both a higher prior variance and asset value and a greater incentive to experiment. The figure also shows that the relationship between the price deviation is not a monotonic function of noise. If there is very little noise, there is no benefit to experimenting since the myopic action reveals almost as much information as the optimal action. This suggests the amount of experimentation may increase with the level of noise, at least up to a certain point. Beyond this point, the signal-to-noise ratio is dominated by the error term, confounding the inference problem and reducing the value of experimentation.

4 Conclusions

This paper analyzes the role of the market maker in intertemporal pricing under imperfect information. We demonstrate that the market maker has the ability and incentive to perform the critical function of price discovery. We explicitly model the decision problem underlying this function and show that dynamic pricing may involve departures from the single period optimal choice. By doing so, the market maker induces order flows that are more informative in a statistical sense than the order flow at the single-period price. Although these strategies are costly in the short run, they are optimal in an intertemporal framework because better current information increases future expected utility. Intertemporal price discovery by a market maker takes the form of price experimentation to invest in the production of information.

We provide necessary and sufficient conditions for the existence of active learning opportunities and present a subclass of demand functions where active learning is optimal. Finally, we examine several examples to illustrate the model. In the first example, the market maker adopts active control strategies even though all agents (including the market maker) have homogeneous expectations regarding the full information value of the asset. In the second example, the market maker trades with agents who have private information, but chooses not to experiment with prices. In a two-sided market, we provide an example where the bid-ask spread is narrower under active learning strategies than under the single period strategy. The numerical example suggests that returns measured immediately following an event may reflect non- trivial experimental pricing and therefore be biased predictors of the security's value.

The paper can be extended in a number of ways. Perhaps the most obvious extension is to make explicit our intuition that competition will not generate sufficient investment in the production of information because free entry undermines the incentives to collect information at cost by following active learning strategies. This analysis may suggest a reason for the dominance of specialist systems over time. Further, an investigation of how experimental pricing affects prices around event announcements is warranted, including the possibility of gaming the market maker's strategy. The decision framework can also be applied to situations where agents attempt to learn the parameters of the model, such as a trader who plans to sell a block but cannot observe the limit order book and dealer inventories, a corporation that periodically issues debt or equity but is uncertain about it's impact on prices, and so on.

References

Berry, D. and B. Fristedt, *Bandit Problems: Sequential Allocation of Experiments.* London: Chapman and Hall (1985).

Blackwell, D., 1953, "Equivalent Comparisons of Experiments." *Annals of Mathematical Statistics*, June, 265-272.

Bronfman, C. and R. Schwartz, 1990, "Trading, Noise, and Market Structure," working paper, New York University.

DeGroot, M., *Optimal Statistical Decisions*, New York: McGraw-Hill (1970).

Demsetz, H., 1968, "The Cost of Transacting." *Quarterly Journal of Economics*, 82, 33-53.

Easley, D. and N. Kiefer, 1988a, "Optimal Learning with Endogenous Data." CAE Working Paper # 88-08, Cornell University.

Easley, D. and N. Kiefer, 1988b, "Controlling a Stochastic Process with Unknown Parameters." *Econometrica*, 56, 1045-1064.

Easley, D. and M. O'Hara, 1987, "Price, Quantity and Information in Securities Markets." *Journal of Financial Economics*, 19, 69-90.

Fusselman, J. and L. Mirman, 1987, "Experimental Consumption for a General Class of Disturbance Densities." Working Paper, Department of Economics, University of Virginia.

Glosten, L. and P. Milgrom, 1985, "Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Agents." *Journal of Financial Economics*, 14, 71-100.

Grossman, S., R. Kihlstrom and L. Mirman, 1977, "A Bayesian Approach to the Production of Information and Learning by Doing." *The Review of Economic Studies*, 44, 533-547.

Handa, P. and R. A. Schwartz, 1989, "Dynamic Price Discovery and the Short-term Behavior of Security Prices," working paper, New York University.

Ho, T. and H. Stoll, 1983, "The Dynamics of Dealer Markets under Competition." *Journal of Finance*, 38, 1053-1074.

Jordan, J., 1982, "Learning Rational Expectations: The Finite State Case." *Journal of Economic Theory*, 36, 257-276.

Kiefer, N. and Y. Nyarko, 1989, "Optimal Control of a Linear Process with Learning," forthcoming, *International Economic Review*.

Kihlstrom, R., L. Mirman and A. Postelwaite, 1984, "Experimental Consumption and the 'Rothschild Effect'." in *Bayesian Models in Economic Theory*, edited by M. Boyer and R. Kihlstrom. Elsevier Science Publishers, 279-302.

Leach, J. C. and A. Madhavan, 1989, "Price Experimentation and Market Structure," working paper, University of Pennsylvania.

Marcet, A. and Sargent, T., 1989, "Convergence of Least Squares Learning in Environments with Hidden State Variables and Private Information." *Journal of Political Economy*, 97, 1306-1322.

McLennan, A., 1984, "Price Dispersion and Incomplete Learning in the Long Run." *Journal of Economic Dynamics and Control*, 7, 331-347.

O'Hara, M. and G. Oldfield, 1986, "The Microeconomics of Market Making." *Journal of Financial and Quantitative Analysis*, 21,
361-76.

Prescott, E., 1972, "The Multiperiod Control Problem Under Uncertainty," *Econometrica*, 40, 1043-1048.

Rothschild, M., 1976, "A Two-Armed Bandit Theory of Market Pricing." *Journal of Economic Theory*, 9, 185-202.

Schreiber, P. and R. Schwartz, 1986, "Price Discovery in Securities Markets." *Journal of Portfolio Management*, 12, 43-48. Schumpeter, J., *Capitalism, Socialism, and Democracy*, New York, Harper: (1942).

APPENDIX

Proof of Lemma 1: Necessity is obvious. For sufficiency, we proceed inductively. The condition of the theorem gives us:

$$\vec{\mu}_1(z_0, p_0, \mu_0) = \vec{\mu}_1(z'_0, p'_0, \mu_0) \tag{38}$$

and the induction hypothesis is:

$$\vec{\mu}_s(\vec{z}_{s-1}, \vec{p}_{s-1}, \mu_0) = \vec{\mu}_s(\vec{z'}_{s-1}, \vec{p'}_{s-1}, \mu_0) \tag{39}$$

for all $s = 1, \ldots, t$, where $\vec{z'}_{s-1}$, and $\vec{p'}_{s-1}$ are the vectors of past realizations of excess demand and prices where the first price is p'_0 rather than p_0. Since actions depend only on the posterior and time, it follows that $p_t = p'_t$ and $z_t = z'_t$ for $t \geq 1$. Applying Bayes' rule (see, e.g., DeGroot, p.148), we obtain:

$$\mu^i_{t+1} = \frac{\mu^i_t h(z_t - f(\beta^i, p_t(\mu_t)); 0)}{\sum_{k=1}^n \mu^k_t h(z_t - f(\beta^k, p_t(\mu_t)); 0)}. \tag{40}$$

By the induction hypothesis and the fact that the realizations at time t (z_t's) are the same, we can write:

$$\mu^i_{t+1} = \frac{\mu^{i'}_t h(z'_t - f(\beta^i, p_t(\mu'_t); 0)}{\sum_{k=1}^n \mu^{k'}_t h(z'_t - f(\beta^k, p_t(\mu'_t)); 0)} \tag{41}$$

which is just $\mu^{i'}_{t+1}$. ∎

Proof of Theorem 1: If controlled learning (experimentation) is futile, setting $p_t \neq p^M_t$ is sub-optimal, and the result follows. Accordingly, we consider functional forms where actions will not lead to differences in posterior distributions. By Lemma 1, invariance of one period forward posteriors in price choice is all we must consider. We need only show that this invariance is equivalent to the stated condition on functional forms.

(Necessity of the condition for invariance:) Let τ denote the true state. Then invariance in the one period forward posterior requires:

$$\frac{\mu^i_0 h(f(p, \beta^\tau) + \epsilon - f(p, \beta^i); 0)}{\sum_{k=1}^n \mu^k_0 h(f(p, \beta^\tau) + \epsilon - f(p, \beta^k); 0)} = \frac{\mu^i_0 h(f(p', \beta^\tau) + \epsilon - f(p', \beta^i); 0)}{\sum_{k=1}^n \mu^k_0 h(f(p', \beta^\tau) + \epsilon - f(p', \beta^k); 0)} \tag{42}$$

for all i, j; for all $p, p' \in I$. It follows that:

$$\frac{h(f(p,\beta^\tau)+\epsilon-f(p,\beta^i);0)}{h(f(p',\beta^\tau)+\epsilon-f(p',\beta^i);0)}=\frac{h(f(p,\beta^\tau)+\epsilon-f(p,\beta^j);0)}{h(f-(p',\beta^\tau)+\epsilon-f(p',\beta^j);0)} \tag{43}$$

for all i, j; for all $p, p' \in I$. Set $i=\tau$ in the left hand side to get:

$$h(f(p,\beta^\tau)+\epsilon-f(p,\beta^j);0)=h(f(p',\beta^\tau)+\epsilon-f(p',\beta^j);0). \tag{44}$$

Since this holds for all ϵ we know:

$$f(p,\beta^\tau)+\epsilon-f(p,\beta^j)=f(p',\beta^\tau)+\epsilon-f(p',\beta^j). \tag{45}$$

Since this holds for all $p, p' \in I$, $f(p,\beta^\tau)-f(p,\beta^j)$ cannot depend on p. It follows then that $f(p,\beta^i)-f(p,\beta^j)$ cannot depend on p, and the condition follows immediately.

(**Sufficiency of the condition for invariance**) If $\frac{\partial f}{\partial p}$ is independent of p, it follows that:

$$f(p,\beta^i)-f(p,\beta^j)=f(p';\beta^i)-f(p';\beta^j)=k(\beta^i,\beta^j) \tag{46}$$

for all $p, p' \in I$; for all $i, j = 1, .., n$. Then

$$\begin{aligned}
h(f(p,\beta^\tau)+\epsilon-f(p,\beta^i);0) &= h(k(\beta^\tau,\beta^i))\\
h(f(p';\beta^\tau)+\epsilon-f(p',\beta^i);0) &= h(k(\beta^\tau,\beta^i))\\
h(f(p,\beta^\tau)+\epsilon-f(p,\beta^i);0) &= h(k(\beta^\tau,\beta^i))\\
h(f(p';\beta^\tau)+\epsilon-f(p',\beta^i);0) &= h(k(\beta^\tau,\beta^i)).
\end{aligned}$$

This implies equal numerators and denominators in the Bayes formulae and thus equal posteriors.∎

Proof of Lemma 2: Our approach follows GKM with certain important modifications necessitated because the unknown parameter, β, enters the utility (expected profit) function. The proof omits some technical details for expositional ease. We begin by constructing the decision set of the agent. For $t \geq 1$, let γ_t denote a function with domain W_{t-1} and range I, and let $\Gamma_t = \{\gamma_t : W_{t-1} \to I\}$. Define $\gamma_T = \{\gamma_t\}$, and $\Gamma_T = \Gamma_1 \times .. \times \Gamma_T$. The interpretation is that γ_T is a vector of contingency plans, one for each period which is measurable with respect to the market maker's information set (actually, measurable with respect to the σ-algebra generated by the z_t's). That is, they are functions of the observable history. Define $T+1$ functionals:

$$\begin{aligned}
\zeta_0 : \Gamma_T \times I &\to I \\
&\vdots \\
\zeta_t : \Gamma_T \times \Re^t \times I &\to I
\end{aligned}$$

by:

$$\begin{aligned}
\zeta_0(\gamma_T, p_0) &= p_0 \\
\zeta_1(\gamma_T, z_0, p_0) &= \gamma_1(z_0, p_0) \\
&\vdots \\
\zeta_t(\gamma_T, z_{t-1}, p_{t-1}) &= \gamma_t(z_{t-1}, \zeta_{t-1}(\gamma_T, z_{t-2}, p_{t-2})))
\end{aligned}$$

where $\zeta_{t-1} = \zeta^k$ for $k = 0, \ldots, t-1$. The interpretation of these functionals is that each realization of the contingency plan (except for the first) is selected as a function of the previous choices (p_t's). Therefore, p_t is actually dependent on previous excess demands (z_t's) and the original price choice (p_0). The relationship of the decision to these variables is specified by the function γ. The rule ζ_t ties the decision p_t to the history of excess demands, the initial price and the contingency plan vector. The prices chosen, for any observable history, are $p_t = \zeta_t(\cdot,\cdot,\cdot)$. The market maker's dynamic optimization problem is to choose an initial price, $p_0 \in I$, and the T- vector of functions $\gamma_T \in \Gamma_T$ to maximize the discounted expected profit. This completes our construction of the action set.

Next, we define the function $\eta_T : I \times \Gamma_T \times \beta \to \Re$ by:

$$\eta_T(p_0, \gamma; \beta^i) = \left[\phi(p_0; \beta^i) + \sum_{t=1}^{T} \delta^t \int_{\Re_t} \phi(p_t; \beta^i) \prod_{\tau=0}^{t-1} h(z_\tau; f(p_\tau; \beta^i) dz_{t-1}\right] \tag{47}$$

where the function $\phi(p; \beta)$ is the expected utility function conditional upon β:

$$\phi(p; \beta^i) \equiv \int_{\Re} u(p, \beta^i, f(p, \beta^i) + \epsilon) h(\epsilon; 0) d\epsilon. \tag{48}$$

Then, we can show that:

$$V_T(\mu_0) = \sup_{p_0, \gamma} \sum_i \eta_T(p_0, \gamma; \beta^i) \mu_0^i. \tag{49}$$

The formal derivation of this equation is omitted here because of its length and the complexity added by state-dependent utility, but it is straightforward using the definition of $V_T(\mu_0)$. Heuristically, the $\gamma_t(.)$ functions specify optimal actions from the set $\equiv I \times \Gamma_T$ for all possible states of the world, which is exactly equivalent to the maximization problem (5) where the actions were prices, concluding the proof of Lemma 2.■

Proof of Theorem 2: Consider the experiments (see definition 1 above) $h(z; p, \beta)$ and $h(z'; p', \beta)$. Construct the random variables z and z' to have a multivariate normal distribution with means $\psi_0 + \psi_1(\beta)\psi_2(p)$, and $\psi_0 + \psi_1(\beta)\psi_2(p')$ respectively and variance σ^2, and set the correlation between z and z' to be $\rho = \frac{\psi_2(p')}{\psi_2(p)}$. Since $\psi_2(p) > \psi_2(p')$, $\rho < 1$ as required. (Remember $\psi_2 > 0$.) Using the projection theorem for the normal distribution, the conditional mean of z' given z is:

$$\begin{aligned} E[z'|z] &= E[z'] + \rho(z - E[z]) & (50)\\ &= \psi_0 + \psi_1(\beta)\psi_2(p') + \frac{\psi_2(p')}{\psi_2(p)} (z - \psi_0 - \psi_1(\beta)\psi_2(p)) & (51)\\ &= \psi_0 \left(1 - \frac{\psi_2(p')}{\psi_2(p)}\right) + \frac{\psi_2(p')}{\psi_2(p')z} & (52)\end{aligned}$$

which is independent of β. Similarly, the conditional variance of z' given z is:

$$\sigma^2[z'|z] = \sigma^2 \left(1 - \left[\frac{\psi_2(p')}{\psi_2(p)}\right]^2\right) \tag{53}$$

which is positive since $\psi_2(p) > \psi_2(p') > 0$. Denote by $\nu(z'; z)$ the conditional distribution of z' given z, which we know to be normal. Now, by construction $h(.;.)$ is the marginal density of z and $h'(.;.)$ is the marginal density of z'. We know that the marginal density of z' is related to the marginal density of z and the conditional density of z' given z as follows:

$$h(z';.) = \int_{\Re} \nu(z';z)h(z;.)dz \tag{54}$$

where we know $\int_{\Re} \nu(z';z)dz' = 1$. Comparing (54) and definition 1, we see that the experiment $h(z;p,\beta)$ is more informative than the experiment $h(z';p',\beta)$, because $\nu(z';z)$ is independent of β.■

Proof of Theorem 3: In order to produce a contradiction suppose $\psi_2^* < \psi_2^M$. It must be that

$$U(p_0^M; \mu_0) \geq U(p_0^*; \mu_0) \tag{55}$$

for $p_0^M \in P_0^M$, and $p_0^* \in P_0^*$. Since the experiment with $\psi_2(p^M)$ is more informative than the experiment with $\psi_2(p^*)$ (by Theorem 2) then by Blackwell's Theorem:

$$E[V_{T-1}(g_1(\mu_0; p_0^M, z_1^M))] \geq E[V_{T-1}(g_1(\mu_0; p_0^*, z_1^*))] \tag{56}$$

Therefore

$$U(p_0^M; \mu_0) + \delta E[V_{T-1}(g_1(p_0^M; \mu))] \geq U(p_0^*; \mu_0) + \delta E[V_{T-1}(g_1(p_0^*; \mu))] \tag{57}$$

which contradicts the definition of ψ_2^*. Therefore $\psi_2^* \geq \psi_2^M$.■

Proof of Theorem 4:

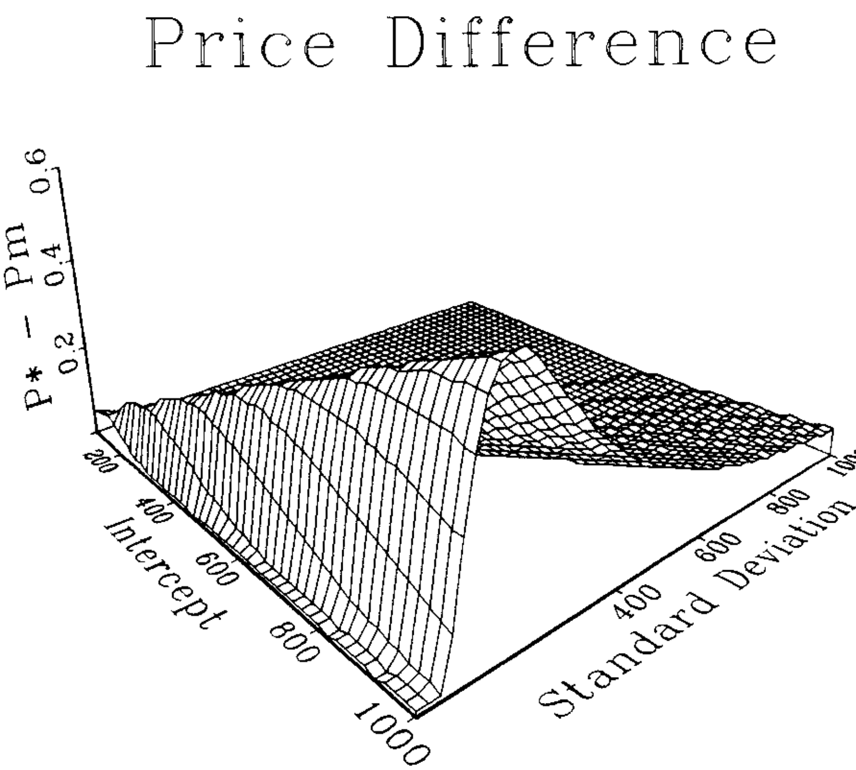
(i) The variance of the initial change in the market maker's inventories is given by:

$$Var[\tilde{z}_0] = Var[\psi_0 + \psi_1(\tilde{\beta})\psi_2(p) + \epsilon_0]$$

Since $\tilde{\epsilon}$ is uncorrelated with $\tilde{\beta}$, we can write this as:

$$Var(\tilde{z}_0) = \psi_1(p)^2 Var[\psi_1(\tilde{\beta})] + \sigma^2 \tag{58}$$

which is higher for $p \in P^*$ than for $p \in P^M$ by Theorem 3. The second part follows directly from applying Theorem 3 and the fact that f is differentiable and strictly decreasing in p.
(iii) With $u(p, \beta, z) = (p - F^{-1}(0))z$, where $F(\cdot) = f(\cdot; \beta)$, active experimentation implies that $U(p_0^M; \mu_0) \geq U(p_0^*; \mu_0)$.∎



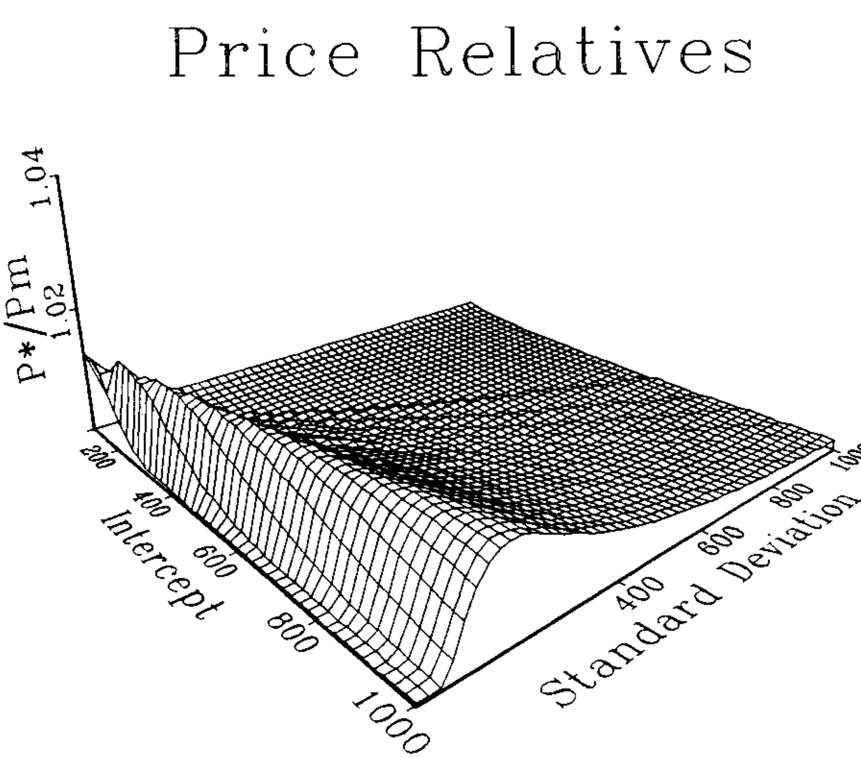


Figure 1: Divergence of Optimal Prices from Myopic Prices