The Scandinavian
Journal of Economics

Scand. J. of Economics 000(0), n/a, 2024
DOI: 10.1111/sjoe.12564

Monopoly pricing with unknown demand

Thomas A. Weber

École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland
thomas.weber@epfl.ch

Abstract

The optimal pricing of goods, especially when they are new and the innovating firm is a monopolist, must proceed without precise knowledge of the demand curve. This paper provides a pricing method with a relative robustness guarantee by maximizing a performance index which amounts to a worst-case ratio of the obtained payoff to the best possible payoff. Assuming monotonicity and complementarity of demand in price and the unknown demand parameter, the performance index is fully determined by its behavior at the boundary of the parameter space. This allows for an efficient computation of an optimal robust price. In the linear case, which can also be used for nonlinear demand with bounded slope, the method provides a simple closed-form solution. A comparison with the standard worst-case payoff criterion reveals substantial improvements in both absolute and relative performance, at only a small cost relative to the maximized expected profit.

Keywords: Monopoly pricing; relative regret; robust estimation

JEL classification: *C*44; *C*61; *D*21; *D*42

1. Introduction

> Where ignorance is bliss,
> ’Tis folly to be wise. [1]
> Thomas Gray

A monopolist wishing to find a profit-maximizing price needs to know demand and cost as a function of the product (or service) quantity released onto the market. While the firm may quite easily deduce its cost function from the observed bills for the required inputs, demand is usually much more difficult to gauge in advance, especially when the firm’s product is fairly new or price experimentation is costly. To deal with the implied model uncertainty, a perhaps naïve but common approach is for the firm to form beliefs about a variety of possible demands (e.g., a family of linear demand curves) and to then maximize expected returns with respect to those beliefs, effectively substituting an average profit function for the considered family

[1] From *Ode on a Distant Prospect of Eton College* (1747).

of profit functions. Unsurprisingly, the corresponding "model-averaging" price provides no performance guarantee for the firm against demand-curve realizations not compatible with the chosen price (which happens almost surely), leading to suboptimal profits because of either low demand (when the chosen price was too high) or low margin (when the chosen price was too low). Model averaging therefore leaves the firm exposed to a significant risk, which decreases the company's attractiveness to investors and thus increases its cost of capital. Compounding the latter, the firm may have been quite unsure about its beliefs in the first place, an ambiguity that implies a range of model-averaging prices (at least one for each possible belief). Hence, ultimately the naïve approach does provide non-trivial performance guarantees over the full range of possible model realizations.[2] This scenario is especially relevant for innovating firms, who by definition tend to be the first and only ones to sell a new product in a market where demand is only poorly known at the outset. The importance of the price-setting decision is heightened when the firm is committed, at least for some significant time, to not adjust its price.

The key question that we address is, therefore, how to devise an easy-to-use method for determining an "optimal robust price" that performs well relative to an entire family of reasonably well-parametrized demand curves? The relative performance should thereby be evaluated as a guarantee of achieving at least a minimum fraction of the profit that could be attained if the exact model parameter were known. In other words, if the firm implements a robust action x leading to a (relative) "performance index" of $\rho(x)$, say, equal to 80 percent, then the firm should be guaranteed to never lose more than 20 percent of what it could have obtained with perfect information. While our proposed method does not need any important prerequisite (other than the boundedness of all model primitives on compact domains) to produce computational results, a few structural assumptions on the parametrization of the demand curves significantly simplify the representation of the firm's performance index ρ, using the general idea of a Chebyshev extremal base. In this manner, the method could provide simple explicit results, allowing for statements about the sensitivity of the outcomes with respect to the size of the parameter space or, in other words, "the cost of robustness".

A strong motivation for our work is the general observation that human decision-makers tend to respond better to relative comparisons than absolute comparisons (Kahneman and Tversky, 1984). Indeed, absolute numbers in

[2] A "trivial" performance guarantee is obtained by the worst-case performance relative to the chosen price. As this *ex post* performance evaluation depends on the particulars of the parameter space (and not on the firm's *ex ante* beliefs), it might well indicate a very low degree of robustness against the model uncertainty.

isolation rarely make sense. For example, how much absolute profit should the firm really expect from its product, given that demand is not even known? And, should the firm care about a given absolute profit difference of \$100,000? Surely, such a gap would be close to insignificant if the firm's expected profit was of the order of \$10 million, whereas it would be rather an important deviation if this expectation was of the order of \$1 million. Indeed, a relative performance guarantee would limit absolute deviations only when they matter in comparison to what could be achieved using better information. The general argument is mirrored in the comparison of net present value (NPV) against internal rate of return (IRR) as a decision criterion, although they can in fact be considered equivalent (Weber, 2014).

In this paper, we present a robust approach to the firm's pricing problem, which provides a relative performance guarantee with respect to a given parametrized family of downward-sloping demand functions, as long as the parametrization satisfies two properties:

D1 monotonicity – in addition to being downward-sloping in price x, demand $D(x, \theta)$ increases in the chosen parameter $\theta \in \Theta = [\theta_1, \theta_2] \subset \mathbb{R}_+$;

D2 complementarity – for any fixed price decrease from x to $\widehat{x} < x$, the demand response $D(\widehat{x}, \theta) - D(x, \theta)$ is non-increasing in $\theta \in \Theta$.

Properties D1 and D2 together ensure that the demand elasticity sweeps in a constant direction as θ goes up. In addition, besides inducing supermodularity of the firm's profit in (x, θ), they also imply monotonicity of the optimal price $x(\theta)$. To find a price that does relatively well under all possible scenarios, the firm considers the performance ratio $\varphi(\widehat{\theta}, \theta)$ as the ratio of its profit for a price $x(\widehat{\theta})$ (at any parameter $\widehat{\theta}$) to its profit for the best possible price $x(\theta)$ (at the true parameter θ). We show that the performance ratio is single-peaked in the candidate parameter $\widehat{\theta} \in \Theta$. This in turn implies that the firm's relative performance index $\rho(\widehat{\theta})$, which quantifies its worst performance ratio $\varphi(\widehat{\theta}, \theta)$ over all possible θ relative to a given $\widehat{\theta}$, can be represented by limiting attention to "boundary" performance ratios, namely $\varphi(\widehat{\theta}, \theta_1)$ and $\varphi(\widehat{\theta}, \theta_2)$. The corresponding "envelope representation" of the firm's performance index can then be used to easily determine the optimal robust parameter $\widehat{\theta}^*$ and an optimal robust price $\widehat{x}^* = x(\widehat{\theta}^*)$, which is guaranteed to yield a performance index of at least $\rho^* = \rho(\widehat{\theta}^*)$ relative to the entire family of demand curves. It also serves as basis for the sensitivity analysis of the robustness with respect to changes of the parameter space Θ.

The power of the proposed method is illustrated using the family of affine demand curves and linear cost. As all results in this leading example can be obtained explicitly, our method is ready to be deployed by practitioners. However, the affine case casts a light well beyond its narrow confines, as elements of the affine demand-curve family can be fitted to the various operating points in a set of arbitrary nonlinear downward-sloping demand curves, as long as they are subject to common Lipschitz bounds limiting the slope of any local demand linearization.

The key advantage of the proposed approach over more standard maximin methods is that it provides performance comparability. The relative performance index also uses the entire range of demand curves (i.e., the full parameter space Θ), whereas worst-case robustness restricts attention (by Property D1) to the least-favorable demand curve $D(\cdot, \theta_1)$. Property D2 (naturally satisfied in the linear case) is inessential to the method; its main purpose is to provide a simple and easy-to-interpret representation of the performance index and thus to simplify computations.

The concept of the demand curve goes back to Jenkin (1870) and was popularized by Marshall (1920) (first published in 1890) who also introduced the concept of demand elasticity.[3] Lerner (1934) uses the relative mark-up (now often referred to as the "Lerner index") to measure a firm's monopoly power. Its special significance is that at a profit-maximizing optimal price the firm's relative mark-up is equal to the inverse of the demand elasticity (see, e.g., Tirole, 1988, p. 66). The applicability of this inverse-elasticity rule naturally depends on the availability of a demand curve, usually identified from empirical demand and supply data.[4] Lewis and Sappington (1988) consider a problem of asymmetric information where a regulator uses screening techniques in order to elicit a firm's private information about its demand so as to induce a socially desirable solution, which proves possible as long as the firm's marginal cost is increasing in its output. Segal (2003) proposes revenue-optimal selling mechanisms in a Bayesian setting when buyers' bids can be used to update the seller's beliefs about the distribution of valuations. An interesting special case arises when the seller faces constant marginal cost, in which case a simple posted price becomes optimal.

In our setting, we assume that the firm has only rudimentary distribution-free knowledge about the demand curve it is facing (e.g., a range of possible slopes with respect to price and a range of possible demands at a given price), which may or may not include access to noisy samples of demand at certain prices. In many cases, especially for new products, considerably less information might be available, so that the firm may resort to the demand characteristics of potential product substitutes

[3]Friedman (1949) provides an early authoritative overview of Marshallian demand curves. The first mention of "supply and demand" in the English-speaking literature originated in the mid-18th century (Groenewegen, 2008).

[4]Standard approaches to demand analysis set out to identify a demand system as an exogenous observer, controlling for both demand and supply shocks to handle the endogeneity problem (see, e.g., Deaton, 1986).

as a proxy. To tackle the inherent residual model uncertainty, the firm needs to allow for the possibility that any particular demand curve used to price its products is in fact incorrect. To effectively deal with this, the firm could form a subjective belief over a set of possible demand curves and maximize expected payoffs. This expectation-centric approach, first axiomatized by von Neumann and Morgenstern (1944), as well as Anscombe and Aumann (1963), has the drawback of keeping the firm exposed to poor sample outcomes. To account for such unfavorable contingencies, Wald (1945) suggested a robust approach for dealing with uncertainty by "minimizing the maximum risk", or rather maximizing the worst-case payoff, which leads to distribution-free optimization. This worst-case approach is equivalent to playing a zero-sum game with nature (Milnor, 1951) and leads to conservative (and therefore also costly) decisions. As an alternative, Savage (1951) proposes the logic of minimizing the maximum regret (relative to an *ex post* optimal decision), a decision rule put into an axiomatic framework by Gilboa and Schmeidler (1989). The last two approaches, based on either maximin outcome or minimax regret, have carved their way into "robust" managerial decision-making and operations research (Bell, 1982; Bertsimas et al., 1990). Lim et al. (2012) use the regret criterion to tackle robust portfolio choice, while Perakis and Roels (2010) use it in revenue management. Bergemann and Schlag (2011) deploy both criteria to treat the robust pricing problem in a mechanism-design setting with a principal and an agent of uncertain type; they characterize solutions and show that robustness tends to lower the monopolist's price to cover for poor model realizations (consistent with our results as well). Handel and Misra (2015) consider the pricing of a new product with unknown demand in a two-period setting, in light of the minimax regret decision criterion. In our robust-optimization framework, it turns out that the maximin payoff solution and the minimax regret solution both exhibit poor performance relative to how well the firm could have done, the main reason being that these solutions focus on minimizing absolute losses and are therefore quite insensitive to the upside potential in the pricing problem or else to the possibility of an unattractive outcome without positive profits. It is for this reason that we consider a relative regret criterion, evaluating for each decision the ratio of the obtained payoff to the best possible payoff. Maximizing the performance index, defined by the worst-case performance ratio, yields a relatively robust solution.

The idea of using a relative performance index in the form of a "competitive ratio" to evaluate, for example, the efficiency of offline versus online algorithms goes back to Sleator and Tarjan (1985). This type of relative measure is application-specific and depends on which exact cost components of an algorithm are being considered (Ben-David and Borodin, 1994). Relative regret, which measures the maximum difference to an optimal solution, divided by the (positive) performance of the best solution, has been employed in certain operations-research applications, usually with computational evaluation of the relative regret criterion over defined "scenarios" (i.e., model realizations; see, e.g., Kouvelis and Yu, 1997). More recently, Goel et al. (2009) have introduced "relative fairness" in the form of a relative-regret criterion that can be employed to pinpoint Lorenz-undominated solutions to resource allocation problems with respect to a general class of social welfare functions. Han and Weber (2023) use a relative performance index to provide a robust solution to the two-type screening problem when the beliefs about the consumer types are unknown. Here we rely on a relative performance index for robust and data-driven single-product monopoly pricing decisions. Under fairly weak assumptions on the class of demand functions, which are satisfied by the class of linear (affine) demand functions, we obtain a simple representation of the robustness measure, which also suggests a balancedness condition at the optimum. We further show that as long as demand curves have bounded slopes and marginal costs are constant, the method can be applied to a non-parametric class of demand curves, with performance results that tend to exceed those for the bounding class of linear demand functions.

The paper proceeds as follows. Section 2 introduces the monopolist's pricing problem together with the ambiguity set of the demand model. A few key assumptions on the properties of demand and cost imply monotone comparative statics of the reference solution under full information. Section 3 defines the firm's robust pricing problem and establishes important properties of the performance ratio, which imply an envelope representation of the robust performance index. The latter yields an implicit characterization of the optimal robust price, as well as monotone comparative statics of the optimal robust parameter in the bounds of the ambiguity set. Here we also compare our solution to the standard maximin approach. In Section 4, we apply our results to the robust pricing problem with linear (affine) demand and constant marginal cost, obtaining a closed-form solution. This solution is then extended to a non-parametric class of nonlinear demand curves, and further to a fully data-driven approach. Section 5 concludes.

2. Model

Consider a monopolist on a market with demand $D(x, \theta)$ where $x \in \mathbb{R}_+$ is the firm's price, and θ is an unknown scalar parameter that lies in the compact interval $\Theta = [\theta_1, \theta_2]$ whose bounds are such that $-\infty < \theta_1 < \theta_2 < \infty$. Given an output quantity $q \geq 0$, the firm's cost is given by $C(q)$. The properties of D and C are now introduced in turn, before we describe the solution to the firm's profit-maximization problem.

2.1. Demand

To exclude trivialities, we assume that demand is not always zero. Whenever it does not vanish, the demand function $D : \mathbb{R}_+ \times \Theta \to \mathbb{R}_+$ is twice continuously differentiable, decreasing in x and increasing in θ, so[5]

(Monotonicity)

$$D(x,\theta) > 0 \quad \Rightarrow \quad D_x(x,\theta) < 0 < D_\theta(x,\theta),$$
$$(x,\theta) \in \mathbb{R}_+ \times \Theta. \tag{D1}$$

The "monotonicity" property D1 reflects the fact that demand is downward-sloping in price and increasing in the scalar parameter.[6] Moreover, as long as it is positive, demand is assumed to exhibit non-decreasing differences (supermodularity):[7]

(Complementarity)

$$D(x,\theta) > 0 \quad \Rightarrow \quad D_{x\theta}(x,\theta) \geq 0,$$
$$(x,\theta) \in \mathbb{R}_+ \times \Theta. \tag{D2}$$

The "monotonicity" and "complementarity" properties taken together imply the monotonicity of price elasticity (defined for positive demand),[8]

$$\varepsilon(x,\theta) = -\frac{xD_x(x,\theta)}{D(x,\theta)} > 0, \qquad (x,\theta) \in \mathbb{R}_+ \times \Theta, \qquad D(x,\theta) > 0, \tag{1}$$

in the parameter θ. Intuitively, this means relative demand variations due to a price change go down when the parameter shifts upwards.

---

[5]Partial derivatives are indicated by indices, e.g., $D_x = \partial D/\partial x$.

[6]In the absence of externalities (such as in certain fulfilled-expectations equilibria) and asymmetric information (where price may become a signal), only so-called Giffen goods have a (locally) upward-sloping demand (Marshall, 1920, p. 109), for which there are virtually no convincing empirical examples with the few on record being rather exotic and historic. As a case in point, potatoes during the Great Famine in Ireland in the 1840s, considered as a classical example of a Giffen good, can be explained by intertemporal effects with a normal downward-sloping demand (Rosen, 1999). The assumption that demand is increasing in the parameter θ is without any significant loss of generality. All results could also be written in terms of a parameter $\vartheta = \phi(\theta)$ where $\phi : \Theta \to \mathbb{R}$ is a smooth monotonic transformation.

[7]The supermodularity property, originally introduced by Topkis (1968), is useful for establishing the monotonicity of solutions to optimization problems in its parameters (Milgrom and Shannon, 1994).

[8]The price elasticity, introduced by Marshall (1920), will also be referred to as "demand elasticity".

Proposition 1 (Elasticity sweeping). *When revenue is positive, the price elasticity is decreasing in θ:*

$$xD(x,\theta) > 0 \quad \Rightarrow \quad \varepsilon_\theta(x,\theta) < 0, \tag{2}$$

for all $(x,\theta) \in \mathbb{R}_+ \times \Theta$.

As a consequence of the "elasticity sweeping" result in Proposition 1, for any fixed positive price x the demand elasticity takes on all values in $[\varepsilon(x,\theta_2), \varepsilon(x,\theta_1)]$, as the parameter θ varies from θ_1 to θ_2.

For convenience and realism, we assume that the consumers' willingness-to-pay (WTP) is fundamentally finite (with the positive bound $\bar{x}$), that is,

(Finiteness)

$$\exists\, \bar{x} > 0 : \qquad D(x,\theta) = 0, \qquad (x,\theta) \in [\bar{x}, \infty) \times \Theta. \tag{F}$$

As the resources in any real economy are limited, the "finiteness" property of demand generally applies. The reason to formalize this rather natural requirement as a recognizable property is that it does exclude the family of constant-elasticity demand curves (which also highlights their lack of realism; see footnote 9). The immediate consequence of (F) is that a price beyond $\bar{x}$ can never be better than charging $\bar{x}$, thus allowing the firm to restrict attention to the compact price domain $\mathcal{X} = [0, \bar{x}]$ in its quest to maximize profits.

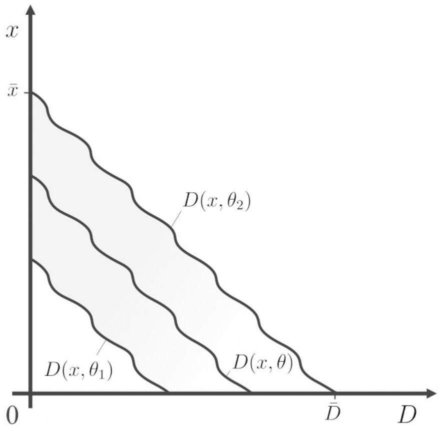
Remark 1 (Boundedness of demand). Because $D(x,\theta)$ is, by assumption, continuous on the compact set $\mathcal{X} \times \Theta$, the smallest upper bound of demand must be attained,[9]

$$\overline{D} = \sup_{(x,\theta) \in \mathcal{X} \times \Theta} D(x,\theta) = D(0,\theta_2), \tag{3}$$

taking into account that by (D1) demand is monotonic in x and θ. The fact that demand remains finite even at zero price excludes unrealistic solutions. Indeed, in the presence of even the smallest positive transaction cost, any agent with non-satiated preferences would want to consume only a finite amount of the product (or service), as long as his or her marginal utility becomes sufficiently small in the limit. Figure 1 shows a family of nonlinear demand curves satisfying (D1), (D2), and (F), featuring a common finite upper bound $\overline{D}$.[10]

[9] Demand continuity rules out certain "standard" demand specifications. For instance, a constant-elasticity demand function implies a diverging D for $x \to 0^+$, although in any real-life situation demand must remain finite.
[10] The family of demand functions used to generate Figures 1–5 is specified in footnote 47.

Figure 1. Unknown demand D for $(x, \theta) \in \mathcal{X} \times \Theta$



2.2. Cost

We assume that the cost function $C : \mathbb{R}_+ \to \mathbb{R}_+$ is twice continuously differentiable, increasing, and convex, so

(Marginal-cost monotonicity)

$$C'(q) \geq 0, \qquad C''(q) \geq 0, \qquad q \in \mathbb{R}_+. \tag{C1}$$

The "marginal-cost monotonicity" property C1 states that the slope of the firm's cost is non-negative and non-decreasing. The implied cost convexity captures a certain complexity cost that prevents a merely sublinear growth of the firm's expenditures in its output. In addition, we neglect the firm's fixed cost, so

(Possibility of inaction)

$$C(0) = 0, \tag{C2}$$

effectively allowing for the "possibility of inaction", at zero output and zero cost. This means that we decouple the firm's pricing decision from the firm's viability and thus from any type of unmodeled entry or exit problem. The firm simply tries to attain the best possible profit without being concerned by an otherwise sunk fixed cost.[11] The preceding two relations together imply that the firm's marginal cost,

$$MC(q) = C'(q), \qquad q \in \mathbb{R}_+, \tag{4}$$

[11]For a ground-breaking discussion of the behavioral "sunk cost fallacy", see Arkes and Blumer (1985).

exceeds its average cost,

$$AC(q) = \begin{cases} C(q)/q, & \text{if } q > 0, \\ C'(0), & \text{if } q = 0, \end{cases} \tag{5}$$

at least weakly.[12] In particular, this means that the company's minimum efficient scale vanishes, with a long-run average cost that is never less than its marginal cost.

Proposition 2 (Zero minimum efficient scale). *The firm's marginal cost (weakly) exceeds its average cost:*

$$MC(q) \geq AC(q), \qquad q \in \mathbb{R}_+. \tag{6}$$

Without any loss of generality, we assume that the upper price bound $\bar{x}$ in (F) weakly exceeds the firm's minimum marginal cost,

(Cost-plus pricing)

$$\bar{x} \geq MC(0), \tag{F'}$$

evaluated at the origin (which is also the lowest production threshold). Condition (F') can be imposed without any loss of generality. The main purpose of choosing a large enough (but finite) upper price bound $\bar{x}$ is to ensure the feasibility of "cost-plus pricing", at least at the point where according to (F) the price is so high that demand must always vanish.[13]

2.3. Profit maximization

Given full knowledge of the demand parameter θ, the firm's profit is

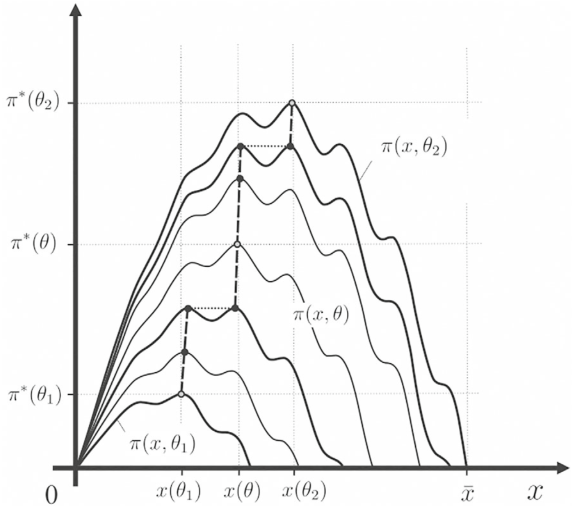
$$\pi(x, \theta) = xD(x, \theta) - C(D(x, \theta)), \qquad (x, \theta) \in \mathbb{R}_+ \times \Theta.$$

As a function of the price x, this objective function is "coercive" in the sense that it is never optimal to charge negative prices or prices above the maximum price $\bar{x}$ in (F) that would lead to zero demand. In order to illustrate this point, Figure 2 shows a (non-concave) profit $\pi(\cdot, \theta)$, for $\theta \in \Theta$, using the demand depicted in Figure 1 and assuming constant marginal cost. The profit vanishes for prices outside $[0, \bar{x}]$, and its maximum $\pi^*(\theta) = \pi(x(\theta), \theta)$ is attained at an interior point $x(\theta) \in (0, \bar{x})$.

[12] The average cost has a continuous completion at $q = 0$, as $\lim_{q \to 0^+} C(q)/q = C'(0)$.

[13] As there are typically no upper bounds on the asking price for a product, an exception may arise when regulatory intervention through price caps renders cost-plus pricing infeasible, as in the California electricity crisis (Sweeney, 2002).

Figure 2. Profit π for $(x, \theta) \in \mathcal{X} \times \Theta$, with maximizer $x(\theta) \in \mathscr{X}(\theta)$



Thus, taking into account the fact that consumers' WTP is bounded and cost-plus pricing is possible, captured by (F) and (F'), the firm can restrict attention to "reasonable prices" in the interval $\mathcal{X} = [0, \overline{x}]$ in its search for the set $\mathscr{X}(\theta)$ of optimal prices that would solve its profit-maximization problem,

$$\mathscr{X}(\theta) = \arg\max_{x \in \mathcal{X}} \pi(x, \theta), \qquad \theta \in \Theta. \tag{7}$$

Given any "selector" $x(\theta) \in \mathscr{X}(\theta)$, the firm's optimal profit is denoted by

$$\pi^*(\theta) = \pi(x(\theta), \theta), \qquad \theta \in \Theta. \tag{8}$$

The first-order necessary optimality condition (Fermat's rule; see, e.g., Aubin and Ekeland, 1984, p. 159) is[14]

$$x \in \mathscr{X}(\theta) \quad \Rightarrow \quad \pi_x(x, \theta) = 0.$$

Meanwhile, the gradient of the firm's profit with respect to price is

$$\pi_x(x, \theta) = D(x, \theta) + (x - MC(D(x, \theta)))\, D_x(x, \theta),$$

[14]For a characterization of all global optima of $\pi(\cdot, \theta)$ on the interval $\mathcal{X}$ given any $\theta \in \Theta$, see Weber (2017). As established in the proof of Proposition 3, prices at the boundary of $\mathcal{X}$ cannot be optimal, so $\mathscr{X}(\theta) = \mathscr{X}(\theta) \setminus \{0, \overline{x}\}$.

which implies the well-known inverse-elasticity rule (see, e.g., Tirole, 1988, p. 66),

$$x \in \mathscr{X}(\theta) \quad \Rightarrow \quad \mu(x,\theta) = \frac{1}{\varepsilon(x,\theta)}, \tag{9}$$

where the firm's relative markup,

$$\mu(x,\theta) = \frac{x - MC(D(x,\theta))}{x}, \qquad (x,\theta) \in (0,\bar{x}) \times \Theta, \tag{10}$$

is also referred to as "Lerner index" of monopoly power (Lerner, 1934). Hence, to solve the profit-maximization problem (7) more effectively, the monopolist can further narrow its choice set by considering only prices x in the rational choice set,

$$\mathcal{X}_0(\theta) = \{x \in \mathcal{X} : x \geq MC(D(x,\theta))\}, \qquad \theta \in \Theta; \tag{11}$$

by construction, these prices (weakly) exceed their induced marginal costs.[15]

Proposition 3 (Properties of $\mathcal{X}_0$). *Let $\theta \in \Theta$ be any given demand parameter.*

(i) The rational choice set is a compact interval of the form $\mathcal{X}_0(\theta) = [x_0(\theta), \bar{x}]$, with its lower bound implicitly defined by

$$x_0(\theta) = MC(D(x_0(\theta),\theta)) \in [MC(0), \bar{x}]. \tag{12}$$

(ii) Provided the optimal profit is positive, the interior of the rational choice set contains the set of profit-maximizing prices:

$$\pi^*(\theta) > 0 \quad \Rightarrow \quad \mathscr{X}(\theta) \subset \operatorname{int} \mathcal{X}_0(\theta). \tag{13}$$

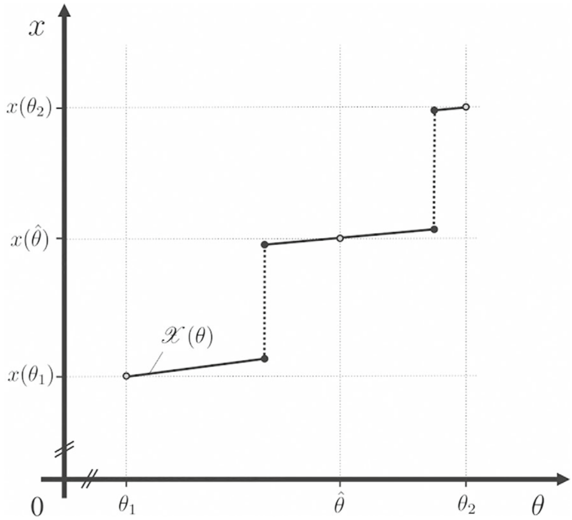
The intuition for the proof of Proposition 3 is that with any price x in the rational choice set $\mathcal{X}_0(\theta)$, the entire interval $[x, \bar{x}]$ must also be in $\mathcal{X}_0(\theta)$. By the assumed monotonicity of the marginal cost and demand, the lowest price in the rational choice set has to satisfy equation (12). The aforementioned coerciveness then implies that for the firm to have positive profits, an optimal price must be strictly above $MC(0)$ and strictly below the upper bound $\bar{x}$ of consumers' WTP (i.e., in the interior of $\mathcal{X}_0(\theta)$).

Remark 2. The lower bound $x_0(\theta)$ of the firm's rational choice set $\mathcal{X}_0(\theta)$ is non-decreasing in $\theta \in \Theta$, as

$$x_0'(\theta) = \frac{C''(D(x_0(\theta),\theta))\; D_\theta(x_0(\theta),\theta)}{1 - C''(D(x_0(\theta),\theta))\; D_x(x_0(\theta),\theta)} \geq 0,$$

[15] Assumptions (F) and (F') imply that $\mathcal{X}_0(\theta) \neq \emptyset$.

Figure 3. Set-valued solution $\mathscr{X}(\theta)$ to the profit-maximization problem in equation (7) for $\theta \in \Theta$



by the demand monotonicity in (D1) and the cost convexity in (C1). This implies that the rational choice sets are successively nested, in the sense that for any $\theta', \theta'' \in \Theta$ with $\theta' < \theta''$ it is $\mathcal{X}_0(\theta') \supseteq \mathcal{X}_0(\theta'')$. In particular, $\mathcal{X}_0(\theta_1)$ contains the rational choice sets for all demand parameters in Θ.

Remark 3. A direct consequence of equation (13) in Proposition 3 is that any optimal price in $\mathscr{X}(\theta)$ must be positive, as long as the optimal profit $\pi^*(\theta)$ does not vanish.

As no assumptions were made about the concavity of the firm's profit function, the solution to the pricing problem (7) is generally set-valued (see Figure 3). Yet, the continuity of $\pi(x, \theta)$ implies that the optimal profit $\pi^*(\theta)$ cannot make any jumps and the solution set $\mathscr{X}(\theta)$ is also well behaved, as summarized hereafter.

Proposition 4 (Solution regularity; Weierstrass 1860/Berge 1963).[16] *(i) $\mathscr{X}(\theta)$ is non-empty for all $\theta \in \Theta$. (ii) The mapping $\mathscr{X}(\cdot)$ is compact-valued and upper semi-continuous on Θ. (iii) The optimal profit $\pi^*(\cdot)$ is continuous on Θ.*

[16]Part (i) is the well-known extreme value theorem (established by Bolzano) usually associated with Karl Weierstrass and his introductory lectures on the theory of analytic functions, held in Berlin for the first time in the winter semester of 1861/62 (Ullrich, 1989).

Part (i) of Proposition 4 is a consequence of the Weierstrass extreme value theorem. Parts (ii) and (iii) follow from the Berge maximum theorem Berge (1963). Because by part (ii) the solution set $\mathscr{X}(\theta)$ is non-empty and compact, both its minimum and maximum do exist. Without any loss of generality, we assume that the firm chooses always the smallest (resp., always the largest) optimal price $x(\theta)$, that is,[17]

$$x(\theta) = \min \mathscr{X}(\theta) \qquad (\text{resp., } x(\theta) = \max \mathscr{X}(\theta))\,. \tag{14}$$

Any selector $x(\theta)$ of the (by Proposition 4) upper semi-continuous map $\mathscr{X}(\theta)$ is generically discontinuous.[18] As an example, Figure 3 depicts the set-valued maximizer $\mathscr{X}(\theta)$ of the firm's profit in Figure 2, from which it would select an optimal price $x(\theta)$ according to the preceding heuristic.

Proposition 5 (Solution monotonicity). *The optimal price $x(\theta)$ is non-decreasing in $\theta \in \Theta$.*

The monotonicity of the optimal price $x(\theta)$ follows from the supermodularity of the profit function, implied by (D2). Because the best price moves in the same direction as the underlying model parameter, in the subsequent analysis one can easily go back and forth between shifts in a candidate parameter $\widehat{\theta}$ and the corresponding qualitative shifts in the optimal decision $x(\widehat{\theta})$.

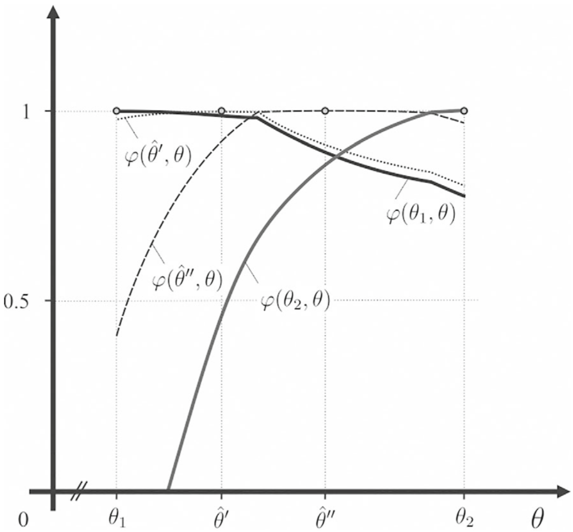
3. Robust identification and pricing

In what follows, we first introduce the performance ratio $\varphi(\widehat{\theta}, \theta)$ to benchmark the firm's pricing decision (optimal for $\widehat{\theta} = \theta$) against the possibility of adverse model realizations $\theta \neq \widehat{\theta}$. The worst-case performance ratio, taken over all θ, determines the performance index $\rho(\widehat{\theta})$ as a function of the candidate

[17] A firm choosing the smallest optimal price errs on the side of consumer surplus, providing the largest possible value to its customers as long as it does not affect profits. A firm opting for the largest optimal price errs on the side of total production cost, keeping output as low as possible for a given optimal profit; on the one hand, this may limit the firm's exposure to contracts but, on the other hand, it tends to increase unit cost (and may limit learning). The technical reason for focusing on either the smallest or the largest element of $\mathscr{X}(\theta)$ is that monotonicity of the set-valued maximizer obtains in the Veinott strong set order (i.e., monotonicity of the set boundary, consistent with Proposition 5).

[18] The implied discontinuities appear generically in the robustness criteria below. For example, the performance ratio $\varphi(\widehat{\theta}, \theta)$ in equation (15) evaluates a decision compatible with a candidate parameter $\widehat{\theta}$ relative to a testing parameter θ, and is thus out of the scope of Proposition 4(iii), which guarantees continuity only in the case where $\widehat{\theta} = \theta$. For more details on the regularity properties of selectors for upper semi-continuous set-valued maps with non-empty compact values, see, for example, Jayne and Rogers (2002).

Figure 4. Performance ratio $\varphi(\widehat{\theta}, \theta)$ for $\widehat{\theta} \in \{\theta_1, \widehat{\theta}', \widehat{\theta}'', \theta_2\}$ and $\theta \in \Theta$



model parameter $\widehat{\theta}$. In Section 3.2, we provide an "envelope representation" of the performance index that implies a simple characterization of any optimal robust parameter $\widehat{\theta}^*$ (which as such maximizes the performance index). Throughout the developments, it is assumed that demand and cost functions satisfy the six key properties (i.e., D1, D2, C1, C2, F, and F') discussed in Section 2.

3.1. Robustness and performance optimization

For demand parameters $\widehat{\theta}$ and θ in Θ, the firm's performance ratio (of $\widehat{\theta}$ with respect to θ) is

$$\varphi(\widehat{\theta}, \theta) = \frac{\pi(x(\widehat{\theta}), \theta)}{\pi(x(\theta), \theta)} \in [0, 1]. \tag{15}$$

The denominator of the performance ratio corresponds to the optimal profit in equation (8), which is always non-negative.[19] Figure 4 shows the performance ratio $\varphi(\widehat{\theta}, \cdot)$ on $\Theta = [\theta_1, \theta_2]$, for various candidate parameters $\widehat{\theta} \in \Theta$, building on the nonlinear example portrayed in Figures 1–3.

[19]In the degenerate case where the firm's optimal profit vanishes (corresponding to what the firm can always achieve by setting the price equal to $\overline{x}$), the performance ratio can be set equal to 1, without any loss of generality.

As a measure of overall robustness, the (relative) performance index,

$$\rho(\widehat{\theta}) = \inf_{\theta \in \Theta} \varphi(\widehat{\theta}, \theta) \in [0, 1], \qquad \widehat{\theta} \in \Theta, \tag{16}$$

is the worst-case performance ratio of $\widehat{\theta}$ on the parameter space Θ.

Remark 4 (Relative regret). The performance index is directly linked to maximum relative regret,

$$\mathrm{RR}(\widehat{\theta}) = \sup_{\theta \in \Theta} \frac{\pi(x(\theta), \theta) - \pi(x(\widehat{\theta}), \theta)}{\pi(x(\theta), \theta)} = 1 - \rho(\widehat{\theta}), \qquad \widehat{\theta} \in \Theta.$$

Thus, minimizing maximum relative regret is equivalent to maximizing the firm's performance index.

Given the performance index in equation (16) as the firm's comparative benchmark against all possible model realizations, the optimal robust parameter,

$$\widehat{\theta}^* \in \arg\max_{\widehat{\theta} \in \Theta} \rho(\widehat{\theta}), \tag{R}$$

provides the best relative performance with respect to the considered family of demand curves. By virtue of the solution to the firm's pricing problem (7) in equation (14), the optimal robust parameter also determines the firm's optimal robust price,

$$\widehat{x}^* = x(\widehat{\theta}^*). \tag{17}$$

To understand the nature of the performance index as a measure of robustness, it is useful to note that the firm's performance ratio $\varphi(\widehat{\theta}, \theta)$ in equation (15) is single-peaked as a function of the "testing" parameter θ, when keeping the "candidate" parameter $\widehat{\theta}$ fixed.

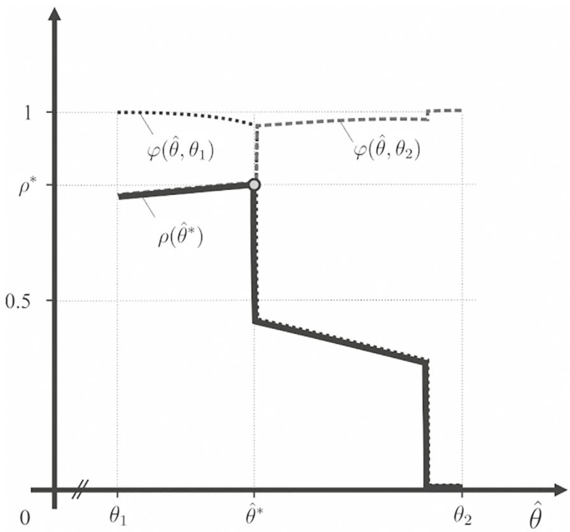
Proposition 6 (Quasi-concavity). *Given any $\widehat{\theta} \in \Theta$, the function $\varphi(\widehat{\theta}, \cdot) : \Theta \to \mathbb{R}$ is non-decreasing to the left of $\widehat{\theta}$ and non-increasing to the right of $\widehat{\theta}$.*

The key idea for obtaining quasi-concavity of the performance ratio is to first establish the log-supermodularity of the firm's profit, which in turn implies that the slope of the performance ratio $\varphi(\widehat{\theta}, \cdot)$ exhibits a single-crossing property, for any given candidate parameter $\widehat{\theta} \in \Theta$. The resulting quasi-concavity of the performance ratio in the testing parameter θ implies that the critical tests against the performance of a candidate parameter must be conducted on the boundary of the parameter space Θ, implying a practically useful representation of the performance index.

Proposition 7 (Envelope representation). *The firm's performance index in equation (16) can be written in the form*

$$\rho(\widehat{\theta}) = \min\left\{\varphi(\widehat{\theta}, \theta_1), \varphi(\widehat{\theta}, \theta_2)\right\}, \qquad \widehat{\theta} \in \Theta. \tag{18}$$

Figure 5. Relative performance index $\rho(\widehat{\theta}) = \min\{\varphi(\widehat{\theta}, \theta_1), \varphi(\widehat{\theta}, \theta_2)\}$ for $\widehat{\theta} \in \Theta$



The preceding envelope representation captures the fact that the performance index depends only on the boundary performance ratios at θ_1 and θ_2. The firm's preference for robustness implies "perfect complementarity" of the relative performance of the candidate parameter $\widehat{\theta}$ with respect to both boundaries of the parameter space.[20] The set $\{\varphi(\cdot, \theta_1), \varphi(\cdot, \theta_2)\}$ can therefore be considered an "extremal base" of the firm's performance index.[21] Figure 5 illustrates the performance index as the lower envelope in equation (18) for the earlier example in Figures 1–4.

3.2. Robust solution

Based on Proposition 7, the optimal robust parameter in equation (16) must maximize the lesser of the performance ratios at the boundary of the parameter space. If demand elasticity is monotonic in the price, then the boundary performance ratios are identical at a robust optimum.

[20]Perfect complementarity or "Leontief preferences" for two commodities bases the obtained utility on the lesser presence of the two goods; at the time, these preferences were introduced chiefly as a computational simplification of the input–output relationships in macroeconomic models (Leontief, 1941).
[21]The idea of an extremal base, common in maximin problems, dates back at least to Chebyshev (1859).

Proposition 8 (Characterization). *If the demand elasticity $\varepsilon(x,\theta)$ is non-decreasing in x,*[22] *then any optimal robust parameter $\widehat{\theta}^*$ in equation (R) is such that*

$$\varphi(\widehat{\theta}^*, \theta_1) = \varphi(\widehat{\theta}^*, \theta_2). \tag{19}$$

The proof of this result proceeds by first establishing that the boundary performance ratios $\varphi(\cdot, \theta_1)$ and $\varphi(\cdot, \theta_2)$ exhibit opposing monotonicities. This implies (by the intermediate value theorem) that the boundary performance ratios must be equal for at least one parameter value. From there, it is shown that equation (19) exactly describes any optimal robust parameter $\widehat{\theta}^*$.

The characterization in Proposition 8 also yields monotone comparative statics of the latter solution ($\widehat{\theta}^*$) to the firm's robust optimization problem (R), as well as the optimal performance index,

$$\rho^* = \rho(\widehat{\theta}^*), \tag{20}$$

with respect to changes in the boundary of $\Theta = [\theta_1, \theta_2]$. Indeed, combining the monotonicity of $\Delta(\widehat{\theta}) = \varphi(\widehat{\theta}, \theta_2) - \varphi(\widehat{\theta}, \theta_1)$ in the candidate parameter $\widehat{\theta}$ (under the hypotheses of Proposition 8) with the observation that the performance ratio $\varphi(\widehat{\theta}, \cdot)$ has opposing slopes to the left and right of $\widehat{\theta}$ (by virtue of Proposition 6), differentiation of equation (19) implies that the optimal robust parameter $\widehat{\theta}^*$ is non-decreasing in both θ_1 and θ_2. The optimal performance index, however, is naturally non-increasing in the diameter of Θ (when successively extending either one of its two boundary points). The following result summarizes these insights.

Proposition 9 (Comparative statics). *If demand elasticity $\varepsilon(x,\theta)$ is non-decreasing in x, then the optimal robust parameter $\widehat{\theta}^*$ is non-decreasing in the boundaries θ_1 and θ_2 of the parameter space Θ. Furthermore, the optimal performance index ρ^* is non-decreasing in θ_1 and non-increasing in θ_2.*

When the lower boundary θ_1 of the parameter space increases, perhaps resulting from an optimistic adjustment of demand forecasts or a positive income shock, the optimal robust parameter $\widehat{\theta}^*$ increases (at least weakly), and so does (by Proposition 5) the optimal robust price in equation (17). All else equal, an increase of the upper boundary θ_2 has the same effect. By contrast,

[22] This assumption corresponds to the standard intuition, originally articulated by Marshall (1920, p. 87) as follows. "The elasticity of demand is great for high prices, and great, or at least considerable, for medium prices; but it declines as the price falls; and gradually fades away if the fall goes so far that satiety level is reached." It is naturally satisfied for linear demands.

the optimal performance guarantee ρ^* improves whenever the ambiguity set Θ shrinks. The corresponding gradient,

$$\rho^*_{\theta_i} = \varphi_\theta(\widehat{\theta}^*, \theta_i), \qquad i \in \{1, 2\}, \tag{21}$$

represents the firm's (relative) "cost of robustness", measured as the change in the performance index induced by a unit deformation of the parameter space. As might have been instinctively obvious from the outset, more parameter ambiguity implies less performance guarantee. Quantifying this intuition, equation (21) can serve as an economic gauge for the value of additional information (e.g., in the form of market studies) to reduce the extent of the parameter space Θ.

Remark 5 (Generalization). The assertions in Propositions 8 and 9 continue to hold when instead of ε merely the product $\varepsilon\mu$ is non-decreasing in x, which is all that is needed in the proofs. Because the relative markup μ in equation (10) is increasing in x, this condition is significantly weaker than the monotonicity of ε in x,[23] and it is satisfied automatically in a neighborhood of $(x(\theta), \theta)$. Indeed, by equation (9) it is $\varepsilon(x(\theta), \theta)\mu(x(\theta), \theta) = 1$, so

$$x'(\theta) = -\left.\frac{(\varepsilon\mu)_\theta}{(\varepsilon\mu)_x}\right|_{(x(\theta),\theta)} \geq 0,$$

where the inequality follows from Proposition 5. Realizing that $\varepsilon_\theta \leq 0$ by Proposition 2, and that in general $\mu_\theta(x, \theta) = -MC(D(x, \theta))D_\theta(x, \theta)/x < 0$ (provided that x and $D(x, \theta)$ are positive), we can conclude that

$$\left.\frac{\partial \varepsilon\mu}{\partial x}\right|_{(x(\theta),\theta)} > 0.$$

Finally, when ignoring the characterization in Proposition 8, computationally we can still quite easily determine the optimal performance index, by suitably discretizing the domain of the candidate parameter $\theta \in \Theta$ and reverting to the envelope representation of ρ in Proposition 7.

3.3. Comparison with other robustness criteria

We now demonstrate that the robust pricing decisions obtained via maximization of the firm's performance index (i.e., through the minimization of relative regret; see Remark 4) tend to be less conservative than the pricing

[23] A further generalization is detailed in footnote 47. While the corresponding requirement no longer involves just the model primitives (but also the range of the solution to the firm's profit-maximization problem), it does illustrate the fact that the characterization in equation (19) reaches beyond the simple assumptions of Proposition 8.

decisions implied by standard (maximin) worst-case analysis (Wald, 1945) or (minimax) regret analysis (Savage, 1951).

3.3.1. Worst-case (maximin) analysis. By the demand monotonicity (D1) and the envelope theorem,[24] the firm's optimal profit is increasing in the demand parameter, as long as the profit is positive:

$$\pi^*(\theta) > 0 \;\Rightarrow\; \frac{d\pi^*(\theta)}{d\theta} = \pi_\theta(x(\theta), \theta)$$
$$= (x(\theta) - MC(x(\theta), \theta))\, D_\theta(x(\theta), \theta) > 0, \tag{22}$$

for all $\theta \in \Theta$.[25] This in turn implies an optimal worst-case action,

$$x^*_{\mathrm{WC}} \in \arg \max_{x \in \mathcal{X}_0(\theta_1)} \pi(x, \theta_1), \tag{23}$$

with $\mathcal{X}_0(\theta_1)$ as in equation (11), for $\theta = \theta_1$. Hence, the worst-case price can be set to

$$x^*_{\mathrm{WC}} = x(\theta_1), \tag{24}$$

and by the solution monotonicity in Proposition 5 it cannot exceed the optimal robust solution $\widehat{x}^*$ in equation (17); that is, $x^*_{\mathrm{WC}} \leq \widehat{x}^*$. Additionally, it comes without much surprise that (by construction) the performance index for the worst-case price, $\rho_{\mathrm{WC}} = \rho(\theta_1) = \varphi(\theta_1, \theta_2)$, can never exceed the optimal performance index $\widehat{\rho}^*$ in equation (20).

3.3.2. Regret (minimax) analysis. For any candidate parameter $\breve{\theta} \in \Theta$ the firm's maximal (absolute) regret is

$$R(\breve{\theta}) = \max_{\theta \in \Theta} \left\{\pi^*(\theta) - \pi(x(\breve{\theta}), \theta)\right\}$$
$$= \max \left\{\pi^*(\theta_1) - \pi(x(\breve{\theta}), \theta_1), \pi^*(\theta_2) - \pi(x(\breve{\theta}), \theta_2)\right\},$$

where the envelope representation follows from the quasi-concavity of $\pi^*(\cdot) - \pi(x(\breve{\theta}), \cdot)$ on Θ, as a consequence of (D1), (D2), (C1), and Proposition 5. Thus, the minimax regret price $\breve{x}^* = x(\breve{\theta}^*)$ is such that

[24]For a formulation and proof of the standard envelope theorem, see, for example, Mas-Colell et al. (1995, pp. 964–966).

[25]Given any $\theta \in \Theta$, in the interior of the firm's rational choice set (i.e., for $x \in \mathcal{X}_0(\theta)$) it is $\pi_\theta(x, \theta) = (x - MC(x, \theta))D_\theta(x, \theta) \geq 0$. But to obtain a positive profit, the firm's demand needs to be positive and its price x needs to strictly exceed the average cost (and so by Proposition 2 also marginal cost), implying that both factors in equation (22) are positive, and the inequality is therefore strict.

$$\check{x}^* \in \{\check{x} \in \mathcal{X} : \check{x}(D(\check{x}, \theta_2) - D(\check{x}, \theta_1)) - (C(\check{x}, \theta_2) - C(\check{x}, \theta_1)) = \pi^*(\theta_2) - \pi^*(\theta_1)\}. \tag{25}$$

Taking into account property F, the definition of $\mathcal{X} = [0, \bar{x}]$, and the intermediate value theorem, the preceding set is non-empty. Accordingly, the minimax regret price $\check{x}^*$ balances the profit shortfalls at the bounds of the parameter space.

4. Application

Based on the developments in Sections 2 and 3, we now provide an illustration for how the proposed method can be used in practice. Throughout this section we assume that the firm's marginal cost is constant but we will otherwise arrive at complete generality, first discussing parametrized linear demand curves (see Section 4.1), then non-parametric nonlinear demand curves (see Section 4.2), and finally a fully data-driven approach with measurement errors (see Section 4.3).[26]

4.1. Linear demand (parametric)

Consider the important special case where demand is linear (affine) and the firm's cost is also linear. More specifically, we assume that the monopolist's model of demand is (piecewise) affine, of the form[27]

$$\widehat{D}(x|a, b) = [a - bx]_+, \qquad x \in \mathcal{X}, \tag{26}$$

where the parameters a and b are known to lie in the respective intervals $[a_0 - \alpha, a_0 + \alpha] \subset \mathbb{R}_{++}$ and $[b_0 - \beta, b_0 + \beta] \subset \mathbb{R}_{++}$, spread around their respective positive nominal values a_0 and b_0 with $\alpha \in (0, a_0)$ and $\beta \in (0, b_0)$ (see Figure 6). Moreover, the firm's production cost is of the form

$$C(q) = cq, \qquad q \in \mathbb{R}_+, \tag{27}$$

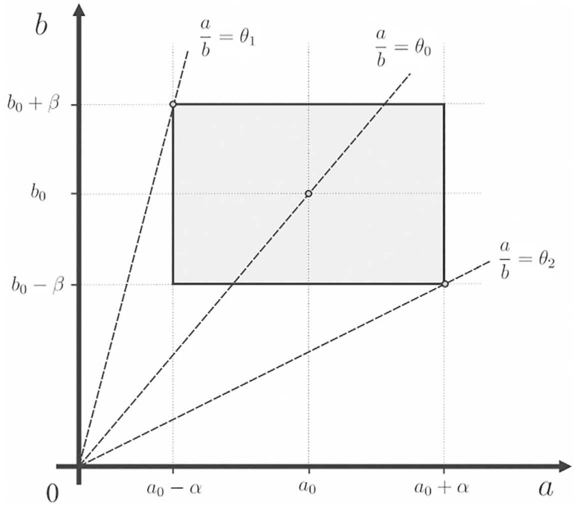
where q denotes the firm's output and c is a known positive constant. To transform this classical set-up with two unknown parameters to the model discussed in the main text, we first note that the firm's profit,

$$\widehat{\pi}(x|a, b) = (x - c)\widehat{D}(x|a, b) = b(x - c)\widehat{D}(x|a/b, 1) = b\widehat{\pi}(x|a/b, 1),$$

[26] Linear production costs (i.e., constant marginal costs), which satisfy properties C1 and C2, allow for an explicit solution of the firm's profit-maximization problem, with conversion of the two-parameter family of linear (affine) demand curves into an equivalent one-parameter family, which exactly matches our theoretical framework.

[27] For any $y \in \mathbb{R}$, we set $[y]_+ = \max\{y, 0\}$.

Figure 6. Reduced scalar demand parameter $\theta = a/b \in [\theta_1, \theta_2]$ as a function of the original demand-parameter vector $(a, b) \in [a_0 - \alpha, a_0 + \alpha] \times [b_0 - \beta, b_0 + \beta]$



is homogeneous (of degree 1) in one of the parameters. As a result, the firm's pricing decision depends solely on the parameter ratio,

$$\theta = a/b.$$

The homogeneous influence of the parameter b on the firm's profit can be viewed as exogenous multiplicative noise that has no direct bearing on the firm's decision-making. Thus, without any loss of generality, we can restrict attention to the (stochastically normalized) objective function $\widehat{\pi}(x, \theta)/b$ eliminating the second parameter, and obtain the "reduced profit",

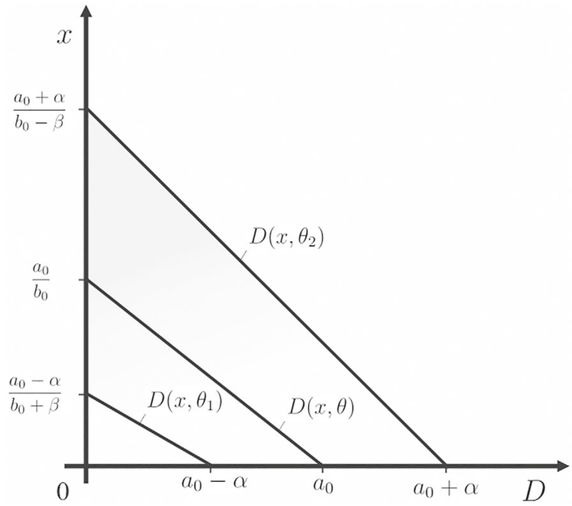
$$\pi(x, \theta) = (x - c)D(x, \theta), \qquad (x, \theta) \in \mathbb{R}_+ \times \Theta,$$

with "reduced demand",

$$D(x, \theta) = [\theta - x]_+, \qquad (x, \theta) \in \mathbb{R}_+ \times \Theta. \tag{28}$$

It is important to note that the noise caused by the uncertainty in $b > 0$ is folded into the uncertainty about the reduced parameter $\theta = a/b$, so that the domain of θ does depend on the domain of b. The corresponding family of demand curves is depicted in Figure 7. The decisions implied by this reduced model are exactly the same as those that would be found optimal in the original

Figure 7. Unknown demand D for $(x, \theta) \in \mathcal{X} \times \Theta$



model.[28] As shown in Figure 6, the relevant parameter interval $\Theta = [\theta_1, \theta_2]$ features the bounds θ_1 and θ_2, which are such that

$$\theta_1 = \frac{a_0 - \alpha}{b_0 + \beta} < \frac{a_0 + \alpha}{b_0 - \beta} = \theta_2. \tag{29}$$

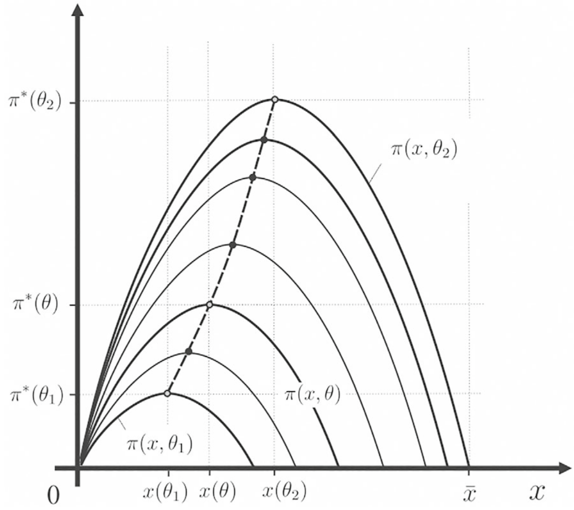
To streamline our presentation, we assume that the maximum absolute deviations α, β from the nominal values a_0, b_0 are such that demand is guaranteed to be positive when the firm charges the smallest viable price, namely marginal cost $x = c$, which merely requires

$$\theta_1 > c. \tag{30}$$

The resulting (reduced) model primitives satisfy all the functional assumptions in the main text (i.e., properties D1, D2, C1, C2, F, and F'). Indeed, the demand $D(x, \theta)$ exhibits monotonicity (D1) and complementarity (D2). In addition, the firm's linear cost $C(q) = cq$ trivially conforms to (C1) and (C2). Finally, the positive constants $\overline{D}$ and $\overline{x}$, with $\overline{D} = \overline{x} = \theta_2 > 0$, satisfy the finiteness condition (F) and – by virtue of equation (30) – also the "cost-plus pricing possibility" in (F'). The family of the firm's profit functions $\pi(\cdot, \theta)$, for $\theta \in \Theta$, is shown in Figure 8.

[28]The firm's performance ratio φ and its performance index ρ remain unaffected by the transition from the original model to the reduced model.

Figure 8. Profit π for $(x,\theta) \in \mathcal{X} \times \Theta$, with maximizer $x(\theta) \in \mathscr{X}(\theta)$



The envelope representation of the firm's performance ratio (in Proposition 7) and the characterization of the optimal robust parameter $\widehat{\theta}^*$ (in Proposition 8) imply the optimal robust price $\widehat{x}^* = x(\widehat{\theta}^*)$, with a performance guarantee given by the optimal performance index $\rho^* = \rho(\widehat{\theta}^*)$. Figure 9 depicts how the optimal robust parameter $\widehat{\theta}^*$ is obtained at the intersection of the two boundary performance ratios $\varphi(\cdot,\theta_1)$ and $\varphi(\cdot,\theta_2)$.

Proposition 10 (Robust pricing with linear demand). *Given the (indeterminate) linear demand in equation (26), together with the linear cost in equation (27), the firm's optimal robust price is*

$$\widehat{x}^* = \frac{1}{2}\frac{\theta_1\theta_2 - c^2}{\overline{\theta} - c}, \tag{31}$$

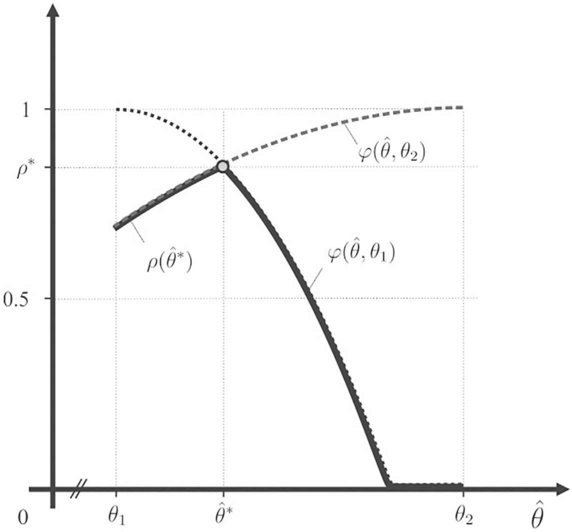
with a corresponding optimal performance index of

$$\rho^* = 1 - \left(\frac{\overline{\theta} - \theta_1}{\overline{\theta} - c}\right)^2 > 0, \tag{32}$$

where θ_1 and θ_2, together with their arithmetic mean $\overline{\theta} = (\theta_1 + \theta_2)/2$, are determined by equations (29) and (30).

By construction, the firm's profit $\widehat{\pi}(\widehat{x}^*|a,b)$ is guaranteed to be at least ρ^* times the *ex post* optimal profit achieved with perfect information about the demand parameters a and b.

Figure 9. Relative performance index $\rho(\widehat{\theta}) = \min\{\varphi(\widehat{\theta}, \theta_1), \varphi(\widehat{\theta}, \theta_2)\}$ for $\widehat{\theta} \in \Theta$



Remark 6 (Relative versus absolute performance). By construction, when implementing the optimal robust price $\widehat{x}^*$, the firm is guaranteed the fraction ρ^* of the *ex post* optimal profit that could have been obtained with complete knowledge about the demand-curve realization (within the considered family of linear demand curves). That is, as a function of the true parameter θ, the firm obtains the profit

$$\widehat{\pi}^*(\theta) = \pi(\widehat{x}^*, \theta) = (\widehat{x}^* - c)\,(\theta - \widehat{x}^*) \geq \rho^*\,\pi^*(\theta) = \frac{\rho^*}{4}(\theta - c)^2 > 0, \quad \theta \in \Theta.$$

This means that when demand is poor, with θ close to θ_1, profit is also low but still positive and still within a factor of ρ^* to how good it could have been with a perfectly adapted price. However, when implementing the worst-case price $x^*_{\mathrm{WC}} = x(\theta_1)$ (see Section 3.3.1) the firm's profit becomes

$$\pi^*_{\mathrm{WC}}(\theta) = \pi(x^*_{\mathrm{WC}}, \theta) = \frac{\theta_1 - c}{2}\left(\theta - \frac{\theta_1 + c}{2}\right) \geq \frac{(\theta_1 - c)^2}{4} > 0, \quad \theta \in \Theta.$$

The situation becomes more delicate when implementing a minimax regret price (see Section 3.3.2), which by equation (25) is determined by $\breve{x}^* \in [c + (\theta_2 - \theta_1)/2, \bar{\theta}]$. Thus, whenever c is close to θ_1, the firm's profit,

$$\breve{\pi}^*(\theta) = \pi(\breve{x}^*, \theta) \in \left[\left(\bar{\theta} - c\right)\left[\theta - \bar{\theta}\right]_+, \frac{\theta_2 - \theta_1}{2}\left[\theta - c - \frac{\theta_2 - \theta_1}{2}\right]_+\right], \theta \in \Theta,$$

vanishes for θ close to θ_1, so the absolute return may be rather poor (e.g., zero, leading to a zero performance index). By contrast, when choosing the optimal robust price $\widehat{x}^*$ the firm is guaranteed a profit of at least $\widehat{\pi}^*(\theta_1) = \rho^* \pi^*_{\mathrm{WC}}(\theta_1) > 0$.

Remark 7 (Comparison with certainty-equivalence pricing). Let $x^*_{\mathrm{CE}} = ((a_0/b_0) + c)/2$ be the certainty-equivalence price that the firm would choose when simply ignoring the model uncertainty,[29] i.e., for $\alpha \to 0^+$ and $\beta \to 0^+$. The worst-case price x^*_{WC} and the certainty-equivalence price x^*_{CE} bracket the optimal robust price $\widehat{x}^*$, in the sense that[30]

$$x^*_{\mathrm{WC}} < \widehat{x}^* < x^*_{\mathrm{CE}}.$$

That is, allowing for demand misspecification (within the family of linear demand curves) leads to lower prices when minimizing relative regret or when maximizing the worst-case payoff.

Remark 8 (Expected-profit comparison). It is instructive to compare the expected profit of the optimal robust price $\widehat{x}^*$ in equation (31) against that of the solutions $\widehat{x}^*_{\mathrm{WC}} = x(\theta_1)$ and $\check{x}^* = x(\overline{\theta})$, implied by the alternative robustness criteria discussed in Section 3.3, namely worst-case payoff and minimax regret, respectively. For this, consider a random parameter $\tilde{\theta}$, distributed on its support $\Theta = [\theta_1, \theta_2]$, with mean $\theta_e = E[\tilde{\theta}] \in [\theta_1, \theta_2]$. Because the profit function $\pi(x, \theta)$ is affine in $\theta \geq \theta_1 > x(\theta_1)$, the expected profit under the worst-case price is

$$\overline{\pi}^*_{\mathrm{WC}} = E[\pi^*_{\mathrm{WC}}(\tilde{\theta})] = \frac{\theta_1 - c}{2}\left(\theta_e - \frac{\theta_1 + c}{2}\right).$$

However, the optimal robust price is $\widehat{x}^* = (1/2)(\theta_1\theta_2 - c^2)/(\overline{\theta} - c) < \theta_1$, as specified in equation (31), so that the corresponding expected profit (for our relative-regret criterion) becomes

$$\overline{\pi}^*_{\mathrm{RR}} = E[\widehat{\pi}^*(\tilde{\theta})] = \frac{\left(\theta_1\theta_2 - c^2 - 2\theta_e(\overline{\theta} - c)\right)(\theta_2 - \theta_1)(\theta_1 - c)}{4\,(\overline{\theta} - c)^2}.$$

Taking the difference between the two expected profits yields

$$\overline{\pi}^*_{\mathrm{RR}} - \overline{\pi}^*_{\mathrm{WC}} \geq 0 \quad \Leftrightarrow \quad \theta_e \geq \check{\theta},$$

[29] With linear demand, the certainty-equivalence price maximizes the firm's expected profits for a uniform distribution of the parameters a and b over their domains (which notably does not correspond to a uniform distribution of the reduced parameter).

[30] While the largest minimax regret price ($\check{x}^* = \overline{\theta}$) always exceeds x^*_{CE}, some $\check{x}^*$ might at times appear below $\widehat{x}^*$.

where

$$\check{\theta} = \theta_1 + \frac{(\theta_2 - \theta_1)(\theta_1 - c)}{4\,(\bar{\theta} - c)} = \bar{\theta} - \frac{(\theta_2 - \theta_1)(\theta_2 - c)}{4\,(\bar{\theta} - c)} \in \left(\theta_1, \theta_1 + \frac{\bar{\theta} - \theta_1}{2}\right).$$

In other words, the relative-regret criterion produces a superior expected profit in the linear demand model whenever the random parameter has a mean θ_e which exceeds the threshold $\check{\theta}$. For the uniform (Laplacian) distribution in particular, it is $\theta_e = \bar{\theta} > \check{\theta}$, and

$$\bar{\pi}^*_{\mathrm{RR}} - \bar{\pi}^*_{\mathrm{WC}} = \frac{(\theta_2 - \theta_1)^2(\theta_1 - c)(\theta_2 - c)}{16\,(\bar{\theta} - c)^2} > 0.$$

Thus, for worst-case pricing to ever outperform the optimal robust price in terms of expected profit, the underlying parameter distribution needs to be significantly skewed, so as to exhibit an expected value $\theta_e < \check{\theta}$, where (as shown above) the threshold $\check{\theta}$ is in the lowest quartile of the Laplacian (uniform) reference distribution. For minimax regret pricing, the exact comparison is intricate (and quite parameter-dependent) due to a generic demand truncation at zero. The latter implies a payoff that is convex in θ. While a zero-profit outcome may occur with positive probability, the expected payoff would tend to increase in uncertainty (i.e., a mean-preserving spread of the parameter distribution) by Jensen's inequality. This in turn can yield an attractive performance in terms of expected profit, for example, under a Laplacian prior.

4.2. Nonlinear demand (non-parametric)

Interestingly, the results obtained for the two-parameter family of demand curves carry over to a setting with nonlinear demand curves, by virtue of the fact that any nonlinear demand curve can be linearized at the "operating point" (corresponding to the chosen price). Hence, not knowing which nonlinear demand curve has realized is essentially the same as not knowing which particular linearization to choose in order to determine an optimal robust price. This equivalence applies as long as one makes sure to include all possible relevant linearizations of the family of nonlinear demand curves, which in turn determines an "equivalent family" of linear demand curves. To motivate the general result, we first embed a parametrized family of nonlinear demand curves into our model, relating it to the results in Section 4.1. We are then ready to determine an optimal robust price for any given (compact) family of continuously differentiable nonlinear demand curves.

4.2.1. Connection to linear-demand model. To motivate the application of our method to nonlinear demand curves, we consider first a parametrized family of continuously differentiable demand curves $D(x,\theta)$ for $\theta \in \Theta$, which satisfy (D1) and (F) – disregarding the complementarity requirement (D2).[31] Given the consumers' maximum WTP $\bar{x}$ in (F),[32] let

$$k_x = \min_{(x,\theta)\in[c,\bar{x}]\times\Theta} |D_x(x,\theta)| \quad \text{and} \quad K_x = \max_{(x,\theta)\in[c,\bar{x}]\times\Theta} |D_x(x,\theta)|,$$

be Lipschitz constants bounding the absolute value of the demand-slope from below and above, respectively. By the continuity of the derivative $D_x(x,\theta)$, together with the Weierstrass extreme value theorem, the constants k_x and K_x are well defined and such that $0 < k_x < K_x < \infty$.

With this, it is clear that all possible linearizations of the given family of nonlinear demand curves, relevant for the price range $[c,\bar{x}]$, must lie in a quadrilateral (i.e., four-sided polygon), as shown in Figure 10, which contains the orthogonal line segments

$$[D(c,\theta_1), D(c,\theta_2)] \times \{c\} \quad \text{and} \quad \{0\} \times \left[c + \frac{D(c,\theta_1)}{K_x}, c + \frac{D(c,\theta_2)}{k_x}\right].$$

The aforementioned demand realizations are described by a two-parameter linear demand family $\widehat{D}(x|a,b)$ as in equation (26) for $(a,b) \in [a_0 - \alpha, a_0 + \alpha] \times [b_0 - \beta, b_0 + \beta]$, with $b_0 = (K_x + k_x)/2$, $\beta = (K_x - k_x)/2$,

$$a_0 = \frac{D(c,\theta_1) + D(c,\theta_2)}{2} + b_0 c, \quad \text{and} \quad \alpha = \frac{1}{2}\left|\frac{D(c,\theta_2) - D(c,\theta_1)}{2} - \beta\right|.$$

By equation (29), this implies the bounds of the parameter space $\Theta = [\theta_1, \theta_2]$,

$$\theta_1 = \frac{\min\{D(c,\theta_1) + K_x c, D(c,\theta_2) + k_x c\}}{K_x}$$

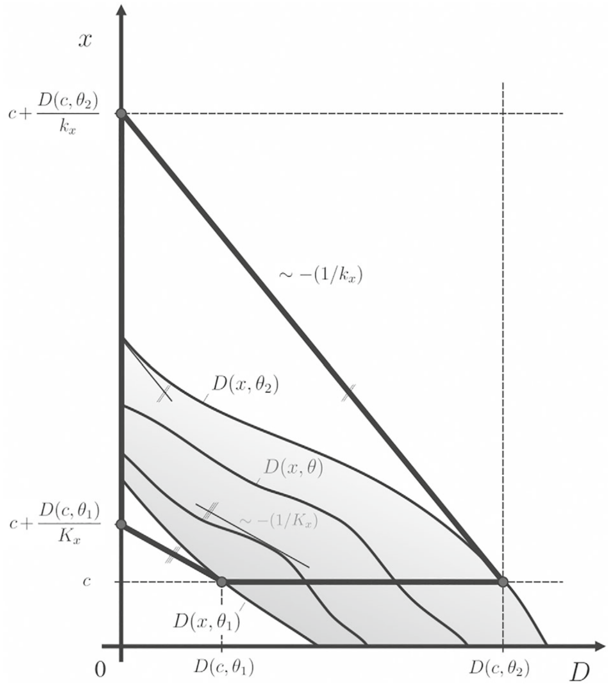
and

$$\theta_2 = \frac{\max\{D(c,\theta_1) + K_x c, D(c,\theta_2) + k_x c\}}{k_x},$$

in the "reduced" linear demand model. The corresponding optimal robust price $\widehat{x}^*$ for our family of nonlinear demand curves is therefore obtained using equation (31) in Proposition 10, with the associated optimal performance index ρ^* in equation (32).

[31]The assumption of constant marginal cost is maintained, as noted at the outset of Section 4.
[32]Without loss of generality, we assume that $\bar{x}$ exceeds the firm's marginal cost c, so (F') is satisfied as well.

Figure 10. Quadrilateral envelope of nonlinear demand-curve realizations



4.2.2. Robust pricing with nonlinear demand. Consider now a compact class $\mathscr{D}$ of continuously differentiable downward-sloping demand curves $D : \mathbb{R}_+ \to \mathbb{R}_+$,[33] so that for all $D \in \mathscr{D}$ and all $x \in \mathcal{X}$:

(Monotonicity)

$$D(x) > 0 \quad \Rightarrow \quad D'(x) < 0, \tag{$\widehat{\text{D1}}$}$$

and

(Finiteness)

$$\exists\, \bar{x} > 0 : \quad D(x) = 0, \quad x \in [\bar{x}, \infty)\,. \tag{$\widehat{\text{F}}$}$$

Properties $\widehat{\text{D1}}$ and $\widehat{\text{F}}$ are the non-parametric equivalents of the monotonicity property D1 and finiteness property F, respectively, as introduced in Section 2.

[33]The reason to require compactness is to exclude any undesirable situation where a sequence of functions of the class $\mathscr{D}$ approximates a demand function that is not in that class (e.g., a discontinuous demand function), so that in the limit our results might not apply.

By the continuous differentiability of all $D \in \mathscr{D}$, the Lipschitz constants

$$k_x = \inf_{D \in \mathscr{D}} \min_{x \in \mathcal{X}} |D'(x)| \quad \text{and} \quad K_x = \sup_{D \in \mathscr{D}} \max_{x \in \mathcal{X}} |D'(x)|,$$

are well defined, with $0 < k_x < K_x < \infty$. If we further denote by

$$\delta_1 = \inf_{D \in \mathscr{D}} D(c) \quad \text{and} \quad \delta_2 = \sup_{D \in \mathscr{D}} D(c)$$

the minimum and maximum demands at marginal-cost pricing (i.e., when the firm chooses the lowest economically viable price $x = c$), then we have the analogous set of primitives as in Section 4.2.1 to enable a transposition of the results in Proposition 10, with

$$\theta_1 = \frac{\min\{\delta_1 + K_x c, \delta_2 + k_x c\}}{K_x} \quad \text{and} \quad \theta_2 = \frac{\max\{\delta_1 + K_x c, \delta_2 + k_x c\}}{k_x},$$

to obtain the optimal robust price $\widehat{x}^*$ and optimal performance index ρ^*.

4.3. Data-driven approach

Based on the results in Sections 4.1 and 4.2 it is now only a small step to obtain a fully data-driven solution to the firm's robust pricing problem. For this, we assume that the firm has access to data in the form of L (possibly noisy) demand realizations D_i for different prices $x_i > 0$, where $i \in \{1, \ldots, L\}$.[34] Without loss of generality, we assume that the tuples (x_i, D_i) are indexed such that $x_i < x_{i+1}$ for all $i \in \{1, \ldots, L-1\}$.[35] In addition, we assume – as everywhere in our model – that demand is downward-sloping, in the sense that $D_i > D_{i+1}$.[36] Let

$$\widehat{k}_x = \min_{i \in \{1,\ldots,L-1\}} \left\{ \frac{D_i - D_{i+1}}{x_{i+1} - x_i} \right\} \quad \text{and} \quad \widehat{K}_x = \max_{i \in \{1,\ldots,L-1\}} \left\{ \frac{D_i - D_{i+1}}{x_{i+1} - x_i} \right\}$$

be the empirical Lipschitz constants, so that $0 < \widehat{k}_x < \widehat{K}_x < \infty$. In addition, we take subsequent pairs of observations to find the empirical upper and lower

[34] If the firm's product is new, these demand realizations could be obtained from a test market or a survey eliciting consumers' WTP and then counting how many respondents would buy at a given price.

[35] If the original dataset contains entries that feature different demands for the same price, we assume only a single price-demand tuple (x_i, D_i) where D_i is averaged over the different observations.

[36] Data points violating this assumption could either be omitted or else "corrected" so as to conform to the piecewise linear upper envelope of the data that does satisfy the monotonicity requirement.

bound for the "demand potential" (at zero price), namely

$$\widehat{m} = \min_{i \in \{1, \ldots, L-1\}} \left\{ D_i + \left(\frac{D_i - D_{i+1}}{x_{i+1} - x_i} \right) x_i \right\}$$

and

$$\widehat{M} = \max_{i \in \{1, \ldots, L-1\}} \left\{ D_i + \left(\frac{D_i - D_{i+1}}{x_{i+1} - x_i} \right) x_i \right\}.$$

Then we can set $a_0 = (\widehat{m} + \widehat{M})/2$, $\alpha = (\widehat{M} - \widehat{m})/2$, $b_0 = (\widehat{K}_x + \widehat{k}_x)/2$, and $\beta = (\widehat{K}_x - \widehat{k}_x)/2$ in Proposition 10 to find the (data-driven) optimal robust price,[37]

$$\widehat{x}^* = \frac{(\widehat{M}\widehat{m}/\widehat{K}_x\widehat{k}_x) - c^2}{(\widehat{m}/\widehat{K}_x) + (\widehat{M}/\widehat{k}_x) - 2c}, \tag{33}$$

by using equation (31) with $\theta_1 = \widehat{m}/\widehat{K}_x$ and $\theta_2 = \widehat{M}/\widehat{k}_x$ as the boundary of the equivalent reduced linear parameter space, consistent with the observed data set. By equation (32) the corresponding (data-driven) optimal performance index is

$$\widehat{\rho}^* = 1 - \left(\frac{(\widehat{M}/\widehat{k}_x) - (\widehat{m}/\widehat{K}_x)}{(\widehat{M}/\widehat{k}_x) + (\widehat{m}/\widehat{K}_x) - 2c} \right)^2. \tag{34}$$

The quality of the data-driven robust price and performance guarantee relies on the "exhaustiveness" of the observed data in capturing the actual extreme slopes (i.e., $\widehat{k}_x$ and $\widehat{K}_x$), as well as the bounds for the demand potential (i.e., $\widehat{m}$ and $\widehat{M}$).

4.4. Performance comparison

Given the successively increasing generality of the three demand models discussed in Sections 4.1–4.3, featuring first a two-parameter family of linear demands, then a class of non-parametric nonlinear demands with known Lipschitz-constants, and finally a fully data-driven approach, the question begs as to how the associated robust solutions compare against each other. In particular, we are interested in conducting a performance benchmarking of the firm's corresponding choices, compared also to the worst-case solution (in Section 3.3.1), as well as certainty-equivalence pricing (in Remark 7).[38]

[37] As in equation (30), the firm's marginal production cost c should be majorized strictly by $\theta_1 = \widehat{m}/\widehat{K}_x$.

[38] A comparison to the minimax regret solution is omitted, as it generically produces a zero performance index.

4.4.1. Set-up. To ensure fair measures across the different levels of generality, we use equivalent boundary conditions for the demand characteristics.

1. The linear demand model is simulated by randomly selecting N demand-curve realizations $\widehat{D}(\cdot|a,b)$ in equation (26) based on parameter tuples (a,b), which are sampled uniformly from the domain $[a_0-\alpha, a_0+\alpha]\times[b_0-\beta, b_0+\beta]$. This implies that the demand potential (at zero price) could be any point in the interval $[m,M]=[a_0-\alpha, a_0+\alpha]$, while the (absolute) slope of the demand curve has the range $[k_x, K_x]=[b_0-\beta, b_0+\beta]$.

2. Nonlinear demand curves are obtained as random realizations of demand curves (in the class $\mathscr{D}$) that start at zero price in the interval $[m,M]$ and have a piecewise constant slope for increasing prices until demand vanishes at a sufficiently high price. The relevant (absolute) slopes are sampled uniformly from the interval $[k_x, K_x]$ at $B\geq 1$ uniformly spaced breakpoints in the interval $\mathcal{X}=[0,\bar{x}]$, where $\bar{x}$ is the maximum WTP guaranteed by (F), which is such that $\widehat{D}(\bar{x}|M,k_x)=0$. Because on a compact interval the class of piecewise linear functions is dense in the class of continuous functions (Shekhtman, 1982), restricting attention to piecewise linear functions is without any loss of generality. As alluded to in Section 4.2, a linearization at a known operating point (i.e., an optimal price) would not affect the optimality of the operating point itself. Thus, our (non-parametric) class of nonlinear functions is fundamentally compatible with the linear class, the latter containing all possible linearizations for all reasonable operating points in $\mathcal{X}$.

3. In our data-driven approach, the randomly generated nonlinear demand curves are sampled at $L\geq 3$ price points x_i, uniformly drawn from the interval $[c,\bar{x}]$,[39] so as to obtain noisy samples $D_i\in[(1-\nu)D(x_i),(1+\nu)D(x_i)]$ of the demand-curve realization $D\in\mathscr{D}$, where $\nu\in[0,1)$ describes the strength of a multiplicative measurement distortion (ranging from 0 percent to any number below 100 percent). The estimators $\widehat{k}_x, \widehat{K}_x$ for the slope and $\widehat{m}, \widehat{M}$ for the demand potential are obtained from the corresponding formulas in Section 4.3, truncating the results to the intervals $[b_0-\beta, b_0+\beta]$ and $[a_0-\alpha, a_0+\alpha]$, consistent with the firm's underlying ambiguity set defining the Lipschitz-domain of demand curves, across all three

[39] There is no economic interest for the firm to sample demand at price points below its marginal cost c.

modeling scenarios. This truncation effectively excludes any outliers that might otherwise arise from the measurement errors, especially in the presence of clustered demand experiments, which would generically break the assumed monotonicity between neighboring demand samples. It also ensures that, in fact, the data-driven approach can only improve over the results in the nonlinear class of demand curve because additional information is used.

Before engaging in numerical analysis, we note that in the absence (or ignorance) of uncertainty, when the demand parameter $\theta_0 \in \Theta$ is unquestioned, the certainty-equivalence price $x^*_{\mathrm{CE}} = x(\theta_0)$ is in fact "optimal". When uncertainty is taken into account, suboptimality of any fixed pricing decision must be anticipated because of the generic mismatch between the chosen decision and the eventual realization of the *ex ante* unknown demand parameter. As an absolute standard, all of our comparisons include the performance of the *ex post* optimal decision that would have been best in hindsight with full information about the demand-curve realization.

4.4.2. Numerical results.[40]

For the performance comparisons to be interesting and meaningful, demand uncertainty needs to be substantial but not to a point where close to "nothing" is known, because then, of course, no method can be expected to perform reasonably well compared to an *ex post* optimal standard. We therefore assume ±20 percent uncertainty about the demand potential, with $a_0 = 100$ and $\alpha = 20$, and ±50 percent uncertainty about the slope of the demand curve, with $b_0 = 2$ and $\beta = 1$.[41] This means that demand potential (at zero price) is in the interval $[m, M] = [a_0 - \alpha, a_0 + \alpha] = [80{,}120]$ and (the absolute value of) the demand gradient with respect to price (which determines its local Lipschitz constant) lies in the interval $[k_x, K_x] = [b_0 - \varepsilon, b_0 + \varepsilon] = [1, 3]$. Finally, the firm's marginal cost is fixed at unity, so $c = 1$.

1. In the (reduced) linear demand model, the scalar parameter θ lies in the space $\Theta = [\theta_1, \theta_2] = [80/3{,}120]$, with nominal value $\theta_0 = a_0/b_0 = 50$. From equation (31) we obtain the optimal robust price $\widehat{x}^* = 1371/62 \approx 22.1129$, and from equation (32) the optimal performance index $\rho^* = 561/961 \approx 0.5838$. The worst-case price x^*_{WC} is equal to $x(\theta_1) = (\theta_1 + c)/2 = 83/6 \approx 13.83$. Finally, the certainty-equivalence price based on the nominal parameter value $\theta_0 = 50$ is $x(\theta_0) = 25.5$.

[40] An m-file for the use with Matlab is available from the author upon request.
[41] The ordinal performance comparison of the different solutions remains largely unaffected when using higher or lower relative dispersions, e.g., $\alpha \in \{10, 30\}$ and $\beta \in \{0.5, 1.5\}$.

2. The independent and identically distributed (i.i.d.) random nonlinear demand curves are obtained by realizations of piecewise linear functions, with $B = 4$ breakpoints at evenly spaced prices in the interval $[c, \bar{x}]$ (with $\bar{x} = \theta_2 = 120$), starting at the point (δ, c) where δ is drawn from a uniform distribution on the interval $[a_0 - \alpha - (b_0 + \beta)c, a_0 + \alpha - (b_0 - \beta)c] = [77{,}119]$.[42] The (absolute) slopes of the (downward-sloping) line segments are i.i.d. draws from a uniform distribution on the interval $[k_x, K_x] = [1, 3]$.[43] As the feasible envelope of linear demands is the same as for the linear model, the optimal robust price $\widehat{x}^*$, as well as all alternatives (x^*_{WC}, $\check{x}^*$, and x^*_{CE}) remain the same as for the linear model.

3. When following a data-driven approach, the firm observes noisy demand realizations D_i at the $L = 5$ random (but ordered) price realizations x_i, for $i \in \{1, \ldots, L\}$. For any given random (piecewise linear) demand curve $D(\cdot)$, an observation D_i is drawn from a uniform distribution on the interval $[(1 - \nu)D(x_i), (1 + \nu)D(x_i)]$ where $\nu = 20$ percent indicates the relative dispersion generated by the multiplicative measurement noise. Based on these measurements, the firm computes extremal slope estimates $\widehat{k}_x, \hat{K}_x$ (truncated to $[k_x, K_x]$), as well as estimates $\widehat{m}, \widehat{M}$ for the bounds of the demand potential (truncated to $[m, M]$). Thus, the firm can choose its optimal robust price $\widehat{x}^*$ contingent on four measured statistics using equation (33). This implies that the firm's robust pricing, contingent on its measurement, adjusts according to the information obtained about each specific demand-curve realization.

Figure 11 shows the performance ratios under the different pricing solutions, for both the linear and the nonlinear (data-driven) model. In the nonlinear case, the data-driven solution provides a significant improvement over both the certainty-equivalence solution and the worst-case solution. The

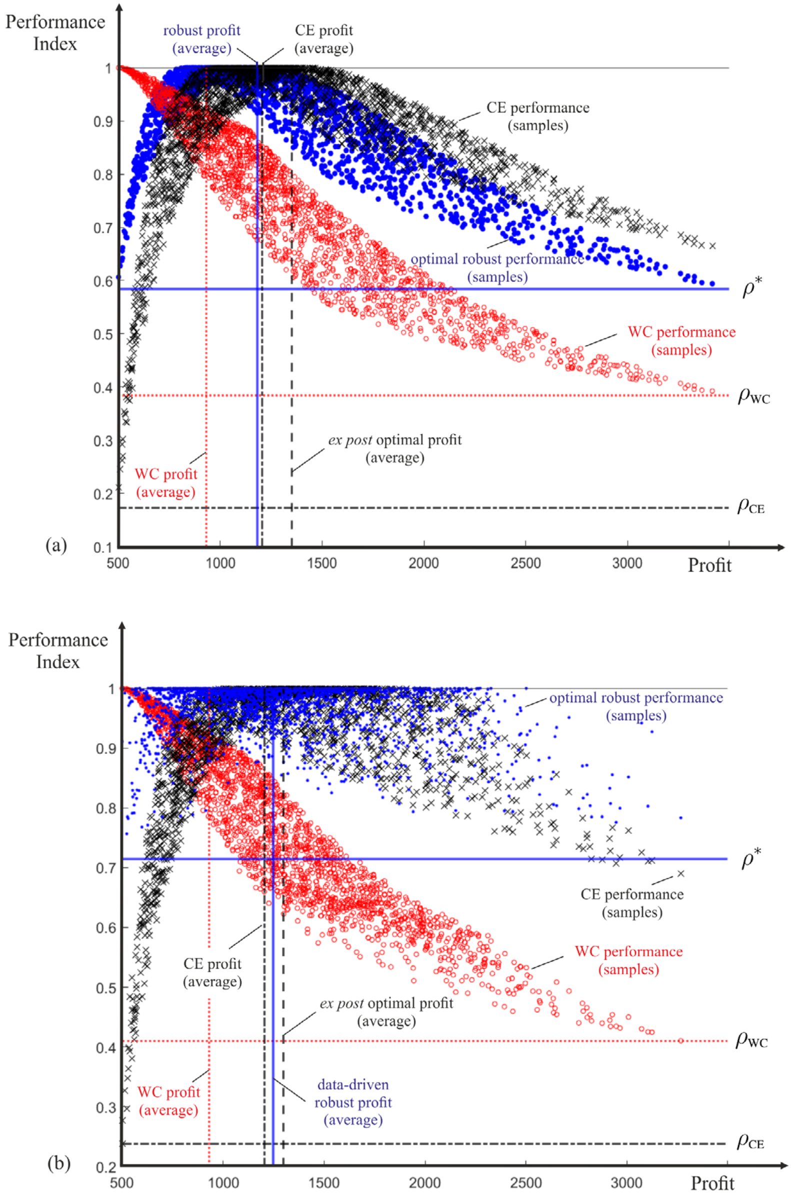
[42] Evenly spacing the breakpoints (as opposed to placing them randomly) avoids long linear demand segments and thus realizations too similar to the linear demand model simulated before. Introducing additional segments (by increasing B) tends to produce a "regression to the mean" in the sense that "extreme" demand curves at the boundary of the feasible set become less likely, as they require bigger clusters of extreme slopes.

[43] The model was also tested with random smooth nonlinear demands, including demand curves with their slopes varying on the entire interval $[k_x, K_x]$, such as

$$D(x) = \frac{1}{2}\left[(K_x + k_x) \pm (K_x - k_x)\frac{x}{\bar{x}}\right][\bar{x} - x]_+.$$

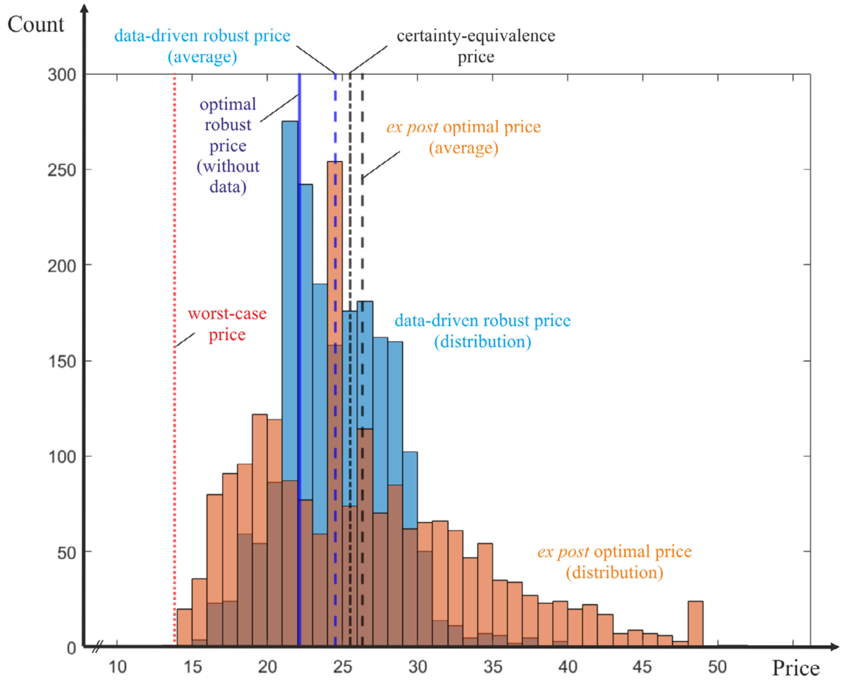
The results have been omitted, as they were even better than for the presented piecewise linear demands (which feature unbounded second derivatives).

Figure 11. Performance ratio versus *ex post* optimal profit: (a) linear model; (b) data-driven nonlinear model



Notes: For worst-case (WC) pricing, certainty-equivalence (CE) pricing, and optimal robust pricing. $N = 2{,}000$ demand-curve realizations.

Figure 12. Distribution of the optimal robust price and the *ex post* optimal price in the data-driven nonlinear model



Notes: Optimal robust price is contingent on up to $L = 5$ demand measurements. $N = 2{,}000$.

underlying reason is that the demand observations generate price adjustments that render the distribution of the optimal robust price similar to the distribution of the *ex post* optimal price (see Figure 12).

Table 1 features a comparison of the respective prices, profits, and performance indices for the various solutions. In terms of head-to-head comparison, for the nonlinear model robust pricing (i.e., choosing $\widehat{x}^*$) yields a higher profit than worst-case pricing in 88.9 percent of all demand-curve realizations (see Remark 8), and a higher profit than certainty-equivalence pricing only in 41.8 percent of all realizations. While the latter record might look quite poor at first sight, it is important to note that the robust pricing almost triples the performance index, from 23.81 percent to 61.89 percent, while on average the profit decreases by less than 2 percent (from 1204.07 to 1180.81). That is, the absolute performance gains obtained from effectively ignoring uncertainty are slight, whereas the relative losses (which occur about 38.2 percent of the time) may be quite devastating.

In the data-driven model, when the robust price is made contingent on a limited number of noisy demand measurements, the robust price increases

Table 1. Comparison of (average) price, profit, and performance index for the different solutions: worst-case, optimal robust, certainty-equivalence, and *ex post* optimal

Model	Linear	Nonlinear	Data-driven
Price			
x^*_{WC}	13.83	13.83	13.83
$\widehat{x}^*$	22.11	22.11	24.55[a]
x^*_{CE}	25.5	25.5	25.5
$\mathbb{E}[x(\tilde{\theta})]$	28.12	26.33	26.33
Profit			
$\mathbb{E}[\pi(x^*_{\mathrm{WC}}, \tilde{\theta})]$	931.18	930.67	930.67
$\mathbb{E}[\pi(\widehat{x}^*, \tilde{\theta})]$	1,183.21	1,180.81	1246.45[a]
$\mathbb{E}[\pi(x^*_{\mathrm{CE}}, \tilde{\theta})]$	1,207.48	1,204.07	1,204.07
$\mathbb{E}[\pi^*(\tilde{\theta})]$	1,352.17	1,298.98	1,298.98
Performance index[b]			
ρ_{WC}	0.3849[c]	0.4106	0.4106
ρ^*	0.5838[c]	0.6189	0.7145[a]
ρ_{CE}	0.1736[c]	0.2381	0.2381

Notes: $N = 2{,}000$; $L = 5$. [a] Values contingent on L demand observations; results averaged over N demand-curve realizations. [b] Values averaged over N demand-curve realizations, unless indicated otherwise. [c] Theoretical performance guarantees: $\rho_{\mathrm{WC}} = \varphi(\theta_1, \theta_2)$; ρ^* obtained from equation (32); $\rho_{\mathrm{CE}} = \varphi(\overline{\theta}, \theta_1)$; φ specified in equation (A15).

on average (from 22.11 to 24.55, close to the certainty-equivalence price of 25.5), with an average absolute performance (of 1,246.45) that is close to the mid-point of the average certainty-equivalence profit (1,204.07, the highest expected profit achievable by any *ex ante* constant price) and the optimal *ex post* profit (of 1,298.98) obtained with perfect information about demand. The data-driven model outperforms the robust model without demand observations 77.35 percent of the time, it is better than worst-case pricing in 92.5 percent of all realizations, and improves over certainty-equivalence pricing in 66.95 percent of all observations. The observed performance index of 71.45 percent means that over all $N = 2{,}000$ demand-curve realizations the obtained profit was within 28.55 percent of the *ex post* optimal profit.

5. Conclusion

The presented robust pricing method provides, by construction, the best relative performance guarantee for any given family of demand functions that satisfies a monotonicity property and a complementarity property. The complementarity property was used to establish the simple envelope representation of the performance index (see Proposition 7). Based on the latter, we obtain a characterization of the optimal robust solution (see

Proposition 8), which captures the intuition that this solution achieves a balanced performance with respect to realizations at either extreme of the parameter space. By sharp contrast, the worst-case solution depends only on the lower bound of the parameter space and is therefore insensitive to its size and thus also to the amount of parameter uncertainty in the problem. In the linear case, the worst-case solution leads to a marked decrease in expected profits compared with the proposed relatively robust approach, given a uniform (Laplacian) belief as reference distribution (see Proposition 10 and Remark 8). However, the minimax regret solution, while potentially attractive in terms of expected profit, can produce generically zero-profit outcomes, which would lead to a zero performance index.

Overall, robust pricing leads to lower prices, which also means that demand uncertainty increases consumer surplus when the firm prices robustly. However, as our examples in Section 4 illustrate, it is remarkable that a slight price decrease may be all that is needed to substantially increase the relative performance index (in our case from 23.81 percent to 71.45 percent) for only a small decrease in expected profits (by 1.93 percent, from 1,204.07 to 1,180.81 in the example). In this, our optimal robust price tends to be substantially larger than the worst-case price (for which the relative performance in the example is 41.06 percent, with an expected profit of only 930.67); this implies better absolute performance in most cases, while at the same time guaranteeing the overall best relative performance. We also show that it is possible to further increase the performance of the optimal robust price by making it "data-driven"; that is, contingent on the outcome of a limited number of (generally noisy) price experiments which produces a distribution of pricing decisions that mimics the distribution of the *ex post* optimal full-information price, as shown in Figure 12.

Relative performance guarantees, notably immune to multiplicative noise, are well adapted to human decision-making and correspond naturally to the mindset of an investor who seeks to maximize the relative returns of however much is the invested amount. The performance guarantee becomes better as the information about the unknown parameter improves, conversely also implying a "price of robustness" when guarding against more contingencies by allowing for a larger parameter space (see Proposition 9). Overall, a relative performance guarantee benchmarks a firm's profit performance against its best possible pricing decision, thus endogenously ratcheting up and down expectations depending on the strength of the realized demand (and doing so before the realization occurs).

Appendix. Proofs

Proof of Proposition 1: Assume that $(x, \theta) \in \mathbb{R}_{++} \times \Theta$ is such that $D(x, \theta) > 0$. From the definition of the price elasticity in equation (1) we obtain

$$\varepsilon_\theta(x,\theta) = -\frac{\partial}{\partial\theta}\left(\frac{xD_x(x,\theta)}{D(x,\theta)}\right)$$
$$= -\frac{x}{D^2(x,\theta)}\left(D(x,\theta)D_{x\theta}(x,\theta) - D_x(x,\theta)D_\theta(x,\theta)\right).$$

Hence, by virtue of the monotonicity property D1 and the supermodularity property D2 the claim follows immediately. □

Proof of Proposition 2: Note first that for zero output the firm's marginal cost weakly exceeds its average cost, because by (C1), together with equations (4) and (5), we have

$$MC(0) = C'(0) = AC(0). \tag{A1}$$

Consider now the function

$$\eta(q) = qC'(q) - C(q), \qquad q \geq 0.$$

By cost convexity (C1) it is non-decreasing, as

$$\eta'(q) = qC''(q) \geq 0, \qquad q \geq 0.$$

But because by the possibility of inaction (C2) it is $\eta(0) = 0$, this yields $\eta(q) \geq 0$, for all $q \in \mathbb{R}_+$, whence

$$MC(q) = C'(q) \geq \frac{C(q)}{q} = AC(q), \qquad q > 0, \tag{A2}$$

concluding our proof, as equations (A1) and (A2) together imply equation (6). □

Proof of Proposition 3: Fix a demand parameter $\theta \in \Theta$.

(i) Consider the rational choice set $\mathcal{X}_0(\theta)$ in equation (11). Note first that $\mathcal{X}_0(\theta)$ is non-empty because by assumption (F') it is $\bar{x} \geq MC(0) = MC(D(\bar{x},\theta))$ in (F), so

$$\bar{x} \in \mathcal{X}_0(\theta).$$

When the price x is increased from 0 to $\bar{x}$, the corresponding induced marginal cost $MC(D(x,\theta))$ cannot increase – as a consequence of the demand monotonicity in (D1) and the cost convexity in (C1). Thus, if a given price $x' \in \mathcal{X} = [0,\bar{x}]$ exceeds its induced marginal cost, then any price $x'' \in \mathcal{X}' = [x',\bar{x}]$ also exceeds its induced marginal cost:

$$x', x'' \in \mathcal{X}, C(D(x',\theta)) \leq x' \leq x'' \Rightarrow C(D(x'',\theta)) \leq C(D(x',\theta)) \leq x''.$$

In other words,

$$x \in \mathcal{X}_0(\theta) \quad \Rightarrow \quad [x,\bar{x}] \subset \mathcal{X}_0(\theta).$$

The largest lower bound of $\mathcal{X}_0(\theta)$ is therefore

$$x_0(\theta) = \inf\{x \in \mathcal{X} : x \geq MC(D(x,\theta))\}.$$

If we denote by $\gamma(x,\theta) = x - MC(D(x,\theta))$ the firm's (absolute) profit margin (for all $x \in \mathcal{X}$), then

$$\gamma(0,\theta) = -MC(D(0,\theta)) \leq 0 \leq \bar{x} - MC(D(\bar{x},\theta)) = \bar{x} - C'(0) = \gamma(\bar{x},\theta).$$

The function $\gamma(\cdot,\theta)$ is continuously differentiable. By the intermediate value theorem (see, e.g., Rudin, 1976, p. 93), there exists a value $x_0(\theta) \in \mathcal{X}$ such that $\gamma(x_0(\theta),\theta) = 0$, so

$$x_0(\theta) = MC(D(x_0(\theta),\theta)) \in \mathcal{X}_0(\theta). \tag{A3}$$

Because $\gamma_x(x,\theta) = 1 - D_x(x,\theta) > 0$ for all $x \in \mathcal{X}$, the function $\gamma(\cdot,\theta)$ is increasing on $\mathcal{X}$, which implies that $x_0(\theta)$ in equation (A3) is uniquely determined. Finally, because $D(x,\theta) \geq 0$ for all $x \in \mathcal{X}$ and $MC(\cdot)$ is non-decreasing by equation (C1), we have that $MC(D(x,\theta)) \geq MC(0)$ for all $x \in \mathcal{X}$ and therefore also $x_0(\theta) \geq C'(0)$, which yields our claim in part (i).

(ii) To begin with, charging a zero price is never optimal for the firm, i.e., $0 \notin \mathscr{X}(\theta)$. Indeed, as a consequence of (C2) and (F) it is

$$\pi(0,\theta) = -C(D(0,\theta)) < 0 = \pi(\bar{x},\theta).$$

Moreover, because demand does not vanish for all prices, there exists a price

$$\xi = \inf\{x \in [MC(0),\bar{x}] : D(x,\theta) = 0\} \in (0,\bar{x}).$$

Thus, taking the left-sided limit,

$$\begin{aligned}\lim_{x\to\xi^-} \pi_x(\xi,\theta) &= D(\xi,\theta) + (\xi - C'(D(\xi,\theta)))\, D_x(\xi^-,\theta)\\ &= (\xi - MC(0))\, D_x(\xi^-,\theta) \leq 0,\end{aligned}$$

we obtain that the gradient π_x either vanishes or is negative at (ξ^-,θ). In case it is negative, the optimal profit must be positive (as $\pi(x,\theta) > 0$ in a left-neighborhood of $x = \xi$). In case the gradient vanishes, the optimal profit must be zero (at zero demand), so any price $x \in [\xi,\bar{x}]$ is optimal. In either case, an optimum does arise at the interior of $(0,\bar{x})$, so that the necessary optimality condition,

$$\pi_x(x,\theta) = (1 - \varepsilon(x^-,\theta)\mu(x^-,\theta))\, D_x(x^-,\theta) = 0,$$

must be satisfied for all $x \in \mathscr{X}(\theta) \subset (0, \bar{x}]$. If the optimal profit $\pi^*(\theta)$ is positive, then demand at an optimal price $x \in \mathscr{X}(\theta)$ must be positive. Hence, by demand monotonicity (D1) it must be that $D_x(x, \theta) < 0$, which implies $\varepsilon(x, \theta) > 0$, so

$$\mu(x, \theta) = \frac{x - MC(D(x, \theta))}{x} = \frac{1}{\varepsilon(x, \theta)} > 0.$$

But this means that $x > MC(D(x, \theta))$, so necessarily

$$x \in (x_0(\theta), \xi) \subset \operatorname{int} \mathcal{X}_0(\theta),$$

which establishes the claim in part (ii).

This completes our proof. □

Proof of Proposition 4: Fix the demand parameter $\theta \in \Theta$.

(i) By the Weierstrass extreme value theorem (see, e.g., Bertsekas, 1995, p. 540), the continuous profit function $\pi(\cdot, \theta)$ attains its extrema on the compact set $\mathcal{X} = [0, \bar{x}]$ for any given $\theta \in \Theta$. That is, the set $\mathscr{X}(\theta)$ in equation (7) is non-empty.

(ii)/(iii) Both claims follow directly from the maximum theorem by Berge (1963, p. 116).

This concludes the proof. □

Proof of Proposition 5: The proof has two parts (I and II). In part I, we establish the supermodularity of the firm's profit for all $(x, \theta) \in \mathcal{X}_0(\theta) \times \Theta$ (provided demand is positive), and from there we obtain in part II the monotonicity of the selected solution $x(\theta) \in \mathscr{X}(\theta)$, for all $\theta \in \Theta$.

Part I. Let $\theta \in \Theta$ and $x \in \mathcal{X}_0(\theta)$. By differentiating the firm's profit with respect to x and θ (at the point (x, θ)) we obtain

$$\pi_{x\theta} = D_\theta + (x - MC(D))\, D_{x\theta} - C''(D) D_x D_\theta. \tag{A4}$$

Because, by assumption, $x \in \mathcal{X}_0(\theta)$, by equation (11) it is $x \geq MC(D(x, \theta))$. Hence, the demand monotonicity in (D1), demand complementarity in (D2), and cost convexity in (C1) together imply that all terms are non-negative (with at least the first term on the right-hand side of equation (A4) being positive), so

$$D(x, \theta) > 0 \quad \Rightarrow \quad \pi_{x\theta}(x, \theta) > 0. \tag{A5}$$

This establishes the claimed (local) supermodularity of π in regions of positive demand.

Part II.[44] Fix $\theta', \theta'' \in \Theta$ with $\theta' < \theta''$. Furthermore, let $x' = x(\theta')$ and $x'' = x(\theta'')$, where $x(\theta) = \min \mathscr{X}(\theta)$ for all $\theta \in \Theta$. By hypothesis, x'' maximizes $\pi(\cdot, \theta'')$, so

$$0 \geq \pi(x', \theta'') - \pi(x'', \theta'').$$

But this implies that

$$0 \geq \pi(x', \theta'') - \pi(\min\{x', x''\}, \theta'').$$

Indeed, for $x' \leq x''$ the right-hand side of the last inequality vanishes, whereas for $x' \geq x''$ the inequality is simply a consequence of the fact that $x'' \in \mathscr{X}(\theta'')$. Thus, by the supermodularity of the firm's profit π established in part I, we have

$$0 \geq \pi(x', \theta'') - \pi(\min\{x', x''\}, \theta'') \geq \pi(x', \theta') - \pi(\min\{x', x''\}, \theta'), \;\; (= 0),$$

which implies that $\min\{x', x''\} \in \mathscr{X}(\theta')$. But if $x'' < x'$, then $x' \neq \min \mathscr{X}(\theta')$, which is a contradiction to the assumption that the firm always uses the minimum of all optimal prices. Hence, it must be the case that $\min\{x', x''\} = x'$, whence

$$x' = x(\theta') \leq x(\theta'') = x'',$$

as claimed.[45] □

Proof of Proposition 6: The proof has two parts (I and II). In part I, we establish the log-supermodularity of the firm's profit for all $(x, \theta) \in \mathcal{X}_0(\theta) \times \Theta$ (provided demand is positive), and from there part II shows the claimed single-crossing property of the slope of the performance index φ.

Part I. Let $\theta \in \Theta$ and $x \in \mathcal{X}_0(\theta)$. Our aim is to show

$$(\log \pi(x, \theta))_{x\theta} = \left(\frac{\pi_x(x, \theta)}{\pi(x, \theta)} \right)_\theta = \frac{\pi(x, \theta)\pi_{x\theta}(x, \theta) - \pi_x(x, \theta)\pi_\theta(x, \theta)}{(\pi(x, \theta))^2} \geq 0. \tag{A6}$$

[44]The monotonicity of $\mathscr{X}(\theta)$ in the Veinott strong set order is implied by Milgrom and Shannon (1994, theorem 4); for convenience, we provide a simple proof following the reasoning by Topkis (1968).

[45]If the firm always chooses the maximum of all optimal prices (i.e., if $x(\theta) = \max \mathscr{X}(\theta)$), $x' \in \mathscr{X}(\theta')$ implies $\pi(x', \theta') - \pi(\min\{x', x''\}, \theta') \geq 0$, and in particular $\pi(\max\{x', x''\}, \theta') - \pi(x'', \theta') \geq 0$. Hence, by the supermodularity of the firm's profit π in part I:

$$(0 =) \quad \pi(\max\{x', x''\}, \theta'') - \pi(x'', \theta'') \geq \pi(\max\{x', x''\}, \theta') - \pi(x'', \theta') \geq 0,$$

which implies $\max\{x', x''\} \in \mathscr{X}(\theta'')$. However, because $x'' = \max \mathscr{X}(\theta'')$, it is necessarily $x'' = \max\{x', x''\}$, so that $x' = x(\theta') \leq x(\theta'') = x''$, yielding the monotonicity of $x(\cdot)$ on Θ.

By equation (A4) it is

$$\pi_{x\theta} = D_\theta + (x - MC(D))\, D_{x\theta} - C''(D) D_x D_\theta$$
$$= \left(1 + \varepsilon \frac{C''(D)D}{x}\right) D_\theta - \mu (D\varepsilon)_\theta,$$

where $\varepsilon = -xD_x/D$ is the demand elasticity and $\mu = (x - MC(D))/x$ the relative markup. Equivalently,

$$\pi_{x\theta} = \left(1 - \varepsilon\mu + \varepsilon \frac{C''(D)D}{x}\right) D_\theta - \mu D \varepsilon_\theta.$$

Hence,

$$\pi\pi_{x\theta} = (x - AC(D)) \left[1 - \varepsilon\mu + \varepsilon \frac{C''(D)D}{x} - \frac{\mu D \varepsilon_\theta}{D_\theta}\right] DD_\theta. \tag{A7}$$

However,

$$\pi_x \pi_\theta = (D + (x - MC(D))D_x)(x - MC(D))D_\theta$$
$$= (x - MC(D))(1 - \varepsilon\mu) DD_\theta. \tag{A8}$$

Thus, combining equations (A7) and (A8) yields

$$\pi\pi_{x\theta} - \pi_x\pi_\theta = (MC(D) - AC(D))(1 - \varepsilon\mu)DD_\theta + (x - AC(D))$$
$$\times \left(\varepsilon \frac{C''(D)D}{x} - \frac{\mu D \varepsilon_\theta}{D_\theta}\right) DD_\theta,$$

where $\varepsilon_\theta \leq 0$ as a consequence of Proposition 1. In contrast, by Proposition 3 we know that without any loss of generality the price x can be considered to lie in the firm's rational choice set $\mathcal{X}_0(\theta)$. Thus, we have $x \geq MC(D(x, \theta))$, whence

$$\pi\pi_{x\theta} - \pi_x\pi_\theta \geq (MC(D) - AC(D)) \left[1 - \varepsilon\mu - \frac{\mu D \varepsilon_\theta}{D_\theta} + \varepsilon \frac{C''(D)D}{x}\right] DD_\theta.$$

Note further that

$$-\varepsilon\mu - \frac{\mu D \varepsilon_\theta}{D_\theta} = \mu \left(\frac{xD_x}{D} - \left(\frac{xD_x}{D} - \frac{xD_{x\theta}}{D_\theta}\right)\right) = \mu \frac{xD_{x\theta}}{D_\theta} \geq 0.$$

As a result,

$$\pi\pi_{x\theta} - \pi_x\pi_\theta \geq (MC(D) - AC(D)) \left[1 + \mu \frac{xD_{x\theta}}{D_\theta} + \varepsilon \frac{C''(D)D}{x}\right] DD_\theta \geq 0,$$

where $MC(D) \geq AC(D)$ by Proposition 2, $C'' \geq 0$ by (D1), and $\varepsilon_\theta < 0$ (when D is positive) by Proposition 1.[46] This establishes the claim in equation (A6) and in turn implies that $\log \pi(x, \theta)$ is supermodular; that is, for any $x', x'' \in \mathcal{X}_0(\theta_1) \supset \mathcal{X}_0(\theta)$ with $x' \leq x''$, for all $\theta \in \Theta$, and $\theta', \theta'' \in \Theta$ with $\theta' < \theta''$, it is

$$[\log \pi(x'', \theta')] - [\log \pi(x', \theta')] \leq [\log \pi(x'', \theta'')] - [\log \pi(x', \theta'')]\,,$$

or, equivalently,

$$\pi(x'', \theta')\pi(x', \theta'') \leq \pi(x'', \theta'')\pi(x', \theta'), \tag{A9}$$

as long as all the profits are positive. This implies the validity of equation (A6).

Part II. Let $\widehat{\theta} \in \operatorname{int} \Theta$, and consider the behavior of $\pi(\widehat{\theta}, \cdot)$ to the left of $\widehat{\theta}$. If $\widehat{\theta}$ lies at the lower boundary of Θ (i.e., if $\widehat{\theta} = \theta_1$), then there is nothing to show. Consider, therefore, the interesting case where $\widehat{\theta} > \theta_1$. In that case, fix $\theta', \theta'' \in \Theta$ with $\theta' < \theta'' < \widehat{\theta}$, and accordingly set $x' = x(\theta')$, $x'' = x(\theta'')$, and $\widehat{x} = x(\widehat{\theta})$. By Proposition 5 it is $x' \leq x'' \leq \widehat{x}$. Thus, by the log-supermodularity of π in equation (A9) (applied to x'' and $\widehat{x}$ instead of x' and x''):

$$\begin{aligned}\varphi(\widehat{\theta}, \theta'') &= \frac{\pi(\widehat{x}, \theta'')}{\pi(x'', \theta'')} \geq \frac{\pi(\widehat{x}, \theta')}{\pi(x'', \theta')} \\ &= \frac{\pi(\widehat{x}, \theta')}{\pi(x', \theta')} \frac{\pi(x', \theta')}{\pi(x'', \theta')} \geq \frac{\pi(\widehat{x}, \theta')}{\pi(x', \theta')} = \varphi(\widehat{\theta}, \theta').\end{aligned}$$

Here we have used the fact that by the optimality of x' (i.e., because $x' \in \arg\max_x \pi(x, \theta')$), we have that

$$\frac{\pi(x', \theta')}{\pi(x'', \theta')} \geq 1.$$

Thus, $\varphi(\widehat{\theta}, \cdot)$ is non-decreasing to the left of $\widehat{\theta}$.

Ignoring the trivial situation where $\widehat{\theta} = \theta_2$, we consider now the case where $\widehat{\theta} < \theta_2$, and we assume that $\theta', \theta'' \in \Theta$ with $\widehat{\theta} < \theta' < \theta''$ are given, so that $\widehat{x} \leq x' \leq x''$ by Proposition 5. Again, by the log-supermodularity of π in equation (A9) it is

$$\begin{aligned}\varphi(\widehat{\theta}, \theta') &= \frac{\pi(\widehat{x}, \theta')}{\pi(x', \theta')} \geq \frac{\pi(\widehat{x}, \theta'')}{\pi(x', \theta'')} \\ &= \frac{\pi(\widehat{x}, \theta'')}{\pi(x'', \theta'')} \frac{\pi(x'', \theta'')}{\pi(x', \theta'')} \geq \frac{\pi(\widehat{x}, \theta'')}{\pi(x'', \theta'')} = \varphi(\widehat{\theta}, \theta''),\end{aligned}$$

[46] In addition, the demand monotonicity in (D1) implies that the price elasticity ε in equation (1) is positive.

where the last inequality following from the fact that $\pi(x'', \theta'')/\pi(x', \theta'') \geq 1$. This means that $\varphi(\widehat{\theta}, \cdot)$ is non-increasing to the right of $\widehat{\theta}$, concluding our proof. $\square$

Proof of Proposition 7: We know that $\varphi(\widehat{\theta}, \theta)$ attains its maximum of 1 for $\theta = \widehat{\theta}$. From Proposition 6 we have that $\varphi(\widehat{\theta}, \theta)$ is non-decreasing in θ to the left of $\widehat{\theta}$ and non-increasing to the right of $\widehat{\theta}$, so that the minimum of $\varphi(\widehat{\theta}, \theta)$ in θ must occur at the boundary of the domain. This implies that

$$\arg \min_{\theta \in [\theta', \theta'']} \varphi(\widehat{\theta}, \theta) \subset \{\theta', \theta''\},$$

for any $\theta', \theta'' \in \Theta$ with $\theta_1 \leq \theta' < \theta'' \leq \theta_2$. Setting $\theta' = \theta_1$ and $\theta'' = \theta_2$ completes our proof. $\square$

Proof of Proposition 8: Fix the demand parameter $\theta \in \Theta$. The proof has three parts (I, II, and III). Part I establishes opposing monotonicities for the boundary performance ratios $\varphi(\cdot, \theta_1)$ and $\varphi(\cdot, \theta_2)$. Based on this, part II finds that the boundary performance ratios must be equal for at least one parameter value, $\widehat{\theta}^* \in \Theta$. Finally, in part III, it is shown that the optimal performance index is achieved exactly for those parameter values where the difference of the boundary performance ratios vanishes.

Part I. We first show that $\varphi(\widehat{\theta}, \theta_1)$ is non-increasing and $\varphi(\widehat{\theta}, \theta_2)$ is non-decreasing in $\widehat{\theta}$. Because, by Proposition 5, the optimal price $\widehat{x} = x(\widehat{\theta})$ is non-decreasing in $\widehat{\theta}$, the desired monotonicities obtain immediately if

$$\pi_x(\widehat{x}, \theta_1) \leq 0 \leq \pi_x(\widehat{x}, \theta_2), \qquad \widehat{x} \in [x_1, x_2]. \tag{A10}$$

However, because $\pi_x(\widehat{x}, \theta) = (1 - \varepsilon(\widehat{x}, \theta)\mu(\widehat{x}, \theta))D(\widehat{x}, \theta)$ and

$$\mu_x(x, \theta) = \frac{MC(\widehat{x}, \theta) + \varepsilon(\widehat{x}, \theta)D(\widehat{x}, \theta)C''(D(\widehat{x}, \theta))}{x^2} \geq 0, \quad \widehat{x} \in [x_1, x_2],$$

a sufficient condition for equation (A10) is

$$\varepsilon_x(\widehat{x}, \theta) \geq 0, \quad \widehat{x} \in [x_1, x_2].$$

In that case $\varepsilon(\widehat{x}, \theta)\mu(\widehat{x}, \theta)$ is non-decreasing in $\widehat{x} \in [x_1, x_2]$.[47] For any $\widehat{x} \in (x_1, x_2)$, it is

[47] The monotonicity of $\varepsilon\mu$ is sufficient (see Remark 5). A careful examination reveals that even equation (A10) is not necessary, as $\widehat{x}$ needs to vary only in the (compact) union of the maximizers, $\widehat{\mathscr{X}} = \cup_{\theta \in \Theta} \mathscr{X}(\theta)$. The family of demands, $\{D(x, \theta) = [10\theta - 2x - (1/3)\sin(5x)]_+ : (x, \theta) \in \mathbb{R}_+ \times \Theta\}$, with $\Theta = [1, 2]$, which was used to generate Figures 1–5 (with $C(q) \equiv q$), illustrates this point. Even though neither $\pi_x(\widehat{x}, \theta_1)$ nor $\pi_x(\widehat{x}, \theta_2)$ are sign-definite for all $\widehat{x} \in [x_1, x_2]$, both gradients in this example satisfy the inequalities in equation (A10) on the non-convex set $\widehat{\mathscr{X}}$; and consequently the performance index ρ has the envelope representation in equation (18); see Figures 4 and 5.

$$\lim_{\hat{x}\to x_1^+}(1-\varepsilon(\hat{x},\theta_1)\mu(\hat{x},\theta_1)) = \lim_{\hat{x}\to x_2^-}(1-\varepsilon(\hat{x},\theta_2)\mu(\hat{x},\theta_2)) = 0,$$

as, for all interior optima, the inverse-elasticity rule in equation (9) necessarily applies. Therefore, by the continuity of $\varepsilon\mu$ and the fact that $\hat{x} = x(\hat{\theta})$, we have

$$\begin{aligned}\pi_x(x(\hat{\theta}),\theta_1) = {} & [(1-\varepsilon\mu)\,D]|_{(x(\hat{\theta}),\theta_1)} \le 0 \le [(1-\varepsilon\mu)\,D]|_{(x(\hat{\theta}),\theta_2)} \\ & = \pi_x(x(\hat{\theta}),\theta_2),\end{aligned} \tag{A11}$$

which shows that $\varphi(\hat{\theta},\theta_1)$ is non-increasing and $\varphi(\hat{\theta},\theta_2)$ is non-decreasing in $\hat{\theta}\in\Theta$.

Part II. Consider the difference of the two boundary performance ratios,

$$\Delta(\hat{\theta}) = \varphi(\hat{\theta},\theta_2) - \varphi(\hat{\theta},\theta_1), \qquad \hat{\theta}\in\Theta, \tag{A12}$$

which is non-decreasing by our result in part I. Note further that $\Delta(\theta_1) = \varphi(\theta_1,\theta_2) - 1 \le 0$ and $\Delta(\theta_2) = 1 - \varphi(\theta_2,\theta_1) \ge 0$. As $\Delta(\cdot)$ is continuous, by the intermediate value theorem (Rudin, 1976, p. 93), there exists $\hat{\theta}^*\in\Theta$ such that $\Delta(\hat{\theta}^*) = 0$.

Part III. By the envelope representation in Proposition 7, the performance index can be written in the form

$$\rho(\hat{\theta}) = \min\left\{\varphi(\hat{\theta},\theta_1),\varphi(\hat{\theta},\theta_2)\right\}, \qquad \hat{\theta}\in\Theta,$$

or equivalently,

$$\rho(\hat{\theta}) = \min\left\{0,\Delta(\hat{\theta})\right\} + \varphi(\hat{\theta},\theta_1) = \min\left\{-\Delta(\hat{\theta}),0\right\} + \varphi(\hat{\theta},\theta_2), \qquad \hat{\theta}\in\Theta.$$

However, this means that

$$\rho(\hat{\theta}) = \begin{cases}\varphi(\hat{\theta},\theta_1), & \text{if } \Delta(\hat{\theta}) \ge 0,\\ \varphi(\hat{\theta},\theta_2), & \text{if } \Delta(\hat{\theta}) < 0.\end{cases}$$

Using our result in part II, $\rho(\hat{\theta})$ is therefore non-decreasing for $\hat{\theta}\in[\theta_1,\hat{\theta}^*]$ and non-increasing for $\hat{\theta}\in[\hat{\theta}^*,\theta_2]$. Thus, necessarily

$$\hat{\theta}^* \in \arg\max_{\hat{\theta}\in\Theta}\rho(\hat{\theta}),$$

so the optimal performance index is

$$\rho^* = \rho(\hat{\theta}^*).$$

We further notice that $\Delta(\hat{\theta}) = 0$ if and only if $\varphi(\hat{\theta},\theta_1) = \varphi(\hat{\theta},\theta_2)$. Hence,

$$\rho(\hat{\theta}) = \rho^* \quad \Leftrightarrow \quad \Delta(\hat{\theta}) = 0.$$

This yields the claimed representation of equation (19), concluding the proof. □

Proof of Proposition 9: From Proposition 8 we know that any optimal robust parameter $\widehat{\theta}^*$ satisfies equation (19). Differentiating this relation with respect to θ_i, for $i \in \{1, 2\}$, yields

$$\Delta'(\widehat{\theta}^*)\frac{\partial \widehat{\theta}^*}{\partial \theta_1} = \varphi_\theta(\widehat{\theta}^*, \theta_1) \tag{A13}$$

and

$$-\Delta'(\widehat{\theta}^*)\frac{\partial \widehat{\theta}^*}{\partial \theta_2} = \varphi_\theta(\widehat{\theta}^*, \theta_2), \tag{A14}$$

where $\Delta(\widehat{\theta}) = \varphi(\widehat{\theta}, \theta_2) - \varphi(\widehat{\theta}, \theta_1)$ is the difference of the boundary performance ratios. As in the proof of Proposition 8 where it was established that $\Delta(\widehat{\theta})$ is non-decreasing, one obtains $\Delta'(\widehat{\theta}) \geq 0$, for all $\widehat{\theta} \in \Theta$. Furthermore, by Proposition 6 we know that

$$\varphi_\theta(\widehat{\theta}^*, \theta_2) \leq 0 \leq \varphi_\theta(\widehat{\theta}^*, \theta_1),$$

so that equations (A13) and (A14) together imply that[48]

$$\frac{\partial \widehat{\theta}^*}{\partial \theta_i} \geq 0,$$

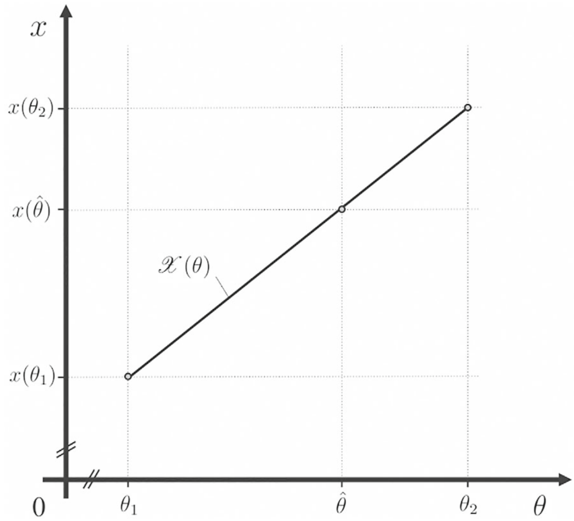
for $i \in \{1, 2\}$, concluding the proof. □

Proof of Proposition 10: To prove the result, which is formulated in the original two-parameter family of linear demand curves, we consider here the implied equivalent family of reduced linear (affine) demand curves in equation (28), which makes the model conform with our theoretical derivations in Sections 2 and 3, as explained in the main text. For any given reduced demand parameter $\theta \in \Theta$, the firm's (unique) optimal price $x(\theta)$ can be found using the inverse-elasticity rule in equation (9) (see Figure A1). Specifically, the solution set $\mathscr{X}(\theta)$ in the profit-maximization problem (7) becomes a singleton with element

$$x(\theta) = \frac{\theta + c}{2}, \qquad \theta \in \Theta.$$

[48]From the characterization of all optimal robust parameters $\widehat{\theta}$, by the relation $\Delta(\widehat{\theta}) = 0$ in Proposition 8, it is clear that $\Delta'(\widehat{\theta}^*) = 0$ implies that the gradient of $\widehat{\theta}^*$ with respect to θ_i also vanishes (for $i \in \{1, 2\}$).

Figure A1. Single-valued solution $\mathscr{X}(\theta) = \{x(\theta)\}$ to the profit-maximization problem in equation (7) for $\theta \in \Theta$ in the linear example



This price produces an optimal profit:

$$\pi^*(\theta) = \pi(x(\theta), \theta) = \frac{1}{4}(\theta - c)^2, \qquad \theta \in \Theta.$$

Thus, if the firm assumes a demand model according to the candidate parameter $\widehat{\theta} \in \Theta$ implying the choice of $x(\widehat{\theta})$ for its price, then the firm's profit takes on the form

$$\pi(x(\widehat{\theta}), \theta) = \frac{1}{4}\left(\widehat{\theta} - c\right)\left(2\theta - \left(\widehat{\theta} + c\right)\right),$$

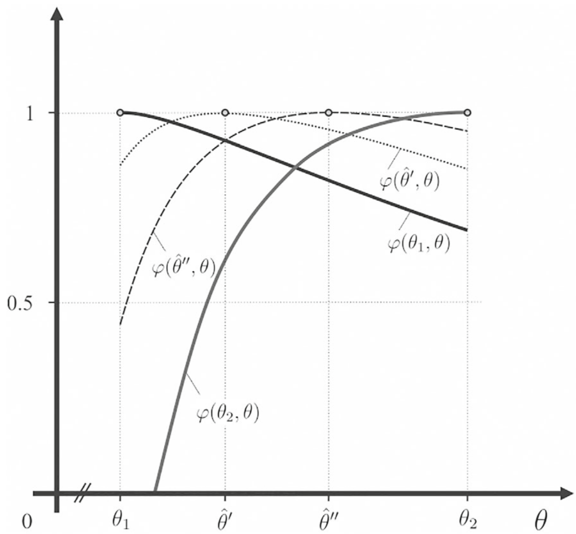
as a function of the actual demand realization, indexed by the testing parameter $\theta \in \Theta$. The corresponding performance ratio in equation (15) becomes

$$\varphi(\widehat{\theta}, \theta) = \frac{\pi(x(\widehat{\theta}), \theta)}{\pi^*(\theta)} = \frac{\left(\widehat{\theta} - c\right)\left(2\theta - \left(\widehat{\theta} + c\right)\right)}{(\theta - c)^2}, \quad \widehat{\theta}, \theta \in \Theta, \tag{A15}$$

leading to the performance index in equation (16), which by Proposition 7 has the envelope representation in equation (18) (see Figure A2). As the demand elasticity $\varepsilon(x, \theta)$ for the family of (reduced) linear demands is increasing in $x > 0$, by Proposition 8, at the optimal value of $\widehat{\theta} = \widehat{\theta}^*$:

$$\varphi(\widehat{\theta}^*, \theta_1) = \frac{2\theta_1 - \left(\widehat{\theta}^* + c\right)}{(\theta_1 - c)^2} = \frac{2\theta_2 - \left(\widehat{\theta}^* + c\right)}{(\theta_2 - c)^2} = \varphi(\widehat{\theta}^*, \theta_2).$$

Figure A2. Performance ratio $\varphi(\widehat{\theta}, \theta)$ for $\widehat{\theta} \in \{\theta_1, \widehat{\theta}', \widehat{\theta}'', \theta_2\}$ and $\theta \in \Theta$



This implies a unique optimal robust parameter,

$$\widehat{\theta}^* = \frac{\theta_1\theta_2}{\overline{\theta} - c}\left(1 - \frac{\overline{\theta}c}{\theta_1\theta_2}\right),$$

where $\overline{\theta} = (\theta_1 + \theta_2)/2$ denotes the barycenter of the parameter space Θ. Equation (30), together with the fact that the arithmetic average $\overline{\theta}$ weakly exceeds the geometric average $\sqrt{\theta_1\theta_2}$ (see, e.g., Bullen, 2003, chapter 2), implies that $c < \sqrt{\theta_1\theta_2} \leq \overline{\theta}$. Hence, the optimal performance index becomes

$$\rho^* = \rho(\widehat{\theta}^*) = \frac{\widehat{\theta}^* - c}{\overline{\theta} - c} = 1 - \frac{\overline{\theta}^2 - \theta_1\theta_2}{(\overline{\theta} - c)^2} = 1 - \left(\frac{\overline{\theta} - \theta_1}{\overline{\theta} - c}\right)^2 > 0, \quad c \in (0, \theta_1),$$

where we have used the fact that $\overline{\theta}^2 - \theta_1\theta_2 = (\overline{\theta} - \theta_1)^2$. The firm's optimal robust price is therefore

$$\widehat{x}^* = \frac{1}{2}\left(\widehat{\theta}^* + c\right) = \frac{\theta_1(\theta_2 - c)^2 - \theta_2(\theta_1 - c)^2}{(\theta_2 - c)^2 - (\theta_1 - c)^2} = \frac{1}{2}\frac{\theta_1\theta_2 - c^2}{\overline{\theta} - c}.$$

This concludes our proof. □

Acknowledgments

The author wishes to thank two anonymous referees, as well as participants of the 2023 Econometric Society conferences in Mumbai, Nairobi, and Sydney, for their helpful suggestions and comments. Open access funding was provided by École Polytechnique Fédérale de Lausanne.

References

Anscombe, F. J. and Aumann, R. J. (1963), A definition of subjective probability, *Annals of Mathematical Statistics 34*, 199–205.
Arkes, H. R. and Blumer, C. (1985), The psychology of sunk cost, *Organizational Behavior and Human Decision Processes 35*, 124–140.
Aubin, J.-P. and Ekeland, I. (1984), *Applied Nonlinear Analysis*, Wiley-Interscience, New York.
Bell, D. E. (1982), Regret in decision making under uncertainty, *Operations Research 30*, 961–981.
Ben-David, S. and Borodin, A. (1994), A new measure for the study of on-line algorithms, *Algorithmica 11*, 73–91.
Berge, C. (1963), *Topological Spaces*, Oliver and Boyd, Edinburgh. [Reprinted by Dover Publications, Mineola, NY, in 1997.]
Bergemann, D. and Schlag, K. (2011), Robust monopoly pricing, *Journal of Economic Theory 146*, 2527–2543.
Bertsekas, D. P. (1995), *Nonlinear Programming*, Athena Scientific, Belmont, MA.
Bertsimas, D. J., Jaillet, P., and Odoni, A. R. (1990), A Priori Optimization, *Operations Research 38*, 1019–1033.
Bullen, P. S. (2003), *Handbook of Means and Their Inequalities*, Kluwer, Dordrecht.
Chebyshev, P. L. (1859), Sur les questions de minima qui se rattachent à la représentation approximative des fonctions, *Mémoires de l'Académie Impériale des Sciences de St. Petersbourg, Sixième Série, Sciences Mathématiques et Physiques*, Vol. 7, 199–291. [Reprinted in A. Markoff and N. Sonin (eds) (1962), *Œvres de P. L. Tchebychef*, Chelsea Publishing, New York, 271–378.]
Deaton, A. (1986), Demand analysis, in Z. Griliches and M. D. Intriligator (eds), *Handbook of Econometrics*, Vol. III, Elsevier, Amsterdam, 1767–1839.
Friedman, M. (1949), The Marshallian demand curve, *Journal of Political Economy 57*, 463–495.
Gilboa, I. and Schmeidler, D. (1989), Maxmin expected utility with non-unique prior, *Journal of Mathematical Economics 18*, 141–153.
Goel, A., Meyerson, A., and Weber, T. A. (2009), Fair welfare maximization, *Economic Theory 41*, 465–494.
Groenewegen, P. D. (2008), Supply and demand, in S. N. Durlauf and L. E. Blume (eds), *The New Palgrave Dictionary of Economics*, Palgrave Macmillan, London.
Han, J. and Weber, T. A. (2023), Price discrimination with robust beliefs, *European Journal of Operational Research 306*, 795–809.
Handel, B. R. and Misra, K. (2015), Robust new product pricing, *Marketing Science 34*, 864–881.
Jayne, J. E. and Rogers, C. A. (2002), *Selectors*, Princeton University Press, Princeton, NJ.
Jenkin, F. (1870), The graphic representation of the laws of supply and demand, and their application to labour, in A. Grant (ed.), *Recess Studies*, Edmonston and Douglas, Edinburgh, 151–185.
Kahneman, D. and Tversky, A. (1984), Choices, values, and frames, *American Psychologist 39*, 341–350.

Kouvelis, P. and Yu, G. (1997), *Robust Discrete Optimization and Its Applications*, Springer, New York.
Leontief, W. W. (1941), *The Structure of American Economy 1919–1929: An Empirical Application of Equilibrium Analysis*, Harvard University Press, Cambridge, MA.
Lerner, A. P. (1934), The concept of monopoly and the measurement of monopoly, *Review of Economic Studies 1*, 157–175.
Lewis, T. R. and Sappington, D. E. M. (1988), Regulating a monopolist with unknown demand, *American Economic Review 78* (5), 986–998.
Lim, A. E. B., Shanthikumar, J. G., and Vahn, G.-Y. (2012), Robust portfolio choice with learning in the framework of regret: single-period case, *Management Science 58*, 1732–1746.
Marshall, A. (1920), *Principles of Economics*, 8th edn, Macmillan, Hampshire.
Mas-Colell, A., Whinston, M. D., and Green, J. R. (1995), *Microeconomic Theory*, Oxford University Press, Oxford.
Milgrom, P. and Shannon, C. (1994), Monotone comparative statics, *Econometrica 62*, 157–180.
Milnor, J. (1951), Games against nature, Research Memorandum RM-679, RAND Corporation, Santa Monica, CA.
von Neumann, J. and Morgenstern, O. (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ.
Perakis, G. and Roels, G. (2010), Robust controls for network revenue management, *Manufacturing and Service Operations Management 12*, 56–76.
Rosen, S. (1999), Potato paradoxes, *Journal of Political Economy 107* (S6), S294–S313.
Rudin, W. (1976), *Principles of Mathematical Analysis*, 3rd edn, McGraw-Hill, New York.
Savage, L. J. (1951), The theory of statistical decision, *Journal of the American Statistical Association 46*, 55–67.
Segal, I. (2003), Optimal pricing mechanisms with unknown demand, *American Economic Review 93* (3), 509–529.
Shekhtman, B. (1982), Why piecewise linear functions are dense in $C[0,1]$, *Journal of Approximation Theory 36*, 265–267.
Sleator, D. D. and Tarjan, R. E. (1985), Amortized efficiency of list update and paging rules, *Communications of the ACM 28*, 202–208.
Sweeney, J. L. (2002), *The California Electricity Crisis*, Hoover Institution Press, Stanford Institute for Economic Policy Research, Stanford, CA.
Tirole, J. (1988) *The Theory of Industrial Organization*, MIT Press, Cambridge, MA.
Topkis, D. M. (1968), Ordered optimal solutions, PhD Dissertation, Stanford University, Stanford, CA.
Topkis, D. M. (1998), *Supermodularity and Complementarity*, Princeton University Press, Princeton, NJ.
Ullrich, P. (1989), Weierstrass' Vorlesung zur 'Einleitung in die Theorie der analytischen Funktionen', *Archive for History of Exact Sciences 40*, 143–172.
Wald, A. (1945), Statistical decision functions which minimize the maximum risk, *Annals of Mathematics 46*, 265–280.
Weber, T. A. (2014), On the (non-)equivalence of IRR and NPV, *Journal of Mathematical Economics 52*, 25–39.
Weber, T. A. (2017), Global optimization on an interval, *Journal of Optimization Theory and Applications 182*, 684–705.

First version submitted March 2023;
final version received February 2024.