Optimal Selling Mechanisms Under Moment Conditions[*,†]

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

April 8, 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, Johannes Höerner, Lucas Maestri, Stephen Morris, 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, Toulouse School of Economics, Universidad Carlos III de Madrid, Princeton University, INSPER, PUC-Rio and EPGE/FGV, 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 by mutually exclusive subsets of the coauthors (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. E-mail:PKLM@fgv.br

[††]FGV/EPGE. 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-Brazil. Moreira acknowledges FAPERJ and CNPq for financial support.

Contents

1 Introduction and related literature

Following Wilson (1987)'s critique, a recent literature in mechanism deisgn theory 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 on the designer's prior assumptions.[1] In line with Wilson (1987)'s critique, mechanisms with good revenue performance over a wide range of possible prior distributions might be preferable in the presence of prior uncertainty. Additionally, this "robust" performance evaluation should take into account any (partial) information the designer has about the mechanism participants.

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.

Consider a seller designing a selling mechanism with 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 high-dimensional space.[2] Instead of formally modelling 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.[3]

Our findings can be summarized as follows. The seller's problem can be recast as a game between the seller and "malevolent" Nature. We show that such a game has an equilibrium. The crucial assumption for existence is that the highest of the moment conditions holds with inequality rather than equality; this ensures that the set of Nature's distributions is compact. The purpose of such a condition is 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

[1]Some notable exceptions are Bergemann and Schlag (2008) and Bergemann and Schlag (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]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)).

[3]Regarding the special casse 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.

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.

We start the characterization of equilibria by studying Nature's best reply. 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. Additionally, since a transfer function in incentive compatible mechanisms are non-decreasing, and we restrict attention to mechanisms in which the transfer of the lowest type is non-negative, we can bound the transfer function below by the non-negative monotonic hull of the aforementioned polynomial: the smallest non-decreasing and non-negative function that is (weakly) above the transfer function. We show that the original mechanism can be substituted by a new mechanism with transfer given by the non-negative monotonic hull of the polynomial without loss of revenue. Indeed, since this new transfer function is weakly above the polynomial function, then it must guarantee a worst-case revenue weakly above the linear combination of the known moments using the coefficients of the polynomial. By implication, when searching for optimal mechanisms, the seller can restrict attention to mechanisms that are non-negative monotonic hulls of polynomials, which are parametrized by the coefficients of the polynomial. And, as discussed above, the worst case revenue by any such mechanism is given by a linear combination of the known moments $(k_1, ..., k_N)$ 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 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 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 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 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 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.

An alternative method to achieve compactness of Nature's strategy set, and thus existence of equilibria, is to assume that there is an upper bound on valuations $\bar{\theta}$. 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, $\bar{\theta}$, 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 $\bar{\theta}$ 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.[4] Auster (2015) 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 (2013) study full implementation under belief restrictions, including moment restrictions.

Lopomo et al. (2009) 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

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

(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 τ the transfer to be 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 θ. If he receives the good with probability x and pays the transfer τ 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, θ, 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 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.[5] We comment more extensively on this in Section 3.1.

We consider moment restrictions $k \in \mathbb{R}_+^N$ such that the set $\mathcal{F}(\cdot)$ is non-empty on a neighborhood of k, that is, we assume that, for some $\varepsilon > 0$, $\mathcal{F}\left(\tilde{k}\right) \neq \emptyset$, whenever $\left|\tilde{k} - k\right| < \varepsilon$. The primitive conditions on the vector $k \in \mathbb{R}_+^N$ for this regularity condition to be satisfied are known as the truncated Stieltjes problem and are discussed in Subsection 2.1. The assumption is not purely technical: if a rational seller is to learn that the distribution belongs to a set of distributions, then this set better be non-empty, otherwise the learning process must have gone astray.

The seller does not assign any value to the good. He evaluates mechanisms with respect to their worst case expected value, and it is this worst case expected revenue that he seeks

[5] 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).

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 $t : [0, \infty) \to \mathbb{R}$ a transfer function.[6] 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

$$\inf_{F \in \mathcal{F}(k)} E_F[t],$$

where $E_F[\cdot]$ is the expectation operator with respect to distribution F.

The seller's problem can be succinctly described as

$$\mathcal{P}_0 \equiv \sup_{M \in \mathcal{M}} \left[\inf_{F \in \mathcal{F}} \int E_F[t] \right], \tag{1}$$

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

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

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 good 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 θ where $q(\cdot)$ is continuous.[7] 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.[8]

[6] 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 $\chi'(\theta) \equiv \{\chi(s) \mid s \in \sigma^*(v^{-1}(\theta))\}$ and $\tau' \equiv \max\{\tau(s) \mid s \in \sigma^*(v^{-1}(\theta))\}$. 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 $\mathbb{R}_+$.

[7] 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 $t'(\cdot) \geq t(\cdot)$.

[8] 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, 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 and Page (2009), Kos and Messner (2013a) and Kos and Messner (2013b) for how to deal with the cases where the distribution are not continuous.

2.1 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 subsection we explore what restrictions one needs to impose on the values of the moments for $\mathcal{F}$ to be non-empty.

Let $F \in \Delta[0,\infty)$ such that $\int_0^\infty \theta^i dF(x) < \infty,\ i = 1, ..., N$ and let

$$H(F) \equiv \left(\int_0^\infty \theta^i dF(\theta) \right)_{i=1,\dots,N}.$$

In order to give meaning to problem (1) we need to ensure that the solution to

$$H_{-N}(F) = k_{-N} \text{ and } H_N(F) \leq k_N \tag{2}$$

exists, and moreover, has a solution in a neighborhood of $k = (k_1, ...k_N)$, where $H_{-N}(F) = (H_1(F), ..., H_{N-1}(F))$ and $k_{-N} = (k_1, ..., k_{N-1})$.

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 vector $k \in \mathbb{R}_+^N$ is:

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

Assumption 1. $D_{2n}^{(0)}(k) > 0,\ 2 \leq 2n \leq N,$ and $D_{2n+1}^{(1)}(k) > 0,\ 3 \leq 2n+1 \leq N.$

Hereafter we invoke Assumption 1. Assumption 1 implies that our revenue problem is well-defined in a neighborhood of k. This result is summarized in the following lemma.

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

This condition are easily computed numerically. For the examples $N = 1$ and $N = 2$, the exact conditions are presented below.

Example 1. In the cases of first two moments:

- $N = 1$ (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 $(k_1)^2 \leq k_2$.[9]

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 $U(H, F)$. We look for a pair (M^*, F^*)—a mechanism, and a 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

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

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 moves first 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.

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

3.1 The mean case

Here we focus on the case $N = 1$. 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 $k_1 = 0$ 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 $\mathcal{Q}$. We write $U(\tau, F)$ for the payoff that the seller obtains when he adopts the deterministic price τ and the buyer's valuation is distributed according to F. Since for a given price τ type θ of the buyer acquires the good only if $\theta \geq \tau$ we have

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

where we define $F(\tau-) = \lim_{\tau' \nearrow \tau} F(\tau')$. 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 follows

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

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

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

Proof. All the missing proofs can be found in the appendix. □

The above result 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.[10]

[10]For an excellent treatment of use of bounds on moments in Econometrics see Manski (1995). It is not hard

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 the above proposition 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, unless otherwise specified, 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 will proceed 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. Then, 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.

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

Lemma 2. *(Local incentive constraints) For any mechanism* $(q, t) \in \mathcal{M}$

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

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

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

For any $\theta \in \mathbb{R}_+$, denote by $M_\theta = (q_\theta, t_\theta)$ the mechanism which allocation rule correspondes to posted price θ. In particular:

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

to see that we could have presented our original problem with 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.

The following lemma provides a uniform upper bound on expected payoff for sufficiently high posted prices. This guarantees that the seller never posts prices above some threshold in equilibrium.

Lemma 3. *(Uniform bounds on revenue) There exist* b, B *such that* $0 < b < B < \infty$ *and*

$$0 \leq \theta (1 - F(\theta)) < b (1 - F(b)),$$

for all $\theta \geq B$ *and all* $F \in \mathcal{F}$.

For every strictly positive posted price b below k_1, there exists B such that any posted price above B yields a smaller revenue than b for every distribution $F \in \mathcal{F}$.[11] 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.[12] We 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 3 implies that high prices are dominated for the seller in the zero-sum game.[13] 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 3 and define

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

to be the set of feasible mechanisms in which no type above B is allotted the good.

Next result formalizes the idea that the seller can restrict his attention to mechanisms that allocate the good to types below B.

Lemma 4. *(Bounded mechanisms) Any mechanism* $M = (q, t) \in \mathcal{M}$ *such that*

$$q(B) < 1$$

is strictly dominated by some mechanism $M' \in \mathcal{M}_B$.

Lemma 4 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.

Nature's strategy set—the set of admissible distributions—is also compact in the weak topology.

[11] The proof of the lemma establishes a stronger result than in the statement of the result.

[12] 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.

[13] Dominance here applies to the auxiliary game between the seller and Nature.

Lemma 5. *(Compactness of $\mathcal{F}$) The set $\mathcal{F}$ is compact in the weak topology.*

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

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 7. *(Optimality of prices) For any $F \in \mathcal{F}$, there exists $\theta_0 \in \mathbb{R}_+$ such that mechanism M_{θ_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 θ_ε, such that the mechanism M_{θ_ε} satisfies*

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

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 with a degree dictated by 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 function $f : \mathbb{R}_+ \mapsto \mathbb{R}$, denote the smallest non-negative and non-decreasing function above $f(\cdot)$ by

$$T^f(x) = \max\{0, \sup\{f(z) \mid z \leq x\}\}.$$

We call T^f the non-negative monotonic hull of f. For any $\lambda = (\lambda_0, ..., \lambda_N) \in \mathbb{R}^{N+1}$, define T^λ as the non-negative monotonic hull of the polynomial $\theta \mapsto \sum_{i=0}^{N} \lambda_i \theta^i$. The following result sheds light on our interest in polynomials.

Lemma 8. *Let $(q, t) \in \mathcal{M}$ and $F_0 \in \mathcal{F}$ be such that*

$$E_{F_0}[t] = \inf_{F \in \mathcal{F}} E_F[t].$$

There exists a non-null $\lambda \in \mathbb{R}^N \times \mathbb{R}_+$ *such that*

$$t(\theta) \geq \sum_{i=0}^{N} \lambda_i \theta^i, \text{ for all } \theta \in \mathbb{R}_-,$$

and this inequality holds as an equality on the support of F_0*. Moreover,* $t(\theta) \geq T^{\lambda}(\theta)$*, for all* $\theta \in \mathbb{R}_-$

The above result establishes that a transfer rules 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, hence the interest in non-negative monotonic hulls of polynomials. Moreover, evaluating the equation above at zero implies $\lambda_0 \leq 0$.

To prove Lemma 8 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

$$E_F[t] - \sum_{i=0}^{N} \lambda_i \int_0^{\infty} \theta^i dF(\theta) \geq E_{F_0}[t] - \sum_{i=0}^{N} \lambda_i \int_0^{\infty} \theta^i dF_0(\theta),$$

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^*, F^*) be a Nash equilibrium of the zero-sum game between Nature and the seller. Let $M^* = (q^*, t^*)$ be a solution to problem $\mathcal{P}_0$, then (M^*, F^*) is also a Nash equilibrium.[14] Since F^* is a solution to the problem

$$\inf_{F \in \mathcal{F}} E_F[t^*],$$

Lemma 8 implies that there exists non-null $\lambda^* \in \mathbb{R}^N \times \mathbb{R}_-$ such that

$$t^*(\theta) \geq T^{\lambda^*}(\theta), \text{ for all } \theta \in \mathbb{R}_+, \tag{3}$$

and this equation 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:

$$\mathcal{P} \equiv \max_{\lambda \in \mathbb{R}_- \times \mathbb{R}^{N-1} \times \mathbb{R}_-} \sum_{i=0}^{N} \lambda_i k_i,$$

[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).

subject to $\int_0^\infty \frac{\dot{T}^\lambda(\tau)}{\tau} d\tau \leq 1$.[15] The following proposition establishes that the solution to $\mathcal{P}$ yields the overall optimal mechanism.

Proposition 3. *If $M^* = (q^*, t^*) \in \mathcal{M}$ solves $\mathcal{P}_0$, then*

$$t^*(\theta) = T^{\lambda^*}(\theta),$$

and

$$q^*(\theta) = \int_0^\theta \frac{\dot{T}^{\lambda^*}(\tau)}{\tau} d\tau,$$

where λ^* solves the problem $\mathcal{P}$.

Notice that for any mechanism $M = (q, t)$ such that

$$t(\cdot) = T^\lambda(\cdot)$$

for some $\lambda \in \mathbb{R}_- \times \mathbb{R}^{N-1} \times \mathbb{R}_-$, the following inequality holds

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

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, $\int_0^\infty \frac{\dot{T}^\lambda(\tau)}{\tau} d\tau \leq 1$, can be interpreted as a resource constraint corresponding to the requirement that the seller has only one good to sell: $\lim_{\theta \to \infty} q(\theta) \leq 1$. 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] Invoking Assumption 1 is tantamount to requiring $k_1^2 \leq k_2$.

Due to Proposition 3 we can pursue the optimal mechanism by studying the non-negative monotonic hull of a polynomial $\lambda_0 + \lambda_1 \theta + \lambda_2 \theta^2$.

Lemma 9. *Fix a feasible pair k_1, k_2 and let $(\lambda_0^*, \lambda_1^*, \lambda_2^*)$ be a solution to problem $\mathcal{P}$. Then, $\lambda_0^* < 0$, $\lambda_1^* > 1$ and $\lambda_2^* < 0$.*

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

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

At the optimum the relevant polynomial intersects the vertical axis below zero and for large θ it goes to minus infinity. Perhaps most important is the argument for $\lambda_2 < 0$, since the result extends to an arbitrary number of moments, namely that the leading coefficient is non-positive. If it were not, the polynomial would be increasing from some θ on. The optimal mechanism constructed as in Proposition 3 would have $q(\theta) < 1$ for every θ. This, however, would contradict Lemma 4. So far we established $\lambda_2 \leq 0$, which is also a consequence of Lemma 8. If the constraint was not binding the seller would expect 0 payoff, as in Subsection 3.1. This would, however, contradict Lemma 3.

Since the seller can guarantee a positive revenue by Lemma 3, 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 when the seller knows the first two moments.

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

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

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

$$\overline{\theta}[1 + \log \overline{\theta} - \log \underline{\theta}] = k_1 \text{ and } \underline{\theta}[2\overline{\theta} - \underline{\theta}] = k_2.$$

The above result provides a characterization of the optimal direct mechanism the seller should use when informed of the mean and the variance. 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 $[\underline{\theta}, \overline{\theta}]$.

The main idea behind the characterization is the following. 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 consequently q^*. This interval, $[\underline{\theta}, \overline{\theta}]$, is precisely the interval the seller randomizes over. Using Lemma 8 to observe that Nature's and the seller's supports coincide, one can characterize Nature's optimal distribution from the seller's indifference condition. If the seller is to randomize over the interval, his payoff, $\theta(1 - F(\theta))$, from offering any price $\theta \in [\underline{\theta}, \overline{\theta}]$

must be the same. In particular

$$F^*(\theta) = 1 - \frac{c}{\theta},$$

where c is a constant. One can further argue that there is no atom at the bottom of Nature's distribution. The seller, thus, sells the object with probability one when he offers price $\underline{\theta}$, resulting in the payoff $\underline{\theta}$; ergo $c = \underline{\theta}$. The boundaries of the interval are then pinned down by the two moment conditions.

The form of the seller's optimal mechanism follows directly from Proposition 3 and Lemma 9: the seller randomizes over interval $[\underline{\theta}, \overline{\theta}]$ with density $h(\theta) = 2\lambda_2 + \lambda_1/\theta$, for some λ_1 and λ_2. Since $\overline{\theta}$ is the highest point at which the polynomial and its non-negative monotonic hull coincide, it must the be the maximizer of the polynomial, therefore $2\lambda_2\overline{\theta} + \lambda_1 = 0$. This condition, along with the requirement $\int h(\theta)d\theta = 1$, pins down λ_1 and λ_2. An interesting implication is that when the seller randomizes with the equilibrium distribution h and Nature assigns the whole probability mass to the interval, the 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.

We should point out 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 able to learn the variance in addition to the mean the information becomes beneficial.

Figure 1 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 1.

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 $\mathcal{P}$. If the highest non-negative coefficient λ_n is strictly positive, then the polynomial would be increasing and convex for θ sufficiently large. But then, so would be q^*. This would contradict Lemma 4 which states that $q^*(\theta) = 1$

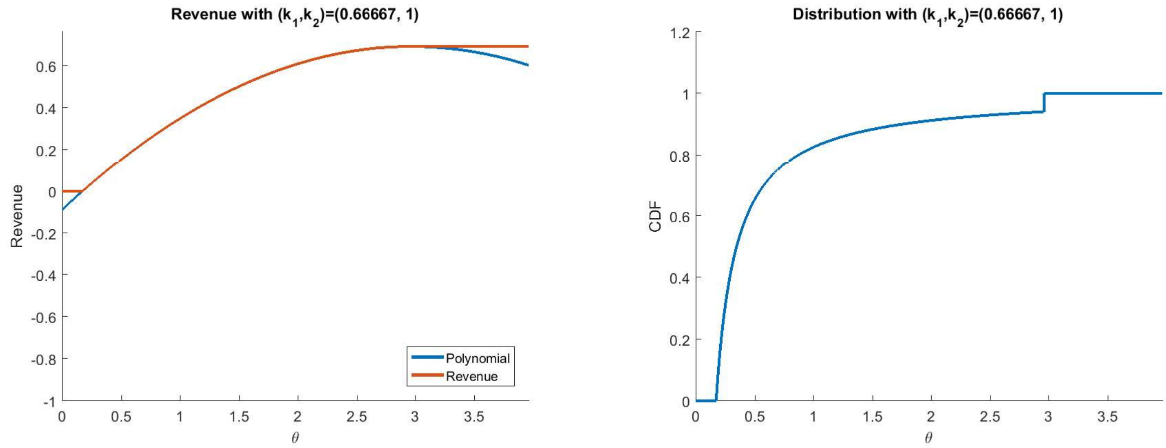

Figure 1: Optimal revenue, supporting polynomial function and distribution for two moments.

for large θ.

Moreover, a polynomial of degree n is increasing on at most $[(n+1)/2]$ intervals, where $[x]$ is the integer part of x. Since in the optimal mechanism q^* is determined by the smallest non-negative increasing function dominating a polynomial of degree n, q^* can also be increasing on at most $[(n+1)/2]$ 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 point zero and at most $[(n+1)/2]$ 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 $h(\theta) = \sum_{i=1}^{n} c_i \theta^{i-2}$, where $c_1, ..., c_n$ are some constants.

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. For that reason we dedicate to it the next several paragraphs.

The analysis of the case of three moments depends on the magnitude of parameter k_3. For sufficiently large k_3, the third moment 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 $\mathcal{F}(k_1, k_2)$ is non-empty. 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 $k_3 < k_3^*$, $M^{(2)} \notin \mathcal{F}(k_1, k_2, k_3)$ and the solutions 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 2 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 2. The distribution is given by

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

for $\left(\alpha, \underline{\theta}, \bar{\theta}\right) \in (0,1) \times \mathbb{R}^2_+$ satisfying $\underline{\theta} < \bar{\theta}$. Similar to the two moments case, the worst-case distribution has the property that the seller is indifferent among all prices $\tau \in \left[\underline{\theta}, \bar{\theta}\right]$. 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.

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. In fact, for any sequence of non-negative random variables $(X_n)_n$ that converges to X, we only have $E[X] \leq \liminf_n E[X_n]$; 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.

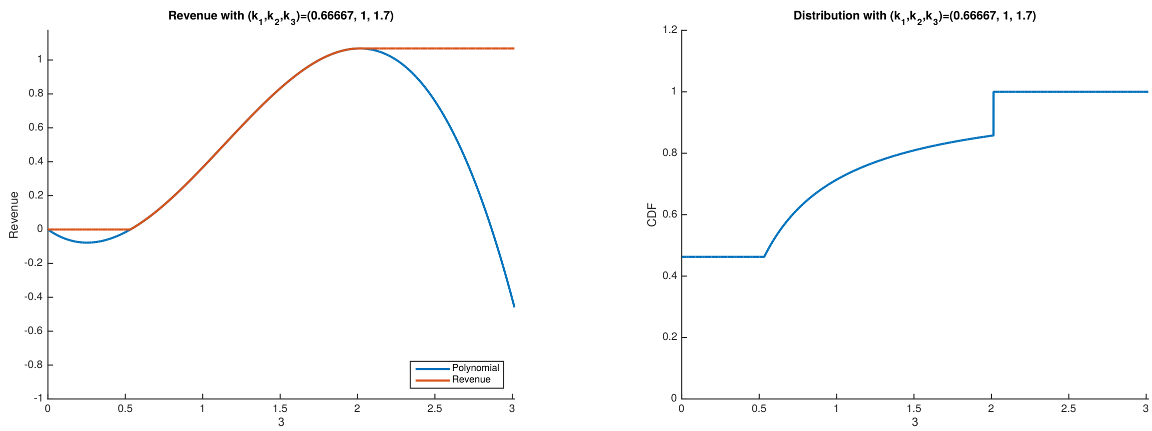

Figure 2: Optimal revenue, supporting polynomial function and distribution for three moments.

5 Discussion

5.1 Bound on valuations

In Section 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 and slightly less trivially when the first condition holds with equality. The result is due to Nature being 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 the seller entails as possible very high valuations. Such extreme pessimism comes from the unbounded support assumption. For many application 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.

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.

For most of this section we focus on the case where the seller knows only the first moment.[17] Now 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

[17] 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.

to τ. This is best done by assigning the rest of the mass to the point $\overline{\theta}$. The 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 θ. 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 the result that even the best mechanism can not guarantee the seller a positive payoff when Nature's support is unrestricted.

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 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*

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

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

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 randomizes over an interval $[\underline{\theta}, \overline{\theta}]$, for some $\underline{\theta}$, with a density of the form $f(\theta) \equiv \frac{dq^*(\theta)}{d\theta} = \lambda_1/\theta$. Nature's distribution can be pinned down from the seller's indifference: $F(\theta) = 1 - \frac{\underline{\theta}}{\theta}$ for $\theta \geq \underline{\theta}$, with a mass point $\underline{\theta}/\overline{\theta}$ at $\overline{\theta}$. The lower bound of the interval is then recovered from the moment condition $\int_{\underline{\theta}}^{\overline{\theta}} \theta dF(\theta) = k_1$.

It is illuminating to directly compute Nature's payoff when the seller uses the above outlined distribution over prices. More precisely, the seller's payoff (the negative of Nature's) when Nature chooses a distribution F is

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

where $\tilde{k}_1$ is the first moment of distribution F. Since h is a density $\lambda_1 > 0$, which makes the second term on the right hand-side negative. Nature, therefore, minimizes the seller's payoff by setting $F(\theta) = 0$ for $\theta < \underline{\theta}$. 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 $\rho(k_1, \bar{\theta})$. 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 $(k_1, \bar{\theta})$. 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 τ, Nature optimally uses a distribution on $\{\tau, 1\}$. The above result is illustrated in Figure 3. 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.

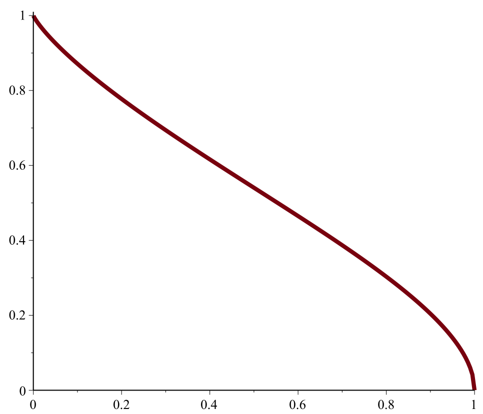

Figure 3: The relative loss of foregoing the option to randomize.

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 $\bar{\theta}$, the seller and Nature randomize over the interval $[\underline{\theta}, \bar{\theta}]$ where $\underline{\theta}$ is defined as the solution to

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

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 $[\underline{\tau}, \bar{\tau}]$ where $\underline{\tau}$ and $\bar{\tau}$ are defined by

$$\underline{\tau}[1 + \ln(\bar{\tau}) - \ln(\underline{\tau})] = k_1,$$
$$\underline{\tau}[2\bar{\tau} - \underline{\tau}] = k_2.$$

The two intervals of randomization coincide when $k_2 = \underline{\theta}[2\bar{\theta} - \underline{\theta}]$. Furthermore, in both cases the seller's payoff is precisely $\underline{\theta}$. Therefore, knowing that valuations are distributed within $[0, \bar{\theta}]$ 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 $k_2 = \underline{\theta}[2\bar{\theta} - \underline{\theta}]$.

This notwithstanding, the problems under the two types of constraints are not equivalent. In fact, there is no pair $\bar{\theta}$ and k_2 such that the k_2-constrained and the $\bar{\theta}$ 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, $\bar{\theta}$ and k_2 be such that the pairs $(k_1, \bar{\theta})$ and (k_1, k_2) induce the same supports for the price distribution, $[\underline{\theta}, \bar{\theta}]$. If the two optimal price distributions are denoted by $q^*_{\bar{\theta}}$ and $q^*_{k_2}$, respectively, then for all $\underline{\theta} < \tau < \bar{\theta}$,*

$$q^*_{k_2}(\tau) > q^*_{\bar{\theta}}(\tau).$$

The above result is illustrated in Figure 4 which shows the price distributions for the pairs $(k_1, \bar{\theta}) = (1/2, 1)$ and $(k_1, \sigma^2) = (1/2, 1.088)$.

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

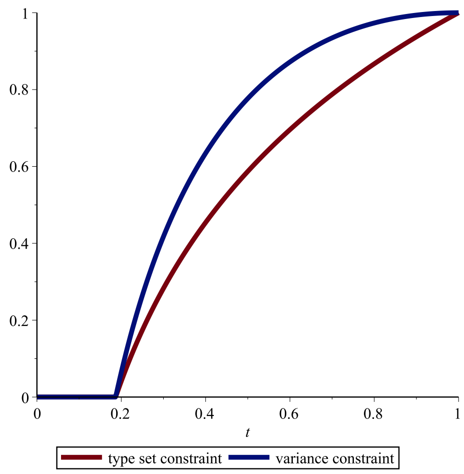

Figure 4: Optimal price distributions: bounded variance (blue), bounded type set (red).

when the seller uses $q^*_{\bar{\theta}}$. Alternatively, mechanism $q^*_{\bar{\theta}}$ 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, it prevent nature from moving probability mass above $\overline{\theta}$. If the seller were to use mechanism $q^*_{\bar{\theta}}$ when his information was about the first two moments, Nature would shift some probability mass to values above $\bar{\theta}$.

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. However, the choice of the seller's mechanism does become non-trivial when he has information both about the mean and the variance—the information is beneficial to the seller. In this case the seller chooses to randomize prices over an interval, which coincides with the support of the worst-case distribution from the seller's perspective. 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. If this is the case, we can sat that knowledge of the third moment is useless for the seller. 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.

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.

7 Appendix

Proof of Proposition 1. Let q be any distribution over prices and denote by F_θ 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

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

Below we argue that $U(q,F_\theta)$ converges to zero as θ increases. 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.

Using the definition of $U(q,F_\theta)$ and integrating by parts:

$$U(q,F_\theta) = \frac{k_1}{\theta}\int_0^\theta \tau dq(\tau) = \frac{k_1}{\theta}\left[\theta q(\theta) - \int_0^\theta q(\tau)d\tau\right].$$

Define $q_\infty \equiv \lim_{\theta\to\infty} q(\theta)$. By definition, for every $\alpha > 0$ there exists a $t_\alpha \ge 0$ such that $q(t) \ge q_\infty - \alpha$ for all $\theta \ge t_\alpha$. Thus for any given $\alpha > 0$ and $\theta > t_\alpha$ we have

$$\int_0^\theta q(\tau)\, d\tau = \int_0^{t_\alpha} q(\tau)\, d\tau + \int_{t_\alpha}^\theta q(\tau)\, d\tau \ge (\theta - t_\alpha)(q_\infty - \alpha).$$

Substituting this bound in the above definition of $U(q,F_\theta)$ yields

$$\begin{aligned} U(q,F_\theta) &= \frac{k_1}{\theta}\left[\theta q(\theta) - \int_0^\infty q(\tau)\, d\tau\right] \\ &\le \frac{k_1}{\theta}\left[\theta q_\infty - (\theta - t_\alpha)(q_\infty-\alpha)\right] = \frac{k_1}{\theta}\left[\theta\alpha + t_\alpha(q_\infty - \alpha)\right]. \end{aligned}$$

For given $\alpha > 0$, this expression converges to $k_1\alpha$ as θ increases. Since $\alpha > 0$ can be chosen arbitrarily close to 0, we conclude that $\lim_{\theta\to\infty} U(q,F_\theta) = 0$. □

Proof of Lemma 3. Consider any $F \in \mathcal{F}$. An application of Markov's inequality to θ^i, for $i = 1, ..., N$, yields

$$\theta^i (1 - F(\theta)) \le k_i. \tag{4}$$

Using (4) for $i = 2$ implies $\theta(1 - F(\theta)) \le k_2/\theta$ and, in particular, $\theta(1-F(\theta))$ converges to zero as θ increases.

First we show that there exists a uniform lower bound on profits for prices in $(0, k_1)$. Now, for every $\beta \in \mathbb{R}_+$ and $T > 0$, by integration by parts we have that

$$\int_\beta^T \theta dF(\theta) = \beta(1-F(\beta)) - T(1-F(T)) + \int_\beta^T (1-F(\theta))\, d\theta,$$

taking the limit $T \to \infty$ gives us

$$\begin{aligned}\int_\beta^\infty \theta dF(\theta) &= \beta(1-F(\beta)) + \int_\beta^\infty (1-F(\theta))\,d\theta \\ &\le \frac{k_2}{\beta} + \int_\beta^\infty \frac{k_2}{\theta^2} d\theta = 2\frac{k_2}{\beta},\end{aligned}$$

where the inequality in the second line uses inequality (4), with $i = 2$.

Finally, we use the following equation, and applying the above inequality,

$$\begin{aligned}k_1 = \int_0^\infty \theta dF(\theta) &= \int_0^b \theta dF(\theta) + \int_b^\beta \theta dF(\theta) + \int_\beta^\infty \theta dF(\theta) \\ &\le bF(b) + \beta(1-F(b)) + 2\frac{k_2}{\beta}\end{aligned}$$

which implies

$$F(b) \le \overline{F}_{b,\beta} \equiv 1 - \frac{k_1 - b - 2\frac{k_2}{\beta}}{\beta - b}.$$

Selecting any $0 < b < k_1$ and β sufficiently large we have $\overline{F}_{b,\beta} < 1$. It follows that

$$0 < b\left(1 - \overline{F}_{b,\beta}\right) < b(1-F(b)). \tag{5}$$

Finally, we provide an uniform (over $F \in \mathcal{F}$) upper bound on profits for sufficiently large prices. Pick B such that $B > b$ and

$$\frac{k_2}{B} < b\left(1 - \overline{F}_{b,\beta}\right). \tag{6}$$

Then, for all $\theta \ge B$ and $F \in \mathcal{F}$,

$$\theta(1-F(\theta)) \le \frac{k_2}{\theta} < b\left(1-\overline{F}_{b,\beta}\right) < b(1-F(b)),$$

where the first inequality is due to (4) for $i = 2$, the second due to (6) and the third follows from (5). □

Proof of Lemma 4. Consider a mechanism $M = (q,t)$ such that $\Delta \equiv 1 - q(B) > 0$. The mechanism $M' = (q', t')$ given by

$$q'(\theta) \equiv 1_{[0,B]}(\theta)\, q(\theta) + \Delta 1_{[b,\infty)}(\theta)$$

and

$$t'(\theta) \equiv \int_0^\theta \tau dq'(\tau)$$

is feasible, where B, b are defined in Lemma 3. The new allocation q' takes the remaining quantity not sold at price $[0, B]$ (either sold for types $\theta > B$ or not sold at all, i.e., $q_\infty < 1$) and sells it at price b. Its profit is given by

$$\int_0^\infty (1-F(\theta-))\,dq'(\theta) = \int_0^\infty (1-F(\theta-))\,dq(\theta) - \int_B^\infty \theta(1-F(\theta-))\,dq(\theta) + \Delta b(1-F(b-)).$$

But notice that, by definition, $\Delta = 1 - q_\infty + q_\infty - q(B)$, which means that

$$\int_0^\infty (1 - F(\theta-))\, dq'(\theta) = \int_0^\infty (1 - F(\theta-))\, dq(\theta) + \int_B^\infty [b(1 - F(b-)) - \theta(1 - F(\theta-))]\, dq(\theta) + (1 - q_\infty)\, b(1 - F(b-)).$$

First notice that the second and third terms are non-negative. Also, the second term is strictly positive if $q_\infty > q(B)$ and the third term is strictly positive if $1 > q_\infty$. Since $\Delta > 0$, one of these two terms is strictly positive and hence $q'(\cdot)$ dominates $q(\cdot)$. $\square$

Proof of Lemma 5. For every function $F \in \mathcal{F}$ and continuity point θ:

$$\theta(1 - F(\theta)) \leq k_1,$$

which means that, for any $\varepsilon > 0$,

$$1 - F(\theta) \leq \varepsilon,$$

for $\theta \geq \frac{k_1}{\varepsilon}$. This means that $\mathcal{F}$ is tight. Hence, the closure of $\mathcal{F}$ is compact in the weak topology and, therefore, it remains to show that $\mathcal{F}$ is closed.

Consider a sequence $(F_n)_{n\in\mathbb{N}}$ converging to F in the closure of $\mathcal{F}$. We must show that $F \in \mathcal{F}$, i.e., that

$$\int_0^\infty \theta^i dF(\theta) = k_i, \text{ for } i = 1, ..., N-1,$$
$$\int_0^\infty \theta^N dF(\theta) \leq k_N.$$

For each $L \in \mathbb{N}$ and $i = 1, ..., N$, the function $\theta \mapsto \min\{L, \theta^i\}$ is bounded and continuous. By definition of the weak topology we have

$$\int_0^\infty \min\{L, \theta^i\}\, dF_n(\theta) \to \int_0^\infty \min\{L, \theta^i\}\, dF(\theta).$$

But since $\min\{L, \theta^i\}$ is dominated by the function $\theta \mapsto \theta^i$ we have that

$$\int_0^\infty \min\{L, \theta^i\}\, dF_n(\theta) \leq k_i,$$

which implies

$$\int_0^\infty \min\{L, \theta^i\}\, dF(\theta) \leq k_i.$$

Taking the limits when $L \to \infty$ and using the monotone convergence theorem yields

$$\int_0^\infty \theta^i dF(\theta) \leq k_i,$$

for $i = 1, ..., N$.

Theorem 25.6 of Billingsley (1995) allows us to embed $\left((F_n)_{n\in\mathbb{N}}, F\right)$ in a single probability space, i.e., we can find a probability space $(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \mu)$ and random variables $\left((X_n)_{n\in\mathbb{N}}, X\right)$ such that the pushforward measure of X_n is F_n for each n and the pushforward measure of X is F. Using the fact that the random variables $(X_n)^N$ and $(X)^N$ are uniformly integrable, we can use the corollary following Theorem 25.12 of Billingsley (1995) to show that, for $i = 1, ..., N-1$

$$\mathbb{E}\left[(X_n)^i\right] \to \mathbb{E}\left[(X)^i\right].$$

But this is equivalent to

$$\int_0^\infty \theta^i dF_n(\theta) = k_i \to \int_0^\infty \theta^i dF(\theta),$$

which implies that $F \in \mathcal{F}$. □

Proof of Lemma 6. Consider a sequence $(M_m)_m$ in $\mathcal{M}$ such that $M_m = (q_m, t_m) \to M_0 = (q_0, t_0)$. For any continuity point of θ of $q_0(\cdot)$:

$$t_m(\theta) = \int_0^\theta \tau dq_m(\tau) \to t_0(\theta) = \int_0^\theta \tau dq_0(\tau).$$

Since $q_0(\cdot)$ and $(q_m(\cdot))_m$ are right-continuous, so are $t_0(\cdot)$ and $(t_m(\cdot))_m$. Right-continuity implies that $t(\cdot) \geq \limsup t_m(\cdot)$. We then have that

$$\begin{aligned}\limsup_{m\to\infty} U(M_m, F) = \limsup_{m\to\infty} \int_0^\infty t_m(\theta)\, dF(\theta) &\leq \int_0^\infty \left[\limsup_{m\to\infty} t_m(\theta)\right] dF(\theta) \\ &\leq \int_0^\infty t(\theta)\, dF(\theta) = U(M, F),\end{aligned}$$

where the first inequality follows from Fatou's lemma (transfers are bounded by the identity function). □

Proof of Lemma 7. Since the function $\theta \to \theta(1 - F(\theta-))$ is upper semi-continuous, Lemma 3 implies a global maximum of this function is achieved in $(0, B)$. Denote a maximizer as θ_0 and maximum be $U_0 \equiv \theta_0(1 - F(\theta_0-))$. Furthermore, for every $M \in \mathcal{M}$:

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

Since the Dirac measure with probability one at point zero is not contained in $\mathcal{F}$, we have that $U_0 > 0$ and $\theta_0 > 0$. For any $\varepsilon > 0$, there exists θ_ε such that $\theta_\varepsilon \in \left(\theta_0 - \frac{\varepsilon}{1-F(\theta_0-)}, \theta_0\right)$ and θ_ε is a continuity point of $F(\cdot)$, which implies

$$F(\theta_\varepsilon) \leq F(\theta_0-).$$

This means that

$$\begin{aligned} U\left(M_{\theta_\varepsilon}, F\right) = \theta_\varepsilon\left(1 - F\left(\theta_\varepsilon\right)\right) &> \left(\theta_0 - \frac{\varepsilon}{1 - F\left(\theta_0 -\right)}\right)\left(1 - F\left(\theta_\varepsilon\right)\right) \\ &\geq \left(\theta_0 - \frac{\varepsilon}{1 - F\left(\theta_0 -\right)}\right)\left(1 - F\left(\theta_0\right)\right) = U_0 - \varepsilon. \end{aligned}$$

□

Proof of Proposition 2. Consider the zero-sum game where the strategy set of the seller is restricted to $\mathcal{M}_B$. By Lemma 5, both players have a compact strategy set. 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

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

for any $F \in V\left(F_0\right)$, where $V\left(F_0\right)$ is an open neighborhood of F_0. From Lemma 7, there exists $\theta_\varepsilon \in \mathbb{R}_+$ such that: θ_ε is a continuity point of F_0 and

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

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

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

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

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

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

is closed by Lemma 6. Hence $\mathcal{M} \backslash \underline{\mathcal{M}}$ is open and contains M_0.

Since the game is payoff secure, it has a Nash equilibrium. Lemma 4 implies that a Nash equilibrium of the restricted game with strategy sets $\left(\mathcal{M}_B, \mathcal{F}\right)$ is a Nash equilibrium of the unrestricted game with strategy sets $(\mathcal{M}, \mathcal{F})$. □

Proof of Lemma 8. Let $(q, t) \in \mathcal{M}$. We proceed in three steps. Let C be the space of non-negative, non-decreasing, right continuous and bounded functions on $\mathbb{R}_+$. It is convenient for this proof to think of Nature's problem as minimization of the seller's revenue over a set of non-negative measures (rather than probability distributions) with an added constraint $\int_0^\infty dF(\theta) = k_0 = 1$.

Consider the problem

$$\inf_{F\in C}\int_0^\infty t\left(\theta\right)dF\left(\theta\right)\ \text{s.t. } H_{-N}(F)=k_{-N} \text{ and } H_N(F)\le k_N, \tag{7}$$

where $H_{-N}(x)=(H_0(F),...,H_{N-1}(F))$, $H_i(F)=\int_0^\infty \theta^i dF(\theta)$ and $k_{-N}=(k_0,...,k_{N-1})$. This is just a different way to write $\inf_{F\in\mathcal{F}}\int_0^\infty t\left(\theta\right)dF(\theta)$. Notice that we modified the functional H presented in text to include the coordinate functional H_0 in order to be consistent with the space C. By assumption, distribution F_0 solves this minimization problem.

Step 1. There exists a non-null $\pi\in\mathbb{R}^N\times\mathbb{R}_+$ such that

$$\int_0^\infty t(\theta)dF(\theta)+\pi\cdot H(F)\ge\int_0^\infty t(\theta)dF_0(\theta)+\pi\cdot H(F_0),$$

for all $F\in C$ and $\pi_N(k_N-H_N(F_0))=0$.

Towards proving the above inequality, define $\mathcal{H}=\{(a,b)\in\mathbb{R}^{N+2}\mid a>\int_0^\infty t(\theta)dF(\theta),$ $b_{-N}=H_{-N}(F)$ and $b_N\ge H_N(F)$, for some $F\in C\}$. Clearly $\mathcal{H}$ is a convex set. In addition, $(\int_0^\infty t(\theta)dF_0(\theta),k)\notin\mathcal{H}$. If $(\int_0^\infty t(\theta)dF_0(\theta),k)$ was in $\mathcal{H}$ there would exist a non-negative measure $\tilde F\in\mathcal{F}$ such that $\int t(\theta)dF_0(\theta)>\int t(\theta)d\tilde F(\theta)$, thus contradicting the optimality of F_0.

By the separation theorem for $\mathbb{R}^{N+2}$, there exists $(c,\pi)\in\mathbb{R}^{N+2}-\{0\}$, where $\pi=(\pi_0,\pi_1,...,\pi_N)$, such that for all $(a,b)\in\mathcal{H}$

$$ca+\pi\cdot b\ge c\int t(\theta)dF_0(\theta)+\pi\cdot k.$$

We claim that $c>0$. First, $c<0$ would lead to a contradiction with the above inequality given that a can be arbitrarily positive. If $c=0$, then the previous inequality becomes

$$\pi_{-N}\cdot H_{-N}(F)+\pi_N b_N\ge\pi\cdot k,$$

for all $F\in C$. Since we assumed that the moment conditions have a solution for parameters in an an open neighborhood of k, then the last inequality must imply $\pi=0$, which is a contradiction. We can then normalize $c=1$.

Choosing $F\in C$ such that $H_{-N}(F)=k_{-N}$ and $H_N(F)\le b_N$, we have

$$\int_0^\infty t(\theta)dF(\theta)+\pi_N b_N\ge\int_0^\infty t(\theta)dF_0(\theta)+\pi_N k_N.$$

Since the above inequality holds for an arbitrarily large b_N, we must have $\pi_N\ge 0$. Hence,

$$\int_0^\infty t(\theta)dF(\theta)+\pi\cdot H(F)\ge\int_0^\infty t(\theta)dF_0(\theta)+\pi\cdot k\ge\int_0^\infty t(\theta)dF_0(\theta)+\pi\cdot H(F_0),$$

for all $F\in C$. Finally, if $F=F_0$, then the first inequality becomes

$$0\ge\pi_N(k_N-H_N(F_0)),$$

and, since the right-hand side is non-negative, it must be zero.

Step 2. $t(\theta) \geq \sum_{i=0}^{N} \lambda_i \theta^i$, for all $\theta \in \mathbb{R}_+$.

We can slightly restate the result from Step 1 to say that it establishes existence of a vector $\lambda = -\pi \in \mathbb{R}^{N+1} - \{0\}$ 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 $F \in C$, which is equivalent to

$$\int_0^\infty \left[t(\theta) - \sum_{i=0}^{N} \lambda_i \theta^i \right] dF(\theta) \geq \int_0^\infty \left[t(\theta) - \sum_{i=0}^{N} \lambda_i \theta^i \right] dF_0(\theta), \tag{8}$$

for every $F \in C$.

Notice that

$$t(\theta) - \sum_{i=0}^{N} \lambda_i \theta^i \geq 0, \text{ for all } \theta \in \mathbb{R}_+.$$

Otherwise, there exists θ_0 such that $t(\theta_0) - \sum_{i=0}^{N} \lambda_i \theta_0^i < 0$, and a non-negative measure $F'(\cdot) \equiv \left(F_0(\cdot) + \mathbf{1}_{[\theta_0,\infty)}(\cdot)\right) \in C$ with

$$\int_0^\infty \left[t(\theta) - \sum_{i=0}^{N} \lambda_i \theta^i \right] dF'(\theta) < \int_0^\infty \left[t(\theta) - \sum_{i=0}^{N} \lambda_i \theta^i \right] dF_0(\theta),$$

which contradicts (8).

Furthermore, since the null measure is in C, it follows that

$$\int_0^\infty \left[t(\theta) - \sum_{i=0}^{N} \lambda_i \theta^i \right] dF_0(\theta) \leq 0.$$

Since the integrand is non-negative, it must be equal to zero on the support of F_0.

Step 3. Since (q,t) is an incentive compatible mechanism, t is monotone. In addition, by step 2, $t(\theta) \geq \sum_{i=0}^{N} \lambda_i \theta^i$, for all $\theta \in \mathbb{R}_+$, therefore $t(\theta) \geq T^\lambda(\theta)$, for all $\theta \in \mathbb{R}_+$. □

Proof of Proposition 3. Suppose (q,t) solves $\mathcal{P}_0$. Lemma 8 implies that $t(\cdot)$ coincides, in the support of F^*, with a polynomial with coefficients $(\lambda_0^*, \lambda_1^*, ..., \lambda_N^*)$, which means

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

We next argue that $t(\cdot) = T^{\lambda^*}(\cdot)$. Lemma 8 implies $t(\cdot) \geq T^{\lambda^*}(\cdot)$. If $t(\cdot) \neq T^{\lambda^*}(\cdot)$, let

$q^{\lambda^*}(\theta) \equiv \int_0^\theta \frac{1}{\tau}\dot{T}^{\lambda^*}(\tau)\, d\tau$. We claim that

$$q(\theta) - q^{\lambda^*}(\theta) \geq \int_0^\theta \frac{1}{\tau^2}\left[t(\tau) - T^{\lambda^*}(\tau)\right] d\tau \geq 0,$$

for all θ, with the inequality holding as an equality if and only if $t(\cdot) = T^{\lambda^*}(\cdot)$.

Incentive compatibility implies that, for any $0 < \underline{\theta} \leq \theta$,

$$q(\theta) = q(\underline{\theta}) + \int_{\underline{\theta}}^\theta \frac{1}{\tau} dt(\tau).$$

Using the dominated convergence theorem,

$$t'(0) = \lim_{\underline{\theta} \searrow 0} \int_0^{\underline{\theta}} \frac{\tau}{\underline{\theta}} dq(\tau) = 0.$$

Since $0 \leq T^{\lambda^*}(\theta) \leq t(\theta)$, we also have $(T^{\lambda^*})'(0) = 0$. Therefore, using integration by parts and right-continuity of $q(\cdot)$

$$q(\theta) = q(0) + \frac{t(\theta)}{\theta} + \int_0^\theta \frac{1}{\tau^2} t(\tau)\, d\tau,$$

and similarly

$$q^{\lambda^*}(\theta) = \frac{T^{\lambda^*}(\theta)}{\theta} + \int_0^\theta \frac{1}{\tau^2} T^{\lambda^*}(\tau)\, d\tau.$$

It follows

$$\begin{aligned} q(\theta) - q^{\lambda^*}(\theta) &= q(0) + \frac{\left[t(\theta) - T^{\lambda^*}(\theta)\right]}{\theta} + \int_0^\theta \frac{1}{\tau^2}\left[t(\tau) - T^{\lambda^*}(\tau)\right] d\tau \\ &\geq \frac{\left[t(\theta) - T^{\lambda^*}(\theta)\right]}{\theta} + \int_0^\theta \frac{1}{\tau^2}\left[t(\tau) - T^{\lambda^*}(\tau)\right] d\tau \\ &\geq \int_0^\theta \frac{1}{\tau^2}\left[t(\tau) - T^{\lambda^*}(\tau)\right] d\tau \geq 0, \end{aligned}$$

with the last inequality holding strictly if $G_+(\theta) \equiv \left\{\theta' \in [0, \theta] \mid t(\theta') > T^{\lambda^*}(\theta')\right\}$ has positive Lebesgue measure. Since both $q(\cdot)$ and $q^{\lambda^*}(\cdot)$ are left-continuous, $G_+(\theta)$ has positive Lebesgue measure as long as $G_+(\theta) \neq \emptyset$.

Now suppose $t(\cdot) \neq T^{\lambda^*}(\cdot)$. The above implies

$$\lim_{\theta \to \infty} q^{\lambda^*}(\theta) \leq \lim_{\theta \to \infty} q(\theta) - \int_0^\theta \frac{1}{\tau^2}\left[t(\tau) - T^{\lambda^*}(\tau)\right] d\tau < 1,$$

where the last inequality is due to the supposition $t \neq T^{\lambda^*}$. Let $\Delta \equiv \lim_{\theta \to \infty} q(\theta) - \lim_{\theta \to \infty} q^{\lambda^*}(\theta) > 0$.

Let b be such that $b(1 - F(b-)) > 0$ for all $F \in \mathcal{F}$; existence of such b was established in

Lemma 3. We define a new mechanism M' with allocation rule

$$q'(\theta) = q^{\lambda^*}(\theta) + \Delta 1_{[b,\infty)}(\theta),$$

and transfer function $t'(\theta) = \int_0^\theta \tau dq'(\tau)$. Since $q'(\cdot)$ is non-decreasing, right-continuous and satisfies $q'(\theta) \in [0,1]$ for $\theta \in \mathbb{R}_+$, it is in $\mathcal{M}$. Additionally, for any $F \in \mathcal{F}$,

$$\int_0^\infty t'(\theta)\, dF(\theta) = \int_0^\infty T^{\lambda^*}(\theta)\, dF(\theta) + \Delta b\,(1 - F(b-)) > \sum_{i=0}^N \lambda_i^* k_i,$$

which is a contradiction with the optimality of $M = (q,t)$.

Finally, we show that λ^* solves the problem $\mathcal{P}$. Consider any vector $\lambda \in \mathbb{R}_- \times \mathbb{R}^{N-1} \times \mathbb{R}_-$ such that

$$\int_0^\infty \frac{\dot{T}^\lambda(\theta)}{\theta} d\theta \leq 1.$$

The pair $M^\lambda = (q^\lambda, T^\lambda)$, where $q^\lambda(\theta) \equiv \int_0^\theta \frac{\dot{T}^\lambda(\theta)}{\theta} d\theta$, is a mechanism in $\mathcal{M}$. But since the equilibrium revenue is equal to $\sum_{i=0}^N \lambda_i^* k_i$ and $T^\lambda(\theta) \geq \sum_{i=0}^N \lambda_i \theta^i$, we have

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

□

Proof of Lemma 9. First notice that $t^*(0) = 0$ implies that $\lambda_0^* \leq 0$. Since the seller obtains strictly positive revenue, it must be the case that $\max\{\lambda_1^*, \lambda_2^*\} > 0$. Now notice that $\lambda_2^* < 0$. If $\lambda_2^* > 0$, the polynomial would be increasing from some θ on, and therefore the optimal mechanism given by Proposition 3 would have $q(\theta) < 1$ for every θ. This would violate Lemma 4. If $\lambda_2^* = 0$, then $\lambda_1^* > 0$, which violates the resource constraint similarly.

Next, $\lambda_1^* > 0$ follows from $\max\{\lambda_1^*, \lambda_2^*\} > 0$. A similar argument as above implies $\lambda_0^* < 0$. If $\lambda_0^* = 0$, then we would have $t^{*'}(0) = \lambda_1^* > 0$, which is impossible. □

Proof of Proposition 4. The proof jumps between the problems $\mathcal{P}$ and $\mathcal{P}_0$. In the text preceding the proposition we argued that the optimal mechanism is determined by a polynomial $\lambda_0 + \lambda_1\theta + \lambda_2\theta^2$, where $\lambda_0 < 0$, $\lambda_1 > 0$, and $\lambda_2 < 0$. The polynomial gives us the function T^λ:

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

where $\underline{\theta}$ is the smaller of the two zeroes of the polynomial and $\overline{\theta}$ is the maximizer of the polynomial; notice that $\underline{\theta}$ and $\overline{\theta}$ are functions of λ_0, λ_1 and λ_2. Proposition 3 implies

$$Q(\theta) = \begin{cases} 0, & \text{if } \theta < \underline{\theta}, \\ 2\lambda_2(\theta - \underline{\theta}) + \lambda_1(\log \theta - \log \underline{\theta}), & \text{if } \underline{\theta} \le \theta \le \overline{\theta}, , \\ 1 & \text{if } \overline{\theta} < \theta, \end{cases}$$

where we invoked Lemma 4 to say that $Q = 1$ for large θ. The mechanism can be interpreted as a randomization over prices with a cdf $H(\theta) = Q(\theta)$. That is, the seller randomizes over an interval $[\underline{\theta}, \overline{\theta}]$ with a density $h(\theta) = 2\lambda_2 + \frac{\lambda_1}{\theta}$.

If the seller is to randomize, he better be indifferent. If F^* is a distribution Nature uses in the equilibrium the seller's payoff from charging a price θ, $\theta(1 - F^*(\theta))$, must be constant on $[\underline{\theta}, \overline{\theta}]$. Consequently,

$$F^*(\theta) = 1 - \frac{c}{\theta},$$

where c is a constant.

Outside of interval $[\underline{\theta}, \overline{\theta}]$ the polynomial and the function T^λ differ. Therefore, Lemma 8 and Proposition 3 imply that the support of Nature's equilibrium distribution F^* coincides with the support of the seller's distribution over prices. The seller can then guarantee himself a payoff arbitrarily close to $\underline{\theta}$ by charging a price just below $\underline{\theta}$. Hence, $c \ge \underline{\theta}$; that is, the seller's equilibrium payoff can be no smaller than $\underline{\theta}$. On the other hand, since F^* is a cdf, $F^*(\theta) \ge 0$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$, which, in turn, implies $c \le \underline{\theta}$. Therefore, $c = \underline{\theta}$: the seller's equilibrium payoff is equal to the lower boundary of the interval over which the seller is randomizing. This establishes that Nature has no mass point at the bottom of the interval.

Nature's equilibrium probability can now be rewritten as $F^* = 1 - \underline{\theta}/\theta$. Notice that

$$\lim_{\tau \nearrow \overline{\theta}} F^*(\tau) = 1 - \underline{\theta}/\bar{\theta},$$

implying that F^* has an atom of mass $\underline{\theta}/\overline{\theta}$ at $\overline{\theta}$. To pin down the upper and the lower bound of the support, we use the two moment conditions. In particular, $\int_{\underline{\theta}}^{\overline{\theta}} \theta dF^*(\theta) = k_1$, boils down to

$$\underline{\theta}\left[1 + \log(\overline{\theta}) - \log(\underline{\theta})\right] = k_1,$$

while $\int_{\underline{\theta}}^{\overline{\theta}} \theta^2 dF^*(\theta) = k_2$ results in

$$\underline{\theta}\left[2\overline{\theta} - \underline{\theta}\right] = k_2.$$

We argue that the above two moment conditions have a unique solution for $\underline{\theta}$ and $\overline{\theta}$. Solving the second moment condition for $\overline{\theta}$ and plugging it into the first moment condition, one obtains

$$\underline{\theta}\left[1 + \log\left(\frac{k_2}{\underline{\theta}^2} + 1\right) - \log 2\right] = k_1.$$

When $\underline{\theta}$ goes to infinity, so does the left-hand side of the above expression. On the other hand,

$$\lim_{\underline{\theta}\to 0} \underline{\theta}\left[1+\log\left(\frac{k_2}{\underline{\theta}^2}+1\right)-\log 2\right] = \lim_{\underline{\theta}\to 0} \frac{\left[1+\log\left(\frac{k_2}{\underline{\theta}^2}+1\right)-\log 2\right]}{\frac{1}{\underline{\theta}}}$$
$$= 2\lim_{\underline{\theta}\to 0} \frac{k_2 x}{x^2+k_2} = 0,$$

where the second equality is obtained after an application of L'Hopital's rule. Lastly, we argue that the left-hand side of the above expression is increasing in $\underline{\theta}$. Let

$$f(x) = x\left[1-\log 2+\log\left(\frac{k_2}{x}+1\right)\right].$$

Then

$$f'(x) = \log\left(\frac{\frac{k_2}{x^2}+1}{2}\right) + \frac{1-\frac{k_2}{x^2}}{1+\frac{k_2}{x^2}} \geq \frac{\frac{\frac{k_2}{x^2}+1}{2}-1}{\frac{\frac{k_2}{x^2}+1}{2}} + \frac{1-\frac{k_2}{x^2}}{1+\frac{k_2}{x^2}} = 0,$$

where we used the inequality $\ln(y) \geq \frac{y-1}{y}$ for $y \geq 0$, which holds strictly with the exception of $y = 1$. Since $f(x)$ is also continuous, there exists a unique x such that $f(x) = k_1$.

The above argument fully pins down Nature's equilibrium strategy. We are left to determine the optimal mechanism for the seller. We already argued that the seller randomizes over prices with a density of the type $2\lambda_2 + \lambda_1/\theta$ over the interval $[\underline{\theta}, \overline{\theta}]$. Since $\overline{\theta}$ is the highest point at which the polynomial and its non-negative monotonic hull coincide, it must be the case that $\overline{\theta}$ is the maximizer. The first-order condition reads:

$$\lambda_1 + 2\lambda_2\overline{\theta} = 0.$$

The other condition to pin down parameters λ_1 and λ_2 comes from the fact that $h(\theta) = \lambda_1 + 2\lambda_2/\theta$ is the density over the seller's prices, hence, $\int h(\theta)d\theta = 1$ or

$$\lambda_1(\overline{\theta}-\underline{\theta}) + 2\lambda_2(\log(\overline{\theta})-\log(\underline{\theta})) = 1.$$

Direct mechanism (q, t) is easily obtained from the seller's randomization. □

Proof of Proposition 5. The same result as is obtained above 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 with just one moment. Therefore, a Nash equilibrium exists in the zero-sum game where Nature's strategy space is constrained to distributions with support contained in $[0, \bar{\theta}]$.

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 $[\underline{\theta}, \bar{\theta}]$, 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 λ_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 θ using distribution $F(\cdot)$ is:

$$\underline{\theta}\left[1+\log(\bar{\theta})-\log(\underline{\theta})\right]=k_1.$$

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

Finally, λ_1 is pinned down from $\int_{\underline{\theta}}^{\bar{\theta}} h(\theta)d\theta=1$: $\lambda_1=\frac{1}{\log(\bar{\theta})-\log(\underline{\theta})}$. □

Proof of Proposition 6. We start by defining the Lambert function and showing some of its useful properties.

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

$$\underline{\theta}(1+\log(\bar{\theta})-\log(\underline{\theta}))=k_1.$$

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

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

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

Remember that the lower branch of the Lambert W function, 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

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

or

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

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 ρ 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 ρ 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 converges 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 α and β 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 α and β are differentiable at all points $k_1 > 0$. That β 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

$$-W_{-1}(-k_1/e) > \frac{k_1}{\left(1 - \sqrt{1-k_1}\right)}. \tag{9}$$

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 rhs goes to 2 as k_1 approaches 0, the lhs diverges to ∞. 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.

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

At k_1' the slope of the lhs 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 rhs 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 lhs 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 (9) the lhs is decreasing faster than the rhs. 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 $\left[\underline{\theta},\bar{\theta}\right]$. The optimal distribution for the seller knowing the mean k_1 and second moment $k_2 = \underline{\theta}\left[2\bar{\theta}-\underline{\theta}\right]$, from Proposition 4, can be written as

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

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

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

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

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

which, in turn, has the same sign as

$$\frac{\left(\ln\theta-\ln\underline{\theta}\right)}{\left(\theta-\underline{\theta}\right)} - \frac{\left(\ln\bar{\theta}-\ln\underline{\theta}\right)}{\left(\bar{\theta}-\underline{\theta}\right)}.$$

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

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

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

References

Auster, S. (2015): "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.

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.

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.

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

Lopomo, G., L. Rigotti, and C. Shannon (2009): "Uncertainty in mechanism design," Working paper.

——— (2011): "Knightian uncertainty and moral hazard," *Journal of Economic Theory*, 146, 1148–1172.

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

Monteiro, P. and F. Page (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 (2013): "Full implementation and belief restrictions," Working paper.

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

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

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

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

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.