Market forces meet behavioral biases: cost misallocation and irrational pricing

Nabil Al-Najjar*

Sandeep Baliga*

and

David Besanko*

Psychological and experimental evidence, as well as a wealth of anecdotal examples, suggests that firms may confound fixed, sunk, and variable costs, leading to distorted pricing decisions. This article investigates the extent to which market forces and learning eventually eliminate these distortions. We envision firms that experiment with cost methodologies that are consistent with real-world accounting practices, including ones that confuse the relevance of variable, fixed, and sunk costs to pricing decisions. Firms follow "naive" adaptive learning to adjust prices and reinforcement learning to modify their costing methodologies. Costing and pricing practices that increase profits are reinforced. In some market structures, but not in others, this process of reinforcement causes pricing practices of all firms to systematically depart from standard equilibrium predictions.

1. Introduction

■ Economic theory offers the unambiguous prescription that only marginal cost is relevant for profit-maximizing pricing decisions. Ongoing fixed costs or previously incurred sunk costs, although relevant for entry and exit decisions, are irrelevant for pricing. This theoretical prescription stands in stark contrast to evidence about real-world pricing practices. In surveys of pricing practices of U.S. companies, Govindarajan and Anthony (1995), Shim (1993), and Shim and Sudit (1995) find that most firms price their products based on costing methodologies that treat fixed and sunk costs as relevant for pricing decisions. Leading textbooks on managerial and cost accounting paint a similar picture. Maher, Stickney and Weil (2004) assert that, when it comes to pricing practices, "[o]verwhelmingly, companies around the globe use full costs rather than variable costs." They cite surveys of U.S. industries "showing that full-cost pricing dominated pricing practices (69.5 percent), while only 12.1 percent of the respondents used a

*Northwestern University; al-najjar@kellogg.northwestern.edu; baliga@kellogg.northwestern.edu; d-besanko@kellogg.northwestern.edu.

variable-cost based approach." (Horngren, Foster, and Datar (2000), another leading accounting textbook, report other surveys in which a majority of managers in the United States, the United Kingdom, and Australia take fixed and sunk costs into account in pricing.

Although suggestive, surveys indicating that firms use full-cost pricing (i.e., confounding the roles of variable, fixed, and sunk costs; see Section 2) do not imply that this has a *material effect* on observed prices. Full-cost pricing may well be a convenient general rule managers use to find the rational pricing point. For instance, managers may be influenced by a combination of biases that "cancel out," leading them to price "as if" they understood the economic reasoning of marginal revenue and marginal cost. Moreover, behavioral biases in critical decisions such as cost allocation and pricing may eventually disappear in response to learning and competition. Pursuit of profit is a powerful incentive for firms to learn to price optimally. Also, competition among managers for promotion and advancement within a firm should favor those who price optimally. Alchian's (1950) classic argument that learning and imitation would propagate good practices suggests that interfirm competition would only reinforce the case for optimal pricing: "[W]henever successful enterprises are observed, the elements common to these observable successes will also be associated with success and copied by others in their pursuit of profits or success. 'Nothing succeeds like success.'"

Yet, important mechanisms for propagating best accounting practices lend little if any support for the use of economics-based pricing principles. For example, textbooks in managerial accounting often list marginal-cost-based pricing as just one of several acceptable methodologies, alongside others that incorporate fixed and sunk costs into pricing. Some managerial accounting texts even argue *against* basing prices on marginal costs.[1]

One way to resolve this issue is through controlled experiments where decision makers face environments that differ only in the presence of irrelevant costs. Offerman and Potters (2006) conduct such an experiment in a Bertrand duopoly context. In their baseline treatment with no fixed or sunk cost, Offerman and Potters find that prices converge to the Bertrand equilibrium. In the sunk cost treatment, subjects face the same demands and marginal costs as in the baseline treatment, but they must pay a sunk entry fee to play the game. In this treatment, Offerman and Potters find that prices are significantly higher: once the sunk entry fee is paid, the average markup over marginal cost is 30% higher than the markup in the baseline treatment.

The goal of this article is explain why cost misallocation, in the form of full-cost pricing, persists, and indeed thrives, despite the forces of learning and competition. Two key ideas underlie our model. First, the presence of irrelevant costs (e.g., sunk cost) triggers a predisposition among firms to use full-cost pricing. Psychological and experimental evidence, as well as a wealth of anecdotal examples, suggests that this confusion is, at least, plausible.[2] Second, when subjected to competitive market pressures, firms adjust their prices and, much less frequently, their "costing methodologies" by reinforcing the practices that yielded the best past results.

Thus, we do not assume optimal pricing rules or market equilibrium at the outset. Rather, we ask whether learning and competitive pressures force firms to overcome their initial biases, leading them to behave as though they played equilibrium strategies with optimal pricing rules. Long-run equilibrium behavior is derived, rather than assumed.

---

[1] To cite one example, Shank and Govindarajan (1989) admonish managers to avoid committing the "Braniff fallacy," the practice (named after the now-defunct Braniff Airlines) of being content to sell seats at prices that merely cover the incremental cost of the seat. They state that "looking at *incremental* business on an *incremental* cost basis will, at best, *incrementally* enhance overall performance. It cannot be done on a *big* enough scale to make a *big* impact. If the scale is that large, then an incremental look is not appropriate!" (emphasis in original). Shank and Govindarajan make it clear that they believe that managers should incorporate fixed and sunk costs into pricing decisions, arguing that "business history reveals as many sins by taking an incremental view as by taking the full cost view."

[2] A particularly vivid example is Edgar Bronfman, former owner of Universal Studios, who criticized the movie industry's pricing model because ticket prices do not reflect differences in the sunk production costs of different movies. "He . . . observed that consumers paid the same amounts to see a movie that costs $2 million to make as they do for films that cost $200 million to produce. 'This is a pricing model that makes no sense, and I believe the entire industry should and must revisit it.'" *Wall Street Journal*, April 1, 1998.

We show that market, forces eradicate irrational pricing in monopoly, perfectly competitive markets, and undifferentiated Bertrand oligopoly (see Section 4). Things are quite different in the case of Bertrand oligopoly with product differentiation. Theorem 1 shows that adaptive learning eventually leads to higher profits for any firm that unilaterally incorporates part of its irrelevant costs in its pricing decisions. We also provide a dynamic model in which firms experiment with new costing practices and reinforce successful ones. We assume that firms' choice of costing methodologies is subject to inertia (they adjust prices more frequently than they experiment with new costing methodologies).[3] Two conceptual problems arise: first, it is implausible that firms know their rivals' distorted costing practices. Second, even if these practices are known, computing payoff functions requires calculating the limit of the adaptive adjustment in prices. Firms that are naive enough to confuse cost concepts are unlikely to have the requisite sophistication and understanding of the game to carry out such computations. Thus, our firms face a learning problem where neither opponents' strategies nor the payoff functions are known.[4] Despite this, Theorem 2 shows that a simple process of experimentation and reinforcement leads all firms to eventually distort their relevant costs by incorporating fixed and sunk costs in their pricing decisions. This explains why full-cost pricing has been so hard to eradicate.

In our model, cost misallocation acts as though it is a commitment to raise prices. But it is a spurious commitment in the sense that it can be easily eliminated through learning. Firms can learn to price "correctly" in our model, and this indeed is what happens under some market structures. This is to be contrasted with the classical analysis of strategic commitment, where sophisticated players make binding decisions to limit their flexibility and foresee the impact of these decisions on their rivals' future behavior. Here, we envision real-world managers whose behavior is shaped by myopic incentives and refined by naive adaptive learning. The driving force in our model is confusion about which costs are relevant for pricing decisions. Because such confusion is unlikely to be observable, it is quite different from the familiar foundation of commitment, namely credible, transparent public actions designed to shape equilibrium outcomes in subsequent stages of the game. See Section 6 for further discussion.

The confusion of relevant and irrelevant costs in human decision making manifests itself in a myriad of ways. This is often referred to as the sunk cost bias or the sunk cost fallacy. See Thaler (1980) for a pioneering study of its role in economic decision making. One pattern, extensively supported by experimental evidence (see Arkes and Ayton, 1999), is for individuals to persist in an activity "to get their money's worth."[5] Under this pattern, individuals deflate the true cost of the activity. In this article, we focus on how the sunk cost bias manifests itself in pricing decisions. We allow firms to be rational, inflate, or deflate relevant cost. Our analysis identifies circumstances under which a systematic bias toward inflating cost appears.

To sum up, our concern is not with the origins of the predisposition to confound relevant and irrelevant costs, or to explain why this or other behavioral biases appear in one-off situations (like the anecdotal examples mentioned earlier). Rather, we examine the claim that the forces of learning and competition will eventually eliminate these biases. Because there is no reason to suspect that managers' innate predispositions to behavioral biases is correlated with industry structure, our model leads to testable predictions about how biases in pricing practices vary across

---

[3] An indirect indication of the inertia involved in the selection of costing systems comes from the empirical literature on the adoption of activity-based costing (ABC). ABC is a set of practices for assigning costs to products in a multiproduct firm that is widely believed to be superior to traditional methodologies for allocating common costs. Beginning in the mid-1980s, academics and consultants began touting the virtues of ABC, but available evidence suggests that the adoption of ABC in the mid to late 1990s was not widespread (see, for example, Brown, Booth, and Giacobbe, 2004; Roztocki and Schultz, 2003). Among the factors that have slowed the adoption of ABC systems include the need to make significant new investments in information technology in order to implement an ABC system and the need to obtain acceptance from key managers to support the transition to an ABC system.

[4] In the pricing game, we assume that each firm knows its own demand and the recent history of its rivals' prices.

[5] For example, in a well-known experiment (see Arkes and Ayton, 1999), theater season tickets were sold at three randomly selected prices. Those charged the lower price attended fewer events during the season. Apparently, those who had "sunk" the most money into the season tickets were most motivated to use them.

market structures. In fact, our prediction that monopolists would eventually rid themselves of the sunk cost bias is consistent with the experimental findings of Offerman and Potters (2006).

The article proceeds as follows. In Section 2, we provide an interpretation of cost accounting practices as they relate to pricing decisions. Sections 3–5 introduce the model and our main results. In Section 5, we specialize our general model to the canonical case of symmetric linear demand. We obtain a closed-form solution that shows how the distortion of relevant cost depends on the degree of product differentiation and the number of firms. Finally, Section 6 compares the implications of our model with available experimental evidence and discusses related literature.

2. Price setting and costing methodologies in practice

■   An economist reading managerial accounting textbooks may be surprised to discover that they mainly consist of a compilation of common company practices. In contrast to the traditional economic treatment of optimal pricing, managerial accounting offers no rigid guidelines as to how various costs should factor into firms' pricing practices.

In this section, we present our understanding and an interpretation of the accounting principles used as a guide in day-to-day costing practices.

☐   **Full-cost pricing.** The most common set of real-world pricing practices falls under the rubrics *cost-based pricing*, *cost-plus pricing*, or *full-cost pricing*. Although they come in a wide range of variations, they all base price on a calculation of an average or unit cost that includes variable, fixed, and sunk costs.[6]

Firms justify full costing using a variety of arguments with typically little or no foundation in economic theory:

• *Simplicity.* Full-cost formulas are thought to be relatively straightforward to implement because they do not require detailed analysis of cost behavior in order to separate the fixed and variable components of various cost items. A related argument is that estimates of marginal cost are often imprecise and indeed even misleading in large, multiproduct firms (Kaplan and Atkinson, 1989).

• *Promote full recovery of all costs of the product.* Pricing based on a full-cost formula is sometimes justified because it provides a clear indicator of the minimum price needed to ensure the long-run survival of the business (Horngren, Foster, and Datar, 2000, henceforth HFD). The idea is that without full-cost pricing, a firm's managers would not be as aware of the "pricing hurdle" that the firm would need to clear in order to generate economic profits for the firm.[7]

• *Competitive discipline.* Some managers believe that full-costing methodologies promote pricing stability by limiting "the ability of salespersons to cut prices" and reducing the "temptation to engage in excessive long-run price cutting." (HFD). Further, full costing may allow firms to better coordinate on price increases: "At a time when all firms in the industry face similar cost increases due to industry-wide labor contracts or material price increases, firms will implement similar price increases even with no communication or collusion among individual firms" (Kaplan and Atkinson, 1989).

☐   **Absence of common standards and the importance of flexibility.** An underlying message of the accounting literature is that firms should adapt their costing methodologies to fit their particular competitive environment and product line idiosyncrasies. That is, an emphasis is placed on *flexibility*. The typical theme is that the appropriate methodology depends on the

---

[6] See Kaplan and Atkinson (1989) for a detailed discussion of full-cost pricing.

[7] This argument may seem strange to economists who might wonder why it is not possible for the firm's managers to price to maximize profits based on standard marginal reasoning and to simply do a "side calculation" to determine whether that profit-maximizing price allows the firm to recover its total costs, including the rental rate on capital.

industry environment. For example, full-cost pricing is considered to be well suited for firms that operate in differentiated product industries (HFD), whereas companies that operate in highly competitive commodity industries are encouraged to set prices based on competitors' prices.[8]

Managerial practice mirrors the textbook emphasis on flexibility. For example, the cost of shared assets is often allocated to individual product lines according to fixed, and more or less arbitrary, percentages. In computing unit costs, there is no universally accepted benchmark for what quantity should go in the "denominator": some firms calculate unit cost at full capacity (in one of its various forms), whereas others use historical output levels, and still others base unit costs on forecasts of future sales. Some companies compute prices by applying a single markup to a chosen cost base, whereas others apply different markup percentages to different cost categories.

☐ **Representational faithfulness.** Given the arbitrariness and flexibility in pricing methodologies, there is no presumption that a firm includes all of its fixed and sunk costs in its computation of unit costs for price-setting purposes. Still, it is reasonable to believe that "*firms do not create costs out of thin air.*" In fact, we shall assume that firms' distortions display a minimal degree of *coherence*, requiring that they only allocate existing fixed and sunk costs.

This can be justified by the accounting principle of *representational faithfulness*, which is one of the two major principles (the other being verifiability) that defines the standard to which external financial statements are expected to adhere (Financial Accounting Standards Board, 1980). This principle requires the "correspondence or agreement between a measure or description and the phenomenon it purports to represent." Although internal financial information is not required to meet external standards of reliability, external reporting systems often have an important impact on accounting information that is used for internal purposes. It therefore seems plausible that a principle designed to keep external accounting information grounded in the economic fundamentals of transactions would also keep internal accounting information similarly grounded.

Another justification may be based on psychological evidence that the sunk cost bias results from decision makers' taste for taking actions that rationalize past choices. Clearly, if no sunk or fixed costs are committed, there is nothing to rationalize and, under this theory, the sunk cost bias should not appear.

☐ **Budgets and variances.** For pricing and other operating decisions, firms typically utilize a budget of forecasts of costs and quantities for a given accounting period. Budgets aid in decision making and also serve as a benchmark against which decisions can be evaluated *ex post*. Budgeted amounts may be based on historical performance, engineering studies of how various cost items might behave, and hypotheses about competitors' actions.[9] For example, a budget might specify a firm's per-unit cost based on the quantity of output produced in the most recent time period.

The budget process is subject to inertia. A firm may change its price and other operating decisions many times before changing the methodology used to determine its budget.

At the end of each accounting period, firms typically compute a *variance*, the difference between the budget amount and the actual result.[10] Variances provide a mechanism by which anticipated and actual performance can be reconciled. For example, suppose a firm has computed its unit costs based on a target output, and the actual output in the accounting period differs from this target output. A variance would then reconcile the unit cost the firm assumed would prevail and the unit cost that actually prevailed. Variance analysis ensures that managers ultimately are held accountable for the actual performance of the firm.

---

[8] This emphasis on flexibility and "sticking with what works" is consistent with theoretical models of reinforcement learning.

[9] Sophisticated firms often build flexibility into their budgeting. This flexibility allows the firm to adjust the values of key variables to reflect unexpected shocks (e.g., abrupt increase in the cost of a raw material, labor strikes, and so on).

[10] See Maher, Stickney, and Weil (2004) for a thorough discussion of variances.

□ **Summary.** The discussion above highlights key features of managerial practice which we will incorporate in our formal model:

(i) Firms choose "costing methodologies" that might confound sunk, fixed, and variable costs. Firms do not necessarily allocate all of fixed and sunk costs to the cost base, but this allocation cannot exceed actual fixed and sunk costs.
(ii) Firms compute the per-unit fixed and sunk cost based on a budgeted output level that remains constant during an accounting period.
(iii) Day-to-day management of the firm's pricing strategy is guided by the firm's budget and costing methodology. Although firms can adjust prices quickly, budgets and costing methodologies are less flexible and change less often than prices.
(iv) Firms periodically reconcile actual and budgeted profits by adding back variances ensuring that, at the end of each accounting period, firms observe their actual economic profit.

3. The model

■ We focus on the most challenging case for us, namely that of Bertrand competition with product differentiation. Analysis of other market structures is more straightforward, and is dealt with in Section 4.

We consider a Bertrand oligopoly in which boundedly rational firms may adopt costing methodologies which cause their pricing decisions to depart from those prescribed by textbook economic theory. This section sets up the basic model.

□ **Demand and cost fundamentals.** The industry consists of N single-product firms engaged in (Bertrand) price competition with differentiated products. (By a slight abuse of notation, N also denotes the set of firms.) Firm n's quantity, denoted q_n, is determined by a demand function $q_n = D_n(p)$, where $p = (p_1, \ldots, p_N)$ denotes the industry's vector of prices.[11] We assume that firm n knows its own demand function, D_n, but not the demand functions of the other firms.

We assume that D_n is differentiable and, wherever $D_n > 0$, satisfies $\frac{\partial D_n}{\partial p_n} < 0$ and $\frac{\partial D_n}{\partial p_m} > 0$ for $m \neq n$. Further, we assume that for any pair of firms n, m, demand is strictly log-supermodular, that is, $\frac{\partial^2 \log D_n(p_n, p_{-n})}{\partial p_n \partial p_m} > 0$. Log-supermodularity is equivalent to the intuitive condition that a firm's demand becomes less price elastic as a rival's price goes up, a property that is satisfied by many common demand systems, including linear, logit, and CES.

Each firm has a constant marginal cost c_n, with $c = (c_1, \ldots, c_N) \geq 0$ denoting the vector of these costs. In addition, each firm has a per-period fixed or sunk cost, denoted $F_n \geq 0$. Our analysis also covers, without any modifications, unsunk flow fixed costs. However, we will use the term sunk cost to avoid redundancy. For example, F_n may represent the per-period capital charge for an asset that has no redeployment value. The calculation of F_n will depend on the specific accounting standards used by the firm (for instance, how depreciation is computed). We abstract from these issues and take F_n as given. In our multiperiod analysis, we also assume that F_n is constant over time.

□ **Economic versus accounting profits.** Firm n's objective measure of success is its *economic profit*, defined as

$$\pi_n^e(p) \equiv (p_n - c_n)D_n(p) - F_n.$$

We model the firm's distorted perception of relevant costs in a way that is both simple and consistent with managerial accounting practices discussed in Section 2.

---

[11] We use the following standard vector notation. For any vector $x \in R^l$, $x = (x_1, \ldots, x_l)$ and $x_{-n} \in R^{l-1}$ denotes the vector $(x_1, \ldots, x_{n-1}, x_{n+1}, \ldots, x_l)$. Also, $x \geq 0$ means that $x_j \geq 0$ for all j; $x > 0$ means that $x_j \geq 0$ for all j and $x \neq 0$; and $x >> 0$ means that $x_j > 0$ for all j.

Introducing the possibility of irrational pricing is delicate, because there is a potentially infinite number of ways a firm can violate the precepts of optimal pricing. Here we introduce irrationality in a tractable and psychologically plausible way that is grounded in common practices. Specifically, we shall assume that a firm may act as though its true marginal cost is inflated by a constant distortion component, s_n, representing the part of sunk cost (mis)allocated as a per-unit variable cost. We require that

$$s_n \geq 0, \qquad \text{and} \qquad s_n = 0 \text{ whenever } F_n = 0.$$

The second part of this requirement says that firms cannot "make up" sunk costs when there are none. This reflects the idea that a firm must be able to "rationalize" its distortion, that it cannot catch itself in a contradiction. Not doing so would lead to too flagrantly illogical choices that even our irrational firms would be able to identify.

A concrete example in which these assumptions are satisfied is as follows. Imagine that firm n bases its cost allocation on a budgeted quantity $\hat{q}_n$ (see Section 2). In practice, $\hat{q}_n$ is based on past realized quantities, and is therefore independent of the firm's current decisions. For now, we take $\hat{q}_n > 0$ as exogenously given. In footnote 20, we explain how its choice can be made endogenous after the full learning model is introduced. Given $\hat{q}_n$, the firm's *accounting profit* is

$$\pi_n(p, s_n) \equiv (p_n - c_n - s_n)D_n(p) - (F_n - s_n\hat{q}_n), \tag{1}$$

where $(F_n - s_n\hat{q}_n)$ denotes the unallocated sunk cost. The accounting profit function π_n represents the manager's *perception* of the relevant payoff function within a given round of short-term price adjustments. Note that in any steady state, budgeted and actual quantities coincide, in which case the difference between economic and accounting profits, $s_n(D_n(p) - \hat{q}_n)$, disappears. A firm that chooses s_n so that

$$0 \leq s_n \leq \frac{F_n}{\hat{q}_n}$$

would satisfy the requirements on s_n stated earlier. Note that we do not require that firms allocate all of their sunk cost. This may be justified in terms of the arbitrariness and flexibility in costing methodologies discussed in Section 2.

☐ **The static price competition game.** Formally, a price competition game $\Gamma(s)$ with cost distortions s is an N-player game, each with strategy set P and payoffs function given by the accounting profit function defined in equation (1).

In $\Gamma(s)$, firm n maximizes $\pi_n(\cdot, s_n)$ given a forecast p_{-n} of its competitors's prices, yielding a first-order condition

$$(p_n - c_n - s_n)\frac{\partial D_n(p)}{\partial p_n} + D_n(p) = 0, \tag{2}$$

provided $D_n(p) > 0$. This gives rise to firm n's reaction function[12]

$$r_n(p, s_n) = \arg\max_{p'_n} \pi_n(p'_n, p_{-n}, s_n).$$

Our assumptions on demand imply that $r_n(\cdot, s_n): P^N \to P$ is differentiable on the interior of P^{N-1} and strictly increasing:

$$\forall n \neq m, \quad \frac{\partial r_n}{\partial p_m} > 0.$$

Let $r(\cdot, s): P^N \to P^N$ denote the vector of reaction functions. A (Nash) equilibrium for $\Gamma(s)$ is a price vector $\bar{p}$ such that $\bar{p} = r(\bar{p}, s)$. Let $\mathcal{E}(s)$ denote the set of equilibria for $\Gamma(s)$, which we

---

[12] Our assumptions on demand imply that this is well defined and single valued.

assume to be compact and contained in the interior of a cube $P = [0, p^+] \subset R^N$ for every vector of distortions s.

Given the assumption that demand is log-supermodular, the price-setting game $\Gamma(s)$ will be supermodular, which implies that a pure strategy equilibrium exists (Milgrom and Roberts, 1990). Our assumptions do not, however, rule out the possibility of multiple equilibria. Note that any equilibrium p such that all firms are active (i.e., $\bar{q} >> 0$) must satisfy $\bar{p} >> c + s$.

4. Main results

■ **Adaptive learning and comparative statics.** We first define adaptive pricing processes in the game $\Gamma(s)$ repeatedly played by myopic players.

Definition 1. An adaptive pricing adjustment for $\Gamma(s)$ starting at $\bar{p}$ is a sequence of prices $\{p_t\}$ such that $p_0 = \bar{p}$ and there is a positive integer γ such that for every t, p_t satisfies[13]

$$r(\inf\{p_{t-\gamma}, \ldots, p_{t-1}\}, s) \leq p_t \leq r(\sup\{p_{t-\gamma}, \ldots, p_{t-1}\}, s).$$

That is, at time t, each firm best responds to some probability distribution on the recent past history of play of their opponents. The class of adaptive pricing adjustment process is quite broad, and includes, as special cases, the Cournot dynamic,[14] a version of fictitious play in which only the past γ rounds of play are taken into account, and sequential best response.

Learning plays two crucial roles in our analysis. First, it is implausible that firms that succumb to the sunk cost fallacy have the sophistication necessary to carry out the complex reasoning justifying Nash equilibrium play. The alternative view, which we take here, is that a Nash equilibrium in the pricing game is the result of a naive adaptive process carried out by boundedly rational players. Second, although a positive cost distortion will shift the equilibrium *set* upward,[15] equilibrium analysis is consistent with the selection of a lower equilibrium price within the new set. In this case, arbitrary equilibrium selection, rather than fundamentals, ends up driving the persistence of cost distortions. Theorem 1 shows that learning can serve as a foundation for meaningful comparative statics when multiple equilibria may be present. That theorem builds on the following.

Proposition 1. Suppose that $\bar{p} \leq r(\bar{p}, s)$. Then, any price adjustment process beginning at $\bar{p}$ will settle at the lowest equilibrium of $\Gamma(s)$ higher than $\bar{p}$. Formally,

$$\lim_{t \to \infty} p_t = \inf\{z \in \mathcal{E}(s); z \geq \bar{p}\}$$

for any adaptive pricing adjustment sequence $\{p_t\}$ for $\Gamma(s)$ starting at $\bar{p}$.

The proof is essentially that of Theorem 3 in Echenique (2002) specialized to the case where best replies are single-valued.[16] The comparative statics implications of this result is that starting with an equilibrium price $\bar{p}$, if one firm distorts its cost, then any adaptive pricing adjustment must converge to the *lowest* equilibrium price vector of $\Gamma(s)$ that is higher than $\bar{p}$. Note that this result yields an unambiguous prediction on the direction of price changes, but is silent on the effect on economic profits, which is our primary concern. This is dealt with by our first theorem.

---

[13] For $t < \gamma$ we set $p_{t-\gamma} = p_0$.

[14] Under the Cournot dynamic, firms best respond to their opponents' last period price (see Vives, 1999 for instance).

[15] In the sense that the smallest (largest) equilibrium price vector of $\Gamma(s')$ is larger than the smallest (largest) equilibrium price vector of $\Gamma(s)$. See Milgrom and Roberts (1990).

[16] As r is a strictly increasing reaction function, the lowest and highest selections from r are equal. Also, the lowest and highest selections from iterated applications of r are also equal. The proof of part 2 of Theorem 3 in Echenique (2002) shows that these selections define the lower and upper bounds of the limit points of any adaptive dynamic. Hence, as r is a best-reply function, this limit is unique. Finally, part 1 of Theorem 3 shows that this limit must be equal to the lowest equilibrium higher than $\bar{p}$.

☐ **Distortion result.**

Theorem 1. There exists $\Delta > 0$ *such that given any firm n, vector of distortions* $\bar{s}_{-n} \geq 0$, *equilibrium price vector* $\bar{p}$ *of* $\Gamma(\bar{s}_{-n}, s_n = 0)$ *with corresponding quantity vector* $\bar{q} >> 0$, *if firm n chooses* $0 < \hat{s}_n < \Delta$, *then for any adaptive pricing adjustment* $\{p_t\}$ *for* $\Gamma(\bar{s}_{-n}, \hat{s}_n)$ *starting at* $\bar{p}$ *there is a time T such that*

$$\pi_n^e(p_t) > \pi_n^e(\bar{p}) \quad \text{for every } t \geq T.$$

Proofs of this and subsequent results are in the Appendix.

To interpret the theorem assume that, unknown to its opponent, firm n "experiments" with a new costing methodology which inflates its relevant costs by some small amount. If all firms adjust their prices adaptively, then firm n's experiment triggers a process of pricing adjustments that leads to higher *economic* profits for this firm. The theorem asserts that this can be accomplished for a range of values of firm n's distortions independent of the distortions of other firms, the initial equilibrium, and the specific adaptive process of price adjustments.

To provide an intuition for the proof, fix an equilibrium $\bar{p}$ of $\Gamma(\bar{s}_{-n}, s_n = 0)$ and consider an arbitrarily small distortion $\hat{s}_n > 0$ that triggers an adaptive adjustment process converging to a new equilibrium price $\hat{p}(s_n)$. The question is whether firm n's economic profits increase as a result. There are two possibilities. In the first case, $\hat{p}(s_n)$ is bounded away from $\bar{p}$ as s_n goes to zero, in which case firm n is easily seen to achieve higher economic profits. Roughly, this happens if $\bar{p}$ is not locally stable, so a slight perturbation leads to a large jump in prices. In the second case, $\hat{p}(s_n)$ is not bounded way from $\bar{p}$ as s_n vanishes. Here, roughly, the net effect on the economic profits of firm n consists of a negative direct effect (due to the distortion of its price, holding the prices of its rivals fixed) and a positive strategic effect (due to its rivals raising their prices). We show that for a small enough distortion of relevant costs, the latter dominates the former.

We note that no assumptions are made about firm n's knowledge about the demand functions, costs, or costing methodologies of its rivals. Nor does any firm need to know, specifically, what new equilibrium would be reached if it distorted its relevant cost.

☐ **Adjustment of costing methodologies.** We now turn to the process through which firms choose costing methodologies. As suggested in the Introduction, new complications arise, namely that firms do not know their rivals' strategies or their own payoff functions. Because firms must learn about their payoffs as well as resolve the strategic uncertainty about their opponents' behavior, the learning model here is of necessity different from the *adaptive* model used in the pricing game.[17]

We introduce a model of *reinforcement* learning through trial and error. In this model, motivated by Milgrom and Roberts (1991), firms know their strategy set and observe their realized payoffs. But they do not know their payoff functions or the strategies used by their opponents. Firms experiment with different costing methodologies from time to time and reinforce those that have done well on average in the past.

A dynamic game of experimentation and cost adjustments. We assume firms can adjust prices much more frequently than they can adjust their costing methodologies. To model this, we first

---

[17] Most adaptive learning models assume that firms know their own payoff function and the actions taken by their opponents. We are aware of only a few models that can handle the case of unobserved opponents' actions and unknown own-payoff function, all of which rely on players' randomly experimenting with different strategies. Milgrom and Roberts (1991) and Hart and Mas-Colell (2000) describe adaptive procedures that eventually lead to the play of only serially undominated strategies and correlated, equilibrium respectively. Friedman and Mezzetti (2001) propose a procedure that leads to the play of a Nash equilibrium in supermodular games (as well as other classes of games with a property called weak finite improvement property). The problem is that supermodularity of the costing methodology game requires stringent, difficult to interpret conditions on the underlying model. In particular, it does not necessarily follow from the supermodularity of the pricing game.

restrict firms to choose from a finite grid of relevant cost distortions $S_n = \{0, \delta, \ldots, K\delta\}$, where K is a positive integer and $\delta > 0$ (S_n is the same for all firms). Firm n chooses $s_n(\tau) \in S_n$ at times $\tau = 1, 2, \ldots$. Within each "interval" $[\tau, \tau + 1]$, we have a sequence of *price adjustment sub-periods* $t = 1, 2, \ldots$, during which firms may adjust prices but not their distortion of relevant costs.[18]

To define the dynamic game of cost adjustments, fix $s(0)$ arbitrarily and let $p(0)$ be any equilibrium of the game $\Gamma(s(0))$.[19] At time τ, firm n is chosen with probability $\frac{1}{N}$, at which point this firm is allowed to reevaluate its costing methodology. With probability $\epsilon > 0$, this firm experiments by picking a distortion $s_n(\tau)$ uniformly from S_n. We call these periods "firm n's experimental periods." With probability $1 - \epsilon$, firm n picks the distortion that generated the highest average payoff in its past experimental periods. Firms experiment independently from each other and over time. All other firms $m \neq n$ set $s_m(\tau) = s_m(\tau - 1)$.

Firm n's payoff at τ is

$$\Pi_n(s(\tau), p(\tau - 1)) \equiv \pi_n^e(p(\tau)),$$

where $p(\tau) = \lim_{t \to \infty} p_t$ and $\{p_t\}$ is any adaptive pricing adjustment sequence for $\Gamma(s(\tau))$ starting at $p(\tau - 1)$.[20] To motivate this definition, suppose that $\tau > 0$ is an experimental period for firm n in which firm n chooses $s_n(\tau) \neq s_n(\tau - 1)$. This triggers a process of price adjustments starting from $p(\tau - 1)$ and converging to a new equilibrium price $p(\tau)$. Our assumption that prices adjust more rapidly than costing methodologies implies that the limiting price $p(\tau)$ obtains "before" period $\tau + 1$ arrives. The payoff function above says that firm n uses the profit from this final price, $\pi_n^e(p(\tau))$, in judging its experiment at τ.

Our model is motivated by Milgrom and Roberts's (1991) study of learning and experimentation in repeated, normal form games. Their analysis is inapplicable in our setting, however, because we are dealing with a *dynamic* game with history-dependent payoffs. In our model, the payoff structure at time τ is a reduced form for an underlying pricing dynamic whose outcome may depend on the initial condition $p(\tau - 1)$.

Learning result.

Theorem 2. *There is $\Delta > 0$ such that for every $0 < \delta < \Delta$ and any grid $S_n = \{0, \delta, \ldots, K\delta\}$, with probability one, there is $\bar{\tau}$ such that for every firm n, in all nonexperimental periods $\tau \geq \bar{\tau}$, $s_n(\tau) > 0$.*

That is, on almost all paths all firms eventually distort their relevant costs. The strength of the results is in how weak the assumptions are. Firms only know their past distortion choices and, in each case, how well they have done in price competition. This rather coarse information does not allow firms to correlate changes in their payoffs with changes in their rivals' actions, or conduct counterfactuals such as "I would do better by choosing s_n' given the actions of my rivals." Nevertheless, this coarse information is sufficient for firms to eventually reject "rational" costing.

☐ **No-distortion results.** Our analysis makes sharp predictions about which environments are unlikely to reinforce the sunk cost bias, and which environments tend to eliminate it.

Monopoly market. The intuition underlying our analysis is that the benefit from distorting relevant costs stems from its favorable strategic effects. This strategic effect is absent in monopoly, so our

---

[18] Thus, formally, we have a double index $\{\tau_t\}_{\tau,t=1}^{\infty}$ and where indices are ordered lexicographically: $\tau_{t'}' > \tau_t$ if and only if $\tau' > \tau$ or $[\tau' = \tau$ and $t' > t]$.

[19] The assumption that $p(0)$ is an equilibrium of $\Gamma(s(0))$ can be relaxed, but not entirely dropped. For example, Theorem 2.10 (ii) in Vives (1999) shows that this assumption is unnecessary under the Cournot dynamic if one starts with an initial price that is lower (higher) than the lowest (highest) equilibrium price of $\Gamma(s = 0)$.

[20] With this notation, we can now endogenize the choice of the reference quantity $\hat{q}_n$ used to determine the per-unit sunk cost $\frac{F_n}{\hat{q}_n}$ that provides an upper bound on s_n. Formally, we set $\hat{q}_n(\tau) = q_n(\tau - 1)$ for $\tau > 1$ and set $\hat{q}_n(0) > 0$ arbitrarily.

model implies that no distortion should persist in a monopoly market. Even though a monopolist is just as likely to be predisposed to confound relevant and irrelevant costs, (nonstrategic) learning will eventually lead the firm to price optimally. This conclusion is consistent with the experimental results of Offerman and Potters (2006), who find that the effects of the sunk cost bias disappear in the monopoly treatment of their experiments.[21]

Homogeneous product benchmarks. Firms' behavior also conforms to standard theoretical predictions in two standard benchmarks: perfect competition and price competition in an oligopoly with homogeneous products. Consider, first, perfect competition, and let p^* be the equilibrium market price. An individual firm is a price taker, and so its output decision has no effect on the market price and hence on other firms' output decisions. A firm that distorts its relevant costs merely moves its own output away from the profit-maximizing quantity and hence reduces its economic profits.

A no-distortion result can also be established for a Bertrand oligopoly with homogeneous products. Suppose the oligopoly consists of N firms with constant marginal costs ordered so that $c_1 \leq c_2 \ldots \leq c_N$. Let $D(p)$ be a demand curve which is continuous and strictly decreasing at all p where $D(p) > 0$ and that there is a "choke price" $\bar{p} < \infty$ such that $D(p) = 0$ for all $p \geq \bar{p}$ and we assume $\bar{p} > c_1$. The demand for firm n is

$$D_n(p) = \begin{cases} D(p_n) & \text{if } p_n \text{ is the lowest price} \\ \dfrac{1}{J} D(p_n) & \text{if } p_n \text{ and } J-1 \text{ other firms set the lowest price} \\ 0 \text{ if } p_n & \text{is not the lowest price.} \end{cases}$$

We also assume that the price firm 1 sets if it is a monopolist is greater than c_2. Otherwise, firm 1 is effectively a monopolist even in the presence of competition and hence the analysis is very similar to the monopoly case above.

Proposition 2. There is no incentive to distort relevant costs under price competition in a homogeneous product oligopoly.

The intuition underlying this result is straightforward. In a homogeneous product Bertrand oligopoly, each firm has such a strong incentive to undercut its competitor to capture the entire market that it is impossible to soften price competition via cost distortion.

Although these textbook benchmarks are useful, a more interesting question in practice is what happens in industries that are near-perfectly homogeneous. To explore this, and a number of other questions, the next section presents an example with linear demand that allows us to study how equilibrium distortions change as product differentiation vanishes and as the number of firms increases without bound.

5. Special case: symmetric linear demand

■   In this section, we examine the important special case of sym metric linear demand. This additional structure enables us to generate comparative statics predictions about how the distortion changes with the number of firms, the degree of product differentiation, and so on. Theorems 1 and 2 illustrated that our main points hold generally. Once we move to the special structure of symmetric linear demand, the analysis becomes much simpler: there is a unique equilibrium in the pricing game for any profile of distortions. The cost distortion game is supermodular, dominance solvable, and its Nash equilibrium is also the unique correlated equilibrium (Milgrom and Roberts, 1990). A broad class of learning models, including our own, converge to the unique equilibrium in the distortion game.

---

[21] See Section 6 for a more detailed discussion.

We assume that the N firms in the industry have a common marginal cost c and face a symmetric system of firm-level demand curves that is consistent with maximization by a representative consumer who has a quadratic net benefit function. The representative consumer chooses quantities $q = (q_1, \ldots q_N)$ to maximize

$$U(q) = aq - \frac{1}{2} bq\Theta q - pq,$$

where a and b are positive constants and

$$\Theta = \begin{pmatrix} 1 & \theta & \ldots & \theta \\ \theta & 1 & \ldots & \theta \\ \vdots & & \ddots & \vdots \\ \theta & \theta & \ldots & 1 \end{pmatrix},$$

where $\theta \in (0, 1)$ parameterizes the extent of horizontal differentiation among the goods. As $\theta \to 0$, the goods become independent, and as $\theta \to 1$, the goods become perfect substitutes. Throughout, we assume that $a > c$.

The system of demand functions implied by the solution to the representative consumer's utility maximization problem is given by

$$q_n = D_n(p) = \alpha - \beta p_n + \gamma \sum_{m \neq n} p_m, \quad n \in \{1, \ldots, N\},$$

where

$$\alpha \equiv \frac{a(1 - \theta)}{b(1 - \theta)[1 + (N - 1)\theta]} > 0,$$

$$\beta \equiv \frac{1 + (N - 2)\theta}{b(1 - \theta)[1 + (N - 1)\theta]} > 0,$$

$$\gamma \equiv \frac{\theta}{b(1 - \theta)[1 + (N - 1)\theta]} \in (0, \beta).$$

For later use, note that

$$2\beta - (N - 1)\gamma = \frac{2(1 - \theta) + (N - 1)\theta}{b(1 - \theta)[1 + (N - 1)\theta]} > 0.$$

The (unique, symmetric) Nash equilibrium in prices *without* distortions is

$$p^0 = \lambda a + (1 - \lambda)c,$$

where

$$\lambda = \frac{(1 - \theta)}{2(1 - \theta) + (N - 1)\theta} \in \left(0, \frac{1}{2}\right).$$

Now, suppose that firm n chooses costing methodology $s_n \in [0, s^+]$, where $s^+ \equiv \frac{F}{\bar{q}}$ is the maximum distortion consistent with a firm's sunk cost and is assumed to be common across all firms. The first-order condition for a firm is

$$\frac{\partial \pi_n(p, s_n)}{\partial p_n} = -\beta(p_n - c) + \alpha - \beta p_n + \gamma \sum_{m \neq n} p_m + \beta s_n = 0, \quad n \in \{1, \ldots, N\}. \quad (3)$$

Given a vector s of distortions, firm n's second-stage equilibrium price $\bar{p}_n(s)$ is found by solving this system of first-order conditions:

$$\bar{p}_n(s) = p^0 + \frac{\beta}{2\beta + \gamma} s_n + \frac{\beta}{2\beta + \gamma} \left(\frac{\gamma}{2\beta - (N - 1)\gamma}\right) \sum_{m=1}^{N} s_m. \quad (4)$$

This implies

$$\frac{\partial \bar{p}_n}{\partial s_n} = \frac{\beta}{2\beta + \gamma} \left[1 + \frac{\gamma}{2\beta - (N-1)\gamma} \right] > 0. \tag{5}$$

$$\frac{\partial \bar{p}_n}{\partial s_m} = \frac{\beta}{2\beta + \gamma} \frac{\gamma}{2\beta - (N-1)\gamma} > 0. \tag{6}$$

Finally, firm n's economic profit is

$$\bar{\pi}_n^e(s) \equiv \pi_n^e(\bar{p}_n(s)) = (\bar{p}_n(s) - c) \left[\alpha - \beta \bar{p}_n(s) + \gamma \sum_{m \neq n} \bar{p}_m(s) \right]. \tag{7}$$

We call the game where the firms choose distortions and payoffs are determined by (7) and (4) the *distortion game.*

Firm n's problem in the distortion game is

$$\max_{s_n \in [0, s^+]} \bar{\pi}_n^e(s) = (\bar{p}_n(s) - c) \left[\alpha - \beta \bar{p}_n(s) + \gamma \sum_{m \neq n} \bar{p}_m(s) \right].$$

We begin by establishing some basic properties of the distortion game:

Proposition 3. The distortion game is supermodular and has a unique, symmetric equilibrium distortion s^*.

Throughout, we assume that the symmetric equilibrium in the first stage is less than the upper bound s^+, and below we show that this distortion is generally positive for $\theta \in (0, 1)$. Given this, the equilibrium distortion s^* satisfies

$$\frac{\partial \bar{\pi}_n^e(s^*)}{\partial s_n} = 0, \ n \in \{1, \ldots, N\},$$

with the induced equilibrium price $p^* = \bar{p}_n(s^*)$, $n \in \{1, \ldots, N\}$. These conditions can be shown to imply

$$s^* = \frac{(p^* - c)(N-1)\xi\gamma}{\beta} \tag{8}$$

$$p^* = p^0 + \frac{\beta s^*}{2\beta - (N-1)\gamma}, \tag{9}$$

where $\xi \equiv \frac{\frac{\partial \bar{p}_n}{\partial s_m}}{\frac{\partial \bar{p}_n}{\partial s_n}} = \frac{\frac{\gamma}{2\beta - (N-1)\gamma}}{1 + \frac{\gamma}{2\beta - (N-1)\gamma}} \in (0, 1)$.

With this derivation in hand, we can establish a baseline set of no-distortion results that mirror those in Section 4:

Proposition 4. (i) As the goods become independent, the equilibrium distortion goes to zero, that is, $\lim_{\theta \to 0} s^* = 0$; (ii) as the goods become perfect substitutes, the equilibrium distortion goes to zero, that is, $\lim_{\theta \to 1} s^* = 0$; (iii) as the number of firms becomes infinitely large, the equilibrium distortion goes to zero, i.e., $\lim_{N \to \infty}$ that is $s^* = 0$.

We now turn our attention to circumstances under which the equilibrium distortion is positive. To do so, we solve (8) and (9) in terms of the primitives of the model: a, b, c, θ, and N.

$$p^* - c = \frac{(p^0 - c)[(2 - \theta) + (N-1)\theta]}{(2 - \theta) + (N-1)\theta(1 - \theta)} \tag{10}$$

$$s^* = \frac{(p^0 - c)(N-1)\theta^2}{[(2 - \theta) + (N-1)\theta(1 - \theta)][(1 - \theta) + (N-1)\theta]}. \tag{11}$$

TABLE 1   Price and Distortions for Different Values of N and θ

θ		$N = 2$	$N = 3$	$N = 5$	$N = 10$	$N = 100$
$\rightarrow 0$	p^0	160.00	160.00	160.00	160.00	160.00
	s^*	0	0	0	0	0
	p^*	160.00	160.00	160.00	160.00	160.00
0.10	p^0	152.11	145.00	132.73	110.00	33.08
	s^*	0.71	1.17	1.65	1.81	0.18
	p^*	152.45	145.65	133.71	111.20	33.24
0.20	p^0	143.33	130.00	110.00	80.59	21.21
	s^*	2.70	3.70	3.95	2.80	0.10
	p^*	144.83	132.22	112.63	82.73	21.31
0.40	p^0	122.50	100.00	74.29	47.50	14.41
	s^*	9.47	9.18	6.29	2.62	0.04
	p^*	128.42	106.43	79.23	49.80	14.45
0.60	p^0	95.71	70.00	47.50	29.35	11.99
	s^*	17.70	12.05	5.76	1.72	0.02
	p^*	108.36	79.64	52.54	30.97	12.01
0.80	p^0	60.00	40.00	26.67	17.89	10.75
	s^*	21.82	9.88	3.40	0.80	0.01
	p^*	78.18	48.89	29.88	18.68	10.76
0.90	p^0	37.27	25.00	17.89	13.61	10.34
	s^*	17.48	6.12	1.80	0.39	≈ 0
	p^*	53.17	30.81	19.64	14.00	10.34
$\rightarrow 1$	p^0	10.00	10.00	10.00	10.00	10.00
	s^*	0	0	0	0	0
	p^*	10.00	10.00	10.00	10.00	10.00

An immediate implication of these expressions is that when there is some degree of (imperfect) product differentiation, firms distort their relevant costs in equilibrium, which in turn elevates the equilibrium price relative to the single-stage Nash equilibrium.

Proposition 5. If $\theta \in (0, 1)$, the equilibrium distortion is positive, that is, $s^* > 0$. Moreover, the induced equilibrium price exceeds the Nash equilibrium price without distortions, that is, $p^* > p^0$.

Table 1 shows calculations of the Nash price, the equilibrium distortion s^*, and the induced equilibrium price p^* for different numbers N of firms and values of the product differentiation parameter, θ.[22] Note that the equilibrium distortion is not monotone in N. For low values of θ, the distortion initially increases in N, but eventually decreases to 0 as N becomes sufficiently large. As one would expect given Proposition 4, the equilibrium distortion is not monotone in θ, either. For any N, it initially increases in θ but eventually begins to decrease, going to zero as the products become perfect substitutes. Thus, for near-perfectly homogeneous products, the cost distortion is very small. There is indirect evidence to support this latter conclusion: Maher, Stickney, and Weilt 2004) note that "cost-based pricing is far less prevalent in Japanese process-type industries (for example, chemicals, oil, and steel)," all of which are products with very weak differentiation. By contrast, in an industry with a small number of firms and an intermediate degree of product differentiation, the equilibrium distortion can be significant. For instance, when $N = 2$ and $\theta = 0.80$, the equilibrium distortion is $21.82, about 36% of the undistorted Nash price of $60.

Table 1 suggests that even though the distortion is nonmonotonic in the number of firms, the induced equilibrium price is. The next proposition confirms this.

Proposition 6. The induced equilibrium price p^* is strictly decreasing in the number of firms.

---

[22] The parameters for these calculations are as follows: $a = 310$, $b = \frac{350}{252}$, and $c = 10$.

Because firms do not internalize the beneficial impact of their distortion of relevant costs on their rivals' profits, they end up with a price lower than the monopoly price.

Proposition 7. When $\theta \in (0, 1)$, the induced equilibrium price is less than the monopoly price.

Let us now turn to a comparative statics analysis of the equilibrium distortions with respect to the other parameters of the model, a and c. To avoid conflating the impact of a and c on the distortion with their effect on the overall price level, we derive the comparative statics result for the percentage distortion above the undistorted Nash equilibrium price, $\frac{s^*}{p^0}$.

Proposition 8. The percentage distortion $\frac{s^*}{p^0}$ is increasing in a and decreasing in c. Hence an increase in demand, as measured by a larger value of a, will result in a greater percentage distortion, whereas an increase in marginal cost will result in a lower-percentage distortion.

This result makes sense. The difference $a - c$ measures the intrinsic value of the industry profit opportunity. The incremental benefit to a firm from suppressing price competition will be greater the more intrinsically profitable the market is, and, as a result, the degree of distortion is greater the greater is a or the lower is c.

As a final point, consider the implications of this analysis for entry. Because distortions increase equilibrium margins, one would expect that entry would be more desirable in that case, and more firms would come into the industry. We can confirm this intuition. Imagine that each firm faces a sunk cost of entry F, and let N^0 and N^* be the equilibrium number of firms in the no-distortion and distortion cases, respectively. Given the equilibrium number of firms, economic profit is zero, which implies

$$\left(\frac{a - p^0}{b}\right) \frac{(p^0 - c)}{1 + (N^0 - 1)\theta} = F \tag{12}$$

$$\left(\frac{a - p^*}{b}\right) \frac{(p^* - c)}{1 + (N^* - 1)\theta} = F. \tag{13}$$

It is straightforward to verify that $N^* > N^0$.

6. Discussion

■ **Experimental results.** As noted in the Introduction, there is strong evidence from the psychology literature that individual decision makers commit the sunk cost fallacy. A well-cited example is a study by Arkes and Blumer (1991) which found that the sunk costs incurred by season ticket holders for a university theater company apparently influenced attendance decisions.

Experimental economists have also studied sunk cost biases, but the evidence from this work is more mixed. Economists' experiments on the sunk cost fallacy differ from those in the psychology literature in that the latter typically do not focus on decision making in market settings in which competition might be expected to create strong pressures for decision makers to "get it right." By contrast, economists' experiments on sunk costs generally place subjects in the context of some sort of market rivalry. For example, Phillip, Battalio & Kogut (1991) study bidding in first-price sealed bid auctions in which subjects paid a sunk admissions fee in order to participate. They found that 95% of the subjects correctly ignored sunk costs in their bidding behavior. Summarizing a variety of experiments in competitive market settings in which sunk costs might effect market outcomes, Smith (2000) writes: "These results showed no evidence of market failure due to the sunk cost fallacy."

The finding that subjects in competitive market experiments do not succumb to the sunk cost fallacy is consistent with the prediction of our theory that it would run counter to the self-interest of a price-taking firm, or a price-setting firm in a homogeneous product oligopoly, to base decisions on sunk costs. However, a more direct test of the implications of our theory would be to explore pricing behavior in a differentiated product oligopoly market in which firms incur

sunk costs. Unfortunately, most of the existing experimental investigations of price competition in a differentiated product oligopoly (e.g., Dolbear et al., 1968) are not relevant to our theory. This is because these experiments typically provided subjects with a table showing profit as a function of own price and the prices of rivals. Participants were simply asked to choose a price, and they did not have to figure out for themselves which costs might be relevant for pricing decisions.

An important exception to this approach is a recent paper by Offerman and Potters (2006). In this study, subjects assumed the role of decision makers in a differentiated product duopoly, and they were given information on firms' demand curves and marginal costs. Subjects were placed in one of three treatments: one in which rights to operate in the market were auctioned to the two highest bidders; one in which entry rights were allocated randomly to two subjects who were required to pay an upfront fee[23]; and one in which entry rights were allocated randomly without payment of a fee. Offerman and Potters find that when an entry fee is charged (either via a fixed fee or by means of an auction), prices are significantly higher than the undistorted (symmetric) Nash equilibrium price, but when no entry fee is charged, prices were close to the undistorted equilibrium price. By contrast, in a monopoly market, prices tend to equal the profit-maximizing monopoly price, irrespective of the level or form of the sunk entry fee.

Our theoretical model gives exactly the same predictions. Our theory implies that a monopolist would have no incentive to succumb to the sunk cost fallacy. In a duopoly with differentiated products but with no sunk costs, firms might ultimately benefit from distorting their relevant costs, but because they lack a compelling justification for distortionary behavior, our theory would predict that they would end up charging the undistorted Nash equilibrium price. When sunk costs are present, firms have the opportunity to distort relevant costs upward, and given the demand curves and marginal costs in the Offerman and Potters environment, it makes sense to do so. Indeed, in this case, the fit between our model and Offerman and Potters' quantitative results is remarkably tight. In the calculations in Table 1 in the previous section, we chose parameters so that when $\theta = 0.80$ and $N = 2$, we have an economic environment that is identical to that in Offerman and Potters's duopoly experiments.[24] As shown in Table 1, the equilibrium price in the two-stage game predicted by our model is 78.18, about 30% higher than the undistorted Nash equilibrium price of 60. In Offerman and Potters's fixed entry fee and auction treatments, the average prices were 75.65 and 72.73, respectively.[25]

Offerman and Potters interpret their results as suggesting that entry fees generate tacit collusion. In fact, their explanation itself relies on a bounded rationality assumption as they study finite repetitions of an oligopoly game which has a unique subgame perfect equilibrium. Moreover, their explanation predicts that tacit collusion should also occur in a Bertrand oligopoly with homogeneous products. This does not coincide with our predictions and hence provides a way of testing our theory against theirs.

In conclusion, we note that Offerman and Potters's experiment is not a formal test of our model. Our model requires separate pricing and costing decisions and a specific timing structure in which adjustments in these decisions occur at different rates. Nevertheless, models should ideally provide robust insights that hold beyond the narrow formalism on which they are based. It is in this sense that we view their results as suggestive of the plausibility of the core intuition of our article.

---

[23] Unbeknownst to the subjects, the fixed entry fees were set equal to the entry fees generated in the auction treatment.

[24] If we set $a = 310$, $b = \frac{350}{252}$, and $c = 10$, then when $\theta = 0.80$ and $N = 2$, the firms' demand curves are

$$D_1(p_1, p_2) = 124 - 2p_1 + 1.6p_2 \text{ and } D_2(p_1, p_2) = 124 - 2p_2 + 1.6p_1,$$

exactly as in the Offerman and Potters experiment.

[25] These calculations are based on the data reported in Table 1 in Offerman and Potters.

☐ **Cournot competition and behavioral biases.** The results for Bertrand competition with product differentiation do not transfer to Cournot competition. In contrast to the Bertrand model, firms in the Cournot quantity competition model benefit from becoming more aggressive by producing *higher* quantities. To develop intuition, assume there are just two firms and firm 1 reduces its output as it suffers from a sunk cost bias. Learning will take place via quantity adjustments, so firm 2 will eventually increase its output and firm 1's economic profit will go down. Thus, in contrast to a price-setting industry, the sunk cost bias is not reinforced in the cost adjustments phase and our model suggests that the sunk cost bias should disappear in a Cournot setting. This provides another refutable implication of our theory that can be tested using experimental or field evidence.

In fact, our model suggests that managers may reinforce a bias to *overlook* relevant costs in a Cournot setting. In an experimental study, Phillip, Battalio, and Kogut (1991) found evidence that subjects treat opportunity costs as different from direct monetary outlays. In practice, a manager whose firm has purchased an input under a long-term contract may treat the cost of that input as sunk when, in fact, it is variable if the input can be resold in the marketplace. A similar bias can arise if the contract price of the input is less than the current market price, and the manager computes marginal cost based on the historical, rather than current, market price.

☐ **Alternative explanations: noisy observation of cost, delegation, and evolution.** Cost and price distortions may arise as a result of firms observing their total cost but not its split between fixed, sunk, and variable components. In this case, rational firms' estimates of relevant costs are likely to include non-zero errors similar to the distortions introduced in this article. In this case, a plausible scenario is that firms' errors have zero mean, in which case no systematic distortion in relevant costs, prices, or profits would appear.

It also is instructive to compare our model with models of strategic delegation (Fershtman and Judd, 1987) and strategic commitments in general. Strategic commitments matter only to the extent that they are observable and irreversible. By contrast, firms' internal accounting methodologies and their confusions about relevant costs are neither observable nor irreversible. Firms in our model have myopic incentives and arrive at their decisions through a process of naive learning. They do not actively fine-tune their incentive schemes in anticipation of more favorable second-stage equilibrium outcomes. Our model has testable implications that separates it from the delegation framework. First, the delegation model would predict higher prices regardless of the presence of sunk cost, a prediction that can be experimentally tested. Second, the logic of delegation has no bite when prices are set by the firms' owners, as in experiments where each firm is a single subject (so there is literally no delegation). In our model, whether the owner of the firm is the same as the agent setting prices is irrelevant.

Related ideas appear in the literature on the evolution of preferences. Samuelson (2001) offers an excellent overview of this literature.[26] Closer to our model is the recent work of Heifetz, Shannon, and Spiegel (2004). They offer an interesting model of the evolution of optimism, pessimism, and interdependent preferences in dominance-solvable games. Like strategic delegation, the evolution of preferences literature underscores the value of commitment, but this time using distortions of players' perceptions of their payoffs as a commitment device. The unobservability of firms' costing practices again separates our model from the evolution of preferences approach. For instance, as Dekel, Ely, and Yilankaya (2007) point out, this approach has little bite when players cannot observe their opponents' preferences. And as in the delegation paradigm, sunk cost plays no role in evolutionary arguments; these arguments work equally well in settings where no sunk cost is present.

---

[26] The special issue of the *Journal of Economics Theory*, 2001, vol. 97 is devoted to this line of research.

7. Concluding remarks

■ There is extensive evidence that real-world decision makers violate the predictions of standard economic theory. Among these violations, the sunk bias in pricing decisions stands out on a number of grounds. First, its impact is not limited to small-stakes decisions: pricing is one of the most critical decisions a company can make. Second, unlike other cognitive biases that disappear once the underlying fallacy is explained, distorted pricing seems to thrive despite the relentless efforts of economics and business educators to stamp it out. Third, although many cognitive biases disappear through learning and training, the survey evidence we report provides no indication that this bias is disappearing over time.

In this article, we provided a theory of why confusion of relevant and irrelevant costs persists in pricing practice.[27] We showed that under conditions rooted in actual cost accounting practices, price competition with product differentiation reinforces managers' innate predisposition to confound relevant and irrelevant costs. And although there is no reason to suspect this predisposition to systematically vary across industries, no similar forces appear in monopoly, perfect competition, or price competition in a homogeneous product oligopoly. Thus, our analysis makes predictions that can be empirically and experimentally tested.

Our theory builds on recent advances in learning in games played by naive players who know little about their environments. As such, this article illustrates how learning theory can be a valuable tool in understanding the nature of behavioral biases in economics.

Appendix

■ The proof of Theorem 1 relies on Proposition 1 and the following result.

Proposition 9. There exists $\Delta > 0$ such that for any firm n, any vector of distortions $s_{-n} \geq 0$, any $s_n \leq \Delta$, and any equilibrium price vector p of $\Gamma(s_{-n}, 0)$ with corresponding quantity vector $q > > 0$, if p^+ is an equilibrium of $\Gamma(s_{-n}, s_n)$ with $p^+ > p$,

$$\pi_n^e(p^+) > \pi_n^e(p).$$

Proof of Theorem 1. Let Δ be as in Proposition 9 and choose any $0 < \hat{s}_n \leq \Delta$. Fix any $\bar{s}_{-n}$ and p as in the statement of the theorem. By Proposition 1, any adaptive price adjustment $\{p_t\}$ starting from $\bar{p}$ for $\Gamma(\bar{s}_{-n}, \hat{s}_n)$ converges to $p = \inf\{z \in \mathcal{E}(\bar{s}_{-n}, \hat{s}_n), z \geq \bar{p}\}$. It is easy to see that in fact $p >> \bar{p}$. The result now follows from Proposition 9.

Proof of Proposition 9. Suppose, by way of contradiction, that for every positive integer k, there is $s_n < 1/k$, a vector of distortions s_{-n}^k, an equilibrium $\bar{p}^k$ of $\Gamma(\bar{s}_{-n}^k, 0)$, and an equilibrium $\hat{p}^k$ of $\Gamma(s_{-n}^k, s_n)$ with $\hat{p}^k > \bar{p}^k$ yet $\pi_n^e(\hat{p}^k) \leq \pi_n^e(\bar{p}^k)$. Passing to subsequences if necessary, we may assume that $s_{-n}^k \to s_{-n}$, $\bar{p}^k \to \bar{p}$, and $\hat{p}^k \to \hat{p}$. Clearly, $\hat{p} \geq \bar{p}$ and $\pi_n^e(\hat{p}) \leq \pi_n^e(\bar{p})$. We note that, by Lemma 2 below, $\hat{p}$ is an equilibrium of $\Gamma(s_{-n}, 0)$. We consider two cases.

Case 1. $\hat{p} > \bar{p}$. In this case, the fact that best responses are strictly increasing implies $\hat{p} >> \bar{p}$. Next we show that $\pi_n^e(\hat{p}) > \pi_n^e(\bar{p})$. From our assumptions on demand and the fact that $q > > 0$.

$$\pi_n(\hat{p}_{-n}, \bar{p}_n, s_n = 0) > \pi_n(\bar{p}, s_n = 0) = \pi_n^e(\bar{p}). \tag{A1}$$

That is, relative to $\bar{p}$, firm n achieves higher profit if its opponents strictly increase their prices while all other prices remain unchanged. Because $\hat{p}$ is an equilibrium, we must also have

$$\pi_n^e(\hat{p}) = \pi_n(\hat{p}, s_n = 0) \geq \pi_n(\hat{p}_{-n}, \bar{p}_n, s_n = 0),$$

which combining with (A1) implies $\pi_n^e(\hat{p}) > \pi_n^e(\bar{p})$. By the continuity of π_n^e, for k large enough, $\pi_n^e(\hat{p}^k) > \pi_n^e(\bar{p}^k)$, a contradiction.

---

[27] The only other model we are aware of that studies the effect of the sunk cost bias in differentiated Bertrand competition is by Parayre (1995). He considers a two-stage game in a linear environment similar to the special case of symmetric linear demand we study in Section 5 (in his model, no naive learning justification of behavior is considered). Parayre assumes that the sunk cost bias has the implication that firms produce in order to use up whatever capacity is available to them. Compared to the standard Bertrand equilibrium, Parayre's model leads to *lower* prices and profits. Our model sheds light on the form of distortion that is likely to persist in a competitive environment. Although the form of distortion we propose will be perpetuated in the long run, our model predicts that Parayre's form of distortion must disappear.

Case 2. $\hat{p} = \bar{p}$. For each k, define

$$z^k = \frac{\hat{p}^k - \bar{p}^k}{\|\hat{p}^k - \bar{p}^k\|}.$$

Let S^N denote the unit sphere and, given $\alpha > 0$, define $B_\alpha = S^N \cap \{x^N \in R_+^N, x_n \le 1 - \alpha\}$. We first need the following lemma.

Lemma 1. There is $\alpha > 0$ such that for all large-enough k, $z^k \in B_\alpha$.

From the definition of derivatives, for every $\epsilon > 0$, there is $\bar{h}$ such that for any $0 < h < \bar{h}$ and any $z \in S^N$,

$$\left| \frac{\pi_n^e(\bar{p}^k + hz) - \pi_n^e(\bar{p}^k)}{h} - [\partial \pi_n^e(\bar{p}^k)]z \right| < \epsilon,$$

where $\partial \pi_n^e(\bar{p}^k)$ is a row vector of partial derivatives of π_n^e evaluated at $\bar{p}^k$. Notice that $\frac{\partial \pi_n^e}{\partial p_n}(\bar{p}^k) = 0$ as $\bar{p}^k$ is an equilibrium and $\frac{\partial \pi_n^e}{\partial p_m}(\bar{p}^k) > 0$ as $\frac{\partial D_n}{\partial p_m} > 0$ for $m \ne n$ and $q_m^k > 0$ for all k large enough. Note that $[\partial \pi_n^e(\bar{p})]z > 0$ for all $z \in B_\alpha$.

Next we note that

$$\liminf_{k \to \infty} \min_{\substack{z \in B_\alpha \\ \bar{p}^k \in (s_{-n}^k, 0)}} \left[\partial \pi_n^e(\bar{p}^k) \right] z \ge \min_{\substack{z \in B_\alpha \\ \bar{p} \in (s_{-n}, 0)}} \left[\partial \pi_n^e(\bar{p}) \right] z.$$

To see this, write

$$\min_{\substack{z \in B_\alpha \\ \bar{p}^k \in (s_{-n}^k, 0)}} \left[\partial \pi_n^e(\bar{p}^k) \right] z = \left[\partial \pi_n^e(\check{p}^k) \right] \check{z}^k. {}^{28}$$

For any pair of subsequences $\{\check{p}^l\}$, $\{\check{z}^l\}$ converging to $\check{p}$ and $\check{z}$, respectively, we have $\check{z} \in B_\alpha$ and $\check{p} \in (s_{-n}, 0)$ (the latter holds by Lemma 2). From the continuity of π_n^e we have

$$\min_{\substack{z \in B_\alpha \\ \bar{p}^l \in (s_{-n}, 0)}} \left[\partial \pi_n^e(\bar{p}^l) \right] z = \left[\partial \pi_n^e(\check{p}^l) \right] \check{z} \to \left[\partial \pi_n^e(\check{p}) \right] \check{z} \ge \min_{\substack{z \in B_\alpha \\ \bar{p} \in (s_{-n}, 0)}} \left[\partial \pi_n^e(\bar{p}) \right] z.$$

Fix $\alpha > 0$ and set $\epsilon = 0.25 \min_{\substack{z \in B_\alpha \\ \bar{p} \in (s_{-n}, 0)}} [\partial \pi_n^e(\bar{p})]z$. Then, given any $z \in B_\alpha$ and $\bar{p}^k \in (s_{-n}^k, 0)$, for all k large enough we have

$$\frac{\pi_n^e(\hat{p}^k) - \pi_n^e(\bar{p}^k)}{\|\hat{p}^k - \bar{p}^k\|} \equiv \frac{\pi_n^e(\bar{p}^k + \|\hat{p}^k - \bar{p}^k\|z^k) - \pi_n^e(\bar{p}^k)}{\|\hat{p}^k - \bar{p}^k\|}$$

$$> \left[\partial \pi_n^e(\bar{p}^k) \right] z^k - \epsilon$$

$$\ge \min_{\substack{z \in B_\alpha \\ \bar{p}^k \in (s_{-n}^k, 0)}} \left[\partial \pi_n^e(\bar{p}^k) \right] z - \epsilon$$

$$\ge 0.5 \min_{\substack{z \in B_\alpha \\ \bar{p} \in (s_{-n}, 0)}} \left[\partial \pi_n^e(\bar{p}) \right] z - \epsilon$$

$$\ge 0.25 \min_{\substack{z \in B_\alpha \\ \bar{p} \in (s_{-n}, 0)}} \left[\partial \pi_n^e(\bar{p}) \right] z > 0.$$

Hence $\pi_n^e(\hat{p}^k) - \pi_n^e(\bar{p}^k) > 0$ for all k large enough. A contradiction.

Proof of Lemma 2. Given k, for $m \ne n$

$$\lim_{h \to 0} \frac{r_m(\bar{p}_{-n}, \bar{p}_n + h, s_m) - \bar{p}_m}{h} = \frac{\partial r_m}{\partial p_n}(\bar{p}, s_m).$$

As $\hat{p}^k \to \bar{p}$, $\bar{p}^k \to \bar{p}$, and $s_{-n}^k \to s_{-n}$, because $\frac{\partial r_m}{\partial p_n}$ is continuous, for every $\epsilon > 0$, there is a $\bar{k}$ such that for all $k > \bar{k}$

$$\frac{r_m\left(\hat{p}_{-n}^k, \bar{p}_n^k + \left(\hat{p}_n^k - \bar{p}_n^k \right), s_m^k \right) - \bar{p}_m^k}{\left(\hat{p}_n^k - \bar{p}_n^k \right)} > \frac{\partial r_m}{\partial p_n}(\bar{p}, s_m) - \epsilon \tag{A2}$$

for all $m \ne n$.

Note further that

$$\hat{p}_m^k = r_m\left(\hat{p}_{-n}^k, \hat{p}_n^k, s_m^k \right) = r_m\left(\hat{p}_{-n}^k, \bar{p}_n^k + \left(\hat{p}_n^k - \bar{p}_n^k \right), s_m^k \right)$$

and hence from (A2), for every $\varepsilon > 0$, for all k large enough and all $m \ne n$,

$$\frac{\hat{p}_m^k - \bar{p}_m^k}{\hat{p}_n^k - \bar{p}_n^k} > \frac{\partial r_m}{\partial p_n}(\bar{p}, s_m) - \epsilon > 0$$

if ϵ is small enough.

---

[28] The minimum is achieved by the compactness of B_α and $(s^k_{-n}, 0)$.

To prove the lemma, suppose, by way of contradiction, that $\{z^k\}$ has as an accumulation point the unit vector z with $z'_n = 1$ and all other entries equal to zero. Passing to subsequences if necessary, we may assume that $z^k \to z'$. This implies that $\frac{\hat{p}^k_m - \bar{p}^k_m}{||\hat{p}^k - \bar{p}^k||} \to 0$ for every $m \neq n$, and $\frac{\hat{p}^k_n - \bar{p}^k_n}{||\hat{p}^k - \bar{p}^k||} \to 1$. Thus, given any $\tau > 0$, for ϵ small and for all k large enough, $|\frac{\hat{p}^k_n - \bar{p}^k_n}{||\hat{p}^k - \bar{p}^k||} - 1| < \tau$, hence $||\hat{p}^k - \bar{p}^k|| < \frac{\hat{p}^k_n - \bar{p}^k_n}{1 - \tau}$. Then for $m \neq n$,

$$\| z^k - z' \| \geq \frac{\hat{p}^k_m - \bar{p}^k_m}{\| \hat{p}^k - \bar{p}^k \|} \geq \frac{\hat{p}^k_m - \bar{p}^k_m}{\dfrac{p^k_n - \bar{p}_n}{1 - \tau}} > (1 - \tau)\left(\frac{\partial r_m}{\partial p_n}(\bar{p}, s_m) - \epsilon\right) > 0,$$

contradicting the assumption that $z^k \to z'$.

Lemma 2. Suppose that $s^k \to s$, and $p^k \to p$ where for each k, p^k is an equilibrium of $\Gamma(s^k)$. Then p is an equilibrium of $\Gamma(s)$.

Proof. If p were not an equilibrium of $\Gamma(s)$, then there would be a firm m and a strategy p'_m such that

$$\pi_m(p_{-m}, p'_m, s_m) > \pi_m(p_{-m}, p_m, s_m).$$

Because π_m is continuous in p and s_m, for all-large enough k,

$$\pi_m\left(p^k_{-m}, p'_m, s^k_m\right) > \pi_m\left(p^k_{-m}, p^k_m, s^k_m\right).$$

This contradicts the fact that p^k is an equilibrium of $\Gamma(s^k)$ for all k.

Lemma 3. For any $s_{-n}, s_n \geq \Delta$, and $p \in \mathcal{E}(s)$,

$$p_0 \equiv \sup\{z \in \mathcal{E}(s_{-n}, 0), z \leq p\} << \inf\{z \in \mathcal{E}(s_{-n}, \Delta), z \geq p_0\} \leq \sup\{z \in \mathcal{E}(s_{-n}, \Delta), z \leq p\}.$$

Proof. First, note that, because best responses are strictly increasing in opponents' prices and cost, $p_0 << p$. Define a sequence of price vectors $\{p_t\}_{t=0}^{\infty}$ such that $p_0 = \sup\{z \in \mathcal{E}(s_{-n}, 0), z \leq p\}$, and for $t > 0$,

$$p_t = (r_m(p_{t-1}, s_m), \quad m \neq n, r_n(p_{t-1}, \Delta)).$$

We show by induction that $p_t << p$ for all t. The claim holds by definition for $t = 0$. Assume that it holds for $t \geq 0$, then because $p_{t,-m} << p_{-m}$ for all m, and best responses are strictly increasing in opponents' prices, $r_m(p_t, s_m) < r_m(p, s_m)$. A similar argument shows that $p_t >> P_0$ for all $t > 0$. Because $[p_0, p]$ is compact, passing to subsequences if necessary, we may assume that $p_t \to P_\Delta \in [P_0, P]$. Because $\{p_t\}$ is an adaptive pricing adjustment starting at p_0, by Proposition 1, $p = \inf\{z \in \mathcal{E}(s_{-n}, \Delta), z \geq p_0\}$. Clearly, $p_\Delta \leq \sup\{z \in \mathcal{E}(s_{-n}, \Delta), z \leq p\}$.

Lemma 4. There is $\phi > 0$ such that for any s and $p \in \mathcal{E}(s)$,

$$\Pi_n(s_{-n}, \Delta, p) - \Pi_n(s_{-n}, 0, p) > \phi.$$

Proof. Consider any $s', p' \in \mathcal{E}(s')$. Assume first that $\hat{s}_n > \Delta$. By Proposition 1,

$$\Pi_n(s'_{-n}, \Delta, p') - \Pi_n(s'_{-n}, 0, p') = \pi^e_n(\sup\{z \in \mathcal{E}(s'_{-n}, \Delta), z \leq p'\})$$

$$-\pi^e_n(\sup\{z \in \mathcal{E}(s'_{-n}, 0), z \leq p'\}).$$

Combining Lemma 3 and Proposition 9, this expression is strictly positive for every $s', p' \in \mathcal{E}(s')$. The conclusion now follows from the facts that s' ranges over a finite set, $\mathcal{E}(s')$ is compact for every s', and π^e_n is continuous. The remaining cases $s'_n = 0, \Delta$ can be dealt with similarly.

Proof of Theorem 2. Choose Δ as in Proposition 9. Fix an arbitrary firm n and define $L_\tau(s)$ to be the sum of payoffs in all periods $\tau' < \tau$ such that τ' was an experimental period and the distortion was chosen. For any $0 < \delta < \Delta$ define

$$X_\tau = L_\tau(\delta) - L_\tau(0).$$

Because experiments are independent across periods and all distortions are picked with equal probability, the strong law of large numbers implies that, almost surely, for all τ large enough, all distortions are played with equal frequencies. Thus, for all large nonexperimental periods τ, $X_\tau > 0$ implies that firm n does not choose $s_n = 0$. Thus, to prove the theorem, it suffices to show that $X_\tau \to \infty$ almost surely.

Define

$$x_\tau = [L_\tau(\delta) - L_\tau(0)] - [L_{\tau-1}(\delta) - L_{\tau-1}(0)]$$

$$y_\tau = x_\tau - E(x_\tau \mid h_{\tau-1}), \qquad z_\tau = E(x_\tau \mid h_{\tau-1})$$

$$Y_\tau = \sum_{s=1}^{\tau} y_s, \qquad Z_\tau = \sum_{s=1}^{\tau} z_s.$$

Clearly,

$$X_\tau = Y_\tau + Z_\tau,^{29} \tag{A3}$$

and $\{Y_\tau\}$ is a martingale.[30] Because each Y_τ has bounded support, by Theorem 6.8.5 in Ash and Doleans-Dade (2000), the random variables $Y_\tau - Y_{\tau-1} = y_\tau$ are orthogonal, and by Theorem 5.1.2 Chung (1974), $\frac{Y_\tau}{\tau} = \frac{\sum_{s=1}^{\infty} y_s}{\tau}, \to 0$ almost surely (i.e. the conclusion of the law of large numbers holds). From this and equation (A3), it follows that

$$\frac{X_\tau - Z_\tau}{\tau} \to 0 \quad \text{a.s.} \tag{A4}$$

On the other hand,

$$z_\tau = E([L_\tau(\delta) - L_\tau(0)] - [L_{\tau-1}(\delta) - L_{\tau-1}(0)] \mid h_{\tau-1})$$

$$= E(\Pi_n(s_{-n}(\tau), \delta, p(\tau-1)) - \Pi_n(s_{-n}(\tau), 0, p(\tau-1)) \mid h_{\tau-1})$$

$$\geq \frac{1}{N} \frac{\epsilon}{K} \phi > 0,$$

where the last two inequalities follow from Lemma 4, and K is the number of distortions in the grid. This implies that $\frac{Z_\tau}{\tau} \geq \frac{1}{N} \frac{\epsilon}{K} \phi$. This fact and equation (A4) imply $X_\tau \to \infty$ almost surely.

Proof of Proposition 2. In any equilibrium of the pricing game, the lowest price offered must be at or just below the second-highest cost including any distortion. Suppose this is not the case and let p^* be the lowest-price. Then, of the two lowest cost firms, at least one is not capturing the entire market. But that firm can price just under p^* and capture the entire market and increase its accounting profit, a contradiction. Also, in any equilibrium, the firm with the lowest cost including distortions must be setting the lowest price. Otherwise, it is making zero profits and can deviate and make positive accounting profit by matching the lowest price. If its cost including distortion is lower than the second-highest cost, it will price slightly below the second-highest cost to capture the entire market. This observation implies that in any equilibrium with no distortions, the lowest price must be at or just below c_2. Also, firm 1 must be setting the lowest price and either capturing the entire market, if $c_2 > c_1$, or sharing it, if $c_1 = c_2$. In either case, firms 2 and above make zero profits and firm 1 makes positive profits only if $c_2 > c_1$.

Suppose firm 2 or higher distorts its relevant costs upward by s. Then, let c^* be the lowest cost including the distortion among all the firms apart from firm 1. By our argument above, any equilibrium of the pricing game must have the price set at or just below c^*. Therefore, firms 2 and higher make zero profits in equilibrium and hence have no incentive to distort their relevant costs upward. Now, suppose firm 1 distorts upward by s. If $c_1 + s < c_2$, then there is no change in the equilibrium in the pricing game and hence no incentive to distort. If $c_1 + s > c_2$, in any equilibrium in the pricing game, firm 1 must make zero profits as firm 2 has the lowest cost and will always undercut the price set by the firm with the second-highest cost including any distortion. Hence, firm 1 has no incentive to distort relevant costs upward. Moreover, there is no incentive for firm 1 to distort downward as it does not alter the pricing equilibrium. Similarly, if firm $n > 1$ distorts downward by s it still makes zero profits, when $c_n - s \geq c_1$, or makes a loss, when $c_n - s < c_1$, and it sets a price just under c_1 which is less than its true cost c_n. ∎

Proof of Proposition 3. We establish existence by showing that $\bar{\pi}_n^e(s)$ is supermodular in s (Milgrom and Roberts, 1990). Differentiating $\bar{\pi}_n^e(s)$ with respect to s_n we have

$$\frac{\partial \bar{\pi}_n^e(s)}{\partial s_n} = \left[-\beta (\bar{p}_n(s) - c) + \alpha - \beta \bar{p}_n(s) + \gamma \sum_{m \neq n} \bar{p}_m(s) \right] \frac{\partial \bar{p}_n}{\partial s_n} + (\bar{p}_n(s) - c) \gamma \sum_{m \neq n} \frac{\partial \bar{p}_m}{\partial s_n}.$$

Using (3), we can write this as

$$\frac{\partial \bar{\pi}_n^e(s)}{\partial s_n} = -\beta \frac{\partial \bar{p}_n}{\partial s_n} s_n + (N-1)\gamma \frac{\partial \bar{p}_m}{\partial s_n} (\bar{p}_n(s) - c), \tag{A5}$$

where $\frac{\partial \bar{p}_n}{\partial s_n}$ and $\frac{\partial \bar{p}_n}{\partial s_m}$ are given by (5) and (6), respectively. This implies

$$\frac{\partial^2 \bar{\pi}_n^e(s)}{\partial s_n \partial s_m} = (N-1)\gamma \left(\frac{\partial \bar{p}_n}{\partial s_m} \right) \left(\frac{\partial \bar{p}_m}{\partial s_n} \right) > 0.$$

---

[29] This is the well-known Doob decomposition. See, for instance, Chung, (1974).
[30] $E(Y_{\tau+1} \mid Y_\tau) = E(y_{\tau+1} + Y_\tau \mid Y_\tau) = E(x_{\tau+1} \mid Y_\tau) - E(E(x_{\tau+1} \mid Y_\tau) \mid Y_\tau) + Y_\tau = Y_\tau.$

Thus, the first-stage game is supermodular in s. To establish uniqueness we establish the contraction condition $\Lambda \equiv \frac{\partial^2 \bar{\pi}_n^e(s)}{\partial s_n^2} + \sum_{n \neq m} \frac{\partial^2 \bar{\pi}_n^e(s)}{\partial s_n \partial s_m} < 0$ for all $s \geq 0$. We first note that

$$\frac{\partial^2 \bar{\pi}_n^e(s)}{\partial s_n^2} = \frac{\partial \bar{p}_n}{\partial s_n} \left[-\beta + (N-1)\gamma \frac{\partial \bar{p}_m}{\partial s_n} \right].$$

Thus, we can write

$$\Lambda = \frac{\partial \bar{p}_n}{\partial s_n} \left\{ -\beta + \left[(N-1)\gamma + (N-1)^2 \gamma \xi \right] \frac{\partial \bar{p}_m}{\partial s_n} \right\},$$

where $\xi \equiv \frac{\frac{\partial \bar{p}_m}{\partial s_m}}{\frac{\partial \bar{p}_n}{\partial s_n}} = \frac{\frac{\gamma}{2\beta - (N-1)\gamma}}{1 + \frac{\gamma}{2\beta - (N-1)\gamma}} \in (0, 1)$. because $\frac{\partial \bar{p}_n}{\partial s_n} > 0$ and $\frac{\partial \bar{p}_m}{\partial s_n} = \frac{\beta}{2\beta + \gamma} \frac{\gamma}{2\beta - (N-1)\gamma}$, $\Lambda < 0$ if and only if

$$[(N-1) + (N-1)^2 \xi] \frac{\gamma^2}{(2\beta + \gamma)(2\beta - (N-1)\gamma)} < 1. \tag{A6}$$

Now, straightforward algebra establishes that

$$\frac{\gamma^2}{(2\beta + \gamma)(2\beta - (N-1)\gamma)} = \left(\frac{\theta}{(N-1)\theta + 2(1-\theta)} \right) \left(\frac{\theta}{(2N-1)\theta + 2(1-\theta)} \right).$$

Furthermore $\xi < 1$, so $(N-1) + (N-1)^2 \xi < (N-1) + (N-1)^2 = N(N-1)$. Thus

$$[(N-1) + (N-1)^2 \xi] \frac{\gamma^2}{(2\beta + \gamma)(2\beta - (N-1)\gamma)}$$

$$< \left(\frac{(N-1)\theta}{(N-1)\theta + 2(1-\theta)} \right) \left(\frac{N\theta}{(2N-1)\theta + 2(1-\theta)} \right).$$

But $N < 2N - 1$, because $N > 1$. Thus

$$\left(\frac{(N-1)\theta}{(N-1)\theta + 2(1-\theta)} \right) \left(\frac{N\theta}{(2N-1)\theta + 2(1-\theta)} \right)$$

$$< \left(\frac{(N-1)\theta}{(N-1)\theta + 2(1-\theta)} \right) \left(\frac{(2N-1)\theta}{(2N-1)\theta + 2(1-\theta)} \right).$$

Because each term in the above expression is positive but less than one, we immediately have

$$\left(\frac{(N-1)\theta}{(N-1)\theta + 2(1-\theta)} \right) \left(\frac{(2N-1)\theta}{(2N-1)\theta + 2(1-\theta)} \right) < 1,$$

which establishes that $\Lambda < 0$. The equilibrium in s_n is thus unique, and because firms are symmetric, the equilibrium distortion will be symmetric across firms.

Proof of Proposition 4.

Proof (i). We begin by noting that $\theta = 0 \Rightarrow \gamma = 0$ and $\beta = \frac{1}{b}$. Using this, it is straightforward to verify that $\lim_{\theta \to 0} \frac{\partial \bar{\pi}_n^e}{\partial s_n} = -\frac{s_n}{2b} < 0$. Hence, no firm has an incentive to choose a positive distortion.

Proof (ii). It is straightforward to show that

$$\lim_{\theta \to 1} p^0 = c$$

$$\lim_{\theta \to 1} \frac{\gamma}{\beta} = \frac{1}{N-1}$$

$$\lim_{\theta \to 1} \frac{\beta}{2\beta - (N-1)\gamma} = 1$$

$$\lim_{\theta \to 1} \xi = \frac{1}{N}.$$

Now let $s^*(1) = \lim_{\theta \to 1} s^*$ and $p^*(1) = \lim_{\theta \to 1} p^*$. Taking limits of each side of (8) and (9), and using the expressions above, we have

$$s^*(1) = \frac{(p^*(1) - c)}{N}$$
$$p^*(1) = c + s^*(1).$$

These can only be satisfied if $s^*(1) = 0$ and $p^*(1) = c$.

Proof (iii). It is straightforward to show that

$$\lim_{N\to\infty} p^0 = c$$

$$\lim_{N\to\infty} (N-1)\xi = 1$$

$$\lim_{N\to\infty} \frac{\gamma}{\beta} = 0.$$

Now let $s^*(\infty) = \lim_{N\to\infty} s^*$ and $p^*(\infty) = \lim_{N\to\infty} p^*$. Taking limits of (8) and using the above expressions implies that $s^*(\infty) = 0$.

Proof of Proposition 5. If $\theta \in (0,1)$, then $p^0 > c$. Furthermore, it is straightforward to show that

$$\frac{(N-1)\theta^2}{[(2-\theta)+(N-1)\theta(1-\theta)][(1-\theta)+(N-1)\theta]} > 0,$$

which implies that $s^* > 0$. When $\theta \in (0,1)$,

$$\frac{(2-\theta)+(N-1)\theta}{(2-\theta)+(N-1)\theta(1-\theta)} > 1.$$

Thus, $p^* - c > p^0 - c$, or $p^* > p^0$.

Proof of Proposition 6. The expression for $p^* - c$ can be written as follows:

$$p^* - c = (a-c)(1-\theta)y(N)z(N),$$

where

$$y(N) = \frac{1}{2-\theta+(N-1)\theta(1-\theta)}$$

$$z(N) = \frac{2-\theta+(N-1)\theta}{2(1-\theta)+(N-1)\theta}.$$

Clearly $y(N)$ is strictly decreasing in N and so too is $z(N)$:

$$z'(N) = -\frac{\theta^2}{[2(1-\theta)+(N-1)\theta]^2} < 0.$$

Hence $p^* - c$ is strictly decreasing in N.

Proof of Proposition 7. The monopoly price p^M satisfies

$$p^M - c = \frac{a-c}{2}.$$

From (10), $p^* - c$ can be written as

$$p^* - c = \frac{(a-c)(1-\theta)[2-\theta+(N-1)\theta)]}{[2-\theta+(N-1)\theta(1-\theta)][2(1-\theta)+(N-1)\theta]}.$$

$p^* < p^M$ if and only if

$$\frac{(1-\theta)[2-\theta+(N-1)\theta]}{[2-\theta+(N-1)\theta(1-\theta)][2(1-\theta)+(N-1)\theta]} < \frac{1}{2}. \tag{A7}$$

This can be rewritten as

$$[2-\theta+(N-1)\theta(1-\theta)][2(1-\theta)+(N-1)\theta] > 2(1-\theta)[2-\theta+(N-1)\theta].$$

Now, for $N \geq 3$, $[2(1-\theta)+(N-1)\theta] \geq 2$ because $\theta \in (0,1)$ and $N-1 \geq 2$. Hence, we have

$$[2-\theta+(N-1)\theta(1-\theta)][2(1-\theta)+(N-1)\theta] > 2[2-\theta+(N-1)\theta(1-\theta)]$$

$$> 2(1-\theta)[2-\theta+(N-1)\theta].$$

For $N = 2$, the inequality (A7) boils down to

$$\frac{1-\theta}{2-\theta} < \frac{2-\theta^2}{4},$$

which can be shown to hold for $\theta \in (0,1)$. Hence, (A7) holds for all $N \geq 2$, establishing $p^* < p^M$.

Proof of Proposition 8. These results follow directly from expression (11).

References

ALCHIAN, A. "Uncertainty, Evolution, and Economic Theory." *Journal of Political Economy,* Vol. 58 (1950), pp. 211–221.

ARKES, H.R. AND AYTON, P. "The Sunk Cost and Concorde Effects: Are Humans Less Rational Than Lower Animals?" *Psychological Bulletin,* Vol. 125 (1999), pp. 591–600.

—— AND BLUMER, C. "The Psychology of Sunk Cost." *Organizational Behavior and Human Decision Processes,* Vol. 35 (1991), pp. 124–140.

ASH, R. AND DOLEANS-DADE, C. *Probability and Measure Theory*, 2d ed. New York: Harcourt/Academic Press, 2000.

BROWN, D., BOOTH, P., AND GIACOBBE, F. "Technological and Organizational Influences on the Adoption of Activity-Based Costing in Australia." *Accounting and Finance,* Vol. 44 (2004), pp. 329–356.

CHUNG, K.L. *A Course in Probability Theory*, 3d ed. New York: Academic Press, 1974.

DEKEL, E., ELY, J.C., AND YILANKAYA, O. "Evolution of Preferences." *Review of Economic Studies,* Vol. 74 (2007), pp. 685–704.

DOLBEAR, F., LAVE, L., BOWMAN, G., LIEBERMAN, A., PRESCOTT, E., REUTER, F., AND SHERMAN, R. "Collusion in Oligopoly." *Quarterly Journal of Economics,* Vol. 82 (1968), pp. 240–259.

ECHENIQUE, F. "Comparative Statics by Adaptive Dynamics and the Correspondence Principle." *Econometrica,* Vol. 70 (2002), pp. 833–844.

FERSHTMAN, C. AND JUDD, K. "Equilibrium Incentives in Oligopoly." *American Economic Review,* Vol. 77 (1987), pp. 927–940.

FINANCIAL ACCOUNTING STANDARDS BOARD. "Qualitative Characteristics of Accounting Information." Statement of Financial Accounting Concepts, No. 2, May 1980, http://www.fasb.org/st/.

FRIEDMAN, J.W. AND MEZZETTI, C. "Learning in Games by Random Sampling." *Journal of Economic Theory,* Vol. 98 (2001), pp. 55–84.

GOVINDARAJAN, V. AND ANTHONY, R. "How Firms Use Cost Data in Pricing Decisions." *Management Accounting,* Vol. 76 (1995), pp. 37–39.

HART, S. AND MAS-COLELL, A. "A Simple Adaptive Procedure Leading to Correlated Equilibrium." *Econometrica,* Vol. 68 (2000), pp. 1127–1150.

HEIFETZ, A., SHANNON, C., AND SPIEGEL, Y. "The Dynamic Evolution of Preferences." *Economic Theory*, Vol. 33 (2007) pp. 251–286.

HORNGREN, C., FOSTER, G., AND DATAR, S. *Cost Accounting*, 9th ed. Upper Saddle River, NJ: Prentice Hall, 2000.

KAPLAN, R. AND ATKINSON, A. *Advanced Management Accounting*, 2d ed. Englewood Cliffs, NJ: Prentice-Hall, 1989.

MAHER, M., STICKNEY, C., AND WEIL, R.L. *Managerial Accounting*, 8th ed. Mason, OH: South-Western, 2004.

MILGROM, P. AND ROBERTS, J. "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities." *Econometrica,* Vol. 58 (1990), pp. 1255–1277.

—— AND ——. "Adaptive and Sophisticated Learning in Normal Form Games." *Games and Economic Behavior,* Vol. 3 (1991), pp. 82–100.

OFFERMAN, T. AND POTTERS, J. "Does Auctioning of Entry Licenses Induce Collusion? An Experimental Study." *Review of Economic Studies,* Vol. 73 (2006), pp. 769–791.

PARAYRE, R. "The Strategic Implications of Sunk Costs: A Behavioral Perspective." *Journal of Economic Behavior and Organization,* Vol. 28 (1995), pp. 417–442.

PHILLIP, O., BATTALIO, R., AND KOGUT, C. "Sunk and Opportunity Costs in Valuation and Bidding." *Southern Economic Journal,* Vol. 58 (1991), pp. 112–128.

ROZTOCKI, N. AND SCHULTZ, S. "Adoption and Implementation of Activity-Based Costing: A Web-Based Survey." *Proceedings of the 12th Annual Industrial Engineering Research Conference (IERC 2003)*, Portland, OR, May 18–20, 2003.

SAMUELSON, L. "Introduction to Evolution of Preferences." *Journal of Economic Theory,* Vol. 97 (2001), pp. 225–230.

SHANK, J. AND GOVINDARAJAN, V. *Strategic Cost Analysis*. Burr Ridge, IL: Irwin, 1989.

SHIM, E. "Cost Management and Activity-Based Costing in a New Manufacturing Environment." Unpublished Dissertation, Rutgers University, 1993.

—— AND SUDIT, E. "How Manufacturers Price Products." *Management Accounting,* Vol. 76 (1995), pp. 37–39.

SMITH, V. "Rational Choice: The Contrast between Economics and Psychology." In *Bargaining in Market Experiments*, ed. V. Smith, Cambridge UK: Cambridge University Press, 2000.

THALER, R. "Towards a Positive Theory of Consumer Choice." *Journal of Economic Behavior and Organization,* Vol. 1 (1980), pp. 39–60.

VIVES, X. *Oligopoly Pricing*. Cambridge, MA: MIT Press, 1999.