MONEY PUMPS IN THE MARKET

Ariel Rubinstein
Tel Aviv University

Ran Spiegler
University College London

Abstract
Agents who employ non-rational choice procedures are often vulnerable to exploitation, in the sense that a profit-seeking trader can offer them a harmful transaction which they will nevertheless accept. We examine the vulnerability of a procedure for deciding whether to buy a lottery: observe another agent who already bought it and buy the lottery if that agent's experience was positive. We show that the exploitation of such agents can be embedded in an inter-temporal market mechanism, in the form of speculative trade in an asset of no intrinsic value. (JEL: D84)

1. Introduction

A common criterion for evaluating non-rational decision procedures is whether they are vulnerable to exploitation, in the sense that a profit-seeking trader could invent a harmful sequence of bilateral transactions which the non-rational agent will accept, despite the fact that it impoverishes him without limit. Such a sequence is sometimes called a "money pump" or a "dutch book" (see Yaari 1985). This criterion suffers from a certain weakness. Carrying out the exploitative transactions requires direct interaction between the trader and the agent. But when an unfamiliar person approaches us with a quaint proposal, our instinct is to treat it with caution (consider your reaction to an invitation to play "Three Card Monte" on the streets of a foreign city). We tend to think strategically about the situation and suspect that there is a "catch," even if we cannot pinpoint it. This is particularly true if the offer has the appearance of a "free lunch" and the other party does not seem to have anything to gain from it.

A harmful transaction might be more effective if it could be embedded and concealed in an *impersonal market mechanism*, which does not trigger the "never trust a stranger" instinct. For illustration, compare the way people apply

Acknowledgments: We thank Kfir Eliaz, Gur Huberman, and Eilon Solan for useful conversations. We also thank the editor and two referees for helpful comments. Spiegler acknowledges financial support from the ISF, Grant No. 87/05.

E-mail addresses: Rubinstein: rariel@post.tau.ac.il; Spiegler: r.spiegler@ucl.ac.uk

Journal of the European Economic Association March 2008 6(1):237–253

adverse-selection reasoning in face-to-face bargaining situations and impersonal, common-value auctions. When a seller approaches an agent with an offer to purchase a product for a low price, the agent instinctively infers that the product's quality cannot be too high. People seem to be less successful at applying this type of strategic thinking in auctions. The "winner's curse" fallacy, often observed in experimental common-value auctions, is a consequence of this failure.

We do not set out to explain why agents are less suspicious of an impersonal market mechanism than of face-to-face transactions. Instead, we simply assume that this tendency exists. Our main objective in this paper is to analyze the ability to exploit non-rational agents on an ongoing basis under a market mechanism, under the assumption that exploitative transactions are better concealed in such an environment.

We focus our investigation on a decision procedure that is applied to a simple class of choice problems under uncertainty: whether or not to buy (at a given price) a lottery that yields a prize of \$1 or \$0, where the winning probability is unknown. The procedure is to observe the experience of another agent who already bought the lottery, and buy it if that other agent's experience was positive (or, more generally, to sample a number of experienced agents and buy the lottery if the empirical winning frequency in the sample is sufficiently high). A risk-neutral trader can manipulate an agent who employs such a procedure—or, for that matter, any procedure that ensures a positive probability of buying the lottery—by offering him a lottery which yields \$1 with probability $h > 0$, for a price that lies strictly between h and 1.

The problem we study is the following: Is it possible to exploit such agents using instead an impersonal market mechanism? In particular, we consider mechanisms in which an asset of no intrinsic value is traded, such that the "lotteries" that agents face are none other than the traded asset's price fluctuations. We model the market in a manner which bears some resemblance to Glosten and Milgrom (1985). In each period, a trader referred to as a *price maker* (PM) posts a price that belongs to a set consisting of three possible prices, $1 > \delta > 0$, and commits to clear the market at that price. A pricing strategy for the PM is a probabilistic rule for switching from one price to another. In addition to the PM, the market is inhabited by agents who wish to buy the asset at the posted price only for the prospect of selling it at a higher price in the next period.

If the agents knew the PM's pricing strategy and reacted rationally, it would be impossible to exploit them. However, we assume that the agents do not know the PM's strategy. Therefore, in each period they face a choice problem to which the described sampling procedure is applicable. When the current price is δ, the procedure leads each agent to buy the asset with the probability h that the PM raises the price from δ to 1. If $h < \delta$, the PM earns a positive expected profit conditional on the price being δ. However, in order to earn this profit on an ongoing basis, the PM must incur expected losses at the other prices: When the price is 0, the procedure implies that a positive fraction of the agents purchase the asset because the price must rise from 0 with positive probability, and when the price is 1, the agents do not buy the asset because the price never rises.

Nevertheless, we show that the PM is able to earn positive expected profits on an ongoing basis, through a suitable choice of the probabilities of switching from one price to another. We characterize the structure of the PM's optimal pricing strategy, and use it to derive an upper bound on the PM's profits. When $\delta > 1/4$, this bound lies below the maximal expected profit that a trader could earn at the agents' expense if he did not have to conceal the exploitation.

We do not insist on interpreting the PM as an actual individual trader. The PM may be viewed as a metaphor for some exogenous market environment. From this alternative point of view, studying the PM's maximal expected profit is of interest because it shows the worst that could happen to agents who follow the sampling procedure in such a market environment—or, to put it differently, the extent to which the market mechanism protects the agents from ongoing exploitation.

Not every boundedly rational choice procedure would lead to the same conclusions. For example, consider an agent who follows the "dumb" rule of buying every lottery with some exogenous probability. Clearly, if a risk-neutral trader can directly interact with such an agent, he can exploit him by offering him the same exploitative lottery that he proposed to the agent who employed the sampling procedure. However, once the exploitation has to be concealed in the market mechanism, it becomes infeasible. On the one hand, the PM will earn a profit at the agent's expense whenever he lowers the price. On the other hand, in order to earn this profit on an ongoing basis, he will have to raise the price back and suffer an equal loss. Because the agent's probability of purchasing the asset is the same for all prices, the expected loss exactly offsets the expected profit.

Although our primary motivation in this paper is to examine "market concealment" of exploitative transactions, the model we analyze may be related to the phenomenon of speculative trade. The agents in our model engage in speculative trade with the PM, because they predict future asset prices on the basis of naive extrapolation from past observations, whereas the PM knows the true transition probabilities. This is quite different from models of speculative trade suggested in the financial economics literature. In Kyle (1985), for example, an informed rational trader interacts with uninformed, rational traders as well as "noise traders," whose behavior follows some exogenous stochastic rule. Although our boundedly rational agents behave probabilistically, they are not "noise traders" in the sense used in the literature, because they follow an explicit choice procedure which causes their behavior to be systematically related to actual price fluctuations.

The plan of the paper is as follows. In Section 2, we present the market model and analyze it. In Section 3, we introduce a variant on this model, in which the PM can condition price fluctuations not only on the current price, but also on his current asset holdings. We provide concluding remarks in Section 4.

2. The Model

Consider a market for a durable, indivisible asset which has no intrinsic value. Trade takes place in a sequence of periods. Market participants consist of a PM and identical agents. In each period, a continuum of agents of measure 1 are born and live for two periods. Each agent can hold only one unit of the asset. He can buy it at the first period of his life. Conditional on buying it in the first period, he must sell it in the second period. In each period, the PM posts a price p from a finite set P, and he commits to buying or selling any quantity needed to clear the market at this price.

The behavior of the PM is a price policy, namely a function $f : P \to \Delta(P)$, where $\Delta(P)$ is the set of probability measures over P. We use the notation $f(p, q)$ for the probability that $f(p)$ assigns to the price q. Our interpretation of f is that if the price in one period is p, the PM chooses the price in the next period to be q with probability $f(p, q)$. In other words f is a Markov process whose state space is P. Given a price policy f, we say that λ_f is an ergodic distribution if, for all p, we have $\lambda_f(p) = \sum_q \lambda_f(q) f(q, p)$. The interpretation of $\lambda_f(p)$ is the long-run proportion of periods in which the price p is realized, according to the price policy f. Of course, at least one ergodic distribution exists for every f.

The model of the agents' behavior is not conventional. Following Osborne and Rubinstein (1998), we label it as the $S(1)$ procedure. Assume that for any price p, each agent independently draws one sample point from $f(p)$, and buys the asset if and only if the outcome is a price strictly higher than p. Thus $h_f(p) = \sum_{q>p} f(p, q)$ is the probability that according to the price policy f, the price which follows p is strictly higher than p. By the $S(1)$ procedure—and because the agents' samples are independent—the proportion of agents who purchase the asset at price p is $h_f(p)$.

The interpretation of the agents' procedure is as follows. An agent enters the market and observes the posted price. Being ignorant of the PM's price policy, he tries to infer the price trend by examining one period from an infinitely long past, in which the price was also p. He decides to buy the asset if and only the price p at the sampled period was followed by a strictly higher price in the next period.

Thus, under the price policy f, the transition from a price p to price q is associated with a payoff for the PM of

$$\pi_f(p, q) = h_f(p) \cdot (p - q).$$

We assume no discounting and thus, we are led to investigating the maximization of the PM's long-run average profits:

$$\Pi(f) = \sum_{p \in P} \left(\lambda_f(p) \sum_{q \in P} f(p, q) \pi_f(p, q) \right),$$

which—employing the definitions of $h_f(p)$ and $\pi_f(p, q)$—is equivalent to

$$\Pi(f) = \sum_{p \in P} \left(\lambda_f(p) \left(\sum_{q > p} f(p, q) \right) \sum_{q \in P} f(p, q)(p - q) \right),$$

where λ_f is the ergodic distribution of f. (When f admits multiple ergodic distributions, some selection is required for the definition of $\Pi(f)$. However, this issue is evaded in this paper, as we shall see shortly.)

Considered from a formal point of view, the PM's optimization problem is a variant on a Markov Decision Problem (MDP), see Derman (1970). The difference between our problem and an MDP is that in our problem, the payoff function $\pi(p, q)$ is a (linear) function of the transition probabilities from p, whereas in an MDP, payoffs are independent of transition probabilities.

We confine ourselves to a three-price case, $P = \{0, \delta, 1\}$. If the set P consisted of fewer than three prices, no transition in f would yield a positive profit because no agent would buy the asset at the high price. It follows that in order for the PM to earn positive expected profits, P must contain at least three prices. Moreover, in the three-price case, the 3×3 matrix f must be irreducible—namely, every state must be reachable from any other state with positive probability—which means that it has a unique ergodic distribution.

Let us write down the payoff function explicitly for the three-price case. For any price policy f, $h_f(1) = 0$ because the probability that the price will rise above 1 is zero. Therefore, $\pi_f(p, q)$ is given by Table 1. The only transition which is associated with a strictly positive payoff is from the intermediate price δ to the lowest price 0.

2.1. The Possibility of Speculative Gains

Let us begin our analysis by showing that there exists a price policy that yields a positive expected profit. Consider the price policy $f_{\alpha,\beta}$ (where $\alpha, \beta \in (0, 1)$), which is represented by Figure 1. The circles represent the states, the arrows represent transitions that receive positive probability, and the transition probabilities are written near the arrows.

TABLE 1. The payoff function, $\pi_f(p, q)$.

	To 0	To δ	To 1
From 0	0	$-h_f(0)\,\delta$	$-h_f(0)$
From δ	$+h_f(\delta)\,\delta$	0	$-h_f(\delta)\,(1-\delta)$
From 1	0	0	0

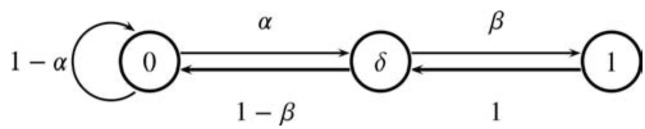

FIGURE 1. A price policy $f_{\alpha,\beta}$.

The ergodic probability distribution induced by $f_{\alpha,\beta}$ is given by the following set of equations:

$$\lambda(0) = \lambda(0)\cdot(1-\alpha) + \lambda(\delta)\cdot(1-\beta),$$
$$\lambda(\delta) = \lambda(0)\cdot\alpha + \lambda(1)\cdot 1,$$
$$\lambda(1) = \lambda(\delta)\cdot\beta,$$

which yields

$$\lambda(0) = \frac{1-\beta}{1-\beta+\alpha+\alpha\beta},$$
$$\lambda(\delta) = \frac{\alpha}{1-\beta+\alpha+\alpha\beta},$$
$$\lambda(1) = \frac{\alpha\beta}{1-\beta+\alpha+\alpha\beta}.$$

The expected profit is thus

$$\Pi(f_{\alpha,\beta}) = \frac{(1-\beta)\alpha(-\alpha\delta) + \alpha\beta\big(-\beta(1-\delta)+\delta(1-\beta)\big) + \alpha\beta\cdot 0}{1-\beta+\alpha+\alpha\beta}. \quad (1)$$

The rationale for the price policy $f_{\alpha,\beta}$ is as follows. Ideally, the PM would like to lure as many agents as possible into buying the asset at the intermediate price δ and then let the price drop to 0. However, in order for agents to purchase the asset with probability β at the intermediate price, this must also be the probability that the price goes up from δ to 1. Thus, the PM faces a trade-off: Lowering β implies a lower proportion of agents who purchase the asset at δ, but a higher expected profit from these agents. Similarly, when the price is 0, the PM balances between the need to raise the price to δ in order to resume the money pump, and the need to make the return from 0 to δ sufficiently unlikely so that sufficiently many agents will choose not to purchase the asset at 0.

PROPOSITION 1. *The price policy $f_{\alpha,\beta}$ yields a positive expected profit.*

Proof. Equation (1) can be simplified to

$$\Pi(f_{\alpha,\beta}) = \alpha \frac{-\alpha\delta + \alpha\beta\delta - \beta^2 + \beta\delta}{1 - \beta + \alpha + \alpha\beta},$$

which is positive if, for example, we set $\beta = \delta/2$ and α is low enough. □

The formula for $\Pi(f_{\alpha,\beta})$ implies that a necessary condition for earning a positive expected profit with $f_{\alpha,\beta}$ is $\beta > \alpha$. That is, in order to make profits, the PM must raise the price more often when it is intermediate than when it is low.

REMARK 1. Our analysis assumes that the agents follow a particular choice procedure. However, we can extend Proposition 1 to a somewhat larger set of procedures. Given a choice procedure C, let $b_{f,C}(p)$ denote the probability that an agent who follows C purchases the asset at the price p, when the PM follows the price policy f. Assume that C is such that these probabilities are well defined. Under the $S(1)$ procedure, $b_{f,C}(p) = h_f(p)$. Suppose that $b_{f,C}(p) = g\big(h_f(p)\big)$, where $g : [0, 1] \to [0, 1]$ is a continuously increasing function satisfying $g(0) = 0$ and $g(1) = 1$. This condition is satisfied, for example, by the procedure in which, for any price p, the agent draws K sample points from $f(p)$ and buys the asset if the drawn price is higher than p in every sample point. A slight change in the proof of Proposition 1 is needed in order to show that given such a procedure, the PM can find a price policy f such that $\Pi(f) > 0$.

In contrast to this class of procedures, consider the "dumb" procedure described in the Introduction, according to which the agents follow some random rule that is independent of p. In this case, every profit that the PM earns when he lowers the price is exactly offset in expectation by the loss that he incurs when raising the price back to its original level. Therefore, the PM cannot exploit the agents on an ongoing basis. In general, any choice procedure C for which $b_{f,C}(p)$ is the same for all prices $p \in P$ prevents ongoing exploitation.

2.2. The Structure of Optimal Price Policies

In this sub-section we characterize the structure of the optimal price policy. We precede the proof with a useful lemma. It is well known that in a standard MDP there is an optimal strategy that is deterministic—that is, a strategy that assigns probability one to a single action at each state. In our model, the optimal policy will not be deterministic. However, our lemma shows that there exists an optimal price policy such that, for every price p, it switches to at most one strictly higher price and at most one weakly lower price.

LEMMA 1. There is an optimal price policy f such that, for every price p, $f(p,q) > 0$ for at most one $q > p$ and at most one $q \leq p$.

Proof. Let f be an optimal price policy. We will show that there is a price policy that has the stated property by applying the classical result (see Derman 1970, chapter 3, Theorem 2) that any MDP admits a deterministic optimal strategy. We cannot apply this result trivially in our setting, because as noted in Section 2, the optimization problem is not an MDP (specifically, $f(p)$ affects $\pi(p,q)$).

We consider an auxiliary optimization problem in which the set of states consists of all p^u and p^{nu} for any $p \in P$. The interpretation is that at p^u the price is supposed to go up, whereas at p^{nu} the price is supposed not to go up. The set of actions available at each state p^u is the set of all prices that are strictly higher than p. Similarly, the set of actions available at each state p^{nu} is the set of all prices that are not strictly higher than p. Given a state p^j $(j = u, nu)$ and a feasible action q, the system switches to q^u with probability $h_f(q)$ and to q^{nu} with probability $1 - h_f(q)$, and yields the payoff $\pi_f(p,q)$.

The auxiliary problem was designed such that any strategy in the original problem yields the same expected payoff when applied to the auxiliary problem. However, the auxiliary problem is an MDP. Therefore, there is an optimal price policy that assigns a deterministic price to every state p^u or p^{nu}. It follows that, in the original problem, there exists an optimal price policy such that for every price p, $f(p,q) > 0$ for no more than one $q > p$ and for no more than one $q \leq p$. □

The intuition for this lemma is as follows. As long as $h(p)$ is held fixed for all prices p, the PM's decision problem is reduced to an MDP. Therefore, the same logic that allows solutions to an MDP to be deterministic implies that in our problem, all the weight that the PM assigns to transitions from p to higher prices can be concentrated on a single transition; and similarly, all the weight that the PM assigns to transitions from p to weakly lower prices can be concentrated on a single transition.

Before proceeding, we introduce some useful notation. Suppose that f is irreducible. Given a state s, define an s-cycle to be a sequence of states of the form $(s, \ldots, s)$, where all the states except the first and the last are not s. Let $C(s)$ be the set of all s-cycles. For every cycle c, $K(c)$ denotes the number of transitions in the cycle, $\sigma(c)$ denotes the sum of payoffs associated with the transitions in the cycle, and $\alpha(c)$ denotes the probability of the cycle according to f. Then, it is possible to write the PM's expected profit as follows:

$$\Pi(f) = \frac{\sum_{c \in C(s)} \alpha(c)\sigma(c)}{\sum_{c \in C(s)} \alpha(c)K(c)}. \tag{2}$$

We will show now that there is always an optimal price policy of the form $f_{\alpha,\beta}$ as defined previously (see Figure 1).

PROPOSITION 2. *There exist α, β such that $f_{\alpha,\beta}$ is an optimal price policy.*

Proof. By Lemma 1, there is an optimal price policy f such that for every price p, $f(p,q) > 0$ for at most one $q > p$ and at most one $q \leq p$. Because f is irreducible, $f(1,1) = 1$ is ruled out, and thus $f(1,1) = 0$.

In order to make a positive expected profit, the PM must assign a positive probability to the transition from δ to 0. By Lemma 1, there must be an optimal price policy f such that $f(\delta,\delta) = 0$.

It remains to show that $f(1,0) = f(0,1) = 0$. Assume that $f(1,0) > 0$ and thus $f(1,\delta) = 0$. Consider a δ-cycle in which a transition from 1 to 0 occurs. Then, the cycle must be of the form $(\delta, \mathbf{a}, 1, 0, \mathbf{b}, \delta)$, where $\mathbf{a}$ and $\mathbf{b}$ are sequences of states, with the restriction that $\mathbf{a}$ does not contain a transition from 1 to 0. Suppose that the PM deviates to a price policy f' that is identical to f except that the probability $f(1,0) = 1$ is shifted to $f'(1,\delta)$. This deviation implies that $h_{f'}(s) = h_f(s)$ for every state s. Then, the probability that f assigns to any cycle of the form $(\delta, \mathbf{a}, 1, 0, \mathbf{b}, \delta)$ is shifted to $(\delta, \mathbf{a}, 1, \delta)$. Note that $K(\delta, \mathbf{a}, 1, \delta) \leq K(\delta, \mathbf{a}, 1, 0, \mathbf{b}, \delta)$, and because the only transition that is associated with a positive payoff is from δ to 0, $\sigma(\delta, \mathbf{a}, 1, \delta) \geq \sigma(\delta, \mathbf{a}, 1, 0, \mathbf{b}, \delta)$. By equation (2), $\Pi(f') \geq \Pi(f)$.

The proof that $f(0,1) = 0$ is identical. □

The intuition for this result is as follows. The only transition that yields a positive payoff for the PM is the transition from δ to 0. Therefore, when the price goes up from 0 or goes down from 1, the PM prefers to switch directly to δ.

2.3. The Cost of "Market Concealment"

There is no closed solution to the problem $\max_{\alpha,\beta} \Pi(f_{\alpha,\beta})$. However, we provide an upper bound on the PM's expected profit. The proof is relegated to the Appendix.

PROPOSITION 3. *No price policy earns an expected profit above $\delta/16$.*

Another upper bound can be obtained through a comparison with a non-market environment. In the Introduction, we described a scenario in which a *non-market manipulator* offers in each period a lottery that yields a prize of 1 with some probability h and a prize of 0 with probability $1 - h$, at a price δ. In each period, agents who face this lottery follow the $S(1)$ procedure. That is, they accept the lottery if and only if the outcome of a random draw from the lottery yields a favorable outcome. The manipulator's expected profit in this case is $h(\delta - h)$. The maximal expected profit is $\delta^2/4$, attained with $h = \delta/2$.

In contrast, the PM in our model may be viewed as a *market manipulator*. He, too, offers a lottery that yields a prize of 1 with some probability h and a prize of 0 with probability $1-h$, at a price δ—but only in those periods for which the price is δ. Because the PM earns a positive profit only in these periods, it follows (independently of the upper bound given by Proposition 3) that the PM's expected profit cannot exceed $\delta^2/4$.

Moreover, the PM's profits are strictly below those enjoyed by the non-market manipulator. The reason is that the prices 1 and 0 are more than mere promises to pay prizes to agents who buy the asset at the price δ. They involve two "costs" . First, in any period for which the announced price is 0 or 1, the PM is unable to offer the profitable lottery. Second, in any period for which the announced price is 0, some agents buy the asset, and when the PM raises the price from 0 to δ, he bears a loss from their purchase decision because he has to pay them δ. Proposition 1 establishes that despite these costs, the PM can choose a price policy that yields positive profits. However, this argument makes it clear that these profits are strictly lower than those of the non-market manipulator.

3. Conditioning Prices on Asset Holdings

Our model has assumed that the state variable in the PM price policy is the price. However, the PM could in principle observe the quantity of the asset in his possession, and condition next period's price on his current holdings as well as on the current price. In the model studied in the previous section this was redundant, because there were many agents whose behavior was statistically independent. As a result, the PM's holdings at the end of each period were uniquely pinned down by the probability distribution over next period's price.

In this section, we present a variant on the model, in which the agents' behavior is coordinated, and therefore the ability to condition on holdings makes a difference. Instead of assuming that the agents' samples from $f(p)$ are independent, suppose that they all observe the same sample point—say, the most recent period in which the price was p—and therefore, in each period all agents take the same action. This means that the quantity held by the PM at the end of each period behaves stochastically.

Formally, let us continue to denote the set of prices by $P = \{0, \delta, 1\}$. Define the set of states S as the set of all p^1 and p^0 where $p \in P$. The state p^z means that the current price is p and the PM currently holds a stock z. A price policy for the PM is a function $f : S \to \Delta(P)$. We use $f(p^z, q)$ to denote the probability that next period's price will be q, conditional on the current state being p^z. Denote $f^u(p^z) = \sum_{q>p} f(p^z, q)$. Assume that, unlike the PM, the agents do not observe the PM's stock and cannot distinguish between periods in which the state is p^1 or p^0. Given a price p, the agents choose whether to purchase the asset according to the $S(1)$ procedure. They sample one period in which the price is p, and choose to buy the asset if and only if the price in the next period is strictly higher.

Let λ_f be an ergodic distribution of f. Then, conditional on the current price being p, the probability that the agents purchase the asset is

$$h_f(p) = \sum_{z=0,1} \frac{\lambda_f(p^z) f^u(p^z)}{\lambda_f(p^0) + \lambda_f(p^1)}.$$

Given a price policy f, and conditional on the current price being p, the probability that the state is p^0 is $h_f(p)$, and the probability that the state is p^1 is $1 - h_f(p)$. Therefore, λ_f is an ergodic distribution if, for every $p \in P$,

$$\lambda_f(p^0) = \sum_{q^z \in S} \lambda_f(q^z) \cdot f(q^z, p) \cdot h_f(p),$$

$$\lambda_f(p^1) = \sum_{q^z \in S} \lambda_f(q^z) \cdot f(q^z, p) \cdot \big(1 - h_f(p)\big).$$

It is easy to see that for every f, there is at least one ergodic distribution.

The PM's payoff associated with the transition from state p^z to state q^w is

$$\pi(p^z, q^w) = q \cdot (z - w).$$

We continue to assume no discounting and thus we shall investigate the maximization of the PM's long-run average profits:

$$\Pi(f) = \sum_{p^z, q} \lambda_f(p^z) \cdot f(p^z, q) \cdot \big(h(q) \cdot \pi(p^z, q^0) + (1 - h(q)) \cdot \pi(p^z, q^1)\big),$$

where λ_f is the ergodic distribution of f.

As in the previous version of the model, the PM cannot earn positive expected profits if the ergodic distribution assigns positive probability to two prices only, say 0 and 1. Because $h(1) = 0$, the state 1^0 is never reached. Therefore, only the states 0^0, 0^1, and 1^1 can have positive probability under the ergodic distribution. But any cycle through these states yields a non-positive sum of payoffs and therefore, the PM's expected profit cannot be positive. It follows that we can restrict attention to irreducible price policies, which are known to induce a unique ergodic distribution.

3.1. An Optimal Price Policy

The following example shows that the PM can exploit the agents' inability to perceive any systematic relation between price movements and the PM's holdings, and earn a higher expected profit than when he cannot condition prices on holdings. Consider the fully deterministic price policy

$$g(0^1, \delta) = g(\delta^1, 1) = g(1^1, \delta) = g(0^0, 0) = g(\delta^0, 0) = g(1^0, 0) = 1.$$

That is, the PM raises the price by one level if his stock is 1—except when the price is 1, in which case he lowers the price to δ. When the PM's stock is 0, he sets the price to 0.

The ergodic distribution of g assigns zero probability to the state 1^0 because $h_g(1) = 0$. Therefore, g induces a five-state Markov process, which is given diagrammatically by Figure 2. Each of the transitions in this diagram has probability $1/2$, except for the transition from δ^1 to 1^1, whose probability is 1. The ergodic distribution assigns equal probability to each of the five states 1^1, δ^0, δ^1, 0^0 and 0^1. To verify that this is the case, note that $h_g(0) = h_g(\delta) = 1/2$. The PM's expected profit is $\delta/5$, because only the transitions from 1^1 and 0^1 to δ^0 entail non-zero payoffs (δ in each case). Observe that this is higher than the upper bound we derived in Section 3.

Our next result shows that the price policy g is optimal.

PROPOSITION 4. *There exists no price policy that generates an expected profit higher than $\delta/5$.*

The price policy g fully exploits the agents' inability to distinguish between p^0 and p^1. The PM raises the market price only when his current stock is 1, and (weakly) lowers the price when his current stock is 0. The agents purchase the asset when they observe a price increase, although the periods in which the price rises are precisely the periods in which the asset is held by the PM. For instance,

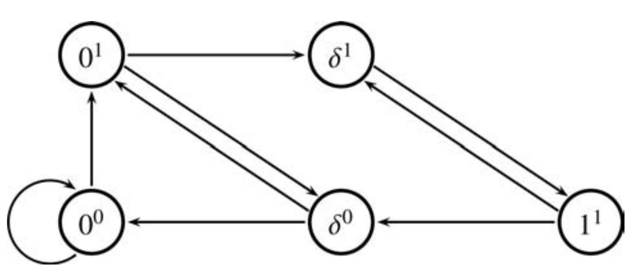

FIGURE 2. The five-state Markov process induced by g.

consider the case of $p = 0$. The agents purchase the asset with probability $1/2$ at this price because the price rises from 0 to δ with probability $1/2$. However, the price rises only from the state 0^1—that is, when the agents have not bought the asset and therefore cannot reap the benefit from the price increase.

3.2. The Cost of "Market Concealment"

As in the case of our original model, the following question is naturally raised: to what extent does the need to support the outcomes of an exploitative lottery as market prices reduces the expected profit that can be earned at the agents' expense? The relevant benchmark, given the model of this section, is as follows. A risk-neutral trader offers in each period a lottery that yields 1 or 0, at a price δ. Imagine that if the agent rejects the offer, the trader gives the lottery to a proxy. In this case, although the outcome of the lottery is made public, it does not involve any monetary transfer. The trader can condition the lottery's outcome on whether he sold it to an agent or gave it to the proxy.

Suppose that the agent decides whether to accept the lottery on the basis of a sampled past realization of the lottery—whether it was sold to an agent or given to the proxy. Let h^a and h^p denote the probabilities that the trader assigns to the outcome 1 when he sells the lottery to an agent or gives it to a proxy, respectively. Then, the trader's expected profit is $\alpha(\delta - h^a)$, where α is the probability that the agent will purchase the lottery, given by $\alpha = \alpha h^a + (1 - \alpha)h^p$. A simple calculation shows that the trader will set $h^p = 1$ and $h^a = 0$, yielding an expected profit of $\delta/2$. Thus, the difference in expected profits due to the need to conceal this trick in the market apparatus is $\delta/2 - \delta/5$.

4. Concluding Remarks

Vulnerability to the invasion of rational agents. Suppose that the population of traders with whom the price maker interacts consisted of both $S(1)$-agents and rational agents who fully understand the price maker's price policy. Then, the rational agents would not purchase the asset at the price δ but they would purchase it at the price 0. The rational agents' strategy inflicts a loss on the PM, and if their proportion is sufficiently large, he is unable to earn a positive expected profit.

In the benchmark situation described in the Introduction, in which the risk-neutral trader constantly offers the same lottery at the price δ, without having to support the lottery's outcome as market prices, he does not suffer from this vulnerability. The rational agents avoid purchasing the lottery, but they do not inflict a loss on the trader. Thus, the trader's vulnerability to the invasion of rational agents is another cost of "market concealment."

The relevant parameter space. The set of prices in our model consists of three prices, 0, δ, and 1. The procedure we ascribed to the boundedly rational agents applies to all values of δ. However, we find that the model is reasonable only for relatively low values of δ. If δ is high, it does not make sense to base one's decision on one random observation of the price that follows δ, because the potential loss is large while the potential gain is small. However, our qualitative results would persist if we replaced the $S(1)$ procedure with an $S(n)$ procedure, in which the agents base their decision on $n > 1$ sample points, although the price maker's profits would be reduced.

The Relation to MDPs. The price maker's maximization problem in this paper is close but not identical to a MDP. As we pointed out in the main text, the difference is that in our model, the payoff associated with a transition is a linear function of the transition probabilities. We solved the maximization problem in two variants of the model. However, we are not aware of a general characterization of the optimal solution for such optimization problems. This also makes it difficult for us to extend the model to the case in which P contains more than three prices.

Related literature. Osborne and Rubinstein (1998) introduced the $S(1)$ procedure in a game-theoretic context. In their model, each player chooses his action after sampling each possible action once. Their focus was on constructing a suitable equilibrium concept for an interaction between agents who employ the $S(1)$ procedure. Equilibrium in a symmetric game is a distribution of actions such that the probability assigned to an action is the probability that an agent who uses the $S(1)$ procedure decides to take this action. Osborne and Rubinstein (2003) apply a variant of the concept to a voting model.

Spiegler (2006a, 2006b) analyzes markets in which profit-maximizing firms compete over consumers who employ the $S(1)$ procedure to evaluate each firm. As in the present paper, the consumers' choice rule makes them vulnerable to exploitation by the firms. The exploitative effect need not disappear as the number of competitors increases.

In its critical exploration of harmful transactions, this paper is related to the literature on Dutch books (e.g., see Yaari 1985). This literature has constructed exploitative transactions for a variety of decision procedures that violate properties such as transitivity, independence or Bayesian updating, and argued that the possibility of such transactions ensures that these violations of rationality will disappear from the market. Laibson and Yariv (2005) study the market performance of Dutch books from a wholly different perspective than ours. They construct a general equilibrium model with dynamically inconsistent consumers. They define a Dutch book in this environment as a temporal sequence of contracts that impoverishes such consumers, and show that the exploitative effect of Dutch books disappear in competitive equilibrium.

Appendix: Proofs

A.1. Proof of Proposition 3

Let f be a price policy and let $\Pi(f, \delta)$ denote the expected profit from f when the intermediate price is δ. The function $\Pi(f, \delta)$ is linear in δ. Moreover, $\Pi(f, 0) \leq 0$. Therefore, it suffices to show that $\max_f \Pi(f, 1) \leq 1/16$.

By Proposition 2, we can restrict attention to price policies of the form $f_{\alpha,\beta}$. An upper bound on $\Pi(f_{\alpha,\beta}, 1)$ is attained if we collapse the two upper states into a single state—that is, as if only one period passes when we move from the state δ to the state 1 and then back to the state δ. The payoff function associated with the modified two-state Markov process is as follows:

	To 0	To δ
From 0	0	$-\alpha$
From δ	β	0

The transition probabilities are the following:

	To 0	To δ
From 0	$1-\alpha$	α
From δ	$1-\beta$	β

The ergodic probabilities of this process are $\lambda(0) : \lambda(\delta) = (1-\beta) : \alpha$. The expected profit is thus (substituting $\delta = 1$)

$$-\frac{1-\beta}{\alpha+1-\beta}\alpha^2 + \frac{\alpha}{\alpha+1-\beta}\beta(1-\beta) = \alpha(1-\beta)\frac{-\alpha+\beta}{\alpha+1-\beta}.$$

Let $\Delta = (\alpha + 1 - \beta)$. Then we have

$$\alpha(1-\beta)\frac{-\alpha+\beta}{\alpha+1-\beta} \leq \frac{(\Delta/2)^2(1-\Delta)}{\Delta} = \frac{\Delta(1-\Delta)}{4} \leq 1/16.$$

A.2. Proof of Proposition 4

Assume that f is a price policy with ergodic distribution λ such that $\Pi(f) > \delta/5$. Our method of proof is to get a contradiction by showing that $\lambda(s) > 1/5$ for $s \in \{1^1, \delta^1, \delta^0, 0^1, 0^0\}$.

Step 0. Without loss of generality, for every state p^z there is at most one $q > p$ for which $f(p^z, q) > 0$, and at most one $q \leq p$ for which $f(p^z, q) > 0$.

Proof. Along the same lines as in Lemma 1.

Step 1. We can assume that $f(\delta^0, \delta) = 0$.

Proof. Assume that $f(\delta^0, \delta) > 0$. By Step 0, we can assume that $f(\delta^0, 0) = 0$. Let $C(\delta^0, p)$ be the set of all δ^0-cycles of the form $(\delta^0, p^z, \ldots, \delta^0)$. Every δ^0-cycle c for which $\alpha(c) > 0$ is in $C(\delta^0, \delta)$ or $C(\delta^0, 1)$ and thus $\sigma(c) \leq 0$. Therefore, by (2), $\Pi(f) \leq 0$, a contradiction.

Step 2. $\lambda(p^1)\, f^u(p^1) = \lambda(p^0)\big(1 - f^u(p^0)\big)$, for $p = 0, \delta$.

Proof. By definition,

$$h(p) = \frac{\lambda(p^0) f^u(p^0) + \lambda(p^1) f^u(p^1)}{\lambda(p^0) + \lambda(p^1)}.$$

By the agents' rule of behavior, $\lambda(p^0) = h(p)\big(\lambda(p^0) + \lambda(p^1)\big)$. The claim follows from these two identities.

Step 3. $\lambda(\delta^0)\, f(\delta^0, 0) > 1/5$ and $\lambda(0^1)\, f(0^1, \delta) > 1/5$.

Proof. In order for the PM to earn a positive expected profit, it must be that $\lambda(\delta^0) > 0$. Therefore, the total probability of cycles in $C(\delta^0)$ is positive.

Consider a cycle $c \in C(\delta^0)$. Observe that $\sigma(c) \leq \delta$. Moreover, $\sigma(c) \leq 0$ unless the cycle contains a transition from δ^0 to 0^z and a transition from 0^1 to δ^z. Therefore, for both pairs $(s, p) = (0^1, \delta)$ and $(\delta^0, 0)$,

$$\begin{aligned}\frac{\delta}{5} < \Pi(f) = \frac{\sum_{c \in C(\delta^0)} \alpha(c)\sigma(c)}{\sum_{c \in C(\delta^0)} \alpha(c) K(c)} &\leq \delta \frac{\sum_{c \in C(\delta^0), \sigma(c) > 0} \alpha(c)}{\sum_{c \in C(\delta^0)} \alpha(c) K(c)} \\ &\leq \delta \frac{\sum_{c \in C(\delta^0, 0)} \alpha(c)}{\sum_{c \in C(\delta^0)} \alpha(c) K(c)} \leq \delta \lambda(s) f(s, p).\end{aligned}$$

The claim thus follows.

Step 4. $\lambda(1^1), \lambda(\delta^1), \lambda(0^1) > 1/5$.

Proof. Because $h(1) = 0$, we have $\lambda(1^0) = 0$ and thus $\lambda(1^1) \geq \lambda(\delta^0) f(\delta^0, 1) + \lambda(\delta^1) f(\delta^1, 1)$. By Step 2, $\lambda(\delta^1) f(\delta^1, 1) = \lambda(\delta^0)(1 - f(\delta^0, 1))$ and therefore $\lambda(1^1) \geq \lambda(\delta^0) > 1/5$.

By Step 1, $f(\delta^0, 0) = 1 - f(\delta^0, 1)$. By Step 2,

$$\lambda(\delta^1) f(\delta^1, 1) = \lambda(\delta^0)\big(1 - f(\delta^0, 1)\big) = \lambda(\delta^0) f(\delta^0, 0),$$
$$\lambda(0^1)\big(1 - f(0^1, 0)\big) = \lambda(0^0)\big(1 - f(0^0, 0)\big) \geq \lambda(0^0) f(0^0, \delta).$$

By Step 3, the right-hand side in both cases is greater than $1/5$.

References

Derman, C. (1970). *Finite State Markovian Decision Processes*. Academic Press.

Glosten, L., and P. Milgrom (1985). "Bid, Ask and Transactions Prices in a Specialist Market with Insider Trading." *Journal of Financial Economics*, 14, 71–100.

Kyle, A. (1985). "Continuous Auctions and Insider Trading." *Econometrica*, 53, 1315–1336.

Laibson, D., and L. Yariv (2005). "Safety in Markets: An Impossibility Theorem for Dutch Books." Working paper, Caltech.

Osborne, M., and A. Rubinstein (1998). "Games with Procedurally Rational Players." *American Economic Review*, 88, 834–849.

Osborne, M., and A. Rubinstein (2003), "Sampling Equilibrium with an Application to Strategic Voting", *Games and Economic Behavior*, 45, 434–441.

Spiegler, R. (2006a). "The Market for Quacks." *Review of Economic Studies*, 73, 1113–1131.

Spiegler, R. (2006b). "Competition over Agents with Boundedly Rational Expectations." *Theoretical Economics*, 1, 207–231.

Yaari, M. (1985), "On the Role of 'Dutch Books' in the Theory of Choice under Risk." Nancy Schwartz Lecture, Kellogg Graduate School of Management, Northwestern University.