# Optimal Selling Mechanisms Under Moment Conditions[*]

Vinicius Carrasco[‡] Vitor Farinha Luz[§] Nenad Kos[¶] Matthias Messner[‖]
Paulo Monteiro[**] Humberto Moreira[††‡‡]

November 7, 2017

**Abstract**

We study the revenue maximization problem of a seller who is partially informed about the distribution of buyer's valuations, only knowing its first $N$ moments. The seller chooses the mechanism generating the best revenue guarantee based on the information available, that is, the optimal revenue is given by maxmin expected revenue. We show that the transfer function in the optimal mechanism is given by non-negative monotonic hull of a polynomial of degree $N$. This enables us to transform the seller's problem into a much simpler optimization problem over $N$ variables. The optimal mechanism is found by choosing the coefficients of the polynomial subject to a resource constraint. We show that knowledge of the first moment does not guarantee strictly positive revenue for the seller, characterize the solution for the cases of two moments and derive some characteristics of the solution for the general case.

*JEL Code*: C72, D44, D82.
*Keywords*: Optimal mechanism design, Robustness, Incentive compatibility, Individual rationality, Ambiguity aversion, Moment conditions.

[*]For helpful discussions and comments we thank Sarah Auster, Eduardo Azevedo, Dirk Bergemann, Luis Braido, Gabriel Carroll, Carlos Da Costa, Alfredo Di Tillio, Nicolas Figueroa, Daniel Garrett, Renato Gomes, Leandro Gorno, Li Hao, Johannes Höerner, Lucas Maestri, Stephen Morris, Michael Peters, Leonardo Rezende, Larry Samuelson, Yuliy Sannikov and Alexander Wolitzky and to Alexey Gorn for his invaluable research assistance. We would also like to thank seminar audiences at Boston College, Pittsburgh University, University of Bonn, University of British Columbia, Simon Fraser University, Toulouse School of Economics, Universidad Carlos III de Madrid, Princeton University, INSPER, PUC-Rio and FGV-EPGE, the 2012 Latin American Workshop of the Econometric Society, the 2014 York-Manchester Workshop in Economic Theory, the 2015 Conference on Economic Design, RUD 2015 in Milan and the 2016 SAET conference.

[†]This paper supersedes two separate papers, (Carrasco et al. (2015b) and Kos and Messner (2015)).

[‡]Department of Economics, PUC-Rio. E-mail: vncarrasco@gmail.com
[§]University of British Columbia. E-mail: vitor.farinhaluz@ubc.ca
[¶]Bocconi University, Department of Economics and IGIER. E-mail: nenad.kos@unibocconi.it
[‖]University of Cologne. E-mail: messner@wiso.uni-koeln.de
[**]FGV EPGE Brazilian School of Economics and Finance. E-mail:PKLM@fgv.br
[††]FGV EPGE Brazilian School of Economics and Finance. E-mail:humberto.moreira@fgv.br

[‡‡]Nenad would like to thank MIUR (Prin Grant 20157NH5TP20157NH5TP) for financial support. Monteiro acknowledges the financial support of CNPq. Moreira acknowledges FAPERJ and CNPq for financial support.

# Contents

# 1 Introduction and related literature

Following Wilson (1987)'s critique, a recent literature has studied the role of beliefs in mechanism design problems (see Bergemann and Morris (2005)). Most of this literature explores the impact of mechanism participants' common knowledge and higher order beliefs on implementable outcomes. On the other hand, rather little attention is paid to the fragility of a mechanisms' revenue performance with respect to changes in the designer's prior.[1]

In statistics, the virtual impossibility of fully quantifying or eliciting prior distributions is a common criticism to Bayesian Analysis. Without a full specification of the prior, the critiques go, a standard Bayesian procedure cannot be conducted. In response to such criticism, the literature on robust Bayesian analysis proposes that one should, in fact, conduct a Bayesian analysis which is robust to prior misspecification. This, in turn, amounts to deriving bounds on the relevant objective function over the range of all priors compatible with the features of the prior decision maker is able to elicit (see, for example, Berger (1982, 1984), Smith (1995) and Betrò and Guglielmi (2000)). In this paper, we follow the robust Bayesian approach and consider a design problem in which a principal, being unable to fully specify the distribution of types of an agent, maximizes expected payoffs under the worst-case distribution of values compatible with the (partial) information elicited about the prior. This is in line with Wilson (1987)'s critique: mechanisms with good performance over a wide range of possible prior distributions should be preferable.

We study the optimal robust trading rule in an environment where a revenue maximizing seller is selling an indivisible good to a single buyer. The seller has very limited information about the buyer's valuation: he is armed merely with the knowledge of a finite number of moments of the value distribution. The seller is pessimistic and evaluates any mechanism by the worst possible performance generated by a value distribution consistent with the known moments. In other words, the seller is a maxmin expected revenue maximizer.[2]

One may interpret the seller as having access to a limited amount of data. In order to avoid the "curse of dimensionality", the seller might prefer to rely on estimates of finitely many moments instead of estimating the density function, which lies in a infinite-dimensional space.[3] Instead of formally modeling the data collection and estimation processes, we look at an extreme case where the seller is certain of a number of moments and nothing else.[4,5]

[1]Some notable exceptions are Bergemann and Schlag (2008, 2011) who study optimal mechanism for sale of an object by a seller under maxmin regret and maxmin preferences. We provide a more substantial overview of the literature in the subsection related literature below.

[2]For an overview and references of partial identification of probability distributions and use of bound on moments in econometrics see Manski (1995) and Manski (2003).

[3]One of the first topics covered in basic Econometrics lectures is how to estimate the mean, which is a very simple procedure. With finite samples, the most common density estimation procedures require the ad-hoc selection of parameters such as bandwidth and the kernel function (Jones et al. (1996)).

[4]Regarding the special case of knowledge of the mean, Carroll (2013) and Wolitzky (2016) provide an interesting interpretation, which arises from the seller's uncertainty about the information acquisition technology available to the buyer.

[5]If instead of coming from (consistent) estimates, the information the seller has is elicited in a different way,

Our analysis relies on recasting the seller's problem as a zero-sum game between the seller and Nature, who chooses a feasible distribution to minimize expected revenue. Finding an optimal mechanism is equivalent to finding Nash equilibria of the induced game. We show that such a game has a Nash equilibrium. The crucial assumption for equilibrium existence is compactness of the set of feasible distributions, which is guaranteed if the seller only knows an upper bound on the highest moment restricted. The role of this inequality condition can be easily illustrated when the seller only knows the first moment. Suppose that Nature can choose any distribution with mean $k_1$ and consider the sequence of binary distributions with support $\{0, a\}$, for $a \in [k_1, \infty)$, and probability mass $\frac{k_1}{a}$ on point $a$. This sequence weakly converges to a distribution that assigns all the mass to zero as $a$ grows without bound. Although all the distributions in the sequence have mean $k_1$, the limiting distribution has mean zero. This generalizes to the observation that the set of distributions defined by a finite number of moment conditions with equality is not compact. To guarantee compactness we assume that the highest moment has an upper bound, rather than require it to hold with equality.

Our equilibrium characterization relies on studying the revenue minimization problem and the corresponding worst-case distributions. We show that, for any incentive compatible mechanism and any distribution in Nature's best-reply correspondence, one can find a polynomial of degree $N$ such that (i) the seller's ex-post payoff is bounded below by the polynomial function and (ii) the two functions are equal on the support of Nature's distribution. The second property implies that worst case revenue is determined by the coefficients of such polynomial. By implication, when searching for optimal mechanisms, the seller can restrict attention to mechanisms with transfer functions that are non-negative monotonic hulls of polynomials; where the polynomials are parametrized by their coefficients. The worst case revenue by any such mechanism is given by a linear combination of the known moments of the distribution, according to the polynomial coefficients. In other words, the seller's problem can be restated as maximization over the polynomial coefficients subject to a resource constraint. The resource constraint corresponds to the restriction that the seller has only one good to allocate. The transformed finite-dimensional problem is much simpler, as the seller is maximizing over $N$ coefficients rather than over the space of all mechanisms.

We fully characterize the optimal mechanisms in several environments. It is instructive to first revisit the environment where the seller only knows the mean of the distribution; thus the moment condition holds with equality. In this environment, for any mechanism chosen by the seller, Nature can construct a sequence of distributions such that the seller's payoff converges to zero along the sequence. This result is closely related to the above made observation that Nature can pick a sequence of distributions with a fixed mean and the limiting distribution that assigns full mass to value zero. The problem that the mean of the limiting distributions would

it might be more natural to assume that either the seller knows one general moment condition in the form $\int_0^\infty h(\theta) dF(\theta) = k$ for some increasing function $h(\cdot)$. This encompasses, for instance, the knowledge of a quantile of the distribution or some information that allows him to impose bounds on the value of some moments (Berger (1984) gives examples of such information). Our model can be easily extended to deal with both cases. The case we consider throughout is, in fact, the most difficult to deal with, as verifying the regularity conditions that allow for the use of the Lagrangian method we adopt is more challenging for equality constraints.

not be in the set is avoided if one assumes that the mean of distribution does not exceeds some value, rather than equals it. However, in that case Nature can trivially guarantee that the seller has payoff zero by placing full mass on valuation zero.

Of greater interest is the case where the seller has information only about the first two moments: he knows the first moment and the upper bound on the second. Our characterization of the general case implies that the transfer rule in the optimal mechanism coincides with the non-negative monotonic hull of a second degree polynomial. We then show that the quadratic coefficient of the polynomial has to be negative, and that the agent is allocated the good with positive probability only on the portion where the polynomial is strictly increasing and non-negative. With a quadratic polynomial such an area is an interval. The seller's optimal mechanism can be interpreted as a non-linear menu or a randomization over an interval. More precisely, the seller randomizes over the interval with a density of the form $h(\theta) = a/\theta + b$ where $a$ and $b$ are some non-zero constants. Interestingly, the seller's mechanism incentivizes Nature to assign probability mass only within the interval where the seller assigns the good and, moreover, the seller's payoff depends only on the first two moments of the distribution Nature uses, as long as its support is contained in the before-mentioned interval. In other words, the seller insures himself against the information about the distribution that he does not have. In the environment with the first two moments, the upper bound on the second moment binds invariably. When the seller has information about more than two moments, however, the bound on the highest moment need not bind. This is illustrated by the case $N = 3$ characterized in subsection 4.2.

We also study the problem where an upper bound $\bar{\theta}$ on the valuations is known. To explore this venue, we characterize the unique seller's optimal mechanism in the environment where the seller knows the first moment and the upper bound on the support. Much as in the case with the first two moments, the seller only assigns the good to the buyer with valuations in some interval. The mechanism can be interpreted as a randomization over prices with a density of the form $h(x) = a/x$, where $a$ is some constant. The problems where there is an upper bound on the valuations, $\overline{\theta}$, and the seller knows the first moment, $k_1$, is closely related to the problem where the seller knows the first moment $k_1$ and an upper bound on the second moment $k_2$. For each $k_1$ and each $\overline{\theta}$ there exists a $k_2$ such that the seller's payoff is the same regardless of which of the two pieces of information he possesses. This is not very surprising, an upper bound on valuations implicitly imposes an upper bound on the variance. This notwithstanding, the two problems are not identical: the seller's pricing scheme in the two cases differs. In particular, the price distribution in the case when the seller knows the upper bound on the valuations first order stochastically dominates the price distribution he uses when he knows the payoff equivalent upper bound on the second moment.

### Related literature

Wilson's critique and Bergemann and Morris (2005) have initiated a large body of literature on robust mechanism design. For an in-depth review see Bergemann and Morris (2013). Our paper is closely related to the work of Bergemann and Schlag (2011). They consider the problem of a seller selling a single good to a buyer. The seller is a maxmin expected utility maximizer with imperfect information about the distribution over the valuations: he knows that the valuations are distributed in an epsilon neighborhood, using the Prokhorov metric, of some distribution. In their environment Nature has a dominant strategy, therefore a deterministic take-it-or-leave-it price is optimal.[6] Auster (2016) analyses a model with common values in which the seller is privately informed, and ambiguity is of the similar form as in Bergemann and Schlag (2011). Garrett (2014) studies a model of cost-based procurement in which the seller is uncertain about the agent's effort cost function. Brooks (2014) explores an environment in which the seller is uninformed about demand, as opposed to the buyers who are well informed. He characterizes a mechanism that maximizes the minimum ratio between expected revenue and expected efficient surplus. Carrasco et al. (2015a) studies the consequences of revenue maximization with knowledge of the first moment of the type distribution in settings with curvature with applications to regulation, taxation and insurance provision.

Azar and Micali (2013), similarly to us, study an environment in which a seller has information only about the mean and the variance of the distribution of buyers' valuations. In their environment the seller has many goods and faces many potential buyers. They focus on proposing a mechanism that works well for a class of distributions, rather than deriving an optimal mechanism. For a related work where the seller only knows a quantile of the distribution of valuations see Azar et al. (2013). More closely related to ours is the work by Pinar and Kizilkale (2015). They explore environments where the seller has information about the mean and the upper bound on the agent's valuations and characterize the optimal pricing policy in the environment with finitely many possible types.

Wolitzky (2016) studies efficiency in a bilateral trade model in which the buyer and the seller know only the mean of each other's valuations. He shows that under some parameters the efficient trade is possible and characterizes when exactly that is the case. Our paper, on the other hand, is concerned with revenue maximization and the value of information for the seller. Also, unlike in Wolitzky (2016), we allow for the information about higher moments. Carroll (2012) studies the problem of providing robust incentives for information acquisition. In his model the decision making maxmin expected utility principal is incentivizing an expert to acquire costly information. Ollár and Penta (2017) study full implementation under belief restrictions, including moment restrictions. Kremer and Snyder (2015) study benefits of preventives versus treatments for a monopolist who can sell a medical product to consumers. By making use of more information about the risk distribution (which is akin to using all the information the

[6] López-Cunat (2000), Bergemann and Schlag (2008) and Bergemann and Schlag (2011) explore the seller's problem when he is minimizing his regret.

seller has about the prior distribution to price in our paper), treatments allow for better price discrimination and lead to larger revenues. The worst case difference in revenues is attained when the distribution of risk (i.e., the demand) satisfies a power law, which is also a property of the worst case demand we derive.

Lopomo et al. (2014) explore robustness of mechanisms under incomplete preferences, as in Bewley (2002). Castro and Yannelis (2012) approach the problem from a different perspective and show that every efficient allocation rule is incentive compatible if and only if the agents have maxmin preferences.

Robustness in the context of moral hazard has been explored in Lopomo et al. (2011), Chassang (2013) and Carroll (2015), to name a few.

Our paper is also related to the growing literature on mechanism design under ambiguity aversion. Though in that literature, unlike in the present paper, the buyers are the ones who are ambiguity averse. See for example Bose et al. (2006), Bose and Daripa (2009) and Bodoh-Creed (2012). More recently Bose and Renou (2014) and Di Tillio et al. (2014) have shown that in such environments the seller might benefit from using non-standard mechanisms.

## 2 Basic definitions

A seller wants to sell a single unit of a good to a buyer (agent). We denote the probability with which the good is transferred to the agent by $x$ and by $\tau$ the transfer to be paid by the agent. Throughout the paper we slightly abuse terminology and refer to $x$ as an 'allocation'.

The buyer is a risk neutral expected utility maximizing agent whose valuation for the good is denoted by $\theta$. If he receives the good with probability $x$ and pays the transfer $\tau$ in exchange, his payoff is $x\theta - \tau$. If, instead, the buyer decides not to participate in the mechanism his payoff is 0.

The seller does not observe the agent's valuation, $\theta$, and does not know which distribution it is drawn from. In particular the seller only knows some moments of the distribution (or one), and that valuations are non-negative. The set of distributions in $\Delta[0,\infty)$ he considers as possible is described as

$$\mathcal{F} = \mathcal{F}(k) \equiv \left\{ F \in \Delta[0,\infty) \mid \int_0^\infty \theta^i dF(\theta) = k_i, \text{ for } i = 1, ..., N-1, \text{ and } \int_0^\infty \theta^N dF(\theta) \leq k_N \right\},$$

where $k_i$ is the value of the $i$-th moment, $k = (k_1, ..., k_N)$ and $N \in \mathbb{N}$. Notice that the restriction on the last moment is with inequality rather than equality. The set where one requires that all the moment conditions hold with equality is rather unruly, its main drawback being that a sequence of distributions in such a set can converge to a distribution whose $N$-th moment is strictly below $k_N$.[7] We comment more extensively on this in subsection 3.1.

We consider moment restrictions on $k \in \mathbb{R}_+^N$ such that the set $\mathcal{F}(\cdot)$ is non-empty on a neighborhood of $k$, that is, we assume that:

**Assumption 1.** There exists $\varepsilon > 0$ such that $\mathcal{F}\left(\tilde{k}\right) \neq \emptyset$, for all $\tilde{k} \in \mathbb{R}_+^N$ with $\left|\tilde{k} - k\right| < \varepsilon$.

The primitive conditions on the vector $k \in \mathbb{R}_+^N$ for this regularity condition in Assumption 1 to be satisfied are known as the truncated Stieltjes problem and are discussed in the Appendix A. Hereafter we invoke Assumption 1, which implies that our revenue problem is well-defined in a neighborhood of $k$. The discussion in the Appendix A implies that the set of moment vectors that satisfy Assumption 1 is open and dense in the set $\{k \in \mathbb{R}_{++}^N \mid \mathcal{F}(k) \neq \emptyset\}$.

The seller does not assign any value to the good. He evaluates mechanisms with respect to their worst case expected revenue that he seeks to maximize. Without loss of revenue, we can restrict attention to "naive" direct mechanisms $(q, t)$, where $q : [0, \infty) \to [0, 1]$ is an allocation rule and $t : [0, \infty) \to \mathbb{R}$ is a transfer function.[8] If $t$ is the transfer function in an incentive compatible mechanism and the seller believes that the buyer's valuation is drawn from a distribution in the set $\mathcal{F}(k)$, then his payoff is

$$\inf_{F \in \mathcal{F}(k)} \int_0^\infty t(\theta) dF(\theta).$$

In what follows we assume that the seller can hedge against ambiguity through randomization, as in Raiffa (1961), which corresponds to $\delta = 1$ in the random-uncertainty-averse representation in Saito (2015). This assumption is commonly made (sometimes implicitly) in the related mechanism design literature, see for example Bergemann and Schlag (2011) and Auster (2016).

The seller's problem can be succinctly described as

$$\sup_{M \in \mathcal{M}} \left[ \inf_{F \in \mathcal{F}} \int_0^\infty t(\theta) dF(\theta) \right], \tag{$\mathcal{P}_0$}$$

where $M = (q, t)$ is a mechanism and $\mathcal{M}$ denotes the set of individually rational and incentive

[7] The set $\underline{\mathcal{F}}(k) \equiv \left\{F \in \Delta[0, \infty) \mid \int_0^\infty \theta^i dF(\theta) = k_i, \text{ for } i = 1, ..., N\right\}$ is not closed in the weak topology and its closure is $\mathcal{F}(k)$. The interested reader should consult Theorems 25.11, 25.12 and the corollary that follows them in Billingsley (1995).

[8] Suppose the buyer knows his own value and that he has possible type $t \in T$ for arbitrary $T$. The valuation function is a mapping $v : T \mapsto \mathbb{R}_+$ and assumed to have image $v(T) = \mathbb{R}_+$. Any mechanism with message space $S$ and outcome rule $(\chi, \tau) : T \mapsto [0, 1] \times \mathbb{R}_+$ an optimal reporting strategy for the buyer is $\sigma^* : T \mapsto S$ such that $v(t)\chi(\sigma^*(t)) - \tau(\sigma^*(t)) \geq v(t)\chi(s) - \tau(s)$ for any $t \in T$ and $s \in S$. The original mechanism can be substituted by a new one with message $S' = \mathbb{R}_+$ and outcome rule $(\chi'(\theta), \tau'(\theta))$ given by $\chi'(\theta) \equiv \left\{\chi(s) \mid s \in \sigma^*\left(v^{-1}(\theta)\right)\right\}$ and $\tau' \equiv \max\left\{\tau(s) \mid s \in \sigma^*\left(v^{-1}(\theta)\right)\right\}$. This new mechanism is incentive compatible, i.e., reporting strategy $\sigma'(t) = v(t)$ is optimal. Moreover, it has the property that $\tau'(\sigma'(t)) \geq \tau(\sigma^*(t))$, for any $t \in T$. In particular, $T$ could be taken as the universal type space generated by underlying uncertainty over values in $\mathbb{R}_+$.

compatible mechanisms:

$$\mathcal{M} \equiv \left\{ (q,t) : [0,\infty) \mapsto [0,1] \times \mathbb{R} \mid \begin{array}{c} \theta q(\theta) - t(\theta) \geq \max\{0, \theta q(\theta') - t(\theta')\}, \text{ for all } \theta, \theta' \in \mathbb{R}_+, \\ t(0) = 0 \text{ and } q(\cdot) \text{ is right-continuous.} \end{array} \right\}.$$

The first requirement is a compact way of writing that the mechanism is incentive compatible and individually rational. With respect to the requirement $t(0) = 0$, notice that any incentive compatible mechanism $M = (q,t)$ such that $t(0) < 0$ is dominated by mechanism $M' = (q, t - t(0))$. The continuity requirement is technical and without loss of revenue. The set of allocation rules $\mathcal{Q} \equiv \{q(\cdot) \mid (q,t) \in \mathcal{M} \text{ for some } t\}$ is endowed with the topology of convergence in distribution: we say $q^n \to q$ if $q^n(\theta) \to q(\theta)$ for every $\theta$ where $q(\cdot)$ is continuous.[9] If $\lim_{\theta\to\infty} q(\theta) = 1$, offering a menu of allocation probabilities to the buyer is payoff-equivalent to using a randomization over prices. In this case, we can treat any allocation rule $q(\cdot) \in \mathcal{Q}$ as a cumulative price distribution.[10]

## 3 Existence of Nash equilibrium

Denote the expected revenue generated by mechanism $M \in \mathcal{M}$ and distribution $F \in \mathcal{F}$ as $U(M,F)$. Instead of directly solving the seller's problem $\sup_{M\in\mathcal{M}} \inf_{F\in\mathcal{F}} U(M,F)$, we solve for a saddle point of the functional $U(M,F)$. We look for a pair $(M^*, F^*)$ —mechanism and distribution of valuations— such that

$$U(M^*, F) \geq U(M^*, F^*) \geq U(M, F^*),$$

for all feasible pairs $(M,F)$. A standard result for zero-sum games states that if such a saddle point exists, then

$$M^* \in \operatorname*{argmax}_{M\in\mathcal{M}} \inf_{F\in\mathcal{F}} U(M,F) \text{ and } U(M^*,F^*) = \sup_{M\in\mathcal{M}} \inf_{F\in\mathcal{F}} U(M,F).$$

One can think of the seller's optimization problem as the problem of finding a subgame perfect equilibrium of a sequential zero-sum game played between the seller and Nature in which the seller chooses the mechanism, Nature chooses the distribution the seller moves first

[9]If mechanism $M = (q,t)$ is incentive compatible, the mechanism $M' = (q', t')$, where $q'$ and $t'$ are the right-continuous Lebesgue-a.e. equal version of $q$ and $t$, is also incentive compatible and, since $t(\cdot)$ is non-decreasing, satisfies $t'(\cdot) \geq t(\cdot)$.

[10]Clearly a randomization over prices is an incentive compatible and individually rational mechanism. On the other hand, starting from any incentive compatible and individually rational mechanism one can obtain a randomization over prices. That is, in an incentive compatible mechanism the allocation rule is monotonic and, therefore, the cumulative distribution over prices can be defined to be equal to the allocation rule. If no type gets the object with probability one, this can be replicated by assigning high probability to high prices. Since both the allocation rule from the original mechanism, as well as randomization over prices assign the object with the same probability to each type the two mechanisms have the same transfers up to a constant. Since any randomization over (non-negative) prices leaves type 0 with zero payoff, the randomization over prices achieves at least as high a payoff as the original mechanism. See Skreta (2006), Monteiro (2009) and Kos and Messner (2013a,b) for how to deal with the cases where the distributions are not continuous.

and Nature's payoff is the negative of the seller's. Instead of solving directly for such a subgame perfect equilibrium we solve for a Nash equilibrium $(M^*, F^*)$ of the simultaneous move version of this zero-sum game, which corresponds to a saddle point of the payoff functional U. The properties of a saddle point imply that the seller's equilibrium strategy in the simultaneous move game, $M^*$, is also his maxmin strategy (i.e. his equilibrium strategy in the subgame perfect equilibrium of the sequential game).

It is instructive to first study the possibility of the seller only knowing the mean, and then the case where seller knows more than one moment.

### 3.1 The mean case

Here we focus on the case $N = 1$ with $k_1 > 0$. Given the definition of $\mathcal{F}(k_1)$, in particular that the first moment condition holds with inequality, the problem is trivial. Nature chooses a distribution with null mean and the seller's maxmin payoff is 0 irrespective of what mechanism he chooses. To show what goes wrong if the $N$-th moment condition was required to hold with equality, we explore this case with $N = 1$ in more detail. Suppose that the seller knows only the mean of the distribution, $k_1$, that is, the first moment condition holds with equality.

The set of distributions over prices is equivalent to the set of allocation rules $\mathcal{Q}$. With some abuse of notation we write $U(\tau, F)$ for the payoff that the seller obtains when he adopts the deterministic price $\tau$ and the buyer's valuation is distributed according to $F$. Since for a given price $\tau$ type $\theta$ of the buyer acquires the good only if $\theta \geq \tau$ we have

$$U(\tau, F) = \tau(1 - F(\tau-)),$$

where we define $F(\tau-) = \lim_{\tau' \uparrow \tau} F(\tau')$. The payoff that the seller obtains by randomizing according to the price distribution $q$ when the buyer's type is drawn from $F$ is, with some abuse of notation, denoted $U(q, F)$, i.e.

$$U(q, F) = \int_0^\infty U(\tau, F) dq(\tau) = \int_0^\infty \tau(1 - F(\tau-)) dq(\tau).$$

Since the seller evaluates each pricing strategy $q$ according to its performance in the corresponding worst case scenario, his problem can be formulated as

$$\sup_{q \in \mathcal{Q}} \inf_{F \in \mathcal{F}(k_1)} U(q, F).$$

**Proposition 1.** *For every price distribution $q \in \mathcal{Q}$, $\inf_{F \in \mathcal{F}(k_1)} U(q, F) = 0$. Consequently,*

$$\sup_{q \in \mathcal{Q}} \inf_{F \in \mathcal{F}(k_1)} U(q, F) = 0.$$

The proof of this proposition and all the missing proofs can be found in the Appendix B.

Proposition 1 shows that the value of the seller's problem is zero. A maxmin seller who only holds information about the mean of the value that the buyer may assign to the good, cannot expect to make any gains from trade. The crucial insight behind the result is that a seller who deems it possible that the buyer might have arbitrarily high valuations, must also believe that among the admissible type distributions there are distributions which place an arbitrarily large fraction of probability mass on 0.

The problem can be cast slightly more widely. Suppose that instead of knowing the mean, the seller only knows the upper and the lower bound on the mean, $\underline{k}_1$ and $\overline{k}_1$, respectively.[11] The value of the seller's problem is still zero: clearly the seller can not gain by having even less information about the distribution of valuations.

On a more technical aside, in the proof of Proposition 1 we construct a sequence of distributions such that each element of the sequence has mean $k_1$, yet the limiting distribution has mean 0. This prevents one from concluding that the sup-inf of the problem is equivalent to the maxmin; in fact, the latter does not exist. While in the case of $N = 1$ it is easy to solve for the value of the problem anyway, we will benefit greatly from the existence of equilibrium when higher order moments are considered.

### 3.2 More than one moment

Hereafter we assume that $N \geq 2$. When the seller knows values of more than one moment, the revenue problem also admits a saddle-point. That is, we can find a pure strategy Nash equilibrium of the zero-sum game involving the choice of a mechanism and a distribution. A direct implication is that the optimization problem ($\mathcal{P}_0$) has a solution. The analysis – which can be found in Appendix A – proceeds by first showing that Nature's strategy set is compact in an appropriate topology and that the seller's strategy set can be restricted to a compact set without loss of generality. The latter result is interesting in and on itself. In particular, we show that high posted prices are dominated for the seller.

**Lemma 1.** *(Uniform bounds on revenue) There exist* $0 < b < B < \infty$ *such that*

$$\sup_{\theta \geq B, F \in \mathcal{F}} \theta \left(1 - F\left(\theta\right)\right) < \inf_{F \in \mathcal{F}} b \left(1 - F\left(b-\right)\right).$$

There exist $b, B$ such that any posted price above $B$ yields a smaller revenue than posted price $b$ for every distribution $F \in \mathcal{F}$.[12] The second moment guarantees the seller a positive payoff at low prices—prices below $k_1$—, but restricts the seller's payoff at very high prices.[13] We

[11] For an excellent treatment of use of bounds on moments in econometrics see Manski (1995). It is not hard to see that we could have presented our original problem with upper bounds (inequalities) on the moments conditions instead of equalities. Indeed, compactness of the distribution set and regularity condition of the Lagrangian are easier to obtain under inequality constraints.

[12] The proof of Lemma 1 establishes a stronger result than in the statement of the result.

[13] At prices above $k_1$ the seller might expect payoff zero. If the second moment is small enough, Nature can put the whole probability mass on valuations below the price.

showed that the seller's payoff is 0 when he only knows the mean by constructing a sequence of distributions that assign increasingly high probability to value 0. However, a simple calculation reveals that the second moment diverges along the sequence. Intuitively, bounding the second moment prevents Nature from attaching enough probability to high values to enable it to attach arbitrarily high probability to value 0.

Lemma 1 implies that high prices are dominated for the seller in the zero-sum game. Even when a mechanism involves intensive margin distortions, no part of the good should be sold for more than $B$. Let $B$ be as given by Lemma 1 and define

$$\mathcal{M}_B \equiv \{(q,t) \in \mathcal{M} \mid q(B) = 1\}$$

of feasible mechanisms in which no marginal probability of allocating the good is sold at prices above $B$.

Next we show that the zero-sum game is payoff secure. The game is payoff secure for a player $i \in \{seller, Nature\}$ if, for any original strategy profile generating payoff $z \in \mathbb{R}^2$ player $i$ has a (potentially different) strategy that guarantees a payoff close to $z_i$ for himself whenever his opponent's strategy is close to the originally chosen one. We then use Reny (1999)'s result to show that a payoff secure zero-sum game has a pure strategy Nash equilibrium. In Appendix A we show the details of the proofs of these results. As we can guarantee that the zero-sum game is payoff-secure, the existence of a (pure) Nash equilibrium follows from Reny (1999).

**Proposition 2.** *The zero-sum game has a Nash equilibrium.*

## 4 Characterization of optimal mechanisms

In this section we show that the optimal mechanism has a transfer function which is the non-negative monotonic hull of a polynomial function of degree $N$. As a consequence, the problem of revenue maximization can be reduced to a simple problem of choosing the parameters of this polynomial function.

For any $\lambda = (\lambda_0, ..., \lambda_N) \in \mathbb{R}^{N+1}$, define $T^\lambda$ as the non-negative monotonic hull of the polynomial $\theta \mapsto \sum_{i=0}^{N} \lambda_i \theta^i$, that is,

$$T^\lambda(\theta) = \max\left\{0, \sup\left\{\sum_{i=0}^{N} \lambda_i \tau^i \mid \tau \leq \theta\right\}\right\}.$$

**Lemma 2.** *Let* $(q,t) \in \mathcal{M}_B$ *and* $F_0 \in \mathcal{F}$ *be such that*

$$\int_0^\infty t(\theta)\, dF_0(\theta) = \min_{F \in \mathcal{F}} \int_0^\infty t(\theta)\, dF(\theta).$$

*Then, there is* $(\lambda_0, \ldots, \lambda_N)$ *satisfying:*

**(i)** $t(\theta) \geq T^{\lambda}(\theta) \geq \sum_{i=0}^{N} \lambda_i \theta^i$, *for all* $\theta \geq 0$, *with equality* $F_0$ *almost surely;*

**(ii)** *the first and last non-zero entry of* $\lambda$ *are negative;*

**(iii)** *the seller's revenue is* $\int_0^\infty t(\theta)\, dF_0(\theta) = \sum_{i=0}^{N} \lambda_i k_i$.

Lemma 2 establishes that any transfer rule from a mechanism in $\mathcal{M}_B$ is bounded below by a polynomial, with the two functions coinciding on the support of the revenue-minimizing distribution. Furthermore, incentive compatibility implies that the transfer rule is monotonic. Moreover, evaluating the equation above at zero implies $\lambda_0 \leq 0$.

To prove Lemma 2 we first slightly restate Nature's problem. We allow Nature to minimize over a set $C$ of all non-negative measures on $\mathbb{R}_+$, with finite moments, but add a restriction that the measures should integrate to 1. The proof then proceeds in steps. First we establish the existence of Lagrangian multipliers $(\lambda_0, \lambda_1, ..., \lambda_N)$ such that

$$\int_0^\infty t(\theta) dF(\theta) - \sum_{i=0}^{N} \lambda_i \int_0^\infty \theta^i dF(\theta) \geq \int_0^\infty t(\theta) dF_0(\theta) - \sum_{i=0}^{N} \lambda_i \int_0^\infty \theta^i dF_0(\theta),$$

for every $F \in C$. Then, a judicious choice of measures $F$ at which one evaluates the left-hand side of the above inequality delivers inequality $t(\theta) \geq \sum_{i=0}^{N} \lambda_i \theta^i$, for all $\theta \in \mathbb{R}_+$, that holds with equality on the support of $F_0$.

Let $M^* = (q^*, t^*)$ be a solution to problem ($\mathcal{P}_0$), i.e., $(M^*, F^*)$ is a Nash equilibrium, for some $F^* \in \mathcal{F}$.[14] Lemma 2 implies that there exists a non-null $\lambda^* \in \mathbb{R}^{N+1}$ such that

$$t^*(\theta) \geq T^{\lambda^*}(\theta), \text{ for all } \theta \geq 0, \tag{1}$$

and this inequality holds as an equality on the support of $F^*$. We now define the optimal revenue the seller can obtain if he restricts attention to mechanisms such that transfers are the non-negative monotonic hull of polynomials, which we refer as problem:[15]

$$\max_{\lambda \in \mathbb{R}^{N+1}} \sum_{i=0}^{N} \lambda_i k_i \tag{$\mathcal{P}$}$$
$$\text{subject to } q^{\lambda}(\theta) \leq 1, \text{ for all } \theta \geq 0,$$

where using the envelope theorem we denote as $q^{\lambda}(\theta) \equiv \int_0^\theta \frac{\dot{T}^{\lambda}(\tau)}{\tau} d\tau$ the implemented allocation by transfer $T^{\lambda}$. Observe that for any $\lambda \in \mathbb{R}^{N+1}$, the mechanism $(q^{\lambda}, T^{\lambda})$ is in $\mathcal{M}_B$ if and only if $q^{\lambda}(\cdot) \leq 1$. The following proposition establishes that the solution to ($\mathcal{P}$) yields the overall optimal mechanism.

**Proposition 3.** *$M^* = (q^*, t^*) \in \mathcal{M}_B$ solves ($\mathcal{P}_0$) if and only if $t^*(\theta) = T^{\lambda^*}(\theta)$ and $q^*(\theta) = q^{\lambda^*}(\theta)$, for all $\theta \geq 0$, for some $\lambda^*$ that solves ($\mathcal{P}$). Moreover, the seller's revenue is $\sum_{i=0}^{N} \lambda_i^* k_i$.*

[14] In a zero sum game, if $(M_1, F_1)$ and $(M_2, F_2)$ are equilibria, so are $(M_1, F_2)$ and $(M_2, F_1)$.

[15] For any $\lambda \in \mathbb{R}^{N+1}$, the function $T^{\lambda}(\cdot)$ is differentiable except on a countable set. We denote the derivative of $T^{\lambda}(\theta)$ by $\dot{T}^{\lambda}(\theta)$. It can be defined arbitrarily for a Lebesgue-measure-zero set.

Notice that for any mechanism $M = (q, t)$ such that $t = T^\lambda$ for some $\lambda \in \mathbb{R}^{N+1}$, the following inequality holds

$$\inf_{F \in \mathcal{F}} \int_0^\infty t(\theta) dF(\theta) \geq \inf_{F \in \mathcal{F}} \int_0^\infty \left[ \sum_{i=0}^{N} \lambda_i \theta^i \right] dF(\theta) = \sum_{i=0}^{N} \lambda_i k_i.$$

Proposition 3 establishes that the optimal mechanism is in fact the non-negative monotonic hull of a polynomial function, and that the inequality holds as an equality at the optimum. The revenue problem can, therefore, be transformed into a finite parametric problem of finding weights $\lambda = (\lambda_0, ..., \lambda_N)$ that maximize the seller's revenue guarantee, subject to the induced mechanism being feasible. Notice that the feasibility assumption, $q^\lambda(\infty) \leq 1$, can be interpreted as a resource constraint corresponding to the requirement that the seller has only one good to sell and therefore $\lim_{\theta \to \infty} q^\lambda(\theta) \leq 1$. In what follows we show how the above result can be used to characterize equilibria.

### 4.1 First two moments

We submit to further scrutiny the environment where the seller has information about the first two moments: $k_1$ and $k_2$. This can be interpreted as the seller having learned the mean and a bound on the variance of the process.[16] Assumption 1 is tantamount to requiring $k_1^2 < k_2$.

First we show how Lemma 2 can be used to characterize the optimal posted price, then we turn to the characterization of the optimal mechanism.

**Optimal posted price.** The appeal of posted prices stems from their simplicity, their empirical relevance and not least from the fact that they are optimal for the seller in a standard setting with ambiguity neutral sellers. While the optimality property of posted prices does not carry over to our environment they nevertheless constitute a natural benchmark. Here we characterize the seller's optimal posted price when he has information about the first two moments of the distribution ($N = 2$).

Lemma 2 implies there exists a $\lambda \in \mathbb{R}^3$, such that $t(\theta) \geq \sum_{i=0}^{2} \lambda_i \theta^i$, for all $\theta \geq 0$, and the inequality holds with equality on the support of $F_0$.[17] Posted price is a mechanism with a transfer function $t$ that is a step function with one step. Given that this step function dominates the quadratic function, they can coincide on at most two points; for a depiction see Figure 1.

For posted prices $p > k_1$, Nature can put the full probability mass on $k_1$, thereby guaranteeing that the seller gets payoff zero. For $p < k_1$ the mechanism and the dominated quadratic function coincide on precisely two points.[18] Moreover, one of the points they coincide on is $p$. Namely,

[16] While we operate with raw moments, the problem is identical to the problem with the central moments.

[17] We are assuming that when the agent is indifferent, he does not buy the good. This guarantees that the transfer function $t$ generates a lower semi-continuous mapping $F \mapsto \int_0^\infty t(\theta) dF(\theta)$.

[18] If they coincided on one point, the support of the distribution would need to be that one point, therefore, the

for every admissible nature's distribution, there exists another nature's distribution that is admissible and shifts all the probability mass in $[0, p]$ to $p$ and gives the seller at most as high a payoff as the distribution we started with.

We are left to determine the other point to which nature assigns mass. Nature's conditions can be written as

$$\alpha p + (1 - \alpha)x = k_1,$$
$$\alpha p^2 + (1 - \alpha)x^2 \leq k_2,$$

where $\alpha$ is the probability nature assigns to $p$ and $x$ is the other point. It is easy to see that the second condition has to hold with equality at the optimum, yielding

$$\alpha(p) = \frac{k_2 - k_1^2}{p^2 - 2k_1 p + k_2}.$$

Finally, the seller is maximizing $(1 - \alpha(p))p$ over $p \in [0, k_1]$, with the solution

$$p^* = k_1 - \frac{k_2 - (k_1)^2}{2[(k_2 - (k_1)^2)(\sqrt{k_2} - k_1)]^{\frac{1}{3}}}.$$

The figure below pictures the optimal posted price and the quadratic function that is tangent to the posted price exactly at the support of the distribution, i.e., $(p^*, x^*)$ solution of the system above.

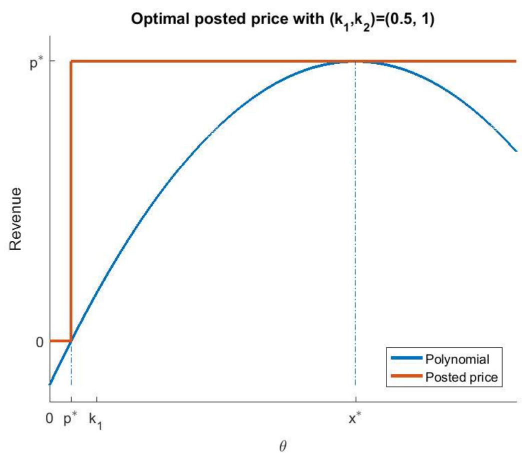


Figure 1: Optimal posted price and supporting polynomial function for two moments.

**The optimal mechanism.** The seller can, however, do better than post a price. Due to Proposition 3 we can pursue the optimal mechanism by studying the non-negative monotonic hull of a polynomial $\lambda_0 + \lambda_1\theta + \lambda_2\theta^2$.

point would have to be $k_1$. Since $p < k_1$, the seller would be trading with probability 1. This is suboptimal for nature as long as $k_2 > k_1^2$, or with other words, as longs as it can choose a distribution with a positive variance.

**Lemma 3.** *Fix a feasible pair* $(k_1, k_2)$ *and let* $(\lambda_0^*, \lambda_1^*, \lambda_2^*)$ *be a solution to problem* $(\mathcal{P})$*. Then,* $\lambda_0^* < 0$*,* $\lambda_1^* > 0$ *and* $\lambda_2^* < 0$*.*

The proof of Lemma 3 is as follows. Lemma 2 implies the existence of $\lambda_0^*, \lambda_1^*$ and $\lambda_2^*$ as well as $\lambda_0^*, \lambda_2^* \leq 0$. Since the seller can guarantee a positive payoff, from Lemma 1, we know that $\lambda_1^* > 0$. Now if $\lambda_2^* = 0$, the allocation rule $q^{\lambda^*}(\cdot)$ would, for $\theta$ sufficiently large, satisfy $\theta \dot{q}^{\lambda^*}(\theta) = \lambda_1^*$ thereby violating the resource constraint $q^{\lambda^*}(\cdot) \leq 1$. Finally, since any mechanism $(q, t) \in \mathcal{M}$ satisfies $t(0) = t'(0) = 0$, we know that $\lambda_1^* < 0$.[19]

Since the seller can guarantee a positive revenue by Lemma 1, the polynomial must be increasing and larger than 0 somewhere. Since the polynomial is quadratic and $\lambda_2^* < 0$ it can be increasing only on one interval. The following result provides a full characterization of the optimal mechanism.

**Proposition 4.** *Suppose* $N = 2$*, then the optimal mechanism is*

$$q^*(\theta) = \begin{cases} 0, & \text{if } \theta < \underline{\theta}, \\ \lambda_1 \ln \theta + 2\lambda_2 \theta - \lambda_1 \ln \underline{\theta} - 2\lambda_2 \underline{\theta}, & \text{if } \underline{\theta} \leq \theta \leq \overline{\theta}, \\ 1, & \text{if } \overline{\theta} < \theta, \end{cases}$$

*and*

$$t^*(\theta) = \begin{cases} 0, & \text{if } \theta < \underline{\theta}, \\ \lambda_2 \theta^2 + \lambda_1 \theta - \lambda_2 \underline{\theta}^2 - \lambda_1 \underline{\theta}, & \text{if } \underline{\theta} \leq \theta \leq \overline{\theta}, \\ \lambda_2 \overline{\theta}^2 + \lambda_1 \overline{\theta} - \lambda_2 \underline{\theta}^2 - \lambda_1 \underline{\theta}, , & \text{if } \overline{\theta} < \theta, \end{cases}$$

*where* $\lambda_1 = \frac{\overline{\theta}}{\overline{\theta}(\overline{\theta} - \underline{\theta}) - (\ln(\overline{\theta}) - \ln(\underline{\theta}))}$, $\lambda_2 = -\frac{1}{2\overline{\theta}(\overline{\theta} - \underline{\theta}) - (\ln(\overline{\theta}) - \ln(\underline{\theta}))}$, *and* $\underline{\theta}$ *and* $\overline{\theta}$ *are given by*

$$\overline{\theta}(1 + \ln \overline{\theta} - \ln \underline{\theta}) = k_1 \text{ and } \underline{\theta}(2\overline{\theta} - \underline{\theta}) = k_2.$$

Proposition 4 provides a characterization of the optimal direct mechanism the seller should use when informed of the mean and the variance. We argued that an optimal mechanism is derived from the non-negative monotonic hull of a quadratic polynomial $\lambda_0 + \lambda_1 \theta + \lambda_2 \theta^2$, with $\lambda_0, \lambda_2 < 0$ and $\lambda_1 > 0$. Such a polynomial can be increasing at most on one interval, and so does the probability of allocating the good $q^*(\cdot)$. This interval, $[\underline{\theta}, \overline{\theta}]$, determines the set of prices used to sell different marginal increases in the probability of allocating the good.

The mechanism has a simple alternative interpretation, the seller commits to a randomization over prices with a density $h(\theta) = 2\lambda_2 + \lambda_1/\theta$ over the interval $[\underline{\theta}, \overline{\theta}]$; with other words, function $q^{\lambda}(\cdot)$ can be seen as a cumulative distribution over prices. When the seller randomizes with the equilibrium distribution $h$ and Nature assigns the whole probability mass to the interval, the

[19] If $\lambda_0^* = 0$, then $t'(0) = \lambda_1$, which would mean that in the optimal mechanism the seller is randomizing over prices all the way down to 0.

seller's payoff is

$$\lambda_1 \tilde{k}_1 + \lambda_2 \tilde{k}_2 - \lambda_1 \underline{\theta} - \lambda_2 \frac{\underline{\theta}^2}{2},$$

where $\tilde{k}_1$ and $\tilde{k}_2$ are the moments of the distribution. This is the case even when Nature's moments do not coincide with $k_1$ and $k_2$. By randomizing with distribution $h$, the seller ensures that Nature does not assign probability mass outside $[\underline{\theta}, \overline{\theta}]$, and moreover, makes his payoff dependent only on the first two moments of Nature's distribution (when Nature assigns mass only to the before-mentioned interval). By making his payoff independent of other moments, the seller—in a sense—insures himself against the information he does not have.

The above results imply that the seller's payoff is zero whenever he has no information about Nature's distribution or when he knows only the mean. However, if he is in addition to the mean able to learn the variance, the information becomes beneficial.

Figure 2 illustrates the optimal transfer function as well as quadratic polynomial determined by the vector $(\lambda_0^*, \lambda_1^*, \lambda_2^*)$ of optimal parameters in ($\mathcal{P}$) when $k_1 = 2/3$ and $k_2 = 1$. The worst-case distribution $F^*$ is also illustrated in Figure 2.

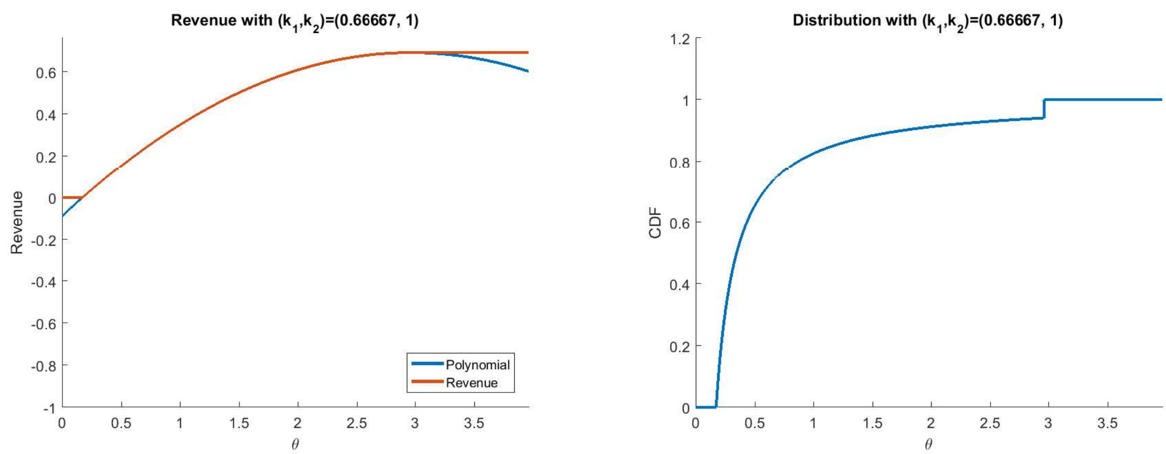


Figure 2: Optimal revenue, supporting polynomial function and distribution for two moments.

## 4.2 Beyond two moments

Much can be said about the seller's optimal mechanism in the case of more than two moments. If the seller knows first $N$ moments, then, for the same reason that $\lambda_2 < 0$ in the two moment case, it follows that the highest non-zero coefficient is negative. Due to Proposition 3 we can characterize the seller's problem by studying problem ($\mathcal{P}$). From Lemma 2, we know that the highest non-negative coefficient $\lambda_m$ is strictly negative.

Moreover, a polynomial of degree $m$ is increasing on at most $[(m+1)/2]$ intervals, where $[x]$ is the integer part of $x$. Since in the optimal allocation $q^*(\cdot)$ is determined by the smallest non-negative increasing function dominating a polynomial of degree $m$, $q^*(\cdot)$ can also be increasing on at most $[(m+1)/2]$ intervals. Accounting for the fact that Nature might put mass on value 0, corresponding to polynomial having value zero at $\theta = 0$ and being decreasing, the seller's optimal mechanism can be represented as randomization over a set that is a union of the point zero and at most $[(m+1)/2]$ disjoint intervals (possibly degenerate).

Another implication of Proposition 4 is that the seller's optimal mechanism can be represented as a randomization over intervals with density $h(\theta) = \sum_{i=1}^{m} i\lambda_i \theta^{i-2}$.

**Three moments.** In the case of two moments the second moment inequality always binds, therefore we might have as well assumed it holds with equality. The problem with three moments is where additional complications arise.

The analysis of the case of three moments depends on the magnitude of parameter $k_3$. For sufficiently large $k_3$, the third moment condition is not binding, in which case the solution, the saddle point and the maxmin value coincide with the two moment solution illustrated in subsection 4.1.

For sufficiently small $k_3$, the new constraint on the set of distributions binds, i.e., the distribution in the saddle point satisfies the three moment conditions as equalities. In this case the optimal transfer function differs from the two moment case: it is the non-negative monotonic hull of a polynomial of degree three.

To be more precise, consider two moment case with fixed $(k_1, k_2) \in \mathbb{R}_+^2$ such that $\mathcal{F}(k_1, k_2)$ satisfies Assumption 1. From Proposition 2 we know that a saddle point $\left(M^{(2)}, F^{(2)}\right)$ exists. Moreover, from subsection 4.1 we know that $F^{(2)}$ satisfies the second moment condition as an equality and has a bounded support. Now define

$$k_3^* \equiv \int_0^\infty \theta^3 dF^{(2)}(\theta),$$

which is the third moment of the worst-case distribution with two moment constraints $(k_1, k_2)$. If $k_3 \geq k_3^*$, $F^{(2)} \in \mathcal{F}(k_1, k_2, k_3)$ and hence $\left(M^{(2)}, F^{(2)}\right)$ is still a saddle point of the three moment problem; the minimax value remains unchanged.

In the case $k_3 < k_3^*$, $F^{(2)} \notin \mathcal{F}(k_1, k_2, k_3)$ and the solution to the maxmin problem as well as the saddle point differ from the two moments case. The solution to the parametric revenue problem ($\mathcal{P}$) now includes a third degree parameter $\lambda_3 < 0$. This implies that the transfer function is the non-negative monotonic hull of a third degree polynomial. Figure 3 illustrates the optimal transfer function as well as the third degree polynomial determined by the vector $(\lambda_0^*, \lambda_1^*, \lambda_2^*, \lambda_3^*)$ of optimal parameters in ($\mathcal{P}$). The worst-case distribution $F^*$ is illustrated in

Figure 3. The distribution is given by

$$F^*(\theta) = \begin{cases} 1-\alpha & , \text{if } \theta \leq \underline{\theta}, \\ 1-\frac{(1-\alpha)\underline{\theta}}{\theta} & , \text{if } \underline{\theta} < \theta \leq \overline{\theta}, \\ 1 & , \text{if } \overline{\theta} < \theta, \end{cases}$$

for $\left(\alpha, \underline{\theta}, \overline{\theta}\right) \in (0,1) \times \mathbb{R}_+^2$ satisfying $\underline{\theta} < \overline{\theta}$. Similar to the two moments case, the worst-case distribution has the property that the seller is indifferent among all prices $\tau \in \left[\underline{\theta}, \overline{\theta}\right]$. However, a new feature of this distribution is the inclusion of a probability mass of size $\alpha \in (0,1)$ at zero; this mass-point allows for the third moment condition to be satisfied while maintaining the seller's indifference among all prices in $[\underline{\theta}, \overline{\theta}]$.

The above reasoning provides a justification for imposing the highest moment condition with inequality rather than equality. When $k_3 > k_3^*$, if the third moment condition was imposed with equality Nature would like to reduce the third moment. But, for any sequence of non-negative random variables $(X_n)_n$ that converges in distribution to $X$, we can only guarantee that $E[X] \leq \liminf_n E[X_n]$; see Theorem 25.11 in Billingsley (1995). Nature can construct a sequence of distributions with first three moments $k_1, k_2$ and $k_3$ such that its weak limit is a distribution whose third moment is smaller than $k_3$. Even more, one can show that in the problem where the third moment holds with equality and $k_3 > k_3^*$ Nature can construct a sequence of distributions such that the seller's payoff converges to the payoff where he only knows the first two moments.

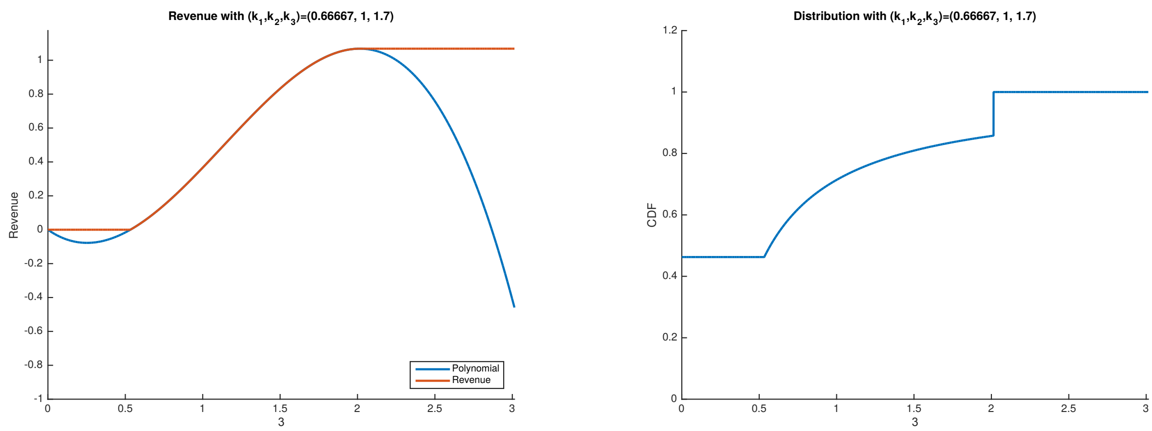


Figure 3: Optimal revenue, supporting polynomial function and distribution for three moments.

# 5 Discussion

## 5.1 Bound on valuations

In subsection 3.1 we established that the seller cannot obtain a strictly positive payoff when he only knows the first moment of Nature's distribution; this holds trivially when he only has an upper bound on the mean, as a Dirac measure with mass at point zero is in $\mathcal{F}$. A slightly more surprising result is that this no-profit result also holds if the seller knows exactly the mean of the distributions he considers as possible. The result holds as Nature is able to attain the given mean by assigning a significant portion of probability mass to zero and very little mass to arbitrarily high valuations. Somewhat counter-intuitively, Nature can hold the seller to zero payoff because he can assign positive mass to arbitrarily high values. Such extreme pessimism comes from the unbounded support assumption. For many applications it is reasonable to assume that the seller along with moment conditions entertains an upper bound on the support of Nature's distribution. The set of distributions he considers possible can then be written as

$$\overline{\mathcal{F}}(k) \equiv \left\{ F \in \Delta[0, \overline{\theta}] \mid \int_0^{\overline{\theta}} \theta^i dF(\theta) = k_i, \text{ for } i = 1, ..., N \right\},$$

where $\overline{\theta} \in \mathbb{R}_{++}$ is the upper bound of Nature's support.[20]

For most of this section we focus on the case where the seller knows only the first moment.[21] With an upper bound on support the seller can make a positive revenue despite only knowing the mean. If the seller uses a posted price $\tau < k_1$, Nature tries to maximize the mass it assigns to values smaller or equal to $\tau$. This is best done by assigning the rest of the mass to the point $\overline{\theta}$. Nature's optimal distribution is, therefore, derived from the moment condition $p(\tau)\tau + p(\overline{\theta})\overline{\theta} = k_1$, where $p(\theta)$ is the probability Nature assign to value $\theta$. The seller's payoff is

$$\frac{k_1 - \tau}{\overline{\theta} - \tau}\tau.$$

As $\overline{\theta}$ goes to infinity, the seller's payoff converges to zero, which is consistent with Proposition 1.

The seller, however, can do better than post a price. Below we characterize the seller's optimal mechanism, which involves randomization over prices. In the maxmin problem Nature gets to choose its preferred strategy after the seller picks a mechanism. When the seller posts a price, Nature has the upper hand due to the informational advantage of knowing what action the seller chose. By randomizing over prices the seller makes it harder for Nature to target its

[20] As a technical side-note, to ensure compactness of $\mathcal{F}$ in the model without the upper bound on valuations, we imposed that the highest moment condition holds with inequality. Bounding the support is another way to secure compactness.

[21] Though, notice that the upper bound on the support restricts the variance of the Nature's distribution. In fact, we will argue later that knowing the mean and the upper bound of the support is much like knowing the mean and the variance.

reply. Stochastic pricing levels the playing field by decreasing Nature's second mover advantage.

**Proposition 5.** *Suppose that the seller knows that Nature's distribution has mean $k_1$ and that its support is contained in the interval $[0, \overline{\theta}]$. Then it is optimal for the seller to commit to a randomization over prices with distribution*

$$q^*(\theta) = \begin{cases} 0, & \text{if } \theta < \underline{\theta}, \\ \frac{\ln(\theta) - \ln(\underline{\theta})}{\ln(\overline{\theta}) - \ln(\underline{\theta})}, & \text{if } \underline{\theta} \leq \theta \leq \overline{\theta}, \end{cases}$$

*where $\underline{\theta}$ is the solution to $\underline{\theta}\left(1 + \ln(\overline{\theta}) - \ln(\underline{\theta})\right) = k_1$.*

A corresponding result for a game with a finite type space was provided in Pinar and Kizilkale (2015). Much like in the previous cases, we characterize the optimal mechanism using the observation that the transfer function must coincide with the non-negative monotonic hull of a polynomial with the power equal to the number of moments, that is, the polynomial is linear. This implies that the seller's optimal mechanism can be interpreted as a randomization over an interval $[\underline{\theta}, \overline{\theta}]$, for some $\underline{\theta}$, with a density of the form $\lambda_1/\theta$. Nature's distribution can be obtained from the seller's indifference over prices in $[\underline{\theta}, \overline{\theta}]$: $F(\theta) = 1 - \frac{\underline{\theta}}{\theta}$ for $\theta \geq \underline{\theta}$, with a mass point $\underline{\theta}/\overline{\theta}$ at $\overline{\theta}$. The lower bound of the interval is recovered from the moment condition $\int_{\underline{\theta}}^{\overline{\theta}} \theta dF(\theta) = k_1$.

It is helpful to directly compute Nature's payoff when the seller uses the above mechanism. More precisely, the seller's payoff (the negative of Nature's) when Nature chooses a distribution $F$ is

$$\int_{\underline{\theta}}^{\overline{\theta}} \theta(1 - F(\theta)) dq^*(\theta) = \lambda_1 \tilde{k}_1 - \int_0^{\underline{\theta}} \lambda_1 (1 - F(\theta)) d\theta,$$

where $\lambda_1 = \frac{1}{\ln(\overline{\theta}) - \ln(\underline{\theta})}$. The second term on the right hand-side is nonnegative and strictly negative if $F(\underline{\theta}) > 0$. Nature, therefore, minimizes the seller's payoff by setting $F(\underline{\theta}) = 0$. Moreover, Nature is indifferent between all distributions that have support contained in $[\underline{\theta}, \overline{\theta}]$ and have the first moment $k_1$.

**Posted prices vs. the optimal mechanism.** Posted prices (deterministic) are the best known mechanism for sale of objects; they are particularly desirable due to their simplicity. The above analysis, however, shows that the seller forgoes some revenue if he resorts to a posted price. Here we take a closer look at how much the seller looses by adopting a posted price rather than the optimal mechanism.

In order to simplify the interpretation of the results we express the loss of revenue that the use of the best posted price implies as a fraction of the maximally achievable payoff. We denote the relative loss by $\rho(k_1, \overline{\theta})$. It is straightforward to show that both the payoff from the optimal mechanism as well as its counterpart in the case of the optimal posted price are homogeneous of degree one in $(k_1, \overline{\theta})$. Therefore, we normalize $\overline{\theta} = 1$ and study how $\rho(k_1, 1)$ varies with $k_1$.

**Proposition 6.** *The relative loss $\rho(\cdot, 1)$ is strictly decreasing. Moreover, it satisfies $\rho(0, 1) = 1$ and $\rho(1, 1) = 0$.*

The seller's relative loss is large (100%) when $k_1$ converges to 0, and it vanishes when $k_1$ converges to the upper boundary of the interval. When $k_1$ is close to 1 most of the probability over valuations must be close to 1 too, therefore there is little loss in using deterministic prices. One might think that the same result obtains when $k_1$ is close to the lower bound of the support, but this is not the case. The relative loss is strictly decreasing and approaches 100% when $k_1$ approaches 0. This is a consequence of the fact that when the seller offers a deterministic price $\tau$, Nature optimally uses a distribution on $\{\tau, 1\}$. The above result is illustrated in Figure 4. More general result in Proposition 3 can be extended to the case of bounded support. We focus here on the case $N = 1$ for brevity.

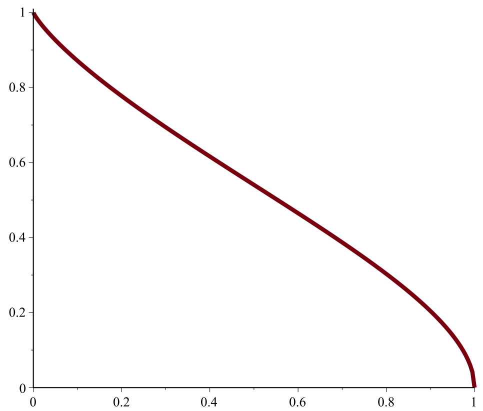


Figure 4: The relative loss of foregoing the option to randomize.

### 5.2 Second moment vs. bounded support

The bound on the support of Nature's distribution implicitly imposes a bound on the second moment. For a fixed first moment, Nature achieves the highest second moment if it assigns all the distribution to value 0 and the upper bound of the support. The problem where the seller knows the first moment and has an upper bound on the second is, thus, very similar to the problem where the seller knows the first moment and the upper bound of support.

More precisely, in the problem where the seller knows that the mean is $k_1$ and the upper bound on the valuations is $\overline{\theta}$, the seller and Nature randomize over the interval $[\underline{\theta}, \overline{\theta}]$ where $\underline{\theta}$ is defined as the solution to

$$\underline{\theta}\left(1 + \ln(\overline{\theta}) - \ln(\underline{\theta})\right) = k_1.$$

If, instead, the seller knows that the mean is $k_1$ and that the second moment is bounded by $k_2$, he randomizes over the interval $[\underline{\tau}, \overline{\tau}]$ where $\underline{\tau}$ and $\overline{\tau}$ are defined by

$$\underline{\tau}(1 + \ln(\overline{\tau}) - \ln(\underline{\tau})) = k_1 \text{ and } \underline{\tau}(2\overline{\tau} - \underline{\tau}) = k_2.$$

The two intervals of randomization coincide when $k_2 = \underline{\theta}(2\overline{\theta} - \underline{\theta})$. Furthermore, in both cases the seller's payoff is precisely $\underline{\theta}$. Therefore, knowing that valuations are distributed within

$[0, \overline{\theta}]$ with a distribution whose mean is $k_1$ is payoff equivalent to knowing that valuations are distributed with the mean $k_1$ and the second moment $k_2 = \underline{\theta}(2\overline{\theta} - \underline{\theta})$.

This notwithstanding, the problems under the two types of constraints are not equivalent. In fact, there is no pair $\overline{\theta}$ and $k_2$ such that the $k_2$-constrained and the $\overline{\theta}$-constrained problem yield the same solution to the seller's problem. The following proposition elaborates on this point. Whenever the two types of constraints lead to the same support for the price and type distributions, the price distribution that corresponds to the case with a bound on the type set first order stochastically dominates the price distribution that corresponds to the case with a constraint on the variance.

**Proposition 7.** *Let $k_1$, $\overline{\theta}$ and $k_2$ be such that the pairs $(k_1, \overline{\theta})$ and $(k_1, k_2)$ induce the same supports for the price distribution, $[\underline{\theta}, \overline{\theta}]$. If the two optimal price distributions are denoted by $q^*_{\overline{\theta}}$ and $q^*_{k_2}$, respectively, then for all $\underline{\theta} < \theta < \overline{\theta}$, $q^*_{k_2}(\theta) > q^*_{\overline{\theta}}(\tau)$.*

The above result is illustrated in Figure 5 which shows the price distributions for the pairs $(k_1, \overline{\theta}) = (1/2, 1)$ and $(k_1, \sigma^2) = (1/2, 1.088)$.

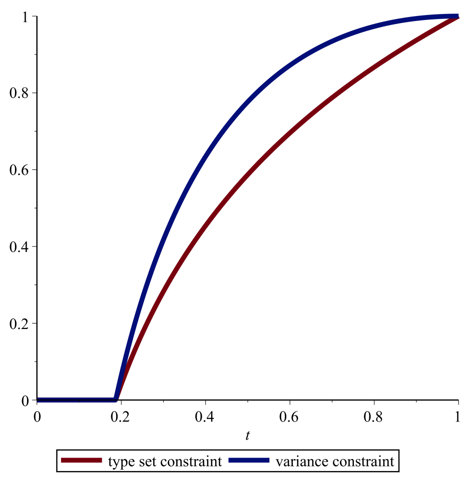


Figure 5: Optimal price distributions: bounded variance (blue), bounded type set (red).

Some intuition for why the seller's optimal mechanism in the two problems differ can be gained from the following. Remember that $q^*_{k_2}$ guarantees that the seller's payoff depends only on the first and the second moment of the distribution; as long as the support of Nature's distribution is contained in the support of the mechanism. Therefore, if the seller used this mechanism when he has information about the first moment and the upper bound, Nature would pick the distribution with the highest second moment, not the distribution it applies when the seller uses $q^*_{\overline{\theta}}$. Alternatively, mechanism $q^*_{\overline{\theta}}$ ensures that the seller's payoff depends only on the first moment of the distribution, as long as the support of Nature's distribution is contained in the support of the mechanism. The upper bound $\overline{\theta}$ serves precisely that, i.e., it prevents Nature from moving probability mass above $\overline{\theta}$. If the seller were to use mechanism $q^*_{\overline{\theta}}$ when his information was about the first two moments, Nature would shift some probability mass to values above $\overline{\theta}$.

## 6 Conclusion

In this paper we consider a seller's problem of designing a robust mechanism with knowledge of a finite number of moments of the distribution from which the buyer's valuations are drawn. We show that the seller who only has information about the mean expects payoff zero regardless of which mechanism he uses, i.e., knowledge of a single moment is useless for a pessimistic seller. When at least two moments of the distribution are known, our method shows that an optimal mechanism guarantees positive revenue and provides a method for finding this optimum. We show that the seller's problem can be reduced to finding an optimal mechanism with the property that its revenue function is the monotonic hull of a polynomial of degree equal to the number of known moments. We fully characterize the solutions for the cases of two and three known moments. If the seller knows the mean and variance of the distribution he uses intensive margin distortions, selling different parts of the good at prices in an interval. Equivalently, the seller might choose a price randomly over this interval. The worst-case distribution from the seller's perspective has as its support exactly the set of prices used by the seller in the optimal mechanism.

When the seller knows three moments, the solution may have a richer structure. Even though the seller still randomizes over an interval, two new features arise. First, the seller's worst-case distribution might be identical to the situation where only the two first moments of the distribution are known. Second, the interval of price randomization does not necessarily coincide with the support of the worst-case distribution but is contained in it. We also discussed the case of bounded support with known mean, when the seller has a positive revenue guarantee, and compared its solution with the one obtained with two known moments. Our results trivially apply to the design of selling mechanisms with divisible or quality-differentiated goods, as in Mussa and Rosen (1978). In this case, the probability of receiving the good can be interpreted as the quantity of good sold or its quality.

Our model also has positive implications for price variability. First, price variation is a necessary feature of optimal pricing with significant prior uncertainty. Such variation allows the seller to obtain significant revenue for a wide range of possible value distributions, which are consistent with minimal knowledge of certain moments of the distribution. Additionally, our characterization shows that there is a tight connection between the (partial) information the seller might have about the values distribution and the optimal randomized-price distribution. This is in contrast with the classical Bayesian model, where the seller may be indifferent among a set of prices, but no restriction is imposed on the relative frequency of such prices. Hence the use of models of price setting with prior uncertainty can provide new insight into markets with seller-generated price variation, such as the presence of sales with varying discount rates.

Although we show that the optimal mechanism is the solution of a finite dimensional optimization problem in the general case when $N$ moments are known, we believe it is possible to construct a simple algorithm that characterizes the optimal distribution step-wise including an extra moment in each step. However, this construction is beyond the scope of this paper. An interesting question for future research is the comparison between the classical Bayesian solution and the one presented here. Of a particular interest is the question what happens with the solution and robust revenue when the number of known moments goes to infinity and whether (and at what speed) both converge to their Bayesian counterparts when we take a sequence of moments consistent with a fixed prior distribution.

# Appendix A

### Non-emptiness of $\mathcal{F}$

An underlying assumption of our model is that the seller knows some moments of the distribution of the valuation, but has not fully identified the distribution. His information gives rise to the set of possible distributions $\mathcal{F}$. Needless to say, for the problem ($\mathcal{P}_0$) to be well defined it must be the case that a distribution with moments $k$ exists. In this appendix we explore what restrictions one needs to impose on the values of the moments for $\mathcal{F}$ to be non-empty.

Let $F \in \Delta[0, \infty)$ be such that $\int_0^\infty \theta^N dF(\theta) < \infty$ and let

$$h(F) \equiv \left(\int_0^\infty \theta^i dF(\theta) - k_i\right)_{i=1,\dots,N-1} \text{ and } g(F) = \int_0^\infty \theta^N dF(\theta) - k_N.$$

In order to give meaning to problem ($\mathcal{P}_0$) we need to ensure that the solution to

$$h(F) = 0 \text{ and } g(F) \leq 0 \tag{2}$$

exists and, moreover, has a solution in a neighborhood of $k = (k_1, \dots .k_N)$.

The existence of an element in $\mathcal{F}(k)$ is known as the truncated Stieltjes problem. Exact characterization of the Stieltjes moment problem (2) can be given in terms of the Hankel determinants:

$$D_{2n}^{(0)}(k) = \begin{vmatrix} 1 & \dots & k_n \\ \vdots & & \vdots \\ k_n & \dots & k_{2n} \end{vmatrix} \text{ and } D_{2n+1}^{(1)}(k) = \begin{vmatrix} k_1 & \dots & k_{n+1} \\ \vdots & & \vdots \\ k_{n+1} & \dots & k_{2n+1} \end{vmatrix}.$$

A necessary and sufficient condition for system (2) to have a solution, for a given $k \in \mathbb{R}_+^N$ is:

$$D_{2n}^{(0)}(k) \geq 0, \text{ for } n \text{ such that } 1 \leq n \leq N/2, \text{ and } D_{2n+1}^{(1)}(k) \geq 0, \text{ for } n \text{ such that } 1 \leq n \leq (N-1)/2.$$

This set is convex and if $k$ belongs to its boundary (interior), the problem has a unique (many) solution(s), i.e., the truncated Stieltjes problem is (in)determinate. The proof of these results can be found in Shohat and Tamarkin (1943) and Karlin and Shapley (1953).

As a consequence, if these determinants are strictly positive, a solution to (2) exists in a neighborhood of $k$, which ensures our Assumption 1, implying that our revenue problem is well-defined in a neighborhood of $k$.

**Lemma 4.** *Assumption 1 holds in an open and dense subset of* $\{k \in \mathbb{R}_+^N \mid \mathcal{F}(k) \neq \emptyset\}$.

These conditions are easily computed numerically. For the examples $N = 1$ and $N = 2$, the exact conditions are the following:

- $N = 1$ (first moment only), the condition for existence of a solution is equivalent to $0 \leq k_1$;
- $N = 2$ (first and second moments), the conditions for existence boil down to $0 \leq k_1$ and $(k_1)^2 \leq k_2$.[22]

## Existence of Nash equilibrium

We start with a standard implication of local incentive compatibility constraints to simplify the seller's problem by choosing only allocation rule $q(\cdot)$. We omit the proof of the following lemma which is standard.

**Lemma 5.** *(Local incentive constraints) For any mechanism $(q, t) \in \mathcal{M}$, we have*

$$t(\theta) = \int_0^\theta \tau dq(\tau),$$

*which implies that, for any $F \in \mathcal{F}$,*

$$U(q, F) = \int_0^\infty \theta (1 - F(\theta-)) dq(\theta).$$

For $\theta \in \mathbb{R}_+$, denote by $M_\theta = (q_\theta, t_\theta)$ the mechanism corresponding to posted price $\theta$, i.e.,

$$q_\theta(\theta') = \mathbf{1}_{[\theta,\infty)}(\theta'),$$

where $\mathbf{1}_A$ is the indicator function of set $A$.

Lemma 1, in the main text, provides a uniform upper bound on expected payoff for sufficiently high posted prices and a uniform lower bound on expected payoff for a sufficiently low, yet positive, price $b$. This guarantees that the seller never posts prices above threshold $B$ in equilibrium.

Next result formalizes the idea that the seller can restrict his attention to mechanisms with support contained in $[0, B]$.

**Lemma 6.** *(Bounded support mechanisms) Any mechanism $M = (q, t) \in \mathcal{M}$ such that $q(B) < 1$ is strictly dominated by some mechanism $M' \in \mathcal{M}_B$.*

*Proof.* Let $B$ be as in Lemma 1. We claim that any mechanism $(q, t) \in \mathcal{M}$ such that $q(B) < 1$ is dominated by another $(\tilde{q}, \tilde{t}) \in \mathcal{M}$ such that $\tilde{q}(B) = 1$. Indeed, the seller's revenue for a feasible distribution $F$ is

$$\int_0^\infty t(\theta) dF(\theta), \text{ where } t(\theta) = \theta q(\theta) - \int_0^\theta q(\tau) d\tau = \int_0^\theta \tau dq(\tau).$$

[22] Notice that the second condition is a simple implication of Jensen's inequality.

Now note that

$$\int_0^\infty t(\theta)\, dF(\theta) = \int_0^\infty \left(\int_0^\theta \tau dq(\tau)\right) dF(\theta) = \int_0^\infty \tau (1 - F(\tau-))\, dq(\tau).$$

Lemma 1 implies that

$$b(1 - F(b-)) > \sup_{F\in\mathcal{F}} \sup_{x\geq B} x(1 - F(x-)).$$

Thus, if we transfer the mass of $(B, \infty)$, $1 - q(B)$, to $b$ revenue improves. Formally, let

$$\tilde{q}(\theta) = \begin{cases} q(\theta) & \text{if} \quad 0 \leq \theta < b \\ q(\theta) + 1 - q(B) & \text{if} \quad b \leq \theta < B \\ 1 & \text{if} \quad B \leq \theta \end{cases}$$

and $\tilde{t}(\theta) = \int_0^\theta \tau d\tilde{q}(\tau)$. Thus,

$$\begin{aligned} \int_0^\infty \tau (1 - F(\tau-))\, d\tilde{q}(\tau) &= \int_0^\infty \tau (1 - F(\tau-))\, dq(\tau) \\ &\quad + b(1 - F(b-))(1 - q(B)) - \int_B^\infty \tau (1 - F(\tau-))\, dq(\tau) \end{aligned}$$

which implies that

$$\begin{aligned} \inf_{F\in\mathcal{F}} \int_0^\infty \tau (1 - F(\tau-))\, d\tilde{q}(\tau) \geq \inf_{F\in\mathcal{F}} \int_0^\infty \tau (1 - F(\tau-))\, dq(\tau) \\ + \inf_{F\in\mathcal{F}} b(1 - F(b-))(1 - q(B)) - \sup_{F\in\mathcal{F}} \int_B^\infty \tau (1 - F(\tau-))\, dq(\tau). \end{aligned}$$

Notice that, by Lemma 1,

$$\begin{aligned} \sup_{F\in\mathcal{F}} \int_B^\infty \tau (1 - F(\tau-))\, dq(\tau) &\leq \int_B^\infty \sup_{F\in\mathcal{F}} \tau (1 - F(\tau-))\, dq(\tau) \\ &< \int_B^\infty \inf_{F\in\mathcal{F}} b(1 - F(b-))\, dq(\tau) \\ &= \inf_{F\in\mathcal{F}} b(1 - F(b-))(1 - q(B)) \end{aligned}$$

which implies that

$$\inf_{F\in\mathcal{F}} \int_0^\infty \tau (1 - F(\tau-))\, d\tilde{q}(\tau) > \inf_{F\in\mathcal{F}} \int_0^\infty \tau (1 - F(\tau-))\, dq(\tau).$$

□

Lemma 6 enables us to restrict the seller's choice to a compact set. For any mechanism in $\mathcal{M}_B$, the allocation rule $q(\cdot)$ can be identified as a probability distribution on $[0, B]$. After endowing the set $\mathcal{M}_B$ with the topology of convergence in distribution in $q(\cdot)$, the set $\mathcal{M}_B$ is compact. Compactness is not automatically guaranteed for Nature's strategy set $\mathcal{F}$ since there is no upper bound on the possible distribution supports. However, the presence of at least two moment restrictions reduces the set enough to guarantee compactness.

**Lemma 7.** *(Compactness of $\mathcal{F}$) $\mathcal{F}$ is compact in the weak topology.*

*Proof.* For every $F \in \mathcal{F}$, the condition $\int_0^\infty \theta^N dF(\theta) \leq k_N$ implies that $\mathcal{F}$ is uniformly integrable. It suffices therefore to prove that $\mathcal{F}$ is closed. Suppose $F_n \to F$ where $F_n \in \mathcal{F}$ for every $n$. Applying Theorem 4.5.2 of Chung (2001) we have, for every $i < N$, $\int_0^\infty \theta^i dF_n(\theta) \to \int_0^\infty \theta^i dF(\theta)$. Thus, $\int_0^\infty \theta^i dF(\theta) = k_i$, $1 \leq i < N$. Finally, Theorem 4.5.1 of Chung (2001) ensures that $\int_0^\infty \theta^N dF(\theta) \leq k_N$. □

A crucial part of arguing existence of equilibrium is to show that the seller's payoff is upper semi-continuous in the allocation rule for every admissible Nature's distribution $F$. This implies that the game is payoff-secure for Nature.

**Lemma 8.** *(Upper semi-continuity of the payoff) $U(\cdot, F)$ is upper semi-continuous, for any $F \in \mathcal{F}$.*

*Proof.* Since $\theta \mapsto \theta(1 - F(\theta-))$ is upper semi-continuous, this is a consequence of Theorem 14.5 (p. 479) of Aliprantis and Burkinshaw (1998). □

The following lemma shows that the revenue maximization problem faced by the seller, for $F \in \mathcal{F}$, has a solution with a posted price. Furthermore, there exists an approximately optimal posted price at a continuity point of $F$, which implies that any distribution close to $F$ in the weak topology is also close to $F$ in terms of expected revenue.

**Lemma 9.** *(Optimality of prices) For any $F \in \mathcal{F}$, there exists $\theta_0 \in \mathbb{R}_+$ such that mechanism $M_{\theta_0}$ satisfies $U(M_{\theta_0}, F) = \sup_{M \in \mathcal{M}} U(M, F)$. Moreover, for every $\varepsilon > 0$, there exists a continuity point of $F(\cdot)$, denoted $\theta_\varepsilon$, such that the mechanism $M_{\theta_\varepsilon}$ satisfies*

$$U(M_{\theta_\varepsilon}, F) > \sup_{M \in \mathcal{M}} U(M, F) - \varepsilon.$$

*Proof.* Since the function $\theta \mapsto \theta(1 - F(\theta-))$ is upper semi-continuous, Lemma 1 implies a global maximum of this function is achieved in $(0, B)$. Denote a maximizer as $\theta_0$ and maximum be $U_0 \equiv \theta_0(1 - F(\theta_0-))$. It is clear that $0 < \theta_0 < B$ and $U_0 > 0$. Furthermore, for every $M = (q, t) \in \mathcal{M}$

$$U(M, F) = \int_0^\infty t(\theta)\, dF(\theta) = \int_0^\infty \theta(1 - F(\theta-))\, dq(\theta) \leq \int_0^\infty U_0 dq(\theta) \leq U_0.$$

Since $\theta \mapsto \theta(1 - F(\theta-))$ is left-continuous, there exists $\theta_\epsilon < \theta_0$ a continuity point of $F$ such that $U(M_{\theta_\varepsilon}, F) = \theta_\varepsilon(1 - F(\theta_\varepsilon)) > U_0 - \varepsilon$. □

## Appendix B - proofs

**Proof of Proposition 1.** Let $q$ be any distribution over prices and denote by $F_\theta$ the binary type distribution that assigns probability $p$ to $\theta > k_1$ and probability $1-p$ to 0, and has mean $k_1$ (thus, $p = k_1/\theta$). Notice that

$$\inf_{F\in\mathcal{F}(k_1)} U(q,F) \leq U(q,F_\theta) = \int_0^\infty (1-F_\theta(\tau-))\tau dq(\tau) = \frac{k_1}{\theta}\int_0^\theta \tau dq(\tau).$$

Below we argue that $U(q,F_\theta)$ converges to zero as $\theta \mapsto \infty$. The above inequality, therefore, implies that the seller's payoff from $q$ must be equal to zero. Since $q$ was chosen arbitrarily, the desired result follows.

To show that $\frac{1}{\theta}\int_0^\theta \tau dq(\tau) \to 0$, take $m > 0$. If $\theta > m$, then

$$\frac{1}{\theta}\int_0^\theta \tau dq(\tau) = \frac{1}{\theta}\left(\int_0^m \tau dq(\tau) + \int_m^\theta \tau dq(\tau)\right) \leq \frac{m}{\theta} + q(\theta) - q(m-) \leq \frac{m}{\theta} + q(\infty) - q(m-).$$

So that, taking limit $\theta \to \infty$, $\frac{1}{\theta}\int_0^\infty \tau dq(\tau) \leq q(\infty) - q(m-)$. Now taking $m \to \infty$ gives us the desired result. □

**Proof of Lemma 1.** We use the following result (see Chung (2001), exercise 11, p. 61):

Suppose $E[|X|] \geq a$ and $E\left[X^2\right] = 1$. Then, $P(|X| \geq \lambda a) \geq (1-\lambda)^2 a^2$, for $\lambda \in [0,1]$.

If $X$ is a non-negative random variable with $E[X] = k_1$ and $E\left[X^2\right] = k_2$, define $Y = \frac{X}{\sqrt{k_2}}$. Thus, $E[Y] = \frac{k_1}{\sqrt{k_2}}$ and $E\left[Y^2\right] = 1$. Choosing $a = \frac{k_1}{\sqrt{k_2}}$ and $\lambda = \frac{1}{2}$, we have

$$P\left(X \geq \frac{k_1}{2\sqrt{k_2}}\right) \geq \frac{k_1^2}{4k_2}.$$

Hence, if $b = \frac{k_1}{2\sqrt{k_2}}$, then $b(1-F(b-)) \geq b^3$, for every $F \in \mathcal{F}$. Now Markov inequality for $\theta \geq B \geq \frac{k_2}{b^3}$ implies

$$P(X \geq \theta) \leq P(X \geq B) \leq \frac{k_2}{B^2},$$

which gives the uniform bound. Hence, $x(1-F(x-)) \leq \frac{k_2}{B} < b^3$, if $x \geq B > \frac{k_2}{b^3} = \frac{8k_2^2\sqrt{k_2}}{k_1^3}$. □

**Proof of Proposition 2.** Consider the zero-sum game where the strategy set of the seller is restricted to $\mathcal{M}_B$. By Lemma 7, both players have compact strategy sets. The payoff function is linear, and hence quasi-concave.

Next we argue that the zero-sum game is payoff secure. Consider an arbitrary $M_0 \in \mathcal{M}$ and $F_0 \in \mathcal{F}$ and fix $\varepsilon > 0$. We start by showing the game is payoff secure for the seller. We need to find $M_1$ such that

$$U(M_1, F) > U(M_0, F_0) - \varepsilon,$$

for any $F \in V(F_0)$, where $V(F_0)$ is an open neighborhood of $F_0$. From Lemma 9, there exists $\theta_\varepsilon \in \mathbb{R}_+$ such that $\theta_\varepsilon$ is a continuity point of $F_0$ and

$$U(M_{\theta_\varepsilon}, F_0) > \sup_{M \in \mathcal{M}} U(M, F_0) - \varepsilon \geq U(M_0, F_0) - \varepsilon.$$

Now consider any sequence $(F_n)_n$ such that $F_n \to F_0$ (convergence in distribution). Since $\theta_\varepsilon$ is a continuity point of $F_0$, it follows that $U(M_{\theta_\varepsilon}, F_n) \to U(M_{\theta_\varepsilon}, F_0)$.

We now show that the game is payoff secure for Nature. We need to find $F_1 \in \mathcal{F}$ such that

$$U(M, F_1) < U(M_0, F_0) + \varepsilon,$$

for any $M \in V(M_0)$, where $V(M_0)$ is an open neighborhood of $M_0$. Take $F_1 = F_0$. The set

$$\underline{\mathcal{M}} \equiv \{M \in \mathcal{M} \mid U(M, F_0) \geq U(M_0, F_0) + \varepsilon\}$$

is closed by Lemma 8. Hence, $\mathcal{M} \backslash \underline{\mathcal{M}}$ is open and contains $M_0$.

Since the game is payoff secure, it has a Nash equilibrium by Reny (1999) - Corollary 3.3, p. 1035. Lemma 6 implies that a Nash equilibrium of the restricted game with strategy sets $(\mathcal{M}_B, \mathcal{F})$ is a Nash equilibrium of the unrestricted game with strategy sets $(\mathcal{M}, \mathcal{F})$. □

**Proof of Lemma 2.** Suppose that $\int_0^\infty t(\theta)\, dF_0(\theta) > 0$ (otherwise, the result is trivial). Let $D = \left\{F \in \Delta[0, \infty) \mid F(0-) = 0 \text{ and } \int_0^\infty \theta^N dF(\theta) < \infty\right\}$ and $C = \{rG \mid r \geq 0, G \in D\}$ be the convex cone generated by $D$. Define $k_0 = 1$ and, for $F \in C$, $h_i(F) = \int_0^\infty \theta^i dF(\theta) - k_i$, $i = 0, 1, \ldots, N-1$ and $g(F) = \int_0^\infty \theta^N dF(\theta) - k_N$. Note that $F \in C$ belongs to $D$ if and only if $h_0(F) = 0$. Suppose now that $F_0 \in C$ solves

$$\begin{aligned} \min_{F \in C} \int_0^\infty & t(\theta)\, dF(\theta) \text{ subject to} \\ h_i(F) = \int_0^\infty & \theta^i dF - k_i = 0, i = 0, \ldots, N-1 \\ g(F) = \int_0^\infty & \theta^N dF - k_N \leq 0. \end{aligned}$$

Using Theorem 9.4, page 182 of Clarke (2013), there exists $(\eta, \mu_0, \ldots, \mu_{N-1}, \mu_N) \neq 0$ such that $\eta \in \{0, 1\}$, $\mu_N \geq 0$, $\mu_N \cdot g(F) = 0$ and, for every $F \in C$,

$$\eta \int_0^\infty t(\theta)\, dF(\theta) + \mu_N g(F) + \sum_{i=0}^{N-1} \mu_i h_i(F) \geq \eta \int_0^\infty t(\theta)\, dF_0(\theta). \tag{3}$$

Suppose that $\eta = 0$. Let $\epsilon > 0$ be as in Assumption 1 and $\tilde{k} = k - \frac{\epsilon}{2}(\text{sign}(\mu_0), \ldots, \text{sign}(\mu_{N-1}), 1)$, then $\mathcal{F}\left(\tilde{k}\right) \neq \emptyset$. For $F \in \mathcal{F}\left(\tilde{k}\right)$, we have

$$\mu_N g(F) + \sum_{i=0}^{N-1} \mu_i h_i(F) = \frac{\epsilon}{2}\left(-\mu_N - \sum_{i=0}^{N-1} \mu_i \,\text{sign}(\mu_i)\right) = \frac{\epsilon}{2}\left(-\mu_N - \sum_{i=0}^{N-1} |\mu_i|\right) < 0,$$

contradicting (3). Hence, $\eta = 1$.

From $\int_0^\infty t(\theta)\, dF_0(\theta) > 0$ we see that $(\mu_0, \ldots, \mu_N) \neq 0$. Using $\int_0^\infty \theta^i dF_0(\theta) = k_i$, we can rewrite (3) as

$$\int_0^\infty \left[ t(\theta) + \sum_{i=0}^N \mu_i \theta^i \right] dF(\theta) \geq \int_0^\infty \left[ t(\theta) + \sum_{i=0}^N \mu_i \theta^i \right] dF_0(\theta), \tag{4}$$

for every $F \in C$. If for some $\theta \geq 0$, $t(\theta) + \sum_{i=0}^N \mu_i \theta^i < 0$, we can choose $F = N\delta_{\{\theta\}}$, contradicting (4) when $N \to \infty$. Thus,

$$t(\theta) + \sum_{i=0}^N \mu_i \theta^i \geq 0, \forall \theta \geq 0.$$

Moreover, $0 \in C$ implies $\int_0^\infty \left[ t(\theta) + \sum_{i=0}^N \mu_i \theta^i \right] dF_0(\theta) = 0$. Thus $\left\{ \theta \mid t(\theta) + \sum_{i=0}^N \mu_i \theta^i > 0 \right\}$ has null $F_0$ measure. If $\lambda = -\mu$, we have $t(\theta) \geq \sum_{i=0}^N \lambda_i \theta^i, \theta \geq 0$, $\lambda_N \leq 0$ and $\lambda \neq 0$. The previous inequality, together with the fact that $T^\lambda$ is the smallest monotonic function dominating the polynomial and the fact that $t$ is monotonic due to incentive compatibility, imply $t(\theta) \geq T^\lambda(\theta) \geq \sum_{i=0}^N \lambda_i \theta^i$, for all $\theta \geq 0$, concluding the proof of **(i)**.

To show **(ii)**, since $\sum_{i=0}^N \lambda_i \theta^i$ is bounded above by $t(B)$, the last non-null entry of $\lambda$ must be negative. To prove the first non-null entry of $\lambda$ should be negative note that if $\alpha = \sup_{\tau \geq 0} \tau (1 - F(\tau-))$, then $q(\theta) = 0$ for $\theta < \alpha$; $\alpha > 0$ due to the seller being able to guarantee himself a strictly positive payoff —see Lemma 1. Thus, $T^\lambda(\theta) = 0$ for $0 \leq \theta < \alpha$. Consequently, either the polynomial $\sum_{i=0}^N \lambda_i \theta^i$ takes a value strictly smaller than 0 at $\theta = 0$, in which case we are done, or it is 0 at $\theta = 0$, implying that it must be decreasing in a right neighborhood of 0, which concludes the proof of **(ii)**. From item **(i)** we have that $t(\theta) = T^\lambda(\theta) = \sum_{i=0}^N \lambda_i \theta^i$ $F_0$ almost surely, which immediately implies item **(iii)**. □

**Proof of Proposition 3.** Let $M^* = (q^*, t^*)$ be a solution of problem ($\mathcal{P}_0$) and $(M^*, F^*)$ be the associated Nash equilibrium. The proof is a consequence of the following lemmas.

**Lemma A.** Let $\Lambda = \left\{ (\lambda_0, \ldots, \lambda_N) \in \mathbb{R}^{N+1} \mid \lambda_0 \leq 0, \lambda_N \leq 0 \right\}$. If $\lambda \in \Lambda$ is such that $q^\lambda(\theta) \leq 1$, for all $\theta \geq 0$, then $\int_0^\infty t^*(\theta) dF^*(\theta) \geq \sum_{i=0}^N \lambda_i k_i$.

*Proof.* The function $q^\lambda$ is increasing and $\theta q^\lambda(\theta) - \int_0^\theta q^\lambda(\tau)\, d\tau = T^\lambda(\theta)$, which implies that $\left(q^\lambda, T^\lambda\right)$ is implementable. By hypothesis, $0 \leq q^\lambda(\theta) \leq 1$, for all $\theta \geq 0$, which ensures its feasibility.

Since $M^*$ is a solution to problem ($\mathcal{P}_0$) for $F_0 = F^*$ we have

$$\int_0^\infty t^*(\theta)\, dF^*(\theta) \geq \int_0^\infty T^\lambda(\theta)\, dF^*(\theta) \geq \int_0^\infty \sum_{i=0}^N \lambda_i \theta^i dF^*(\theta) = \sum_{i=0}^N \lambda_i k_i.$$

□

Lemma A implies that $\int_0^\infty t^*(\theta)dF^*(\theta)$ is an upper bound for the value of problem ($\mathcal{P}$). Moreover, by Lemma 2 (iii) there exists $\lambda^*$ such that $\int_0^\infty t^*(\theta)dF^*(\theta) = \sum_{i=0}^N \lambda_i^* k_i$. To show that the upper bound is tight we need to argue that the mechanism induced by $\lambda^*$, $\left(q^{\lambda^*}, T^{\lambda^*}\right)$, is feasible. The next lemma provides this result.

**Lemma B.** $\left(q^{\lambda^*}, T^{\lambda^*}\right)$ is feasible and generates payoff $\sum_{i=0}^N \lambda_i^* k_i$ for the seller.

*Proof.* As we argued in the proof of Lemma A, $(q^{\lambda^*}, T^{\lambda^*})$ is implementable. It remains to prove that $q^{\lambda^*}(\theta) \leq 1$, for all $\theta \geq 0$. From

$$t^*(\theta) = \theta q^*(\theta) - \int_0^\theta q^*(\tau)\, d\tau \geq T^{\lambda^*}(\theta) = \theta q^{\lambda^*}(\theta) - \int_0^\theta q^{\lambda^*}(\tau)\, d\tau,$$

we get

$$u(\theta) - \int_0^\theta \frac{u(\tau)}{\tau} d\tau \leq 0, \tag{5}$$

where $u(\theta) = \theta\left(q^{\lambda^*}(\theta) - q^*(\theta)\right)$. If $0 = t^*(\theta) \geq T^{\lambda^*}(\theta) \geq 0$, for $\theta < \alpha$, then $q^{\lambda^*}(\theta) = 0$ for $\theta < \alpha$ as well. Gronwall inequality applied to (5) gives $u(\theta) \leq 0$. Thus, $q^{\lambda^*}(\theta) \leq q^*(\theta) \leq 1$. □

We are left to show that mechanisms $(q^*, t^*)$ and $\left(q^{\lambda^*}, T^{\lambda^*}\right)$ must coincide at every point (not only $F^*$ a.s.). We start with a technical result.

**Lemma C.** Any $(q,t) \in \mathcal{M}$ satisfies $q(\theta) = q(0) + \frac{t(\theta)}{\theta} + \int_0^\theta \frac{t(\tau)}{\tau^2} d\tau$.

*Proof.* If $(q,t) \in \mathcal{M}$, we know that $t(\theta) = \int_0^\theta \tau dq(\tau)$ and hence

$$\frac{t(\theta)}{\theta} = \int_0^\theta \frac{\tau}{\theta} dq(\tau) = \int_0^\infty \frac{\tau}{\theta} \mathbf{1}_{[0,\theta]}(\tau)\, dq(\tau).$$

Since the integrand is dominated by 1 and converges pointwise to 0, dominated convergence theorem implies $\frac{t(\theta)}{\theta} \to 0$ as $\theta \to 0$. Now, notice that

$$q(\theta) = q(\underline{\theta}) + \int_{\underline{\theta}}^\theta \frac{1}{\tau} dt(\theta) = q(0) + \frac{t(\theta)}{\theta} - \frac{t(\underline{\theta})}{\underline{\theta}} + \int_0^\theta \frac{t(\tau)}{\tau^2} d\tau.$$

Taking limit $\underline{\theta} \to 0$ gives us the result. □

**Lemma D.** If $(q,t) \in \mathcal{M}_B$ solves problem ($\mathcal{P}_0$), then there exists $\lambda \in \mathbb{R}^{N+1}$ such that $t = T^\lambda$.

*Proof.* Consider an optimal mechanism $M \equiv (q,t) \in \mathcal{M}_B$. From Lemma 2, there exists $\lambda \in \mathbb{R}^{N+1}$ such that

$$t(\theta) \geq T^\lambda(\theta)$$

and

$$\inf_{F \in \mathcal{F}} \int_0^\infty t(\theta)\, dF(\theta) = \sum_{i=0}^N \lambda_i k_i.$$

If $t \neq T^{\lambda}$, then $t(\theta) > T^{\lambda}(\theta)$ for some $\theta \geq 0$. Since the function $\theta \mapsto t(\theta) - T^{\lambda}(\theta)$ is right-continuous, non-decreasing and upper semi-continuous, we have that the set $\{\theta \mid t(\theta) > T^{\lambda}(\theta)\}$ has positive Lebesgue measure. Using Lemma C, we have that

$$\begin{aligned} q(\theta) - q^{\lambda}(\theta) &= q(0) + \frac{t(\theta)}{\theta} + \int_0^{\theta} \frac{t(\tau)}{\tau^2} d\tau - \left( q^{\lambda}(0) + \frac{T^{\lambda}(\theta)}{\theta} + \int_0^{\theta} \frac{T^{\lambda}(\tau)}{\tau^2} d\tau \right) \\ &\geq q(0) + \int_0^{\theta} \frac{t(\tau) - T^{\lambda}(\tau)}{\tau^2} d\tau, \end{aligned}$$

which implies that $\lim_{\theta \to \infty} q^{\lambda}(\theta) \leq 1 - q(0) - \int_0^{\infty} \frac{t(\tau) - T^{\lambda}(\tau)}{\tau^2} d\tau < 1$. Since the mechanism $(q^{\lambda}, T^{\lambda})$ is feasible and also generates the optimal payoff for the seller, Lemma 6 gives a contradiction. □

For the reciprocal result, notice that the above argument just showed that any solution to problem $(\mathcal{P})$ attains the value of zero-sum game between the Nature and the seller, which concludes the proof. □

**Proof of Proposition 4.** In Lemma 3, we show that the optimal mechanism is described by the monotonic hull of a second degree polynomial $\lambda_0 + \lambda_1 \theta + \lambda_2 \theta^2$ with coefficients $\lambda_0, -\lambda_1, \lambda_2 < 0$. The function $T^{\lambda}$, in this case, becomes

$$T^{\lambda}(\theta) = \begin{cases} 0 & \text{, if } \theta < \underline{\theta}, \\ \lambda_0 + \lambda_1 \theta + \lambda_2 \theta^2 & \text{, if } \underline{\theta} < \theta < \overline{\theta}, \\ \lambda_0 + \lambda_1 \overline{\theta} + \lambda_2 \overline{\theta}^2 & \text{, if } \overline{\theta} < \theta, \end{cases}$$

where $\underline{\theta}$ is the lowest zero of the polynomial and $\overline{\theta}$ is its maximizer. The probability of allocating the good $q^{\lambda}(\cdot)$ is directly obtained from $q^{\lambda}(\underline{\theta}) = 0$ and $\theta \dot{q}^{\lambda}(\theta) = \dot{T}^{\lambda}(\theta)$, for $\theta \in (\underline{\theta}, \overline{\theta})$. Notice that we can solve for $\lambda_0, \lambda_1, \lambda_2$ using conditions: (i) polynomial is zero at $\underline{\theta}$, (ii) polynomial has zero derivative at $\overline{\theta}$, (iii) $q^{\lambda}(\overline{\theta}) = 1$ by Lemma 6. This gives us

$$\lambda_2 = \frac{1}{2\left(\overline{\theta} - \underline{\theta} - \overline{\theta} \ln\left(\frac{\overline{\theta}}{\underline{\theta}}\right)\right)}, \tag{6}$$

$$\lambda_1 = -\frac{\overline{\theta}}{\overline{\theta} - \underline{\theta} - \overline{\theta} \ln\left(\frac{\overline{\theta}}{\underline{\theta}}\right)}, \tag{7}$$

$$\lambda_0 = \frac{\underline{\theta}\left(2\overline{\theta} - \underline{\theta}\right)}{2\left(\overline{\theta} - \underline{\theta} - \overline{\theta} \ln\left(\frac{\overline{\theta}}{\underline{\theta}}\right)\right)}. \tag{8}$$

Using Proposition 3, we can find the optimal pair $(\underline{\theta}, \overline{\theta})$ such that the implied coefficients solve $(\mathcal{P})$, i.e.,

$$\max_{0 \leq \underline{\theta} \leq \overline{\theta}} \lambda_0 + \lambda_1 k_1 + \lambda_2 k_2$$

subject to (6)-(8).

We know that the optimal mechanism necessarily has $0 < \underline{\theta} < \overline{\theta}$, which means that $(\underline{\theta}, \overline{\theta})$ is a local interior optimum of this maximization problem.

The necessary optimality conditions, for $\underline{\theta}$ and $\overline{\theta}$ respectively, are

$$4\left(\overline{\theta}-\underline{\theta}\right)\left(\overline{\theta}-\underline{\theta}-\overline{\theta}\ln\left(\frac{\overline{\theta}}{\underline{\theta}}\right)\right)=\frac{2\left(\overline{\theta}-\underline{\theta}\right)}{\underline{\theta}}\left(\underline{\theta}\left(2\overline{\theta}-\underline{\theta}\right)-2\overline{\theta}k_1+k_2\right), \tag{9}$$

$$4\left(\underline{\theta}-k_1\right)\left(\overline{\theta}-\underline{\theta}-\overline{\theta}\ln\left(\frac{\overline{\theta}}{\underline{\theta}}\right)\right)=2\left(-\ln\left(\frac{\overline{\theta}}{\underline{\theta}}\right)\right)\left(\underline{\theta}\left(2\overline{\theta}-\underline{\theta}\right)-2\overline{\theta}k_1+k_2\right). \tag{10}$$

Dividing (9) by (10), we find that $k_1=\underline{\theta}\left(1+\ln\left(\frac{\overline{\theta}}{\underline{\theta}}\right)\right)$. Substituting $k_1$ back into (9), we find that $k_2=\underline{\theta}\left(2\overline{\theta}-\underline{\theta}\right)$. □

**Proof of Proposition 5.** The same result as is obtained for the case without the bounds on the support can be obtained for the environment here. In fact, notice that the bound on the support guarantees that Nature has a compact set of strategies even when only the first moment is considered. Therefore, a Nash equilibrium exists in the zero-sum game where Nature's strategy space is constrained to distributions with support contained in $[0,\overline{\theta}]$.

The seller's optimal mechanism can be obtained from a linear polynomial $\lambda_0+\lambda_1\theta$, with $\lambda_0<0$ and $\lambda_1>0$. The seller has positive transfers in the interval $[\underline{\theta},\overline{\theta}]$, where linearity of the transfer function implies $\theta q'(\theta)=t'(\theta)=\lambda_1$. This means the seller randomizes over this interval with a density $q'(\theta)\equiv f(\theta)=\lambda_1/\theta$, for some constant $\lambda_1$. The seller's indifference yields Nature's distribution $F(\theta)=1-\frac{\underline{\theta}}{\theta}$, for $\theta\geq\underline{\theta}$, with a mass point $\underline{\theta}/\overline{\theta}$ at the top. The moment condition pins down $\underline{\theta}$. The expected value of $\theta$ using distribution $F(\cdot)$ is:

$$\underline{\theta}\left(1+\ln(\overline{\theta})-\ln(\underline{\theta})\right)=k_1.$$

The left hand-side of this equation is increasing in $\underline{\theta}$ and, hence, it has a unique solution for $\underline{\theta}$.

Finally, $\lambda_1$ is pinned down from $\int_{\underline{\theta}}^{\overline{\theta}}h(\theta)d\theta=1$, i.e., $\lambda_1=\frac{1}{\ln(\overline{\theta})-\ln(\underline{\theta})}$. □

**Proof of Proposition 6.** We start with a brief introduction to Lambert function. The domain of the lower branch of the Lambert W function, $W_{-1}$, is the interval $[-1/e,0)$ and its defining equation is

$$W_{-1}(x)e^{W_{-1}(x)}=x.$$

Implicit differentiation yields

$$x(1+W_{-1}(x))\frac{dW_{-1}(x)}{dx}=W_{-1}(x),$$

for all $x\in(-1/e,0)$, or equivalently

$$\frac{dW_{-1}(x)}{dx}=\frac{W_{-1}(x)}{x(1+W_{-1}(x))}.$$

In the analysis of the optimal mechanism we encountered the moment condition

$$\underline{\theta}(1 + \ln(\overline{\theta}) - \ln(\underline{\theta})) = k_1.$$

Here we argue that $\underline{\theta} = \frac{k_1}{W_{-1}(-k_1/e\overline{\theta})}$.

Dividing both sides of the above equation by $\underline{\theta}$ and evaluating them with the exponential function we can rewrite the condition as follows

$$-\frac{k_1}{\underline{\theta}} e^{-\frac{k_1}{\underline{\theta}}} = -\frac{k_1}{e\overline{\theta}}.$$

Since $W_{-1}$ is defined on the interval $(-1/e, 0)$ and describes the corresponding part of the inverse of the function $z$, defined by $z(W) = We^W$. Using this function we can rewrite the above equation as

$$z\left(-\frac{k_1}{\underline{\theta}}\right) = -\frac{k_1}{e\overline{\theta}}$$

or

$$W_{-1}\left(-\frac{k_1}{e\overline{\theta}}\right) = -\frac{k_1}{\underline{\theta}},$$

which finally yields

$$\underline{\theta} = -\frac{k_1}{W_{-1}\left(-\frac{k_1}{e\overline{\theta}}\right)}.$$

Now we return to the actual proof. We argued above that the seller obtains payoff $\frac{k_1-\tau}{\overline{\theta}-\tau}\tau$ by posting a price $\tau < k_1$. Maximizing the seller's payoff yields the optimal posted price $\tau^* = \overline{\theta} - \sqrt{\overline{\theta}(\overline{\theta} - k_1)}$ and the seller's payoff from the optimal posted price $\left(\sqrt{\overline{\theta}} - \sqrt{\overline{\theta} - k_1}\right)^2$.

The relative loss $\rho$ is therefore

$$\rho(k_1, 1) = \frac{\underline{\theta} - \left(1 - \sqrt{1 - k_1}\right)^2}{\underline{\theta}}.$$

Using $\underline{\theta} = -k_1/W_{-1}(-k_1/e\overline{\theta})$ we can rewrite $\rho$ as

$$\rho(k_1, 1) = 1 + \frac{W_{-1}(-k_1/e)\left(1 - \sqrt{1 - k_1}\right)^2}{k_1}.$$

Since $W_{-1}(-1/e) = -1$, it follows that $\rho(1, 1) = 0$.

As for the behavior of $\rho(k_1, 1)$, when $k_1$ goes to 0 we can write

$$\rho(k_1, 1) = 1 + \frac{\alpha(k_1)}{\beta(k_1)}$$

where

$$\alpha(k_1) = \frac{\left(1-\sqrt{1-k_1}\right)^2}{k_1} \qquad \text{and} \qquad \beta(k_1) = \frac{1}{W_{-1}(-k_1/e)}.$$

Using L'Hopital's rule it can be shown that $\alpha(k_1) \to 0$ as $k_1 \to 0$. Moreover, since $\lim_{x\uparrow 0} W_{-1}(x) = -\infty$ it follows that also $\beta(k_1)$ tends to zero when $k_1$ goes to zero. To compute the limit of the ratio of $\alpha$ and $\beta$ we therefore apply L'Hopital's rule once more. Doing so yields

$$\lim_{k_1\to 0} \frac{\alpha(k_1)}{\beta(k_1)} = \lim_{k_1\to 0} \frac{\alpha'(k_1)}{\beta'(k_1)} = \frac{1/4}{-\infty} = 0.$$

In order to see that we are entitled to apply L'Hopital's rule notice that both $\alpha$ and $\beta$ are differentiable at all points $k_1 > 0$. That $\beta$ is differentiable at all $k_1$ follows from the fact that the (lower branch of the) Lambert $W$ function is differentiable on $(-1/e, 0)$. It is also straightforward to verify that $\beta'(k_1) \neq 0$ for all $k_1 > 0$.

It remains to be shown that $\rho(\cdot, 1)$ is decreasing. Computing the derivative of $\rho(\cdot, 1)$ yields

$$\frac{\partial \rho(k_1, 1)}{\partial k_1} = \frac{W_{-1}(-k_1/e)\left[1-\sqrt{1-k_1}\right]\left[W_{-1}(-k_1/e)\left(1-\sqrt{1-k_1}\right)+k_1\right]}{\left[1+W_{-1}(-k_1/e)\right]k_1^2\sqrt{1-k_1}}.$$

Since $W_{-1} \leq -1$ this expression is negative if and only if

$$-W_{-1}(-k_1/e) > \frac{k_1}{\left(1-\sqrt{1-k_1}\right)}. \tag{11}$$

Both sides of this inequality are decreasing functions of $k_1$ and both go to 1 as $k_1$ goes to 1. Moreover, while the r.h.s. goes to 2 as $k_1$ approaches 0, the l.h.s. diverges to $\infty$. Thus, the above inequality holds at least in a neighborhood of 0. Suppose that contrary to our hypothesis there is some $k_1' \in (0, 1)$ at which the two sides of the inequality cross (are equal), i.e.

$$-W_{-1}(-k_1'/e) = \frac{k_1'}{1-\sqrt{1-k_1'}}.$$

At $k_1'$ the slope of the l.h.s. is

$$\frac{d(-W_{-1}(-k_1'/e))}{dk_1} = -\frac{W_{-1}(-k_1'/e)}{k_1'\left(1+W_{-1}(-k_1'/e)\right)} = -\frac{1}{\sqrt{1-k_1'}\left(1-\sqrt{1-k_1'}\right)}.$$

The slope of the r.h.s. is instead

$$\frac{-2+k_1'+2\sqrt{1-k_1'}}{2\left(1-\sqrt{1-k_1'}\right)^2\sqrt{1-k_1'}}.$$

Dividing this expression by the slope of the l.h.s. we obtain

$$1 - \frac{k_1'}{2\left(1-\sqrt{1-k_1'}\right)} < 1.$$

Thus, at every point of intersection of the two sides of (11) the l.h.s. is decreasing faster than the r.h.s. Consequently, there can be at most one point of intersection. Since the two curves meet at $k_1 = 1$ it follows that there can be no intersection point in $(0, 1)$. □

**Proof of Proposition 7.** Both distributions, by assumption, have support $[\underline{\theta}, \overline{\theta}]$. The optimal distribution for the seller knowing the mean $k_1$ and second moment $k_2 = \underline{\theta}\left[2\overline{\theta} - \underline{\theta}\right]$, from Proposition 4, can be written as

$$q^*_{k_2}(\theta) = \frac{\overline{\theta}\left(\ln\theta - \ln\underline{\theta}\right) - \left(\theta - \underline{\theta}\right)}{\overline{\theta}\left(\ln\overline{\theta} - \ln\underline{\theta}\right) - \left(\overline{\theta} - \underline{\theta}\right)},$$

while the optimal distribution when the seller knows the mean $k_1$ and an upper bound on the support of $\overline{\theta}$, from Proposition 5, is given by

$$q^*_{\overline{\theta}}(\theta) = \frac{\ln\theta - \ln\underline{\theta}}{\ln\overline{\theta} - \ln\underline{\theta}}.$$

The difference $q^*_{k_2}(\theta) - q^*_{\overline{\theta}}(\theta)$ has the same sign as

$$\left(\ln\theta - \ln\underline{\theta}\right)\left(\overline{\theta} - \underline{\theta}\right) - \left(\ln\overline{\theta} - \ln\underline{\theta}\right)\left(\theta - \underline{\theta}\right),$$

which, in turn, has the same sign as

$$\frac{\ln\theta - \ln\underline{\theta}}{\theta - \underline{\theta}} - \frac{\ln\overline{\theta} - \ln\underline{\theta}}{\overline{\theta} - \underline{\theta}}.$$

The last expression is positive if the function $\theta \mapsto \left(\ln\theta - \ln\underline{\theta}\right)/\left(\theta - \underline{\theta}\right)$ is decreasing in $[\underline{\theta}, \overline{\theta}]$. Differentiating this expression yields

$$\frac{d}{d\theta}\left[\frac{\ln\theta - \ln\underline{\theta}}{\theta - \underline{\theta}}\right] = \frac{1 - \frac{\underline{\theta}}{\theta} + \ln\left(\frac{\underline{\theta}}{\theta}\right)}{\left(\theta - \underline{\theta}\right)^2} \leq 0,$$

with this inequality being strict for $\theta > \underline{\theta}$. □

## References


Aliprantis, C. and O. Burkinshaw (1998): *Problems in real analysis*, Elsevier, 2nd edition ed.

Auster, S. (2016): "Robust contracting under common value uncertainty," Working paper.

Azar, P., C. Daskalakis, S. Micali, and S. M. Weinberg (2013): "Optimal and efficient parametric auctions," in *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, SIAM, 596–604.

Azar, P. D. and S. Micali (2013): "Parametric digital auctions," in *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, ACM, 231–232.

Bergemann, D. and S. Morris (2005): "Robust mechanism design," *Econometrica*, 73, 1521–1534.

——— (2013): "An introduction to robust mechanism design," *Foundations and Trends in Microeconomics*, 8, 169–230.

Bergemann, D. and K. Schlag (2008): "Pricing without priors," *Journal of the European Economic Association*, 6, 560–569.

——— (2011): "Robust monopoly pricing," *Journal of Economic Theory*, 146, 2527–2543.

Berger, J. (1982): "The Robust Bayesian Viewpoint," Technical Report, Department of Statistics, Purdue University.

——— (1984): "An Overview of Robust Bayesian Analysis," *Test*, 3, 5–124.

Betrò, B. and A. Guglielmi (2000): *Robust Bayesian Analysis*, Spring Science, vol. 152, chap. Methods for Global Prior Robustness under Generalized Moment Conditions, 273–293.

Bewley, T. (2002): "Knightian decision theory: part I," *Decisions in Economics and Finance*, 25, 79–110.

Billingsley, P. (1995): "Probability and measure," *Wiley series in probability and mathematical statistics.*

Bodoh-Creed, A. (2012): "Ambiguous beliefs and mechanism design," *Games and Economic Behavior*, 75, 518–537.

Bose, S. and A. Daripa (2009): "A dynamic mechanism and surplus extraction under ambiguity," *Journal of Economic Theory*, 144(5), 2084–2114.

Bose, S., E. Ozdenoren, and A. Pape (2006): "Optimal auction with ambiguity," *Theoretical Economics*, 1, 411–438.

Bose, S. and L. Renou (2014): "Mechanism design with ambiguous communication devices," *Econometrica*, 82, 1853–1872.

Brooks, B. (2014): "Surveying and selling: belief and surplus extraction in auctions," Working paper.

Carrasco, V., V. Luz, P. Monteiro, and H. Morreira (2015a): "Robust mechanism: the curvature case," Working paper.

Carrasco, V., V. F. V. Luz, P. Monteiro, and H. Moreira (2015b): "Robust selling mechanisms," Working paper.

Carroll, G. (2012): "Robust incentives for information acquisition," Working Paper.

——— (2013): "Notes on informationally robust monopoly pricing," Mimeo, Stanford University.

——— (2015): "Robustness and linear contracts," *The American Economic Review*, 105, 536–563.

Castro, L. D. and N. Yannelis (2012): "Uncertainty, efficiency and incentive compatibility," Working paper.

Chassang, S. (2013): "Calibrated incentive contracts," *Econometrica*, 81, 1935–1971.

Chung, K. L. (2001): *A course in probability theory*, Academic Press, 3rd edition ed.

Clarke, F. (2013): *Functional Analysis, Calculus of Variations and Optimal Control*, Springer.

Di Tillio, A., N. Kos, and M. Messner (2014): "The design of ambiguous mechanisms," IGIER Working Paper.

Garrett, D. F. (2014): "Robustness of simple menus of contracts in cost-based procurement," *Games and Economic Behavior*, 87, 631–641.

Jones, M. C., J. S. Marron, and S. J. Sheather (1996): "A brief survey of bandwidth selection for density estimation," *Journal of the American Statistical Association*, 91, 401–407.

Karlin, S. and L. S. Shapley (1953): *Geometry of moment spaces*, 12, American Mathematical Soc.

Kos, N. and M. Messner (2013a): "Extremal incentive compatible transfers," *Journal of Economic Theory*, 148, 134–164.

——— (2013b): "Incentive compatibility in non-quasilinear environments," *Economics Letters*, 121, 12–14.

——— (2015): "Selling to the mean," Working Paper.

Kremer, M. and C. Snyder (2015): "Preventives Versus Treatments," *The Quarterly Journal of Economics*, 130, 1167–1239.

López-Cunat, J. M. (2000): "Adverse selection under ignorance," *Economic Theory*, 16, 379–399.

Lopomo, G., L. Rigotti, and C. Shannon (2011): "Knightian uncertainty and moral hazard," *Journal of Economic Theory*, 146, 1148–1172.

——— (2014): "Uncertainty in mechanism design," Working paper.

Manski, C. F. (1995): *Identification problems in the social sciences*, Harvard University Press.

——— (2003): *Partial identification of probability distributions*, Springer Science & Business Media.

Monteiro, P. (2009): "Abstract types and distributions in independent private value auctions," *Economic Theory*, 40, 497–507.

Mussa, M. and S. Rosen (1978): "Monopoly and product quality," *Journal of Economic Theory*, 18, 301–317.

Ollár, M. and A. Penta (2017): “Full implementation and belief restrictions,” *The American Economic Review*, 107, 2243–2277.

Pinar, M. C. and C. Kizilkale (2015): “Robust screening under distribution ambiguity,” Working paper.

Raiffa, H. (1961): “Risk, Ambiguity, and the Savage Axioms: Comment,” *Quarterly Journal of Economics*, 75, 690–694.

Reny, P. J. (1999): “On the existence of pure and mixed strategy Nash equilibria in discontinuous games,” *Econometrica*, 67, 1029–1056.

Saito, K. (2015): “Preferences for flexibility and randomization under uncertainty,” *The American Economic Review*, 105, 1246–1271.

Shohat, J. and J. Tamarkin (1943): *The problem of moments*, vol. I, American Mathematical Society.

Skreta, V. (2006): “Mechanism design for arbitrary type spaces,” *Economic Letters*, 91, 293–299.

Smith, J. (1995): “Generalized Chebyshev inequalities: theory and applications in decision analysis,” *Operations Research*, 43, 807–825.

Wilson, R. (1987): *Game theoretic analysis of trading processes*, Cambridge University Press: Cambridge.

Wolitzky, A. (2016): “Mechanism design with maxmin agents: theory and an application to bilateral trade,” *Theoretical Economics*, 11, 971–1004.