Robust Monopoly Pricing: The Case of Regret*

Dirk Bergemann† Karl Schlag‡

First Version: May 2003
This Version: June 2005

Abstract

We consider a robust version of the classic problem of optimal monopoly pricing with incomplete information. The robust version of the problem is distinct in two aspects: (*i*) the seller minimizes regret rather than maximizes revenue, and (*ii*) the seller only knows that the true distribution of the valuations is in a neighborhood of a given model distribution.

We characterize the robust pricing policy as the solution to a minimax problem for small and large neighborhoods. In contrast to the classic monopoly policy which is a single deterministic price, the robust policy is always a random pricing policy, or equivalently, a multi-item menu policy. The responsiveness of the robust policy to an increase in risk is determined by the curvature of the static profit function.

KEYWORDS: Monopoly, Optimal Pricing, Regret, Robustness
JEL CLASSIFICATION: C79, D82

*The first author gratefully acknowledges support by NSF Grants #SES-0095321, #CNS-0428422 and a DFG Mercator Research Professorship at the Center of Economic Studies at the University of Munich. We would like to thank Peter Klibanoff, Stephen Morris, David Pollard, Phil Reny, John Riley and Thomas Sargent for helpful suggestions. We are grateful to seminar participants at the California Institute of Technology, Columbia University, the University of California at Los Angeles, the University of Wisconsin and the Cowles Foundation Conference "Uncertainty in Economic Theory" at Yale University for many helpful comments. Sahar Malik provided excellent research assistance.

†Department of Economics, Yale University, 28 Hillhouse Avenue, New Haven, CT 06511, dirk.bergemann@yale.edu.

‡Department of Economics, European University Institute, schlag@iue.it

1 Introduction

1.1 Motivation

In this paper, we investigate a robust version of the classic monopoly problem of selling a product under incomplete information. The optimal pricing policy is the most elementary instance of a revenue maximizing design problem. In recent years, the theory of mechanism design has found increasingly widespread applications in the real world, favored partly by the growth of the electronic marketplace and trading platforms on the internet. Many selling procedures, such as auctions and exchanges implement key elements stemming from the theoretical models. Naturally, with an increase in the use of optimal design models, the performance of these mechanisms becomes an important issue.

Within the narrow context of the Bayesian models, the question of performance permits a straightforward answer. Given the prior of the designer, the optimal mechanism achieves the maximal feasible expected revenue or utility possible. The logical next step is then to ask how well the recommended policies perform outside of the narrow confines of the model given by the prior. There are currently two approaches to address this question. The first approach is often referred to as *worst case analysis* whereas the second approach is often referred to as *model misspecification* or *robust analysis*. The worst case analysis largely disregards any information contained in a prior or a set of priors. It does not compute expected revenues but focuses on the worst case performance of a policy. It then typically *compares* the worst case performance of the objective function to some ideal performance measure. The second approach begins with a model (distribution) of the world but acknowledges that the model might be misspecified. It then considers the policy which maximizes the true objective function of the designer for the worst possible distribution in the neighborhood of the model distribution. The later distribution is therefore referred to as worst case distribution or *worst case demand*.

In this paper, we investigate the performance of the optimal selling policy by enriching the canonical model in two key aspects. First, instead of a given true distribution regarding the valuations of the buyers, in our set-up the seller only knows that the true distribution is in a neighborhood of a given model distribution. The enlargement of the set of possible priors represents the model misspecification. Second, the objective function of the seller is formulated as a regret minimization rather than a revenue maximization problem. The regret is the difference between the actual valuation of the buyer for the object and the actual revenue obtained by the seller. The regret of the seller can be positive for two reasons: (i) the buyer has a low valuation relative to the price and hence does not purchase the object, or (ii) he has a high valuation relative to the price and hence the seller could have obtained a higher revenue. For a given neighborhood of possible distributions, we then characterize the pricing policy which minimizes maximal regret. The main objective of the paper is to describe how the robust policies depend on the model distribution and the size of the risk as represented by the size of the neighborhood. As part of the analysis, we also determine how the regret varies with the amount of risk faced by the seller.

By pairing the robustness analysis with the notion of regret rather than revenue we combine the attractive features of the worst case analysis with those of the robustness analysis. In particular, for any given neighborhood, the seller applies the information contained in the model distribution and its neighborhood in the regret minimization. In addition, at the worst case demand, the pricing policy which minimizes regret also maximizes revenue. Thus the regret minimization problem has a direct decision theoretic link to the original objective function of the seller, namely revenue maximization. We also show that for small neighborhoods, the regret minimizing policy is also a nearly optimal revenue maximizing policy for any given distribution among all distributions of the given neighborhood.

Yet, for large neighborhoods, in particular when the neighborhood contains the set of all possible distributions, the regret minimization problem still leads the seller to adopt a non-trivial policy. In contrast, a worst case revenue maximizing policy under a large neighborhood would simply set the price equal to the lowest possible realization. The resulting policy recommendation would hence be extremely pessimistic and would not suggest any trade-offs. In contrast, by solving the regret minimization problem, the seller weighs the benefit of lowering price to increase the probability of sale against the increased revenue conditional on making a sale to a high value customer.

The notion of regret contains a benchmark against which the realized revenue is measured and offers a trade-off which determines the optimal policy. In this respect, the notion of regret shares features with the notion of competitiveness central in the worst case analysis. In the context of the optimal selling problem studied here, the competitiveness of a policy is the ratio of actual revenue against the value of the buyer, where the valuation constitutes clearly the maximal feasible revenue. Yet, the notion of competitive ratio lacks a decision theoretic foundation and it is difficult to relate the competitive ratio to an associated Bayesian revenue maximizing problem.

We obtain the robust pricing policy as the solution to a saddlepoint problem in which the seller minimizes regret whereas malevolent nature maximizes regret. The resulting minimax regret problem thus represents a zero sum game between seller and nature. The buyers acts optimally given their valuations and the prices.

We begin our analysis with the case of large neighborhoods. With large neighborhoods, nature is not restricted by a model distribution and can choose any feasible distribution. In the equilibrium of the zero sum game, the optimal pricing policy for the seller has to resolve the conflict between the regret which arises with low prices against the regret associated with high prices. If she offers a low price, nature can cause regret with a distribution which puts substantial probability on high valuation buyers. On the other hand, if she offers a high price, nature can cause regret with a distribution which puts substantial probability at valuations just below the offered price. It then becomes evident that a single price will always expose the seller to substantial regret. Consequently, she can decrease her exposure by offering many prices. This can either be achieved by a random pricing policy or, alternatively, by a menu pricing policy. With a probabilistic pricing policy, the seller diminishes the likelihood that the nature will be able to cause large regret. Equivalently, the seller can offer a menu of prices and quantities. The quantity element in the menu can either represent a true quantity in the case of a divisible object or a probability of obtaining the indivisible object.

The intuition regarding the nature of the robust policy is easy to establish in comparison to the optimal revenue maximizing policy for a given distribution. An optimal policy for a given distribution of valuations is always to offer the entire object at a fixed price. In contrast a robust policy will offer many prices (with varying quantities). With a single price, the risk of missing a trade at a valuation just below the given price is substantial (especially in the presence of malevolent nature). On the other hand, if the seller were simply to lower the price, she would miss the chance of extracting revenue from higher valuation customers. She resolves this conflict by offering smaller trades at lower prices to the low valuation customers. The size of the trade is simply the probability by which a trade is offered or the quantity offered at a given price. In the game against nature, the seller will have to be indifferent between offering small and large trades. In terms of the virtual utility, the key notion in optimal mechanisms, this requires that the seller will receive zero virtual utility over a range of valuations. The resulting implication on the distribution of valuations determine the worst case demand chosen by nature. Similarly, the worst case demand by nature cannot be a pure strategy as the seller would then have a best response which would guarantee her zero regret. But for nature to choose a nongeneric distribution, it must be indifferent over a set of valuations. As nature seeks to maximize the regret, the indifference condition requires that the regret is constant across a set of valuations. In consequence, the density of the pricing policy has to be inverse proportional to the valuations. The indifference condition then fixes the random pricing policy of the seller in a transparent fashion. The complete equilibrium analysis also has to locate the support of the distribution and locate mass points, but the basic trade-offs determine the nature of pricing and worst case demand.

The derivation of the equilibrium is indeed elementary when nature is unconstrained and the neighborhood is assumed to be so large that it encompasses the set of feasible distributions. While the determination of the equilibrium turns into a substantially more challenging task for small neighborhoods, the basic elements in the pricing and demand strategies derived for large neighborhoods, will appear again for small model misspecifications. More precisely, the equilibrium strategies for small neighborhoods will be *scaled down* versions of the strategies for large neighborhoods. Indeed, an additional attractive feature of our robust policy is that an increase in risk will simply lead to a scale change in the policies by means of a change in the support of the distribution, but will leave the densities intact (up to constant factor).

We define the notion of a neighborhood through the usual metric of weak convergence, often called Prohorov metric. In the Prohorov metric two distributions are close to each other if they permit with *large* probability *small* changes in the valuations and with *small* probability *large* changes in the valuations. After describing robust policies for small and large neighborhoods in terms of regret minimization, we finally return to revenue maximization. Here we show that the robust policies resulting from regret minimization have a second and important link to revenue maximization. We already mentioned that in equilibrium, the regret minimizing policy is equivalent to the revenue maximizing policy at the *worst case demand*. The second link extends this relationship to small amounts of model specification in the following sense. The regret minimizing policy will be a nearly optimal revenue maximizing policy for every distribution in a small neighborhood. In this sense, the regret minimizing policy is also a robust policy for revenue maximization with model misspecification.

1.2 Related Literature

The basic ideas of robust analysis were first formalized in the context of statistical inference, in particular with respect to the classic Neyman-Pearson hypothesis testing. The statistical problem is to distinguish on the basis of a sample between two known alternative distributions. The model misspecification and consequent concern of robustness comes from the fact that each one of the two distributions might be misspecified. Huber (1964), (1965) first formalized robust estimation as the solution to a minmax problem and an associated zero sum game.[1] The basic ideas of robust statistics were introduced in the optimal control literature as robust control methods since the early '80s (see Zhou, Doyle & Glover (1996)). The techniques of robust control were recently introduced into dynamic macroeconomic models by Hansen and Sargent and their co-authors to investigate robust intertemporal decision making (see the survey by Hansen & Sargent (2004)).

The robust policy in our model is the result of a minmax regret problem. The seller could therefore also be interpreted as an ambiguity averse seller in the sense of Gilboa & Schmeidler (1989) if we were to maximize revenue rather than minimize regret. A recent paper by Bose, Ozdenoren & Pape (2004) investigates the nature of the optimal auction in the presence of an ambiguity averse seller as well as ambiguity averse bidders. As we consider a monopoly pricing problem the ambiguity aversion (or robustness concern) is of no consequence for the behavior of the buyers.

The worst case analysis and the notion of competitiveness is central in many optimal design problems analyzed in computer science (see specifically Goldberg, Hartline & Wright (2001) for an application to auctions with multiple goods and the recent survey to online design problems by Borodin & El-Yaniv (1998)). In auction theory, Neeman (2003) analyzes the competitiveness of the second price auction. A recent article by Prasad (2003) presents negative result, an in particular shows that the standard optimal pricing policy of the monopolist is not robust to small model misspecifications.

The idea of a minimax regret rule was advocated by Savage (1954) and appears to have originated in Wald (1950). A decision theoretic axiomatization of regret was provided by Milnor (1954). A recent contribution by Hayashi (2005) provides an axiomatization for minimax regret for a subset of priors. Stoye (2005) provides axioms which are necessary conditions for interim regret.[2] The notion of regret was investigated in the context of mechanism design by Linhart & Radner (1989) as well as by Selten (1989). Linhart & Radner (1989) analyze minimax regret strategies in a bilateral bargaining framework. In their framework, the valuation of the buyer and the cost function of the seller depend on a choice variable q which may represent quantity or quality. In contrast to the incomplete

[1] See also Huber (1981) and Hampel, Ronchetti, Rousseeuw & Stahel (1986) for more up-to-date surveys on robust statistics.

[2] Recently, the notions of regret and interim regret have also received attention in the econometrics literature, e.g. Chamberlain (2000) and Manski (2004).

information environment here, the bulk of the analysis in Linhart & Radner (1989) is concerned with bilateral trade under complete information. In addition, they largely restrict their analysis to deterministic strategies, even though mixed strategies will typically lead to lower regrets. In Selten (1989) first and second price auctions are considered under a modified form of regret for the bidders.

The reminder of the paper is organized as follows. In Section 2 we present the model, the notion of regret and the neighborhoods. In Section 3 we characterize the robust pricing policy with large risk. In Section 4, we characterize the robust policy in the presence of small risks. Here, an important special case is the Dirac function as a model distribution. In Section 5 we show that a policy which minimizes regret is also a nearly optimal revenue maximizing policy for *all* distributions contained in a small neighborhood of the model distribution. Section 6 provides a discussion of some related and open issues. Section 7 concludes. Section 8 collects auxiliary results and the proofs.

2 Model

Demand The seller faces a single potential buyer with value v for a unit of the object. The value v of the object is private information for the buyer and unknown to the seller. The marginal cost of production is constant and normalized to zero. The buyer wishes to buy at most one unit of the object but we allow for purchases of intermediate quantities $x \in (0,1)$. The net utility of the buyer of purchasing a quantity x at price p_x is given by

$$u(v, x, p_x) = vx - p_x.$$

With continuous allocations, the optimal policy of the seller extends from a price to menu policy.

The valuation v of the buyer is drawn from a probability distribution on $[0,1]$.

Robustness In contrast to the standard model of incomplete information, in our robust version the seller is uncertain about the true distribution over the buyer's valuations. The risk (or ambiguity in the language of Ellsberg (1961)) is represented by a model distribution $F_0(v)$ and the requirement that the true distribution $F_v(v)$ is in a neighborhood of the model distribution $F_0(v)$. The magnitude of the risk is quantified by the size of the neighborhood around the model distribution.

Neighborhoods We describe ε neighborhoods of the model distribution $F_0(v)$ by the Prohorov neighborhood, denoted by $\mathcal{P}_\varepsilon(F_0)$, and associated metric:

$$\mathcal{P}_\varepsilon(F_0) = \{F_v \,|F_v(A) \leq F_0(A^\varepsilon) + \varepsilon,\ \forall A\}, \tag{1}$$

where the set A^ε denotes the closed ε neighborhood of any Borel measurable set A. Formally, the set A^ε is given by

$$A^\varepsilon = \left\{ v \in [0,1] \;\middle|\; \inf_{y \in A} d(x,y) \leq \varepsilon \right\},$$

where $d(x,y) = |x-y|$ is the distance on the real line. The Prohorov metric has evidently two components. The additive term ε in (1) allows for a *small* probability of *large* changes in the valuations relative to the model distribution whereas the larger set A^ε permits *large* probabilities of *small* changes in the valuations. The notion of a Prohorov neighborhood is illustrated in the graphic below for a Dirac function δ_{v_0} at some v_0 with $0 < v_0 < 1$.

Insert Figure 1: The Prohorov Neighborhood of a Dirac Function

The distributions which are in the Prohorov neighborhood of the Dirac function δ_v are described by all those distribution function which place at least probability $1-\varepsilon$ in the ε neighborhood of the point v given by $[v-\varepsilon, v+\varepsilon]$. The remaining ε probability can be placed anywhere by nature.[3]

We shall refer to the case of $\varepsilon = 1$ as the case of *large risk*. With $\varepsilon = 1$, the neighborhood is not anchored by any model distribution F_0 at all. Similarly, we refer to the case of small ε as the case of *small risk*.

Regret In the standard monopoly problem with incomplete information, the seller maximizes the expected revenue for a given prior distribution F over valuations. In contrast, we analyze the problem in which the seller seeks to minimize the expected regret. The *regret* of the monopolist at a given price p and valuation v of the buyer is defined as:

$$r(p,v) \triangleq v - p\mathbb{I}_{\{v \geq p\}}, \tag{2}$$

[3] The band around the model distribution δ_v in the above figure is not exactly the Prohorov neighborhood, but contains the Prohorov neighborhood. The inequality of the Prohorov distance (1) has to hold for all measurable sets A. It therefore imposes additional constraints on the shape of nearby distributions. In the above example of the Dirac distribution it implies that for all sets A such that $A \cap [v-\varepsilon, v+\varepsilon] = \varnothing$, the probability assigned to the set A under a nearby distribution F has to be less or equal to ε. A distribution F in the neighborhood of δ_v can therefore not be simultaneously at the upper bound of the band for values smaller than $v-\varepsilon$ and at the lower bound for values larger than $v+\varepsilon$.

where $\mathbb{I}_{\{v \geq p\}}$ is the indicator function specifying:

$$\mathbb{I}_{\{v \geq p\}} = \begin{cases} 0, & \text{if} \quad v < p, \\ 1, & \text{if} \quad v \geq p. \end{cases}$$

The regret of the monopolist charging price p facing a buyer with value v is the difference between (a) the profit the monopolist could make if she were to know the value v of the buyer before setting her price and (b) the profit she makes without this information. The regret is non-negative and can only vanish if $p = v$. The regret of the monopolist is strictly positive in either of two cases: (i) the value v exceeds the price p, the indicator function is then $\mathbb{I}_{\{v \geq p\}} = 1$, and the regret is the difference between possible revenue and actual price or (ii) the value v is below the price p, the indicator function is then $\mathbb{I}_{\{v \geq p\}} = 0$, and the regret is the foregone surplus due to a high price.

The strategy space of the seller is the set of all probabilistic pricing policies with support on the positive real line. The random variable associated with the mixed strategy is denoted by $\tilde{p}$, the distribution function by $F_p \in \Delta\mathbb{R}_+$ and the density by f_p. The expected regret from adopting a probabilistic pricing policy F_p when facing a buyer with value drawn according to F_v is given by:

$$r(F_p, F_v) \triangleq \mathbb{E}\left[r(\tilde{p}, \tilde{v})\right] = \int \int r(p, v)\, dF_p(p)\, dF_v(v).$$

In the presence of robustness concerns, the seller chooses a pricing policy which minimizes over all policies F_p the maximum regret over all distributions F_v in the neighborhood of a model distribution F_0:

$$\inf_{F_p \in \Delta\mathbb{R}_+} \sup_{F_v \in \mathcal{P}_\varepsilon(F_0)} \int \int r(p, v)\, dF_p(p)\, dF_v(v).$$

The robust policy $F_p^* \in \Delta\mathbb{R}_+$ *attains minimax regret* if

$$F_p^* \in \underset{F_p \in \Delta\mathbb{R}_+}{\arg\min} \sup_{F_v \in \mathcal{P}_\varepsilon(F_0)} r(F_p, F_v).$$

Correspondingly, we refer to F_v^* which maximizes the regret as *worst case demand*:

$$F_v^* \in \underset{F_v \in \mathcal{P}_\varepsilon(F_0)}{\arg\max} \inf_{F_p \in \Delta\mathbb{R}_+} r(F_p, F_v).$$

3 Large Risk

We begin our robust analysis with the case of large risk and hence $\varepsilon = 1$. Under large risk, the specific model distribution F_0 does not matter for the model misspecification. The relevant neighborhood is the entire set of possible distributions on the interval $[0,1]$. In other words, the neighborhood is not anchored by any model distribution F_0 at all. The support of the valuations is the only information that the seller has. While the case of large risk is interesting in its own rights, the analysis here also will be useful in two additional aspects. First it will help us understand the differences between revenue maximization and regret minimization. Second, the resulting equilibrium profiles will reappear in a scaled down version in the case of small risk.

3.1 Regret and Saddlepoint

We begin with a general saddlepoint characterization of the optimal policies for arbitrary neighborhoods. The expected regret from adopting a probabilistic pricing policy F_p when facing a buyer with value drawn according to F_v equals

$$r\left(F_p, F_v\right) = \int_v \int_p \left(v - p\mathbb{I}_{\{p \leq v\}}\right) dF_p\left(p\right) dF_v\left(v\right),$$

which we may rewrite as

$$r\left(F_p, F_v\right) = \int_v v dF_v\left(v\right) - \int_p \left(\int_{v:v\geq p} p dF_v\left(v\right)\right) dF_p\left(p\right).$$

The first integral in the above expression is simply the expected value of the random variable $\tilde{v}$. It also represents the expected profit the seller could obtain if she would learn the value of the buyer before she sets her price. The double integral is the expected profit of the seller under her actual pricing strategy. If we define the expected profit of the seller as:

$$\pi\left(p, F_v\right) \triangleq \int_{v:v\geq p} p dF_v\left(v\right),$$

then we can express the regret of the seller simply as

$$r\left(F_p, F_v\right) = \int_v v dF_v\left(v\right) - \int_p \pi\left(p, F_v\right) dF_p\left(p\right). \tag{3}$$

It is now apparent that if the seller were to know the true distribution over the buyer's valuations, which is true for the special case of $\varepsilon = 0$, then the seller minimizes regret if and only if she maximizes expected profit.

Consider now the case of true risk in which the neighborhood around a model distribution is given by $\mathcal{P}_\varepsilon\left(F_0\right)$ with $\varepsilon > 0$. Three questions arise that we will solve in this order: (i) how can the minimax regret solution be formally derived, (ii) does a minimax regret solution always exist and (iii) is the minimax regret solution related to Bayesian decision-making?

Following Savage (1954) we obtain the minimax regret strategies by solving the associated saddle point problem. For this we consider a zero-sum (normal-form) game between the seller and nature where the seller chooses the probabilistic pricing policy F_p and nature chooses the distribution of buyer values from the set $\mathcal{P}_\varepsilon(F_0)$. In this game the payoff of the seller is the negative of regret while the payoff to nature is regret itself. It is well known that if $\left(F_p^*, F_v^*\right)$ is a Nash equilibrium of this game then F_p^* attains minimax regret as defined above and similarly F_v^* attains maximin regret and is hence a worst case demand. Moreover, we know that the equilibrium payoff equals the value of maximin regret and equals the negative of minimax regret. Formally, a Nash equilibrium of this zero-sum game can be characterized as a solution to the saddle point problem of finding $\left(F_p^*, F_v^*\right)$ that satisfy:

$$r\left(F_p^*, F_v\right) \leq r\left(F_p^*, F_v^*\right) \leq r\left(F_p, F_v^*\right), \ \forall F_p \in \Delta\mathbb{R}_+, \forall F_v \in \mathcal{P}_\varepsilon(F_0). \tag{SP}$$

If $\left(F_p^*, F_v^*\right)$ solve this saddle point problem then the equilibrium regret r^* is:

$$r^* \triangleq r\left(F_p^*, F_v^*\right) = \inf_{F_p \in \Delta\mathbb{R}_+} \sup_{F_v \in \mathcal{P}_\varepsilon(F_0)} r\left(F_p, F_v\right).$$

We prove existence of a solution to this saddlepoint problem and thus existence of a minimax regret pricing policy using results from Reny (1999). From (SP) we see that F_p^* minimizes regret against F_v^*. Together with (3) we observe that the solution F_p^* to minimax regret is also a profit maximizing strategy against the worst case demand F_v^*. More specifically, any price in the support of F_p^* maximizes profits against the worst case demand. In particular, a Bayesian decision maker cannot say that the minimax regret solution is implausible if she cannot rule out F_v^* as a subjective prior.

Theorem 1

1. *A solution $\left(F_p^*, F_v^*\right)$ to the saddlepoint condition (SP) exists.*
2. *If F_p attains minimax regret and $\left(F_p^*, F_v^*\right)$ satisfies (SP) then $\left(F_p, F_v^*\right)$ also satisfies (SP).*
3. *If $\left(F_p^*, F_v^*\right)$ satisfy (SP), then (i) F_p^* attains minimax regret; (ii) F_v^* is worst case demand and (iii) $\operatorname{supp}\left(F_p^*\right) \subseteq \arg\max_p \pi\left(p, F_v^*\right)$.*

In addition, with large risk, i.e. $\varepsilon = 1$, the saddlepoint condition (SP) is equivalent to a condition involving the pure rather than mixed strategies of seller and nature:

$$r\left(F_p^*, v\right) \leq r\left(F_p^*, F_v^*\right) \leq r\left(p, F_v^*\right), \forall p \in \mathbb{R}_+, \quad \forall v \in [0,1].$$

The saddlepoint result allows us to connect minimax regret behavior to payoff maximizing behavior under a prior as follows. When minimax regret is derived from the equilibrium characterization in (SP) then any price chosen by a monopolist who minimizes maximal regret, is at the same time a price which maximizes expected profits against a particular demand, namely the worst case demand. We shall use the above result to establish robust strategies for the seller and worst case demand by nature.

3.2 Random Pricing

We now consider the robust pricing strategy in the case of large risk. The robust pricing strategy of the monopolist minimizes her regret. The regret arises qualitatively from two, very different exposures. If the valuation of the buyer is very high, then the regret may arise from having offered a price too low relative to the valuation. We might refer to this as the *upward exposure*. On the other hand, by having offered a price too high, the buyer risks to have a valuation below the price and the regret of the seller arises from not selling at all. Correspondingly, we may refer to this as the *downward exposure*. At every given price p, the seller faces both a downward and an upward exposure. In this context, a deterministic price policy will always leave the seller exposed to substantial regret and the regret can be reduced significantly by offering a probabilistic pricing policy. If the seller is to be indifferent in her pricing policy against the worst case demand, then the marginal revenue must be zero over the range of prices which the seller offers. In the language of optimal monopoly pricing this means that the virtual utility of different prices has to be constant and equal to zero:

$$p - \frac{1 - F_v^*(p)}{f_v^*(p)} = 0.$$

For nature to be indifferent between valuations, it must be that the regret:

$$r\left(v, F_p^*(p)\right) = v - \int_{p \leq v} p dF_p^*(p),$$

is constant for those valuations, and differentiating with respect to v we obtain:

$$1 - p f_p^*(p) = 0,$$

or

$$f_p^*(p) = \frac{1}{p}.$$

For very low valuations of the buyer, the possible downward risk is small and in consequence the seller will offer only prices above a certain positive threshold.

Theorem 2 (Large Risk)

1. *The minimax regret strategy is given by* F_p^**:*

$$F_p^*(p) = \begin{cases} 0 & \text{if } \; 0 \le p < \frac{1}{e} \\ 1 + \ln p & \text{if } \; \frac{1}{e} \le p \le 1 \end{cases} . \tag{4}$$

2. *The worst case demand is given by* F_v^**:*

$$F_v^*(v) = \begin{cases} 0 & \text{if } \; 0 \le v < \frac{1}{e} \\ 1 - \frac{1}{ev} & \text{if } \; \frac{1}{e} \le v \le 1 \\ 1 & \text{if } \; v = 1 \end{cases} . \tag{5}$$

3. *The minimax regret* r^* *is*

$$r^* = 1 - \mathbb{E}[\tilde{p}^*] = \frac{1}{e} .$$

The above minimax regret strategy F_p^* and worst case demand F_v^* are unique solutions to the regret problem. The indifference conditions for nature determine *exactly* the randomization of the seller (up to the size and position of the mass point). In contrast, the indifference condition of the seller determines only a class of distribution functions, whose single parameter is determined by the boundary of the support. The density $f_v^*(v)$ is diminishing less rapidly than the density $f_p^*(p)$ because an increase in price by the seller is made less appealing by the decrease in sales, whereas an increase in the value leads to a direct increase in the regret. The equilibrium as represented by the distribution functions F_p^* and F_v^* are depicted below.

Insert Figure 2: Optimal Pricing and Worst Case Demand With Large Risk

The equilibrium pricing policy F_p^* is one of many revenue maximizing policies against the worst case demand F_v^*. As the seller is indifferent between all deterministic prices $p \in \left[\frac{1}{e}, 1\right]$, there are indeed many revenue maximizing prices. Yet even with the multiplicity of revenue maximizing prices, the regret strategy F_p^* is uniquely determined as any other strategy would allow nature to establish a larger regret. The trade-off between downward and upward risk thus leads to an interesting prediction of optimal behavior under regret even though there is maximal ambiguity about the true distribution over valuations. If the seller were constrained to offer a deterministic price then she would set $p^* = \frac{1}{2}$ and the resulting regret would be $\frac{1}{2}$ as well.

3.3 Menu Pricing

So far, our analysis assumed that the seller can only offer an indivisible object at some price p. We now extend the instruments of the seller and allow her to offer a menu of items. The equilibrium policies with menus rather than single prices can be directly derived from the random pricing policies studied earlier and thus little new analysis will be necessary. The equilibrium use of menus allows us to understand the selling policies from a different and perhaps more intuitive point of view. The optimality of menus also emphasizes the role of robustness concerns in the optimal selling policies as would never be used in the standard setting.

If the allocative decision regards an indivisible object, or $x \in \{0,1\}$, then a specific item on the menu assigns a probability of receiving the object at a corresponding price. If on the other hand, the allocative decision regards a continuous variable, or $x \in [0,1]$, then a menu offers a variety of quantities at different prices. We observe that with the multiplicative utility vx used here, the notions of probability and quantity are mathematically interchangeable. In a direct mechanism, a menu is a pair $(x(v), p(v))$ which maps a reported type v into a quantity $x(v)$ and price $p(v)$. We transform an equilibrium random pricing policy into a menu policy by defining the quantity assigned in the direct mechanism through:

$$x^*(v) \triangleq F_p^*(v), \tag{6}$$

and the corresponding nonlinear prices as:

$$p^*(v) \triangleq \int_0^v y dF_p^*(y). \tag{7}$$

By standard arguments this assignment of quantities to values defines an incentive compatible mechanism. Vice versa, we can also start with an incentive compatible direct mechanism $(x^*(v), p^*(v))$ and transform this into a mixed pricing policy F_p^* by means of (6). As an aside, we should mention that if we start with any incentive compatible menu, then $x^*(v)$ may not be right continuous, a necessary property for a distribution function. But it is immediate that for every incentive compatible $x^*(v)$, there is a right continuous version, which is at least weakly preferred in terms of expected payoffs and regret to the original version. For this reason, it is without loss of generality to focus on right continuous menus here.

Proposition 1 (Saddlepoint for Menus)
F_p^* attains minimax regret with random pricing policies if and only if $(x^*(v), p^*(v))$ attains minimax regret with menu policies. In particular, a menu that attains minimax regret with menu policies exists and the value of minimax regret among random pricing policies is the same as under menu policies.

The probabilistic pricing policy now translates into a multi-item menu offered by seller. In fact, the menu will be the unique equilibrium offer by the seller. This has to be contrasted to the optimal offer by the seller in the case without risk in which the model distribution constitutes the true distribution. In this case, there always exist an optimal policy in the form of a single item menu in which the entire object is offered at a fixed price.[4] The optimality of a single item menu even in the case of a divisible good comes from the fact that if the marginal cost of production is constant, then a lower priced item on the menu would lower the revenue which can be extracted from higher types. The worst case demand has the property that any increasing menu which offers the entire object at the upper boundary point is revenue maximizing or regret minimizing against the worst case demand. Yet, in the presence of robust concerns there is a unique multi-item menu which is the optimal choice of the seller. This underlines the impact of robustness on the policy selection of the seller. By combining the insights of Theorem 1 and Proposition 1, we obtain the following characterization of the robust menu for large risk.

Corollary 1 (Menues with Large Risk)
The robust menu $(x^*(v), p^*(v))$ is given by:

$$x^*(v) = \begin{cases} 0 & \text{if } \; 0 \leq v < \frac{1}{e} \\ 1 + \ln v & \text{if } \; \frac{1}{e} \leq v \leq 1 \end{cases},$$

and

$$p^*(v) = \begin{cases} 0 & \text{if } \; 0 \leq v < \frac{1}{e} \\ v - \frac{1}{e} & \text{if } \; \frac{1}{e} \leq v \leq 1 \end{cases}.$$

It is a noteworthy feature of the robust menu that the price in the direct mechanism is linear in the valuation v.

[4] The optimality of a single price policy under constant marginal cost was initally derived by Riley & Zeckhauser (1983).

4 Small Risk

Next we describe the robust policy and some of its properties when the risk, represented by the size ε of the neighborhood, is small. We begin the analysis with an arbitrary model distribution. We then illustrate the robust policy for two specific model distributions, the uniform density and the Dirac function. The example of the Dirac function, representing a mass point, is of independent interest as in this context the standard revenue maximizing policy dramatically fails to be revenue robust. It thus represents a leading example for the robustness analysis in the next section.

4.1 Small Neighborhoods

We shall consider small neighborhoods for a general model distribution F_0. We denote the revenue maximizing price for the model distribution F_0 by p_0 :

$$p_0 \triangleq \arg\max_p \pi(p, F_0).$$

For the remainder of this Subsection we shall assume that (i) p_0 is the unique maximizer of the profit function, (ii) the profit function, $\pi(p, F_0)$ at the model distribution F_0 is strictly concave near p_0, (iii) the density f_0 is continuously differentiable near p_0, and (iv) that either $f_0(v) > 0$ in the neighborhood of $v = 0$ or that $F_0(0) > 0$. We comment on the role of these regularity assumptions after the statement of the result.

The basic insights into the nature of robust policies which we gained from the analysis of large risk can be transferred into the analysis of small risk. The new element with small risk is that nature will be constrained in its choice of the worst case demand. The nature of the constraint is described by the size of the neighborhood, namely ε. First, we will therefore characterize the robust policy of the seller for a given neighborhood. Second, we will investigate the response in the robust policy to an increase in the size of the neighborhood. It is one of the conceptual advantages of model misspecification that we can describe risk by a one dimensional parameter, namely the size of the neighborhood, and in turn consider comparative static results as a function of the neighborhood size.

In the case of small risk, the robust policy of the seller has to respond to a smaller set of possible distributions. Yet, the basic trade-offs which we discovered in the large risk environment will again be present, only on a smaller scale. With malevolent nature, the robust policy of the seller will again be to offer many prices. But as nature is constrained, the equilibrium prices will now be close to the revenue maximizing price for the model distribution. On the other hand, the indifference conditions which we derived for the case of large risk are still valid in the case of small risk. Thus the general features of random pricing policy of the seller will be similar, and in particular the density function identical, in the case of small or large risk. The only new aspect is the determination of the support of the pricing policy and the location of the mass point.

The existence of a mass point in the pricing strategy of the seller naturally creates an incentive for nature to place a positive probability event just below the mass point of the seller. This in turn might suggest that an equilibrium of the robust pricing game fails to exist. In the case of large risk, the existence of equilibrium could be guaranteed as the seller offered such low prices that nature could create more regret by offering valuations above the lowest price. Under small risk, we might guess intuitively that even the lowest price offered by the seller is not very far away from p_0, the optimal price for the model distribution. In consequence, the price might not be low enough to dissuade nature from "undercutting" by placing probability just below the mass point of the seller. The stability of the equilibrium strategies will be restored by using the constraints on the worst case demand. In particular, the mass point in the pricing strategy of the seller will be placed precisely at the point where nature is constrained by the neighborhood to place any additional probability just below the mass point of the seller. In consequence, the probability mass will appear in the interior of the support rather than at the boundary points of the support of the pricing strategy.

We describe the robust policy of the seller in terms of the distribution $F_p^*(p)$ of the pricing policy.

Theorem 3 (Small Risk)
If ε is sufficiently small then the robust pricing policy F_p^* is given by:

$$F_p^*(p) = \begin{cases} 0 & \text{if } 0 \leq p < a \\ \ln \frac{p}{a} & \text{if } a \leq p < b \\ 1 - \ln \frac{c}{p} & \text{if } b \leq p \leq c \\ 1 & \text{if } c < p \leq 1 \end{cases},$$

where $0 < a < b < c < 1$ and $a < p_0 < c$.

The support of the pricing policy F_p^* is hence given by $[a, c]$. The mass point in the pricing policy is located at b. The earlier equivalence result between random pricing and menu pricing remains valid. The above result therefore also characterizes the optimal menu policy once we recall that:

$$x^*(v) = F_p^*(v),$$

and the pricing schedule of the menu is obtained from incentive compatibility, or:

$$p^*(v) = \int_0^v p dF_p^*(p).$$

The robust menu offered by seller then has three important characteristics, which can be described with reference to the mass point b: (i) low volume offers are made for buyers with low valuations, or $v < b$, (ii) a much higher offer is made for all buyers with valuation $v \geq b$, and (iii) even higher volume offers are made to buyers with large values $v > b$. We may think of a standard offer given by the quantity offered at $v = b$, and given by $x^*(b)$. In addition, the seller offers low volume downgrades and high volume upgrades. The expanded menu relative to the optimal single item menu for the model distribution seeks to minimize the exposure of the seller. Obviously, the seller looses revenue on the high value buyers from making offers to the low value buyers by granting the high value buyers a larger information rent. The size of the information rent is kept small by offering menu items to the low value buyers only of substantially lower volume. This is the source of the gap in the quantities offered in the menu.

A natural comparison to a robust decision maker is a risk averse decision maker. In particular, we could ask how the behavior of a risk-averse seller would differ from the behavior of a robust seller. Clearly, a risk averse seller would never find a probabilistic pricing policy optimal. Similarly, she would never offer a menu consisting of lotteries. In contrast, if the good were divisible, then a risk averse seller might indeed offer a menu consisting of different quantities. The difference with respect to the robust seller would then be in the shape of the menu. In particular, if a risk averse seller were to face a continuous demand function (as expressed by F_v^*), then the optimal menu would also be continuous. Yet, with a robust seller, we saw that the optimal menu is discontinuous (at a single jump point) and essentially offers two (or three) classes of distinct service.[5]

It remains to describe the comparative static of the pricing policy and the regret of the seller as a function of the size of the neighborhood. The behavior of the regret and the expected price to an increase in the risk can be explained intuitively by the first order

[5] To the best of our knowledge, the problem of a risk averse seller in an optimal pricing environment (in contrast to auctions where results do exist) has not been investigated in the literature. The comparitive results reported here rely on preliminary work by the authors based on standard optimal control techniques.

effects. For a small level of risk, we may represent the regret through a linear approximation

$$r^* = r_0 + \varepsilon \frac{\partial r^*}{\partial \varepsilon},$$

where r_0 is the regret at the model distribution. For a small level of risk, the marginal change in regret can then be computed by holding the price policy of the seller at the optimal price p_0 without risk. Suppose then for the moment that $p_0 \leq \frac{1}{2}$. If the risk increases marginally, the constraints on the choice of a worst case demand are relaxed. What precisely then can nature do given the specification of neighborhood. First nature can place the density $f_0(p_0)$ slightly below p_0 to marginally increase regret by $p_0 f_0(p_0)$, then nature can shift each value up by ε to marginally increase regret by 1 and finally shift mass from 0 to 1 to marginally increase regret by $1 - p_0$. The first two changes correspond to small changes in valuation with large probability, the third to large changes in the valuation with small probability. So the overall marginal effect of an increase in ε near $\varepsilon = 0$ is:

$$p_0 f_0(p_0) + 1 + (1 - p_0).$$

If instead the optimal price without risk would be $p_0 > \frac{1}{2}$, then the only modification would affect the third element as nature would move mass for 0 to just below p_0, so that the marginal increase would be

$$p_0 f_0(p_0) + 1 + p_0.$$

The optimal response of the seller to an increase in risk is now to find a pricing policy which minimizes the additional regret

$$\varepsilon \frac{\partial r^*}{\partial \varepsilon}$$

coming from the increase in risk. Of course, the cost of adjusting the price to minimize the marginal regret is that it changes the regret relative to the model distribution F_0. Locally, the cost of changing the policy variable away from the optimum is given by the second derivative of the objective function. With small risk, the curvature of the regret is identical to the curvature of the revenue function. The rate at which the robust price responses to an increase in risk is then simply the ratio of the response of the marginal regret to a change in price divided by the curvature of the revenue function, or

$$\frac{\partial \mathbb{E}}{\partial \varepsilon}[\widetilde{p}] = \frac{\frac{\partial}{\partial p}\left[\frac{\partial r^*}{\partial \varepsilon}\right]}{\frac{\partial^2}{\partial p^2}\pi(p_0, F_0)}.$$

The next proposition shows that the above intuition can be made precise and shows its implication for the net utility of the buyer.

Proposition 2 (Risk and Pricing)
For small risk, the expected price $\mathbb{E}[\tilde{p}]$ *responds to an increase in risk by:*

$$\frac{\partial}{\partial \varepsilon}\mathbb{E}[\tilde{p}]\,|_{\varepsilon=0} = \begin{cases} -1 + \frac{1+f_0(p_0)}{2f_0(p_0)+p_0 f_0'(p_0)} & \text{if} \quad p_0 \leq \frac{1}{2} \\ -1 + \frac{f_0(p_0)-1}{2f(p_0)+p_0 f'(p_0)} & \text{if} \quad p_0 > \frac{1}{2} \end{cases}, \tag{8}$$

and the expected regret $r\left(F_p^*, F_v^*\right)|_{\varepsilon=0}$ *responds by:*

$$\frac{\partial}{\partial \varepsilon} r\left(F_p, F_v\right)|_{\varepsilon=0} = \begin{cases} 2 - p_0 + p_0 f_0\left(p_0\right) & \text{if} \quad p_0 \leq \frac{1}{2} \\ 1 + p_0 + p_0 f_0\left(p_0\right) & \text{if} \quad p_0 > \frac{1}{2} \end{cases}.$$

The net utility of a buyer with value v *is increasing in* ε *for all* v *if* $\frac{\partial}{\partial \varepsilon}\mathbb{E}[\tilde{p}]\,|_{\varepsilon=0} < 0$.

We construct the robust pricing policy by means of the implicit function theorem, for which we need the differentiability of the density function near p_0. The worst case demand makes the seller indifferent among all prices $p \in [a, c]$. With the strict concavity of the profit function around p_0, we know that there do not exist other local optima nearby which the seller may choose to avoid malevolent nature. The final regularity assumption of either $F_0(0)$ or $f_0(v) > 0$ near $v = 0$ guarantees that nature always removes initially density from $v = 0$ to increase regret. The nature of the robust policy would not change if there were no mass or density at $v = 0$, it would merely change the position of the midpoint $1/2$ in the above comparative static result.

In this paper, we report the robust policies for the case of small and large risk. We expect that the solution to the case of intermediate risk shares all the features of the small and large risk cases. In fact for the class of linear densities we have obtained the complete characterization of robust policies, confirming our intuition. The basic obstacle to completely solve the case of intermediate risk is that the implicit function technique does not apply anymore, while the guess and verify method used for large risk cannot account for the constraints imposed on nature through the size of the neighborhood.

The robust response of the seller to an increase in risk is perhaps even more informative when we consider the menu policy. In a menu, the seller is offering many different choices to the buyers. An immediate question therefore is how the choice set for the buyers changes with an increase in the risk. We define the size of the menu simply as the set of quantities offered by the seller (and accepted by some buyers) in equilibrium.

Proposition 3 (Risk and Menus)
For small risk:

1. *The size of the menu is increasing in ε.*

2. *The price per unit $p^*(v)/x^*(v)$ is decreasing in ε.*

As the risk increases, the seller seeks to minimize her exposure by offering more choices to the buyers and hence increasing the probability of a sale, even if the sale is not "big" in terms of the sold quantity. For every given valuation v, the seller also increases the size of the deal offered. As larger deals are offered to buyers with lower valuations, it follows that the seller is willing to concede a larger information rent to buyers with higher valuations. In consequence, the average price per unit is decreasing as well. Jointly, these three properties imply that the seller is offering her products more aggressively and to a larger number of buyers with an increase in risk. We observe that the monotonicity in the unit price holds even as the previous proposition showed that the expected price may be increasing. The resolution of this apparent conflict comes from the fact that the seller is offering larger quantities in response to an increase an risk.

4.2 Linear Density

We now illustrate the equilibrium behavior with the uniform model distribution:

$$F_0(v) = v,$$

and the revenue maximizing price p_0 under the model distribution is given by:

$$p_0 = \frac{1}{2}.$$

We graphically represent the optimal behavior of the seller and nature for a small neighborhood.

INSERT FIGURE 3: OPTIMAL PRICING AND WORST CASE DEMAND WITH LINEAR MODEL DENSITY

The interior curve in the above graph identifies the model distribution. Constraints induced by small changes in values constrain the distribution function of F_v^* to be within an ε bandwidth of the model distribution. The large changes of values, occurring with probability of most ε move the smallest valuation to the largest valuation, namely 1. The strategy of nature is then to place as little probability as necessary below the range of the prices offered by the seller and to shift values above the range as high as possible. Inside the range of prices offered by the seller, nature uses a density function which maintains the virtual utility of the seller at 0. In turn, the seller sets the density to make nature indifferent between all values above the mass point and all values below the mass point. Given the mass point set by the seller, nature shifts as much mass as possible below this point. We observe that even with the small neighborhood of $\varepsilon = 0.04$, the impact of the risk on the pricing policy is rather large and leads to a wide spread in the prices offered by the seller.

As an implication from Proposition 2, we find that in the class of linear densities the change in expected price as well as the change in the mass point is strictly positive if and only if the density is strictly decreasing. This can be contrasted with the robust behavior under max min of revenues where it is immediate to see that any increase in ε uncertainty has a downward effect on prices.

4.3 Dirac Function

The second class of model distributions which we shall study in some more detail is the class of Dirac functions, δ_{v_0}, for some $0 < v_0 < 1$. The example of a Dirac function will be informative for at least two reasons. First, the Dirac function describes the case of certainty in the absence of any model misspecification. The changes in the robust pricing policies relative to the optimal policy can therefore be traced directly to the model specification. The simplicity of the model distribution in fact allows us to explicitly derive the robust pricing policy and the worst case demand as functions of the size ε of the neighborhood. Second, and perhaps more importantly, the case of Dirac distribution is the simplest example of an environment for which the standard pricing policy clearly fails robustness. In particular, the optimal price $p = v_0$, will result in zero revenue if the true distribution is not given by δ_{v_0} but rather $\delta_{v_0-\varepsilon}$. In the next section we shall argue that the robust policy derived here is also robust when we consider revenues rather than regret.

We described the neighborhood of the Dirac function in Section 2 to illustrate the Prohorov distance. Compared to the case of large risk, the seller now faces a dramatically smaller window of risk, yet on this smaller scale the trade-offs are almost exactly the same ones as in the case of large risk. The main difference concerns the placement of the remaining ε probability. In the case of large risk, the choice of nature was completely unrestricted. With small neighborhoods, most of the probability has be assigned to valuations near v_0.

As the remaining ε probability is unconstrained, its placement will depend on the size of the downward and upward exposure. The relative size of each exposure is determined by the location of v_0. For $v_0 \leq \frac{1}{2}$, the upward risk looms large as nature can place ε probability outside of the reach of the pricing policy by placing at 1 and hence creating a regret of $1 - v_0$. In contrast for $v_0 > \frac{1}{2}$, the downward exposure creates the larger regret and hence nature is tempted to place the remaining ε probability just below the support of the seller's strategy, for a regret arbitrarily close to v_0.

Proposition 4 (Dirac Function)
For ε sufficiently small:

1. *the minimax regret pricing policy F_p^* is given by*

$$F_p^*(p) = \begin{cases} 0 & \text{if} \quad 0 \leq p < (v_0 - \varepsilon)(1 - \varepsilon) \\ 0 & \text{if} \quad (v_0 - \varepsilon)(1 - \varepsilon) \leq p < (v_0 - \varepsilon) \text{ and } v_0 \leq \frac{1}{2} \\ \ln \frac{p}{a} & \text{if} \quad (v_0 - \varepsilon)(1 - \varepsilon) \leq p < (v_0 - \varepsilon) \text{ and } v_0 > \frac{1}{2} \\ 1 - \ln \frac{v_0 + \varepsilon}{p} & \text{if} \quad v_0 - \varepsilon \leq p < v_0 + \varepsilon \\ 1 & \text{if} \quad v_0 + \varepsilon \leq p \leq 1 \end{cases};$$

2. *the response in expected price satisfies $\frac{\partial}{\partial \varepsilon} \mathbb{E}[\tilde{p}]|_{\varepsilon=0} = -1$;*

3. *the expected regret r^* is increasing in ε with*

$$\frac{\partial r^*}{\partial \varepsilon}|_{\varepsilon=0} = \begin{cases} 3 - v_0 & \text{if} \quad v_0 \leq \frac{1}{2} \\ 2 + v_0 & \text{if} \quad v_0 > \frac{1}{2} \end{cases};$$

4. *the net utility of any buyer is continuously increasing in ε.*

The equilibrium strategies are illustrated in Figure 4.

INSERT FIGURE 4: OPTIMAL PRICING AND WORST CASE DEMAND WITH DIRAC MODEL DISTRIBUTION

We observe that as $v_0 \leq \frac{1}{2}$, the maximal regret is achieved by placing all the mass at the top of the interval.

The expected price is decreasing initially at the rate of -1. The decrease in the price is a direct response to the possibility that the true value might not be v_0 but rather $v_0 - \varepsilon$.

At $\varepsilon = 0$, the price completely and exclusively responds to this risk that the true value might not be v_0 but rather $v_0 - \varepsilon$. The marginal regret converges to $2 + v_0$ and $3 - v_0$ as ε converges to zero for $v_0 \leq \frac{1}{2}$ and $v_0 > \frac{1}{2}$, respectively. As the seller is worried that the true value might be $v_0 - \varepsilon$ rather than v_0, she essentially lowers her price to $v_0 - \varepsilon$. At that point, malevolent nature puts much of the weight on $v_0 + \varepsilon$. The resulting regret is 2ε, and hence the marginal increase is 2. The additional component, which varies across the two situations comes from the small probability that there is large change. With $v_0 \leq \frac{1}{2}$, nature places the additional probability on the largest possible value, 1, and the resulting regret per $1 - (v_0 - \varepsilon)$, which yields $1 - v_0$ as ε goes to zero. Conversely, for $v_0 > \frac{1}{2}$, nature places the ε probability just below $v_0 - \varepsilon$, and as ε goes to zero, the additional component is v_0. Yet we observe that the regret and in particular the marginal regret is continuous across the two regimes. The comparative static results for the menu are identical to the ones obtained for the concave and differentiable environment, and the proof is analogous.

Proposition 5 (Robust Menus with Dirac Function)
For ε sufficiently small:

1. *The size of the menu is increasing in ε.*
2. *The price per unit $p^*(v)/x^*(v)$ is decreasing in ε.*

The equilibrium menu is illustrated in Figure 5.

Insert Figure 5: Robust Menu With Dirac Model Distribution

In the case of a model distribution without uncertainty, it is natural to investigate the relationship between a deterministic and probabilistic pricing policy. The deterministic solution to the minmax regret problem naturally does not satisfy the saddlepoint characterization and if we restrict the seller to deterministic strategies then we observe that:

$$\min_{p} \max_{F_v \in \mathcal{P}_\varepsilon(F_0)} r(p, F_v) > \max_{F_v \in \mathcal{P}_\varepsilon(F_0)} \min_{p} r(p, F_v).$$

We should perhaps add that the above inequality is true for general distributions and not only for the Dirac distribution. To see this suppose, that there is indeed a saddle point in which the seller uses a single price. For p to be optimal for the seller, there must be a positive probability of valuations at or above p. But now nature has an incentive to move some of these valuations to slightly below p, contradicting the saddle point hypothesis.

On the other hand, it is immediate that the respective local responses of the randomized and deterministic pricing policy converge as ε goes to zero. It follows that for small values of ε, the loss from concentrating on deterministic policies is minor. Yet as the risk grows, the loss due to a deterministic strategy increases at an increasing rate. Similarly, the difference between the expected price under the randomized strategy and the deterministic strategy is increasing at an increasing rate.

Corollary 2 (Deterministic Versus Random Strategies)
For ε sufficiently small:

1. *The optimal deterministic strategy is $p_d^* = v_0 - \varepsilon$.*
2. *The difference between deterministic and stochastic price, $\mathbb{E}[\tilde{p}^*] - p_d^*$, is an increasing and convex function in ε.*
3. *The difference between deterministic and stochastic regret, $r_d^* - r^*$, is an increasing and convex function in ε.*

5 Revenue Robustness

We finally return to the issue of revenue robust policies and the link to minmax regret strategies. We define a general *pricing rule* $\hat{p}$ as a mapping from a model distribution F_0 for demand and a neighborhood of size ε into a probability distribution over prices:

$$\hat{p} : \Delta[0,1] \times (0,1) \to \Delta[0,1],$$

such that $\hat{p}(F_0, \varepsilon)$ is the mixed pricing policy chosen under prior F_0 and a neighborhood of size ε. For instance, a pricing strategy that attains minimax regret for all F_0 and all ε is such a decision rule.

Definition 1 (Revenue Robust)
A pricing rule $\hat{p}$ is called revenue robust at F_0 *if for each $\gamma > 0$ there is $\varepsilon > 0$ such that for all priors $F \in \Delta[0,1]$:*

$$F \in \mathcal{P}_\varepsilon(F_0) \Rightarrow \pi(p^*(F), F) - \pi(\hat{p}(F_0, \varepsilon), F) < \gamma.$$

A pricing rule $\hat{p}$ is revenue robust *if it is revenue robust at every F_0.*

The notion of revenue robustness provides a formal definition of robustness for revenue maximization in the spirit of Huber (1964). It generalizes the definition of α-robustness of Prasad (2003). The robust policy is allowed to depend on the size ε of the neighborhood. Similarly, Hansen & Sargent (2004) use the term robustness for maximizing the minimum utility within an ε neighborhood. A stronger notion of robustness would require the robust policy to be independent of ε. Under this stronger notion, typically only the optimal policy at F_0 would be a candidate for a robust policy. We next show that the minimax regret policy is a revenue robust policy.

Theorem 4 (Robustness)
If $\hat{p}$ attains minimax regret at F_0 for all sufficiently small ε then $\hat{p}$ is revenue robust at F_0.

6 Discussion

Neighborhoods The local notion of robustness is naturally tied with the corresponding neighborhood notion.[6] The Prohorov metric allows nearby distributions to be different either through a *small* probability of *large* changes relative to the model distribution or through a *large* probability of *small* changes in the valuations. The first advantage of the Prohorov metric is that it applies to both discrete as well as continuous distributions. In contrast, the Kullback-Leibler distance only usefully defines neighborhoods for continuous distributions. In addition, the Prohorov metric is actually a metric for weak convergence of probability measures.[7]

A related model to represent neighborhoods is given by the contamination "neighborhood", where the neighborhood of a model distribution F_0 is described by $\mathcal{N}_\varepsilon(F_0)$:

$$\mathcal{N}_\varepsilon(F_0) = \{F_v \,|F_v = (1-\varepsilon) F_0 + \varepsilon G \text{ for some } G \in \Delta\mathbb{R}_+\}. \tag{9}$$

The contamination neighborhood is actually not a neighborhood in the sense of the weak topology. Yet it is somewhat easier to handle than the Prohorov neighborhood as the additive term εG only allows for a small probability of a large change, but not for a large probability of a small change. For a given ε, the contamination neighborhood is then strictly smaller than the Prohorov neighborhood. The difference between the two neighborhoods

[6] The issue of an appropriate neighborhood also appears in stability analysis of evolutionary games, see Oechssler & Riedel (2001) and Oechssler & Riedel (2002).

[7] There is also a close relationship between the Prohorov norm and the variational norm, see Shiryaev (1995), p.360.

is most apparent when we look at our earlier example of a Dirac distribution. The contamination neighborhood essentially reduces to the horizontal bands but omits the vertical bands as displayed below.

Insert Figure 6: Contamination Neighborhood With Dirac Model Distribution

In consequence, for a given neighborhood size ε, the robust policy has to secure itself against fewer possible distributions and the robust policy will be *closer* to the pricing policy before robustness concerns arise.

Interim Regret The notion of regret we used here is an ex-post criterion as we compare the realized revenue with the revenue she could have realized for the given realization of the random variable $\widetilde{v}$. This suggests a weaker version of regret in form of an interim regret.[8] Interim regret is the difference between the expected revenue if the seller knew the true distribution and the expected revenue she actually obtains:

$$R\left(F_p, F_v\right) = \sup_p \pi\left(p, F_v\right) - \int \pi\left(p, F_v\right) dF_p\left(p\right).$$

The resulting min max problem would be given by:

$$\inf_{F_p \in \mathcal{P}_\varepsilon(F_0)} \sup_{F_v} R\left(F_p, F_v\right). \tag{R}$$

Interestingly, the robust pricing policy coincide for the case of large risk in the case of interim or ex post regret. In the case of small risk, the equilibrium policies sill have the same shape as under ex post regret but typically the support of the pricing policy and the worst case demand is located lower with interim regret. The reason is that with interim regret, the benchmark is the expected revenue and hence the missed opportunities due to upward exposure matter less.

The notion of interim regret also permits an alternative interpretation with a continuum of buyers. Our earlier notion of regret investigates the impact of not knowing the value of a single buyer. Consider now a continuum of buyers. There are two possible benchmarks with which expected payoffs are compared. One is where the monopolist is able to price discriminate among buyers if she knew their values. Here our original notion of regret

[8] The notion of interim regret is called "regret risk" in Chamberlain (2000) and is the basis of the analysis in Manski (2004).

applies and the monopolist prices as if facing a single buyer. However, if the monopolist is not able to price discriminate, but instead has to set a single price then the original notion of regret may appear less plausible as a benchmark. Without the ability to price discriminate, a monopolist who knows F_v will charge $p^* \in \arg\max_{p\in\mathbb{R}_+} \pi(p, F_v)$ and an alternative notion of regret with many buyers and without price discrimination would be interim regret.

Competitive Analysis In computer science the optimality of a selling mechanism is often evaluated by its competitive ratio, see Borodin & El-Yaniv (1998). The evaluation criterion is the ratio between realized revenue and feasible revenue. The optimal policy or mechanism is defined as the one which maximizes the competitive ratio among all possible distributions $F_v(v)$. In our setting the competitive ratio is given by

$$\frac{p\mathbb{I}_{\{p\leq v\}}}{v}.$$

The competitive ratio is obtained by random policies of seller and nature. A random strategy of nature now requires that for all v in the support of the distribution F_v the above ratio is constant in expectations, or:

$$\frac{\int_{p\leq v} pf_p(p)\,dp}{v}. \tag{10}$$

A similar condition can be established based for the competitive pricing policy of the seller, which requires that for all p in the support of the distribution F_p, the competitive ratio is constant in expectations, or:

$$\frac{p}{\int_p^1 vf_v(v)\,dv}. \tag{11}$$

The conditions (10) and (11) then yield the characterization of competitive pricing and worst case demand. In fact, the competitive pricing policy has the same density (up to a constant factor) as the robust pricing policy. Yet in contrast to the robust pricing policy, the competitive pricing policy has as its support the entire interval, $[\underline{v}, 1]$ of possible valuations. The worst case demand for the competitive ratio is given by $f(p) = ae^{-p}$ for some constant factor α. The competitive strategies also display the mass points at the same locations as the robust policies, namely a mass point in the strategy of seller at $v = \underline{v}$ and a mass point in the strategy of nature at $v = 1$.

Thus, while the notion of competitive ratio does not address the case of small risks, in the case of large risk, the resulting competitive policy is very close to our robust policy with large risk.

7 Conclusion

In this paper we analyzed robust pricing policies by a monopolist. We began by considering globally robust policies as solution to minimax regret strategies with an unconstrained adversary. We then considered locally robust policies by restricting the strategy space by the adversary to contain only nearby distributions. The monopolist anticipates that her prior is only a noisy forecast of the true distribution. The magnitude of the change in expected price is approximately indirectly proportional to the curvature of the profit function at the optimal price.

The problem of optimal monopoly pricing is in many respects the most elementary mechanism design problem. It would be of interest to extend the insights and apply the techniques developed here to a wide class of design problems, such as the discriminating monopolist (as in Mussa & Rosen (1978) and Maskin & Riley (1984)) and optimal auctions. The monopoly setting has the simplifying feature that the buyers have complete information about their payoff environment. Given their know valuation and known price, each buyer simply had to make a decision as to whether or not to purchase the object. With the complete information by the buyer, there was no need to look for a robust purchasing policy. A substantial task would consequently arise by considering multi-agent design problems with incomplete information such as auctions, where it becomes desirable to "robustify" both the decisions of the buyers and the seller. The recent result by Segal (2003) and Chung & Ely (2003) regarding the sufficient conditions for the existence of dominant strategies for the bidders in optimal auctions might offer a first step in this direction. The complete solution of these problems poses a rich field for future research.

8 Appendix

The appendix contains some auxiliary results and the proofs for the results in the main body of the text.

Proof of Theorem 1. (1.) We apply Corollary 5.2 in Reny (1999). Clearly we have a compact Hausdorff game. Reciprocal upper semi continuity follows directly as we are investigating a zero sum game. So all we have to ensure is payoff security. Payoff security for the monopolist means that we have to show for each $\left(F_p^*, F_v^*\right)$ with $F_v^* \in \mathcal{P}_\varepsilon\left(F_0\right)$ and for every $\delta > 0$ that there exists $\gamma > 0$ and $\bar{F}_p$ such that $F_v \in \mathcal{P}_\gamma\left(F_v^*\right)$ implies $r\left(\bar{F}_p, F_v\right) \leq r\left(F_p^*, F_v^*\right) + \delta$.

Let $\gamma \triangleq \delta/4$ and let $\bar{F}_p$ be such that $\bar{F}_p\left(p\right) \triangleq F_p^*\left(p+\gamma\right)$. Then using the fact that $F_v\left(v\right) \geq F_v^*\left(v-\gamma\right) - \gamma$ we obtain

$$\int_0^1 v dF_v\left(v\right) \leq 2\gamma + \int_0^1 v dF_v^*\left(v\right).$$

Using the fact that $F_v\left(v\right) \leq F_v^*\left(v+\gamma\right) + \gamma$ we obtain

$$\pi\left(\bar{F}_p, F_v\right) \geq \pi\left(F_p^*\left(p+\gamma\right), \min\left\{F_v^*\left(v+\gamma\right)+\gamma, 1\right\}\right) \geq \pi\left(F_p^*, F_v^*\right) - 2\gamma$$

and hence $r\left(\bar{F}_p, F_v\right) \leq r\left(F_p^*, F_v^*\right) + \delta$.

To show payoff security for nature we have to show for each $\left(F_p^*, F_v^*\right)$ with $F_v^* \in \mathcal{P}_\varepsilon\left(F_0\right)$ and for every $\delta > 0$ that there exists $\gamma > 0$ and $\bar{F}_v \in \mathcal{P}_\varepsilon\left(F_0\right)$ such that $F_p \in \mathcal{P}_\gamma\left(F_p^*\right)$ implies $r\left(F_p, \bar{F}_v\right) \geq r\left(F_p^*, F_v^*\right) - \delta$.

Here we set $\bar{F}_v \triangleq F_v^*$. Given $\gamma > 0$ consider any $F_p \in \mathcal{P}_\gamma\left(F_p^*\right)$. All we have to show is that $\pi\left(F_p, F_v^*\right) \leq \pi\left(F_p^*, F_v^*\right) + \delta$ for sufficiently small γ. Note that $F_p\left(p\right) \leq F_p^*\left(p+\gamma\right) + \gamma$ implies

$$\begin{aligned}
\pi\left(F_p, F_v^*\right) &\leq \gamma + \int \left(p+\gamma\right)\left(\int_p^1 dF_v^*\left(v\right)\right) dF_p^*\left(p+\gamma\right) = \gamma + \int p\left(\int_{p-\gamma}^1 dF_v^*\left(v\right)\right) dF_p^*\left(p\right) \\
&= \gamma + \pi\left(F_p^*, F_v^*\right) + \int p\left(\int_{[p-\gamma,p)} dF_v^*\left(v\right)\right) dF_p^*\left(p\right) \\
&\leq \gamma + \pi\left(F_p^*, F_v^*\right) + \int \int_{[p-\gamma,p)} dF_v^*\left(v\right) dF_p^*\left(p\right).
\end{aligned}$$

Given continuity of $\int \int_{[p-\gamma,p)} dF_v^*\left(v\right) dF_p^*\left(p\right)$ in γ the claim is shown.

(2.) The proof is a standard result of zero-sum games and is easily verified directly using (12).

(3.) Let $\left(F_p^*, F_v^*\right)$ be a solution to (SP). Parts (i) and (ii) follow directly from

$$\begin{aligned} r\left(F_p^*, F_v^*\right) &= \max_{F_v \in \mathcal{P}_\varepsilon(F_0)} r\left(F_p^*, F_v\right) \geq \inf_{F_p \in \Delta\mathbb{R}_+} \sup_{F_v \in \mathcal{P}_\varepsilon(F_0)} r\left(F_p, F_v\right) = r^* \\ &\geq \sup_{F_v \in \mathcal{P}_\varepsilon(F_0)} \inf_{F_p \in \Delta\mathbb{R}_+} r\left(F_p, F_v\right) \geq \min_{F_p \in \Delta\mathbb{R}_+} r\left(F_p, F_v^*\right) = r\left(F_p^*, F_v^*\right) \end{aligned} \tag{12}$$

as $\inf\sup \geq \sup\inf$ holds generally. Part (iii) follows from the fact that (SP) and linearity of r in p implies $\operatorname{supp}\left(F_p^*\right) \subseteq \arg\min_p r\left(p, F_v^*\right)$ and thus $\operatorname{supp}\left(F_p^*\right) \subset \arg\max_p \pi\left(p, F_v^*\right)$ as $r\left(p, F_v^*\right) = \int v dF_v^*\left(v\right) - \pi\left(p, F_v^*\right)$. ■

Proof of Theorem 2. Let $\tilde{v}^*$ have cdf $F_v^* \in \Delta\left[\frac{1}{e}, 1\right]$ with density $\frac{1}{ev^2}$ and where $\Pr\left(\tilde{v}^* = 1\right) = \frac{1}{e}$. Hence $\Pr\left(\tilde{v}^* \geq v\right) = \frac{1}{ev}$ for $v \in \left[\frac{1}{e}, 1\right]$ and

$$\begin{aligned} r\left(p, F_v^*\right) &= \frac{1}{e} + \int_{\frac{1}{e}}^{1} v\frac{1}{ev^2}dv - p\frac{1}{ep} = \frac{1}{e} \text{ for } p \in \left[\frac{1}{e}, 1\right], \\ r\left(p, F_v^*\right) &= \frac{1}{e} + \int_{\frac{1}{e}}^{1} v\frac{1}{ev^2}dv - p > \frac{1}{e} \text{ for } 0 \leq p < \frac{1}{e}, \\ r\left(F_p^*, v\right) &= v - \int_{\frac{1}{e}}^{v} p\frac{1}{p}dp = \frac{1}{e} \text{ for } v \in \left[\frac{1}{e}, 1\right], \\ r\left(F_p^*, v\right) &= v < \frac{1}{e} \text{ for } 0 \leq v < \frac{1}{e} . \end{aligned}$$

Hence, $\left(F_p^*, F_v^*\right)$ satisfies the conditions of Theorem 1. Note that $r^* = r\left(F_p^*, 1\right)$ so $r^* = 1 - \mathbb{E}\left[\tilde{p}\right]$. ■

The proof of Proposition 1 follows directly from the following lemma.

Lemma 1 (Equivalence)

1. *For any mixed pricing policy $F_p\left(v\right)$ the menu $\left(x\left(v\right), p\left(v\right)\right)$ is incentive compatible.*
2. *If $\left(x\left(v\right), p\left(v\right)\right)$ is incentive compatible, then there exists a mixed pricing policy F_p such that $\pi\left(F_p, v\right) \geq p\left(v\right)$ for all $v \in [0,1]$.*

Proof. First we show that if $g : [0,1] \to [0,1]$ is non decreasing then

$$vg\left(v\right) - \int_0^v sdg\left(s\right) - \int_0^v g\left(s\right)ds \equiv 0.$$

Let h be the left hand side of this equation. Clearly, $h\left(0\right) = 0$. Since g is non decreasing and bounded, h is differentiable almost everywhere which implies that $h' = 0$ almost everywhere.

Consider some $\bar{v} \in [0,1]$. If g is continuous at $\bar{v}$ then so is h. Assume that g is not continuous at $\bar{v}$. Then

$$\bar{v}g(\bar{v}) - \int_0^{\bar{v}} sdg(s) = \lim_{v<\to\bar{v}} vg(v) + \bar{v}\left(g(\bar{v}) - \lim_{v<\to\bar{v}} g(v)\right) - \lim_{v<\to\bar{v}} \int_0^{v} sdg(s) - \bar{v}\left(g(\bar{v}) - \lim_{v<\to\bar{v}} g(v)\right)$$

so h is continuous at $\bar{v}$ and thus $h \equiv 0$.

For the proof we can use a standard result on incentive compatibility, see Proposition 23.D.2 in Mas-Collel, Whinston & Green (1995). Part (1) follows immediately from the fact that F_p is nondecreasing and that $vF_p(v) - \pi(F_p, v) = \int_0^v F_p(s)\,ds$ given our calculations above.

For part (2), notice that $x(v) \in [0,1]$ and that incentive compatibility implies that $x(v)$ is non decreasing and $vx(v) - p(v) = \int_0^v x(s)\,ds$. Moreover, we can limit attention to menus where x is right continuous as otherwise there exists a right continuous incentive compatible menu $(\hat{x}(v), \hat{p}(v))_{v\in[0,1]}$ such that $\hat{p}(v) \geq p(v)$ for all v. As we consider x that is right continuous, F_p such that $F_p(v) \triangleq x(v)$ for all v is a well defined mixed pricing policy and we obtain $p(v) = vx(v) - \int_0^v x(s)\,ds$. Our calculations above then imply that $\pi(F_p, v) = p(v)$. ■

In order to derive the equilibrium policies in the case of small risk we present a characterization of the Prohorov distance that builds on the following celebrated result of Strassen (1965).

Theorem (Strassen (1965)).
F and G have Prohorov distance ε if and only if there exist random variables X and Y such that X has distribution F, Y has distribution G and $\Pr(|Y - X| \leq \varepsilon) \geq 1 - \varepsilon$.

The two cumulative distributions F, G are close if and only if they are associated to two random variables that realize similar values with high probability. Our characterization describes the Prohorov distance in terms of cumulative distribution functions only. In order to stay within epsilon distance of a given distribution function G one may first alter any value by at most ε, this creates a probability measure F_1, and then take ε mass away from some values, the removal is described by a measure F_3, and move it to other values, described by a measure F_2.

Lemma 2 (Decomposition)
Consider $\varepsilon > 0$ and probability measures F and G. $F \in \mathcal{P}_\varepsilon(G)$ if and only if there exists a probability measure F_1 and positive additive measures F_2 and F_3 such that:

$$G(x-\varepsilon) \le F_1(x) \le G(x+\varepsilon), F_2, F_3 \le \varepsilon \quad \textit{and} \quad F \equiv F_1 + F_2 - F_3.$$

Proof. ($\Leftarrow$) Suppose F can be decomposed into F_1, F_2 and F_3. We then want to show that $F(A) \le G(A^\varepsilon) + \varepsilon$. To this purpose, it is clearly sufficient sufficient to consider only closed sets A.

(a) We first prove the claim for $A = [x, y]$ with $0 \le x \le y \le 1$. Given a probability measure H let $H^-(\hat{v}) \triangleq \lim_{v < \rightarrow \hat{v}} H(v)$. Then

$$F_1([x, y]) = F_1(y) - F_1^-(x) \le G(y+\varepsilon) - G^-(x-\varepsilon) = G([x, y]^\varepsilon).$$

Since $F_2([x, y]) \le \varepsilon$ and $F_3([x, y]) \ge 0$ we obtain

$$F([x, y]) = F_1([x, y]) + F_2([x, y]) - F_3([x, y]) \le G([x, y]^\varepsilon) + \varepsilon.$$

(b) Next we consider $A = [x_1, y_1] \cup [x_2, y_2]$ with $y_1 + 2\varepsilon < x_2$ which implies $[x_1, y_1]^\varepsilon \cap [x_2, y_2]^\varepsilon = \emptyset$. Using the results obtained in part (a) together with the fact that $A^\varepsilon = [x_1, y_1]^\varepsilon \cup [x_2, y_2]^\varepsilon$ holds for the $^\varepsilon$ operator it follows that

$$F_1(A) = F_1([x_1, y_1]) + F_1([x_2, y_2]) \le G([x_1, y_1]^\varepsilon) + G([x_2, y_2]^\varepsilon) = G(A^\varepsilon).$$

Since $F_2(A) \le \varepsilon$ and $F_3(A) \ge 0$ the claim is proven.

(c) The arguments in part (b) are easily generalized for any set A that can be decomposed into a finite union of disjoint closed intervals of distance greater than 2ε so $A = \cup_{k=1}^m [x_k, y_k]$ with $x_k \le y_k < x_{k+1} + 2\varepsilon$ for $k \le m-1$.

(d) Finally we show that we do not have to prove the statement for more general sets A.

Notice that if $A_1^\varepsilon = A_2^\varepsilon$, $A_1 \subset A_2$ and $F(A_2) \le G(A_2^\varepsilon) + \varepsilon$ then $F(A_1) \le G(A_1^\varepsilon) + \varepsilon$. So we can restrict attention to proving the claim for closed sets A such that $A^\varepsilon = A_1^\varepsilon$ and $A \subseteq A_1$ implies $A = A_1$. Consider $x, y \in A$ such that $x < y \le x + 2\varepsilon$. Then $\{A \cup [x, y]\}^\varepsilon = A^\varepsilon$ and hence $[x, y] \subseteq A$. It now follows easily that A belongs to the class of sets investigated in part (c).

($\Rightarrow$) Consider probability measures F and G with $\|F - G\| \le \varepsilon$. Extend G to $[-\varepsilon, 1+\varepsilon]$ such that $G(x) = 0$ for $-\varepsilon \le x < 0$ and $G(x) = 1$ for $1 < x \le 1+\varepsilon$. Given the above result of Strassen (1965), there exist random variables X and Y such that X has distribution F, Y has distribution G and $\Pr(|Y - X| \le \varepsilon) \ge 1 - \varepsilon$.

Let Z_1 be the random variable with cdf F_1 such that $Z_1 \triangleq X$ if $|Y - X| \le \varepsilon$ and $Z_1 \triangleq Y$ if $|Y - X| > \varepsilon$. Let $\varepsilon' \triangleq \Pr(|Y - X| > \varepsilon)$ so $\varepsilon' \le \varepsilon$. Then $G(x - \varepsilon) \le F_1(x) \le G(x + \varepsilon)$.

Let Z_2 be the random variable with cdf $\hat{F}_2$ such that $Z_2 \triangleq 0$ if $|Y - X| \leq \varepsilon$ and $Z_2 \triangleq X$ if $|Y - X| > \varepsilon$. Let Z_3 be the random variable with cdf $\hat{F}_3$ such that $Z_3 \triangleq 0$ if $|Y - X| \leq \varepsilon$ and $Z_3 \triangleq Y$ if $|Y - X| > \varepsilon$. Then $X = Z_1 + Z_2 - Z_3$ and $\hat{F}_2(0), \hat{F}_3(0) \geq 1 - \varepsilon'$. Let $F_i \triangleq \hat{F}_i - (1 - \varepsilon')$ for $i = 2, 3$. Then F_2, F_3 are positive additive measures with $F_2, F_3 \leq \varepsilon'$ and the proof is complete. ■

Proof of Theorem 3. We start by assuming $p_0 > \frac{1}{2}$. The proof proceeds in three steps. First we show the existence of the parameters a, b and c and use these to construct the worst case demand F_v^*. Second, we decompose the worst case demand by using Lemma 2 to show that it is close to F_0. Third we use this decomposition to verify that we have a saddle point.

Step 1. We start by showing that for sufficiently small ε there exist parameters a, b, c such that $a < b < c$ and $a < p_0 < c$ such that

$$\begin{aligned} F_0(a - \varepsilon) - \varepsilon &= 1 - \frac{b^2 f_0(b + \varepsilon)}{a}, & (13)\\ F_0(b + \varepsilon) &= 1 - \frac{b^2 f_0(b + \varepsilon)}{b}, & (14)\\ F_0(c - \varepsilon) &= 1 - \frac{b^2 f_0(b + \varepsilon)}{c}. & (15) \end{aligned}$$

Concerning existence of b note that $b = p_0$ solves (14) if $\varepsilon = 0$. As

$$\frac{d}{db}\left(1 - F_0(b + \varepsilon) - b f_0(b + \varepsilon)\right)|_{\varepsilon=0} = -2f(p_0) - p_0 f'(p_0) < 0,$$

due to strict concavity of profits at p_0, the implicit function theorem implies that a solution $b = b(\varepsilon)$ to (14) (with $b > 0$) exists for ε in a neighborhood of 0.

To prove existence of c, define

$$h(v) \triangleq 1 - \frac{b^2 f_0(b + \varepsilon)}{v} - F_0(v - \varepsilon)\, for v > 0.$$

Then $h(b) = F_0(b + \varepsilon) - F_0(b - \varepsilon)$ with

$$h'(b) = f_0(b + \varepsilon) - f_0(b - \varepsilon),$$

and

$$h''(b) = -\frac{2 f_0(b + \varepsilon)}{b} - f_0'(b - \varepsilon) \approx -\frac{2 f_0(p_0) + p_0 f_0'(p_0)}{p_0} < 0.$$

We note that $h(b) > 0$ by our assumptions on F_0. Looking at the Taylor approximation of h near $v = b$ for small ε we obtain that there exists $c > b$ such that $h(c) = 0$ with $c \to p_0$

as $\varepsilon \to 0$. As for existence of a, analogous calculations for $h(v) + \varepsilon$ show that there exists $a < b$ such that $h(a) + \varepsilon = 0$ with $a \to p_0$ as $\varepsilon \to 0$.

We can describe the local behavior of the parameters a, b,and c by appealing to the implicit function theorem. Since $2f_0(p_0) + p_0 f_0'(p_0) > 0$ we know that b is differentiable and by implicitly differentiating (14) we obtain:

$$b'(0) = -\frac{f_0(p_0) + p_0 f_0'(p_0)}{2f(p_0) + p_0 f'(p_0)}. \tag{16}$$

Next we show that a is differentiable. Since

$$\frac{b^2 f_0(b+\varepsilon) - a^2 f_0(a-\varepsilon)}{b-a} = (b+a) f_0(b+\varepsilon) + a^2 \frac{f_0(b+\varepsilon) - f_0(a-\varepsilon)}{b-a} \approx 2p_0 f_0(p_0) + p_0^2 f_0'(p_0),$$

we find that $b^2 f_0(b+\varepsilon) > a^2 f_0(a-\varepsilon)$ near $\varepsilon = 0$. Hence we can implicitly differentiate (13) to obtain

$$a'(\varepsilon) = -a \frac{a + a f_0(a-\varepsilon) + b f_0(b+\varepsilon)}{b^2 f_0(b+\varepsilon) - a^2 f_0(a-\varepsilon)} \tag{17}$$

so

$$\lim_{\varepsilon \to 0} \left(\frac{b-a}{a} a'(\varepsilon) \right) = -\frac{1 + 2f_0(p_0)}{2f_0(p_0) + p_0 f_0'(p_0)}.$$

In particular we obtain that

$$\lim_{\varepsilon \to 0} a'(\varepsilon) = -\infty. \tag{18}$$

Similarly for c, we find that:

$$c'(\varepsilon) = -c \frac{c f_0(c-\varepsilon) + b f_0(b+\varepsilon)}{b^2 f_0(b+\varepsilon) - c^2 f_0(c-\varepsilon)} \tag{19}$$

and hence

$$\lim_{\varepsilon \to 0} \left(\frac{c-b}{c} c'(\varepsilon) \right) = \frac{2f_0(p_0)}{2f_0(p_0) + p_0 f_0'(p_0)}.$$

In particular,

$$\lim_{\varepsilon \to 0} c'(\varepsilon) = \infty. \tag{20}$$

It now follows from (18) and (20) that $a < p_0 < c$.

Step 2. We now construct the worst case demand on the basis of a, b and c. Consider F_v^* given by

$$F_v^*(v) \triangleq \begin{cases} \max\{0, F_0(v-\varepsilon) - \varepsilon\}, & \text{if} \quad v \in [0,a] \\ 1 - \frac{b^2 f_0(b+\varepsilon)}{v}, & \text{if} \quad v \in (a,c) \\ F_0(v-\varepsilon), & \text{if} \quad v \in [c,1) \\ 1 & \text{if} \quad v = 1 \end{cases},$$

where the definitions of a and c imply that F_v^* is continuous at a and c. It follows that F_v^* is a probability measure.

We now show that $F_v^* \in \mathcal{P}_\varepsilon(F_0)$ by using Lemma 2. Consider F_1^* defined by

$$F_1^*(v) \triangleq \begin{cases} F_0(v-\varepsilon), & \text{if } v \in [0,a] \\ \max\{F_v^*(v), F_0(v-\varepsilon)\}, & \text{if } v \in (a,b) \\ F_v^*(v), & \text{if } v \in [b,1] \end{cases}.$$

Then F_1^* is a probability measure with $F_0(v-\varepsilon) \leq F_1^*(v)$. By definition of b we obtain $F_v^*(b) = F_0(b+\varepsilon)$ and $F_v^{*\prime}(b) = \frac{d}{dv}F_0(v+\varepsilon)|_{v=b}$. Moreover, given $F_v^{*\prime\prime}(v) = -\frac{2b^2 f_0(b+\varepsilon)}{v^3}$ and $\frac{d^2}{(dv)^2}F_0(v+\varepsilon) = f_0'(v+\varepsilon)$, strict concavity of profits near p_0 implies that $F_0''(v) < F_v^{*\prime\prime}(v)$ for $v \in [a,c]$ and ε sufficiently small. Thus, for sufficiently small ε, as a and c are close to p_0, we obtain $F_1^*(v) \leq F_0(v+\varepsilon)$ with equality if and only if $v = b$. So $F_0(v-\varepsilon) \leq F_1^*(v) \leq F_0(v+\varepsilon)$.

Consider F_2^* defined by

$$F_2^*(v) \triangleq \begin{cases} 0, & \text{if } v \in [0,a] \\ \varepsilon - \max\{F_0(v-\varepsilon) - F_v^*(v), 0\}, & \text{if } v \in (a,b] \\ \varepsilon, & \text{if } v \in (b,1] \end{cases}.$$

Then

$$\frac{d}{dv}\left(F_v^*(v) - F_0(v+\varepsilon)\right) = \frac{b^2 f_0(b+\varepsilon)}{v^2} - f_0(v+\varepsilon) \geq 0 \quad \text{for } v \leq b,$$

as

$$\frac{d}{dv}\left(v^2 f_0(v+\varepsilon)\right) = v^2 f_0'(v+\varepsilon) + 2v f_0(v+\varepsilon) > 0,$$

holds for ε sufficiently small and hence F_2^* is weakly increasing with $F_2^*(1) = \varepsilon$. Since F_2^* is also right continuous we obtain that F_2^* can be extended to a non additive probability measure.

Let F_3^* be defined by

$$F_3^*(v) \triangleq \min\{F_0(v-\varepsilon), \varepsilon\}, \text{ if } v \in [0,1],$$

so $F_3^*(v)$ is a non additive probability measure and $F_3^*(1) = \varepsilon$. Since $F_v^* = F_1^* + F_2^* - F_3^*$ we obtain from Lemma 2 that $F_v^* \in \mathcal{P}_\varepsilon(F_0)$.

Step 3. Next we show that $\left(F_p^*, F_v^*\right)$ is a saddle point. For the monopolist we verify easily that $\pi(p, F_v^*) = b^2 f_0(b+\varepsilon)$ for $p \in [a,c]$. Similar to the calculations following the definition of F_1^* it is easily shown that there exists $\zeta > 0$ such that $1 - \frac{b^2 f_0(b+\varepsilon)}{v} < F_v^*(v)$ holds for all $v \in [p_0 - \zeta, p_0 + \zeta] \setminus [a,c]$ and all sufficiently small ε. Thus, for sufficiently small

ε we obtain $[a,c] = \arg\max_{p\in[p_0-\zeta,p_0+\zeta]} \pi(p, F_v^*)$ and together with the upperhemicontinuity of profits that $[a,c] \subseteq \arg\max_p \pi(p, F_v^*)$.

Consider now the incentives of nature. Note that

$$r\left(F_p^*, F_v^*\right) = r\left(F_p^*, F_1^*\right) + r\left(F_p^*, F_2^*\right) - r\left(F_p^*, F_3^*\right),$$

where the definition of regret r is naturally extended to arbitrary measures. Notice that it is best to choose F_2^* and F_3^* such that $F_2^*(1) = F_3^*(1) = \varepsilon$. In the following we show that each term is maximized separately starting with F_2^*. If nature could put all mass on a single value v, by construction of F_v^* nature would be indifferent over $v \in [a,b)$ and over $v \in [b,c]$. Since $r\left(F_p^*, v\right)$ is monotone increasing on $[0,a]$ and $[c,1]$ it follows that $\arg\max_v r\left(F_p^*, v\right) \subseteq [a,b) \cup \{1\}$. For sufficiently small ε, $r\left(F_p^*, a\right) \approx p_0$ while $r\left(F_p^*, 1\right) \approx 1 - p_0$ so given $p_0 > \frac{1}{2}$ we obtain $[a,b) = \arg\max_v r\left(F_p^*, v\right)$.

Concerning F_3^* let $\bar{v} = \inf\{v : F_0(v-\varepsilon) \geq \varepsilon\}$. We have to show that $r\left(F_p^*, \tilde{v}\right) \leq r\left(F_p^*, \hat{v}\right)$ for $\tilde{v} \leq \bar{v} \leq \hat{v}$. Given the above it is sufficient to consider only $\tilde{v} = \bar{v}$ and $\hat{v} = c$ where $r\left(F_p^*, c\right) = c - \mathbb{E}(\tilde{p})$. Let $\gamma \triangleq 2\sup_{v>0} \frac{v}{F_0(v)}$. For v sufficiently small, $\gamma \geq \frac{v}{F_0(v)}$ and hence $r\left(F_p^*, \tilde{v}\right) = \tilde{v} \leq \varepsilon + \gamma F_0(\tilde{v} - \varepsilon) = \varepsilon(1+\gamma)$. On the other hand, we show in the proof of Proposition 2 that $\frac{\partial}{\partial\varepsilon} c|_{\varepsilon=0} = \infty$ and $\frac{\partial}{\partial\varepsilon}\mathbb{E}(\tilde{p})|_{\varepsilon=0} < \infty$ so $\frac{\partial}{\partial\varepsilon} r\left(F_p^*, c\right)|_{\varepsilon=0} = \infty$ and hence $r\left(F_p^*, \tilde{v}\right) < r\left(F_p^*, c\right)$ for ε sufficiently small.

Finally, consider F_1^*. More mass cannot be allocated to regret maximizing values $[a,b)$ as $F_1^*(b) = F_0(b+\varepsilon)$, weight on values below a and above c are shifted up as far as possible as $F_v(v) = F_0(v-\varepsilon)$ for $v < a$ and $c < v < 1$ and allocation of F_1^* for $F_1^* \in (F_1^*(b), F_0(c-\varepsilon))$ will not influence regret as $r\left(F_p^*, v\right)$ is constant on $[b,c]$.

The case of $p_0 \leq \frac{1}{2}$ proceeds in an analogous manner. It is easily shown that there exist parameters a, b, c such that $a < b < c$ and $a < p_0 < c$ such that

$$\begin{aligned} F_0(a-\varepsilon) - \varepsilon &= 1 - \frac{b^2 f_0(b+\varepsilon)}{a} \\ F_0(b+\varepsilon) &= 1 - \frac{b^2 f_0(b+\varepsilon)}{b} + \varepsilon \\ F_0(c-\varepsilon) - \varepsilon &= 1 - \frac{b^2 f_0(b+\varepsilon)}{c}. \end{aligned} \tag{21}$$

Let the worst case demand F_v^* be defined by

$$F_v^*(v) \triangleq \begin{cases} \max\{0, F_0(v-\varepsilon) - \varepsilon\}, & \text{if} \quad v \in [0,a] \\ 1 - \frac{b^2 f_0(b+\varepsilon)}{v}, & \text{if} \quad v \in (a,c) \\ \max\{0, F_0(v-\varepsilon) - \varepsilon\}, & \text{if} \quad v \in [c,1) \\ 1 & \text{if} \quad v = 1 \end{cases}.$$

F_v^* is now decomposed such that $F_v^* = F_1^* + F_2^* - F_3^*$ where

$$F_1^*(v) \triangleq \begin{cases} F_0(v-\varepsilon), & \text{if} \quad v \in [0,a] \\ 1 - \frac{b^2 f_0(b+\varepsilon)}{v} + \varepsilon, & \text{if} \quad v \in (a,c) \\ F_0(v-\varepsilon), & \text{if} \quad v \in [c,1) \\ 1 & \text{if} \quad v = 1 \end{cases},$$

$$F_2^*(v) \triangleq \begin{cases} 0, & \text{if} \quad v \in [0,1) \\ \varepsilon, & \text{if} \quad v = 1 \end{cases},$$

$$F_3^*(v) \triangleq \min\{F_0(v-\varepsilon), \varepsilon\}, \text{ if } v \in [0,1].$$

Lemma 2 can be applied to show that $F_v^* \in \mathcal{P}_\varepsilon(F_0)$. In contrast to the previous case of $p_0 > \frac{1}{2}$, now $v = 1$ maximizes $r\left(F_p^*, v\right)$ so that F_2^* puts all mass at $v = 1$. For the case of $p_0 = \frac{1}{2}$ Proposition 2 can be used to show that $r\left(F_p^*, 1\right) = 1 - \mathbb{E}[\tilde{p}] > r\left(F_p^*, a\right) = a$. As in the case where $p_0 > \frac{1}{2}$, $F_1^*(v) \leq F_0(v+\varepsilon)$ with tangency only at $v = b$ so F_1^* again maximizes weight on $[a,b)$. $[a,b)$ is now only a local maximum of $r\left(F_p^*, v\right)$ but nevertheless it still follows easily that F_1^* maximizes regret (use the fact that $F_0(b+\varepsilon) < F_0(c-\varepsilon)$). ■

Proof of Proposition 2. We obtain that

$$\mathbb{E}[\tilde{p}] = \int_a^c p\frac{1}{p}dp + b\left(1 - \int_a^c \frac{1}{p}dp\right) = c - a + b\left(1 - \ln\frac{c}{a}\right).$$

As a, b, and c are differentiable as shown in Proposition 3, we have:

$$\frac{\partial}{\partial\varepsilon}\mathbb{E}[\tilde{p}] = \frac{b-a}{a}a'(\varepsilon) + \frac{c-b}{c}c'(\varepsilon) + \left(1 - \ln\frac{c}{a}\right)b'(\varepsilon).$$

Inserting the value for $a'(\varepsilon), b'(\varepsilon)$ and $c'(\varepsilon)$ from (16), (17) and (19) respectively, we obtain for $p_0 > \frac{1}{2}$:

$$\frac{\partial}{\partial\varepsilon}\mathbb{E}[\tilde{p}]\,|_{\varepsilon=0} = -1 + \frac{f_0(p_0) - 1}{2f(p_0) + p_0 f'(p_0)}.$$

The same operations yield the result for $p_0 < \frac{1}{2}$.

Next we consider the behavior of the regret with an increase in risk. Given our assumptions on $f(\cdot)$ there exists κ such that $F_0(\kappa - \varepsilon) = \varepsilon$ so $f_0(\kappa-\varepsilon)\kappa'(\varepsilon) = 1 + f_0(\kappa - \varepsilon)$. We can then write the regret as:

$$\begin{aligned} r(F_p, F_v) = r(a, F_v) \quad &= \quad \int_\kappa^a v f_0(v-\varepsilon)\,dv + \int_a^c v\frac{b^2 f_0(b+\varepsilon)}{v^2}dv + \int_c^1 v f_0(v-\varepsilon)\,dv \\ &\quad + (1 - F_0(1-\varepsilon)) - a(1 - F_0(a-\varepsilon) + \varepsilon). \end{aligned}$$

We calculate

$$\begin{aligned}
\frac{\partial}{\partial\varepsilon} r\left(a, F_v\right) &= -\int_{\kappa}^{a} v f_0'\left(v-\varepsilon\right) dv - \kappa f_0\left(\kappa-\varepsilon\right)\kappa'\left(\varepsilon\right) + \int_{a}^{c} \frac{b^2 f_0'\left(b+\varepsilon\right)}{v} dv - \int_{c}^{1} v f_0'\left(v-\varepsilon\right) dv \\
&\quad + f_0\left(1-\varepsilon\right) - a f_0\left(a-\varepsilon\right) - a \\
&\to -\int_0^1 v f_0'\left(v\right) dv + f_0\left(1\right) - p_0 f_0\left(p_0\right) - p_0 = 1 - p_0 f_0\left(p_0\right) - p_0 \quad \text{as } \varepsilon \to 0,
\end{aligned}$$

as

$$\int_0^1 v f_0'\left(v\right) dv = \left[v f_0\left(v\right)\right]_0^1 - \int_0^1 f_0\left(v\right) dv = f_0\left(1\right) - 1.$$

Similarly we find

$$\frac{\partial}{\partial a} r\left(a, F_v\right) = 2\frac{a^2 f_0\left(a-\varepsilon\right) - b^2 f_0\left(b+\varepsilon\right)}{a} + \varepsilon,$$

so

$$\begin{aligned}
\frac{\partial}{\partial a} r\left(a, F_v\right) a'\left(\varepsilon\right) &= -\frac{2a^2 f_0\left(a-\varepsilon\right) - 2b^2 f_0\left(b+\varepsilon\right) + \varepsilon a}{a} a \frac{a + a f_0\left(a-\varepsilon\right) + b f_0\left(b+\varepsilon\right)}{b^2 f_0\left(b+\varepsilon\right) - a^2 f_0\left(a-\varepsilon\right)} \\
&\to 2\left(p_0 + 2 p_0 f_0\left(p_0\right)\right) \text{ for } \varepsilon \to 0,
\end{aligned}$$

as $\lim_{\varepsilon\to 0} \frac{b-a}{\varepsilon} = \infty$ from the local behavior of $a\left(\varepsilon\right)$ and $b\left(\varepsilon\right)$ near $\varepsilon = 0$.

By a similar argument, we can evaluate the behavior of the regret at the price $p = a$

$$\lim_{\varepsilon\to 0}\left(\frac{\partial}{\partial c} r\left(a, F_v\right) c'\left(\varepsilon\right)\right) = -2 p_0 f_0\left(p_0\right)$$

and $p = b$:

$$\lim_{\varepsilon\to 0} \frac{\partial}{\partial b} r\left(a, F_v\right) = 0$$

so

$$\frac{\partial}{\partial\varepsilon} r\left(a, F_v\right)|_{\varepsilon=0} = 1 + p_0 + p_0 f_0\left(p_0\right) = 2 + p_0 - F_0\left(p_0\right).$$

The proof for $p_0 \leq \frac{1}{2}$ is analogous and we omit the details. ■

Proof of Proposition 3. Clearly, (1) holds for ε sufficiently small. Next we verify (2). Assume $a < v < b$. Then $x^*\left(v\right) = \ln\frac{v}{a}$ and $p^*\left(v\right) = \int_a^v y\frac{1}{y}dy = v - a$ so given $a' < 0$ for ε small we obtain $\frac{\partial}{\partial\varepsilon}x^*\left(v\right) > 0$, $\frac{\partial}{\partial\varepsilon}p^*\left(v\right) > 0$ and

$$\frac{\partial}{\partial\varepsilon}\frac{p^*\left(v\right)}{x^*\left(v\right)} = \frac{\left(v-a\right)\frac{1}{a} - \ln\frac{v}{a}}{\left(\ln\frac{v}{a}\right)^2} a' < 0$$

as $\frac{d}{dv}\left(\left(v-a\right)\frac{1}{a} - \ln\frac{v}{a}\right) = \frac{1}{a} - \frac{1}{v} > 0$. Thus, $x^*\left(v\right)v - p^*\left(v\right)$ is strictly increasing in ε.

Assume $b < v < c$. Then $x^*(v) = 1 - \ln \frac{c}{v}$ and $p^*(v) = v - a + \left(1 - \ln \frac{c}{a}\right) b = \mathbb{E}[\tilde{p}] + v - c$ so $\frac{\partial}{\partial \varepsilon} x^*(v) < 0$, $\frac{\partial}{\partial \varepsilon} p^*(v) < 0$ and

$$\frac{\partial}{\partial \varepsilon} \frac{p^*(v)}{x^*(v)} = \frac{\frac{\partial}{\partial \varepsilon} \mathbb{E}[\tilde{p}]}{1 - \ln \frac{c}{v}} + \frac{\frac{1}{c}\left(\mathbb{E}[\tilde{p}] + v - c\right) - \left(1 - \ln \frac{c}{v}\right)}{\left(1 - \ln \frac{c}{v}\right)^2} c' < 0$$

where we use the fact that $c'(\varepsilon)$ is large and $\frac{d}{dv}\left(\frac{1}{c}\left(\mathbb{E}[\tilde{p}] + v - c\right) - \left(1 - \ln \frac{c}{v}\right)\right) = \frac{1}{c} - \frac{1}{v} < 0$ for ε small.

We obtain

$$\frac{\partial}{\partial \varepsilon} u(v) = \left(v - \frac{p^*(v)}{x^*(v)}\right) \frac{\partial}{\partial \varepsilon} x^*(v) - x^*(v) \frac{\partial}{\partial \varepsilon} \frac{p^*(v)}{x^*(v)} = \frac{c - v}{c} c'(\varepsilon) - \frac{\partial}{\partial \varepsilon} \mathbb{E}[\tilde{p}].$$

Since incentive compatibility implies that $x^*(v) v - p^*(v)$ is continuous in v and since x^* has an upwards jump at $v = b$ we obtain

$$\frac{p^*(b)}{x^*(b)} > \lim_{v \to b^-} \frac{p^*(v)}{x^*(v)}.$$

Clearly, $\frac{p^*(v)}{x^*(v)} > \frac{p^*(b)}{x^*(b)}$ for $v > b$ holds from above using right continuity of x^*.■

Proof of Proposition 4. Assume $v_0 \leq \frac{1}{2}$ and ε sufficiently small. Note that F_p^* is well defined. Let F_v^* be defined by

$$f_v^*(v) \triangleq \frac{v_0 - \varepsilon}{v^2}, \quad v \in (v_0 - \varepsilon, v_0 + \varepsilon),$$

with two upper mass points:

$$\Pr(v = v_0 + \varepsilon) \triangleq 1 - \varepsilon - \frac{2\varepsilon}{v_0 + \varepsilon} \text{ and } \Pr(v = 1) \triangleq \varepsilon.$$

Then F_v^* is also well defined and $F_v^* \in \mathcal{P}_\varepsilon(F_0)$.

It suffices to verify that $\left(F_p^*, F_v^*\right)$ constitutes a saddlepoint. It is clear that $r\left(F_p^*, v\right)$ is constant on $[v_0 - \varepsilon, v_0 + \varepsilon]$ and that $v = 1$ maximizes $r\left(F_p^*, v\right)$. Similarly, it is immediate to see that the seller is indifferent over all prices in the interval $[v_0 - \varepsilon, v_0 + \varepsilon]$. Since $\pi(v_0 - \varepsilon, F_v^*) = v_0 - \varepsilon > \varepsilon = \pi(1, F_v^*)$ it follows that the seller is choosing a best response to F_v^*.

The expected price is now given by

$$\mathbb{E}[\tilde{p}] = (v_0 - \varepsilon)\left(1 - \ln \frac{v_0 + \varepsilon}{v_0 - \varepsilon}\right) + 2\varepsilon,$$

so

$$\frac{\partial \mathbb{E}[\tilde{p}]}{\partial \varepsilon} = -1 + \frac{2\varepsilon}{v_0 + \varepsilon} + \ln \frac{v_0 + \varepsilon}{v_0 - \varepsilon}.$$

The regret is given by

$$r\left(F_p^*, F_v^*\right) = (1-\varepsilon)\, r\left(F_p^*, v_0+\varepsilon\right) + \varepsilon r\left(F_p^*, 1\right) = (1-\varepsilon)(v_0+\varepsilon) + \varepsilon - \mathbb{E}[\tilde{p}],$$

so

$$\frac{\partial r\left(F_v^*, F_p^*\right)}{\partial \varepsilon}|_{\varepsilon=0} = 3 - v_0.$$

Assume now that $\frac{1}{2} < v_0 > 1$. Clearly, F_p^* is well defined. Let F_v^* be defined by

$$f_v^*(v) \triangleq \frac{(1-\varepsilon)(v_0-\varepsilon)}{v^2}, \quad v \in ((1-\varepsilon)(v_0-\varepsilon), v_0+\varepsilon),$$

with one upper mass point

$$\Pr(v = v_0+\varepsilon) \triangleq (1-\varepsilon)\frac{v_0-\varepsilon}{v_0+\varepsilon}.$$

Note that F_v^* is well defined and that $F_v^* \in \mathcal{P}_\varepsilon(F_0)$.

The argument showing that $\left(F_p^*, F_v^*\right)$ forms a saddle point is as before with the exception, due to $v_0 > \frac{1}{2}$, that nature now places mass ε on the interval $((1-\varepsilon)(v_0-\varepsilon), v_0-\varepsilon)$ instead of mass ε on $v = 1$. The rest follows directly from

$$\begin{aligned}
\mathbb{E}[\tilde{p}] &= (v_0-\varepsilon)\left(1 - \ln\frac{v_0+\varepsilon}{(1-\varepsilon)(v_0-\varepsilon)}\right) + 2\varepsilon + \varepsilon(v_0-\varepsilon), \\
r(F_v, F_p) &= (1-\varepsilon)((v_0+\varepsilon) - \mathbb{E}[\tilde{p}]) + \varepsilon(v_0-\varepsilon),
\end{aligned}$$

which concludes the proof.■

Proof of Theorem 4. Assume that $\hat{p}$ attains minimax regret but is not robust. So there exists $\gamma > 0$ such that for all $\varepsilon > 0$ there exists F_ε such that $F_\varepsilon \in \mathcal{P}_\varepsilon(F_0)$ but

$$\pi(p^*(F_\varepsilon), F_\varepsilon) - \pi(\hat{p}(F_0,\varepsilon), F_\varepsilon) \geq \gamma. \tag{22}$$

Assume that $(\hat{p}(F_0,\varepsilon), G_\varepsilon)$ is a saddle point of the regret problem (SP) given $\varepsilon > 0$. Then

$$r(\hat{p}(F_0,\varepsilon), G_\varepsilon) = \sup_{F \in \mathcal{P}_\varepsilon(F_0)} r(\hat{p}(F_0,\varepsilon), F),$$

and hence

$$\hat{p}(F_0,\varepsilon) = p^*(G_\varepsilon).$$

We can rewrite the rhs of (22) as follows:

$$\begin{aligned}
& \pi(p^*(F_\varepsilon), F_\varepsilon) - \pi(\hat{p}(F_0,\varepsilon), F_\varepsilon) \qquad (23)\\
= \;& \pi(p^*(F_\varepsilon), F_\varepsilon) - \pi(p^*(G_\varepsilon), G_\varepsilon) + \pi(p^*(G_\varepsilon), G_\varepsilon) - \pi(p^*(G_\varepsilon), F_\varepsilon).
\end{aligned}$$

Using (SP) we also obtain

$$0 \leq r\left(p^*\left(G_\varepsilon\right), G_\varepsilon\right) - r\left(p^*\left(G_\varepsilon\right), F_\varepsilon\right) = \int v dG_\varepsilon(v) - \int v dF_\varepsilon(v) + \pi\left(p^*\left(G_\varepsilon\right), F_\varepsilon\right) - \pi\left(p^*\left(G_\varepsilon\right), G_\varepsilon\right)$$

so that:

$$\pi\left(p^*\left(G_\varepsilon\right), G_\varepsilon\right) - \pi\left(p^*\left(G_\varepsilon\right), F_\varepsilon\right) \leq \int v dG_\varepsilon(v) - \int v dF_\varepsilon(v).$$

Entering this into (23) we obtain from (22) that:

$$\pi\left(p^*\left(F_\varepsilon\right), F_\varepsilon\right) - \pi\left(p^*\left(G_\varepsilon\right), G_\varepsilon\right) + \int v dG_\varepsilon(v) - \int v dF_\varepsilon(v) \geq \gamma. \tag{24}$$

Since $F_\varepsilon, G_\varepsilon \in \mathcal{P}_\varepsilon\left(F_0\right)$ and since $h(v) = v$ is a continuous function and the Prohorov norm metrizes the weak* topology we obtain that

$$\int v dG_\varepsilon(v) - \int v dF_\varepsilon(v) < \gamma/2,$$

if ε is sufficiently small.

Now we will show that $\pi\left(p^*(F), F\right)$ is continuous with respect to the Prohorov neighborhood at $F = F_0$. Consider F, G such that $G \in \mathcal{P}_\varepsilon(F)$.Consider $p_\varepsilon \triangleq p^*(F) - 2\varepsilon$. Then the claim follows as from the following relationship of the probabilities:

$$G\left(v < p_\varepsilon\right) \leq G\left(p_\varepsilon\right) \leq F\left(p_\varepsilon + \varepsilon\right) + \varepsilon = F\left(p^*(F) - \varepsilon\right) + \varepsilon \leq F\left(v < p^*(F)\right) + \varepsilon.$$

The second inequality follows from the hypothesis of the Prohorov neighborhood, We use implies that

$$\begin{aligned} \pi\left(p^*(G), G\right) \geq \pi\left(p_\varepsilon, G\right) &= p_\varepsilon\left(1 - G\left(v < p_\varepsilon\right)\right) \geq p_\varepsilon\left(1 - F\left(v < p^*(F)\right)\right) \\ &= \pi\left(p^*(F), F\right) - 2\varepsilon\left(1 - F\left(v < p^*(F)\right)\right) \\ &\geq \pi\left(p^*(F), F\right) - 2\varepsilon. \end{aligned}$$

Since $\pi\left(p^*(F), F\right)$ is continuous in F at $F = F_0$ we find that

$$\pi\left(p^*\left(F_\varepsilon\right), F_\varepsilon\right) - \pi\left(p^*\left(G_\varepsilon\right), G_\varepsilon\right) < \gamma/2$$

if ε is sufficiently small. This yields the desired contradiction.■

References

Borodin, A. & R. El-Yaniv. 1998. *Online Computation and Competitive Analysis.* Cambridge: Cambridge University Press.

Bose, S., E. Ozdenoren & A. Pape. 2004. Optimal Auctions with Ambiguity. Technical report University of Texas, Austin and University of Michigan.

Chamberlain, G. 2000. "Econometrics and Decision Theory." *Journal of Econometrics* 95:255–283.

Chung, K.-S. & J. Ely. 2003. "Implementation with Near-Complete Information." *Econometrica* 71:857–871.

Ellsberg, D. 1961. "Risk, Ambiguity and the Savage Axioms." *Quarterly Journal of Economics* 75:643–669.

Gilboa, I. & D. Schmeidler. 1989. "Maxmin Expected Utility with Non-Unique Prior." *Journal of Mathematical Economics* 18:141–153.

Goldberg, A., J. Hartline & A. Wright. 2001. "Competitive Auctions and Digital Goods." Intertrust Technical Report STAR-TR-99-01.

Hampel, F. R., E. M. Ronchetti, P.J. Rousseeuw & W.A. Stahel. 1986. *Robust Statistics - The Approach Based on Influence Functions.* New York: John Wiley and Sons.

Hansen, L. Peter & T. Sargent. 2004. Misspecification in Recursive Macroeconomic Theory. Technical report University of Chicago, New York University, and Hoover Institution.

Hayashi, T. 2005. Regret Aversion and Opportunity Dependence. Technical report University of Texas.

Huber, Peter. J. 1981. *Robust Statistics.* New York: John Wiley and Sons.

Huber, P.J. 1964. "Robust Estimation of a Location Parameter." *Annals of Mathematical Statistics* 35:73–101.

Huber, P.J. 1965. "A Robust Version of the Probability Ratio Test." *Annals of Mathematical Statistics* 36:1753–1758.

Linhart, P.B & R. Radner. 1989. "Minimax - Regret Strategies for Bargaining over Several Variables." *Journal of Economic Theory* 48:152–178.

Manski, C. 2004. "Minimax-Regret Treatment Choice with Missing Outcome Data." *Journal of Econometrics* forthcoming.

Mas-Collel, A., M.D. Whinston & J.R. Green. 1995. *Microeconomic Theory.* Oxford: Oxford University Press.

Maskin, E. & J. Riley. 1984. "Monopoly with Incomplete Information." *Rand Journal of Economics* 15:171–196.

Milnor, J. 1954. Games Against Nature. In *Decision Processes*, ed. R.M. Thrall, C.H. Coombs & R.L. Davis. New York: Wiley.

Mussa, M. & S. Rosen. 1978. "Monopoly and Product Quality." *Journal of Economic Theory* 18:301–317.

Neeman, Zvika. 2003. "The Effectiveness of English Auctions." *Games and Economic Behavior* 43:214–238.

Oechssler, J & F. Riedel. 2001. "Evolutionary Dynamics on Infinite Strategy Spaces." *Economic Theory* 17:141–162.

Oechssler, J. & F. Riedel. 2002. "On the Dynamic Foundation of Evolutionary Stability in Continuous Models." *Journal of Economic Theory* 107:223–252.

Prasad, K. 2003. "Non-Robustness of some Economic Models." *Topics in Theoretical Economics* 3:1–7.

Reny, P. J. 1999. "On the Existence of Pure and Mixed Strategy Nash Equilibria in Discontinuous Games." *Econometrica* 67:1029–1056.

Riley, J. & R. Zeckhauser. 1983. "Optimal Selling Strategies: When to Haggle, When to Hold Firm." *Quarterly Journal of Economics* 98:267–290.

Savage, L.J. 1954. *The Foundations of Statistics.* 1st ed. New York: Wiley.

Segal, I. 2003. "Optimal Pricing Mechanism with Unknown Demand." *American Economic Review* 93:509–529.

Selten, R. 1989. Blame Avoidance as Motivating Force in the First Price Sealed Bid Private Value Auction. In *Economic Essays in Honor of Werner Hildenbrand*, ed. G. Debreu, W. Neuefeind & W.Trockel. Heidelberg: Springer Verlag pp. 333–344.

Shiryaev, A. 1995. *Probability.* New York: Springer Verlag.

Stoye, J. 2005. Statistical Decisions under Ambiguity: An Axiomatic Analysis. Technical report Northwestern University.

Strassen, V. 1965. "The Existence of Probability Measures with Given Marginals." *Annals of Mathematical Statistics* 36:423–439.

Wald, A. 1950. *Statistical Decision Functions.* New York: Wiley.

Zhou, K., J.C. Doyle & K. Glover. 1996. *Robust and Optimal Control.* London: Prentice Hall.

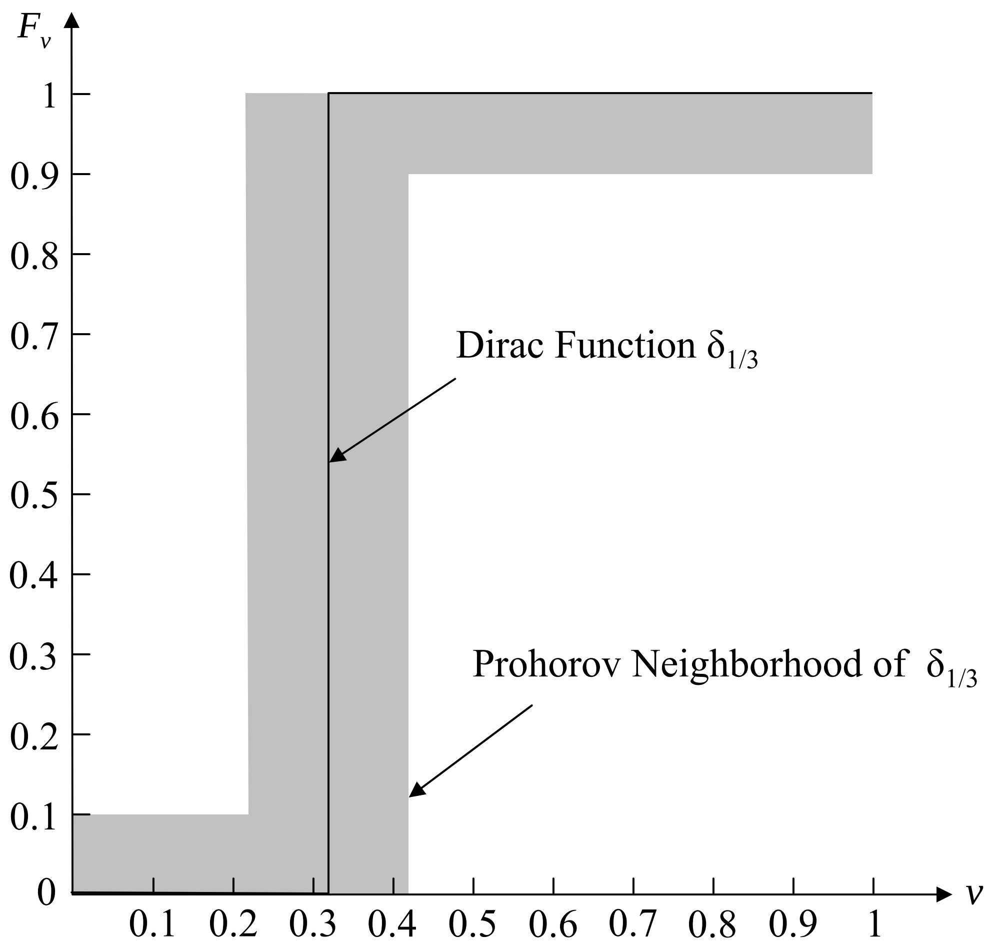

Figure 1. Prohorov Neighborhood of Dirac Function

($\delta_{1/3}$, $\varepsilon = 1/10$)

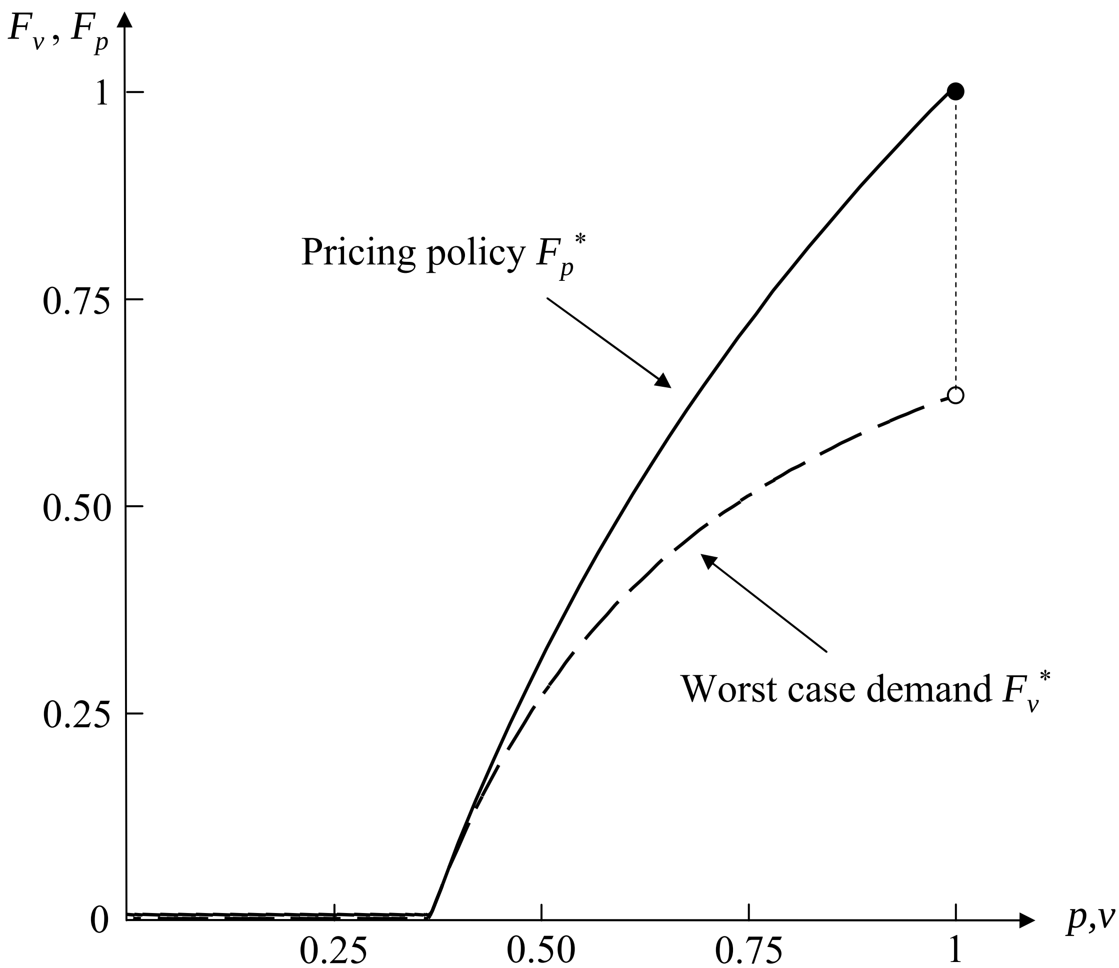

Figure 2. Optimal Pricing and Worst Case Demand with Large Risk

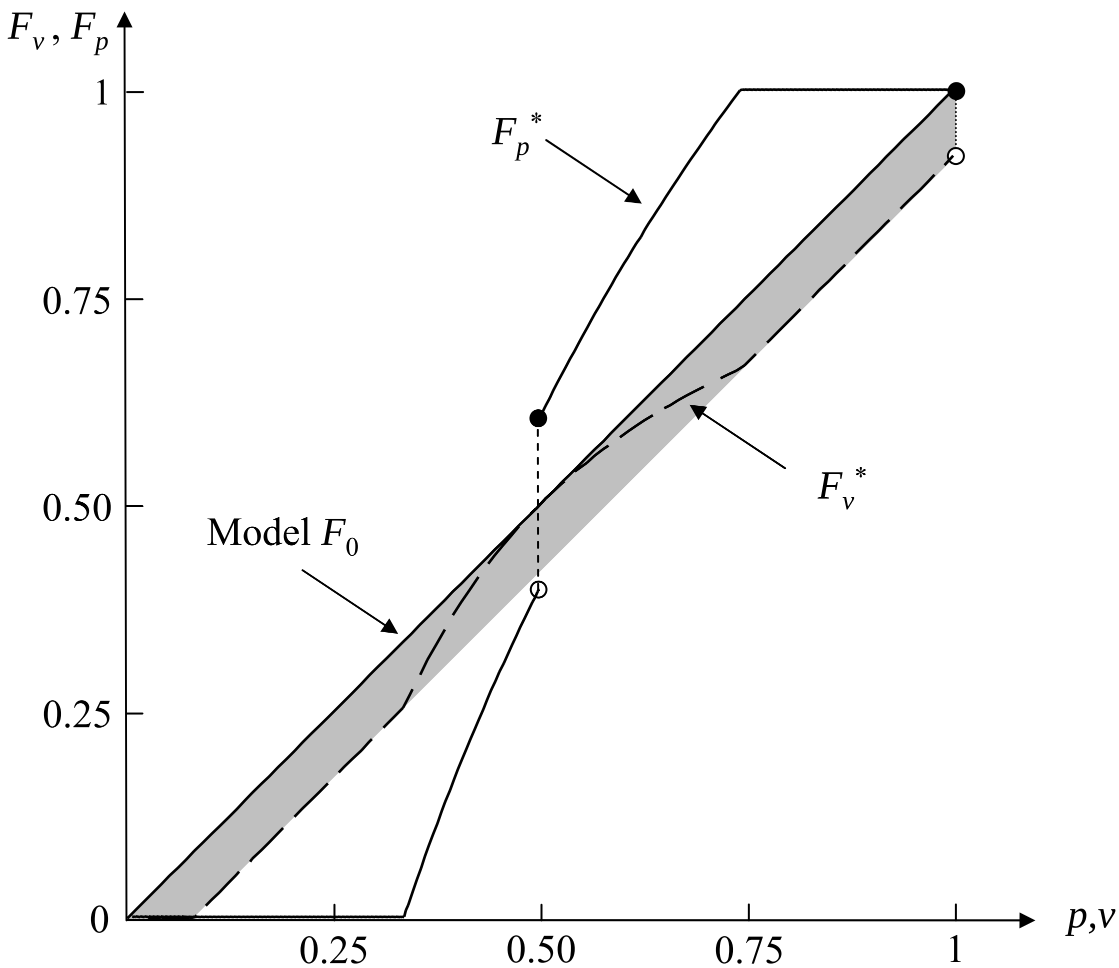

Figure 3. Optimal Pricing and Worst Case Demand with Uniform Model Density

(ε = 0.04)

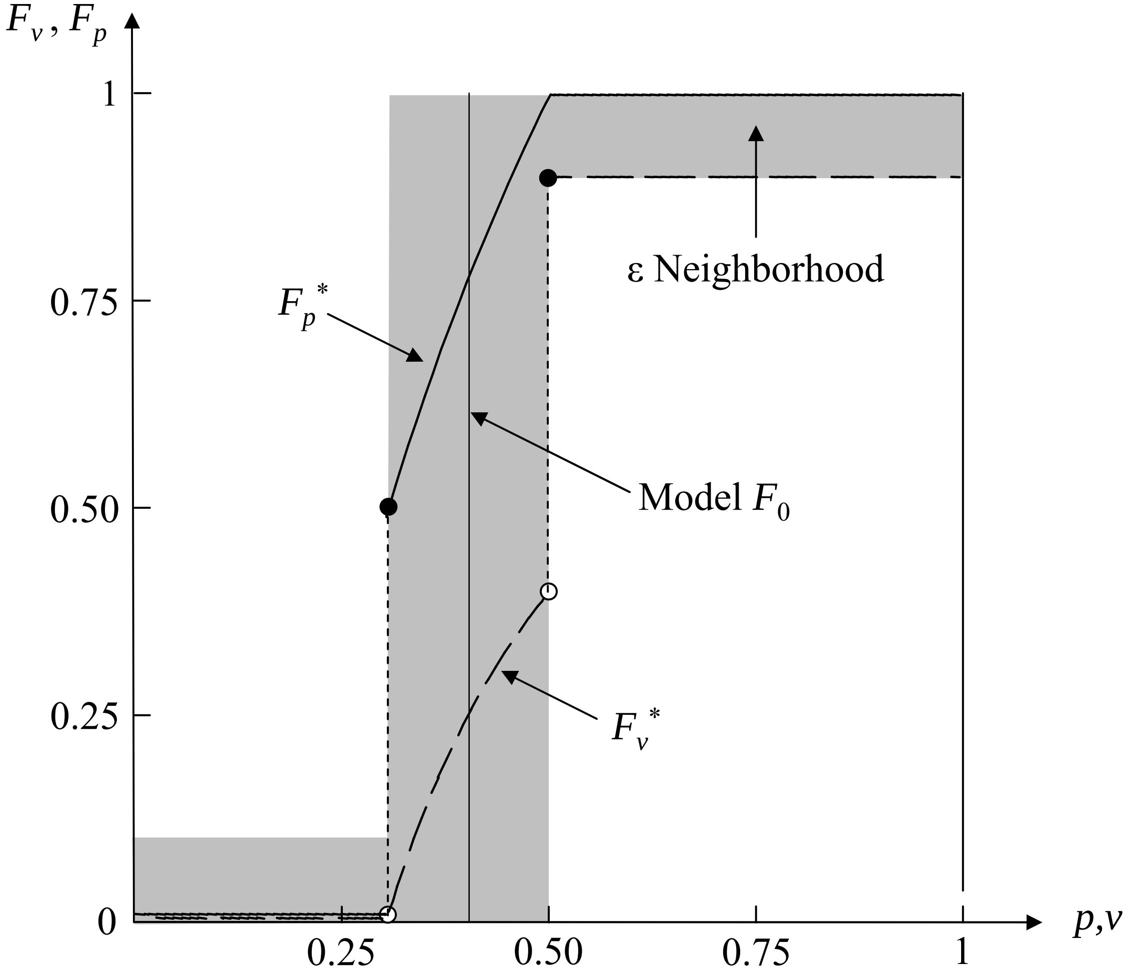

Figure 4. Optimal Pricing and Worst Case Demand with Dirac Demand

($\delta_{4/10}$, $\varepsilon = 1/10$)

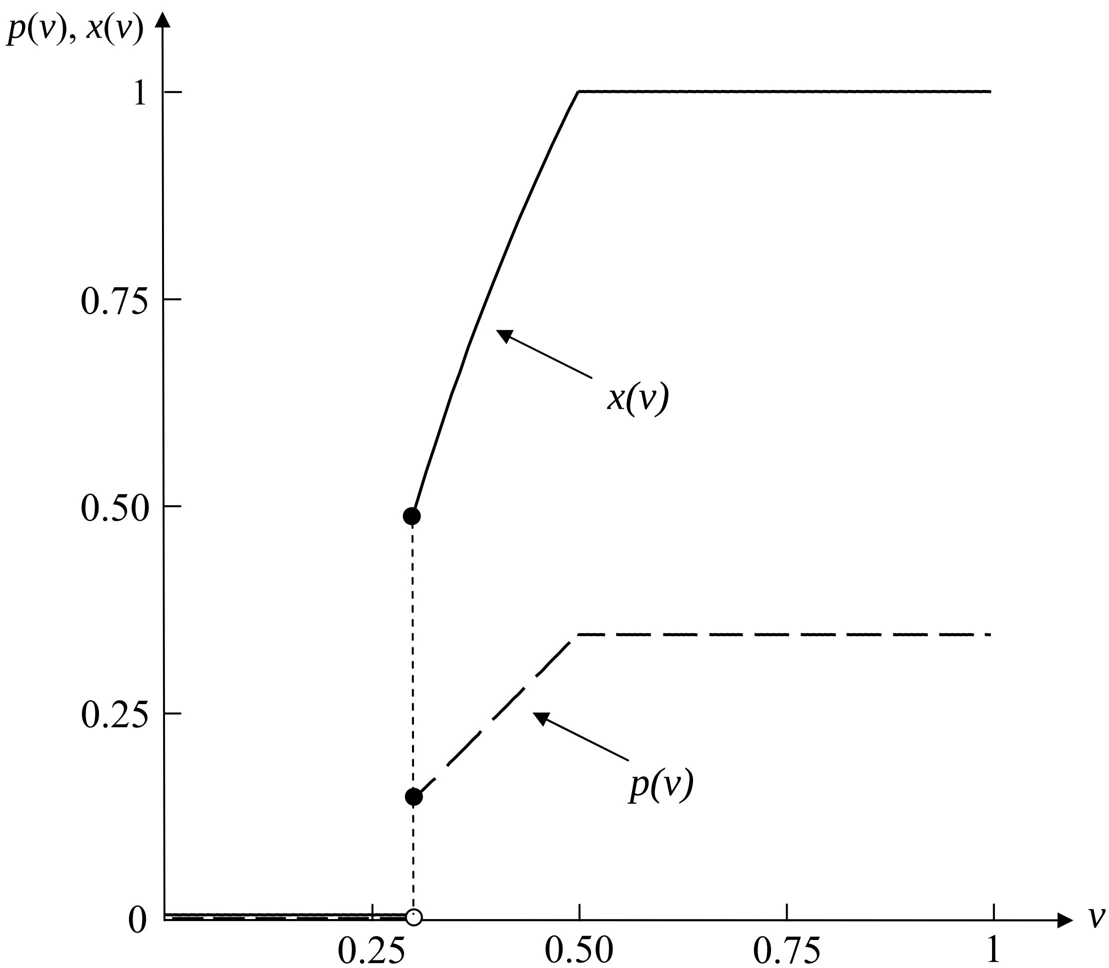

Figure 5. Robust Menu with Dirac Demand ($\delta_{4/10}$, ε= 1/10)

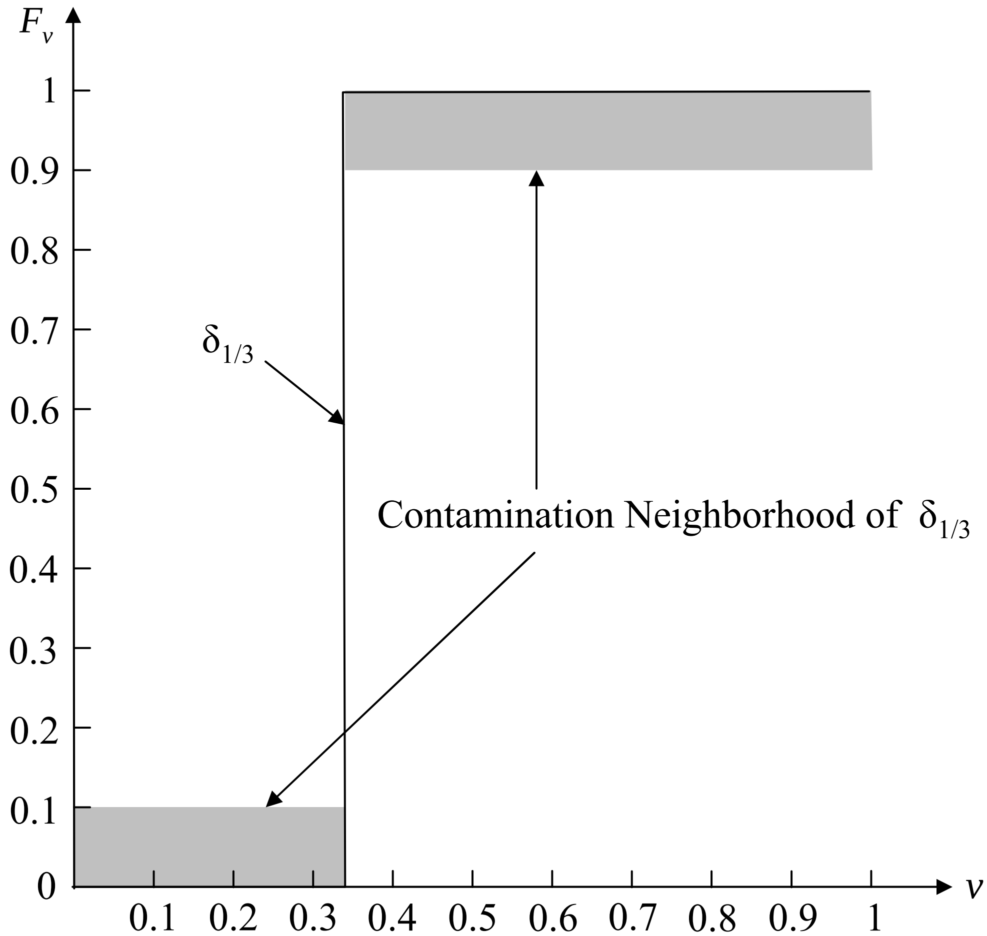

Figure 6. Contamination Neighborhood of Dirac Demand

($\delta_{1/3}$, $\varepsilon = 1/10$)