Multidimensional Mechanism Design: Finite-Dimensional Approximations and Efficient Computation

Alexandre Belloni, Giuseppe Lopomo, Shouqiang Wang

Fuqua School of Business, Duke University, Durham, North Carolina 27708
{abn5@duke.edu, glopomo@duke.edu, sw55@duke.edu}

Multidimensional mechanism design problems have proven difficult to solve by extending techniques from the one-dimensional case. This paper considers mechanism design problems with multidimensional types when the seller's cost function is not separable across buyers. By adapting results obtained by Border [Border, K. 1991. Implementation of reduced form auctions: A geometric approach. *Econometrica* **59** 1175–1187], we transform the seller's problem into a representation that only involves “interim” variables and eliminates the dimensionality dependence on the number of buyers. We show that the associated infinite-dimensional optimization problem posed by the theoretical model can be approximated arbitrarily well by a sequence of finite-dimensional linear programming problems.

We provide an efficient—i.e., terminating in polynomial time in the problem size—method to compute the separation oracle associated with the Border constraints and incentive compatibility constraints. This implies that our finite-dimensional approximation is solvable in polynomial time.

Finally, we illustrate how the numerical solutions of the finite-dimensional approximations can provide insights into the nature of optimal solutions to the infinite-dimensional problem in particular cases.

Subject classifications: linear programming; large-scale systems; algorithms; infinite dimensional; bidding; auctions.

Area of review: Special Issue on Computational Economics.

History: Received August 2008; revision received November 2009; accepted December 2009. Published online in

Articles in Advance June 3, 2010.

1. Introduction

This paper revisits the classic problem of a monopolist offering a product that can be sold at different quality levels. The literature on this subject dates back to Mussa and Rosen (1978), who studied a model in which the seller's cost function is separable across buyers, and the buyers' private information is represented by a one-dimensional parameter (see also Maskin and Riley 1984). In his seminal paper on optimal auctions (Myerson 1981), Myerson dealt with the case in which the seller has a single object (hence, the cost function is not separable across buyers), and showed that under mild regularity assumptions, any auction in a large class that includes several popular formats, e.g., the English auction, the Dutch auction, and the first-price auction, maximizes the expected profit among all feasible selling procedures. Myerson's results immediately extend to the case in which the seller can choose the object's quality, as long as one maintains the assumption that each buyer's private information is represented by a one-dimensional variable.

This paper considers the case in which all buyers have multidimensional types and the seller's cost function is not separable across buyers. We show that the associated infinite-dimensional optimization problem posed by the theoretical model can be approximated arbitrarily well

by a sequence of finite-dimensional linear programming problems. In addition, we develop an efficient computational method to solve the resulting linear programming problem efficiently (in theory and practice).

The analysis of the multidimensional case has proven significantly more complex, and has produced results, under the assumption that costs are separable across buyers, that seem unable to explain the use of popular selling procedures in terms of their optimality for the seller. Rochet and Choné (1998) provide general characterizations of optimal mechanisms in a multidimensional version of Mussa and Rosen's model, which had been previously investigated by Wilson (1993) and Armstrong (1996). Manelli and Vincent (2007) study the closely related problem of a monopolist selling several objects to a single buyer, which has a privately known value for each.

The main technical difficulty that emerges in these papers is that, for generic parameter values, “non-local” incentive compatibility constraints are binding, hence it is optimal for the seller to “bunch” types of consumers in various regions, i.e., to induce them to choose the same configuration of quality levels. Rochet and Choné (1998) deal with this issue by developing a “sweeping operator” that allows them to obtain an implicit characterization of optimal solutions. Manelli and Vincent (2007) show that a large variety

Table 1. Summary of the relation with literature.

One-dimensional types	Multidimensional types
Separable cost Mussa and Rosen (1978), Maskin and Riley (1984)	Wilson (1993), Armstrong (1996), Rochet and Choné (1998), Manelli and Vincent (2007)
Nonseparable cost Myerson (1981)	Present paper

of mechanisms can be optimal for the seller, depending on her prior beliefs about the buyer's type.

The main difference between the model studied in the present paper and the models described above is that in our case the seller's cost function is not separable across buyers. In particular, we assume that the seller has just one object, as in Myerson's optimal auction problem, but can set the object's quality equal to any "grade" within a finite set. Each buyer has a privately known willingness to pay for each quality level. To the best of our knowledge, this problem has not been studied in detail. Table 1 summarizes how this paper relates to the previous literature.

Our contribution can be described as follows. First, we reformulate the seller's problem using "interim" variables only. This reformulation also eliminates the dimensionality dependence on the number of buyers, a property that is desirable in practical applications where the number of buyers tends to be significantly larger than the number of quality levels. This approach hinges on the adaptation of results obtained by Border (1991) to our environment with multiple quality levels.

Next we establish that the infinite-dimensional optimization problem posed by the theoretical model can be approximated arbitrarily well by a sequence of finite-dimensional linear programming problems, under standard assumptions on the buyers' valuations. Indeed, this result cannot hold for arbitrary specifications of the buyers' valuation, because the set of all feasible mechanisms includes discontinuous and nonsmooth functions, for which finite-dimensional approximations may perform poorly. Our proof hinges on regularity assumptions, such as smoothness conditions on the probability distribution that represents the seller's belief about the buyers' values, and consists of using an isoperimetric inequality to establish a notion of stability under small perturbations.

Having provided a theoretical justification for the finite-dimensional approximation, we approach the computational problem associated with such approximation. The proposed reformulation is based on interim variables and can still be cast as a linear programming problem. To deal with the exponentially large constraint sets, we resort to the implementation of a cutting-plane method that dynamically incorporates constraints as needed, and thus never deals with all of them at once. We provide an algorithm, which terminates in polynomial time in the problem size, to compute

the separation oracle associated with the Border constraints and incentive compatibility constraints. Because it is well-known, this implies that the finite-dimensional approximation is solvable in polynomial time (e.g., Grötschel et al. 1981, 1988; Padberg and Rao 1980).

The computation of convergent approximations also provides an opportunity to detect structural patterns in the solution of finite-dimensional approximations. These patterns can help in formulating well-educated guesses about the nature of simple mechanisms that can approximate the optimal solution to the original infinite-dimensional problem.

The paper is organized as follows. Section 2 introduces the theoretical model and poses the infinite-dimensional problem. In §3 we discuss the finite-dimensional approximation and establish the approximation results. Section 4 describes a cutting-plane algorithm for the finite-dimensional approximation. This section also shows that the separation oracles can be implemented in polynomial time, and finite termination of the overall method is achieved. The implementation and computational results are discussed in §5. Section 6 is devoted to the discussion of a particular example. Proofs and technical results are deferred to the appendix.

2. The Model

The seller of a single object faces N potential buyers. The seller can set the object's quality equal to any "grade" $j = 1, \dots, J$. For notational convenience we identify the set of buyers and the set of quality grades with their cardinalities, that is, $N = \{1, \dots, N\}$ and $J = \{1, \dots, J\}$. Each buyer $i \in N$ has utility function

$$u_i = \sum_{j \in J} v_j^i q_j^i - m_i,$$

where q_j^i denotes the probability that the object of grade j is awarded to buyer i , v_j^i denotes the buyer's willingness to pay for it, and m_i denotes his payment to the seller. Buyer i 's type $v^i = (v_1^i, \dots, v_J^i)$ is obtained as the realization of a random variable, distributed independently of the other buyers' values $v^{-i} := (v^1, \dots, v^{i-1}, v^{i+1}, \dots, v^N)$, with support $V := \prod_{j \in J} [\underline{v}_j, \bar{v}_j]$ and probability distribution function F . Buyer i observes the realization of his type v^i privately. Finally, the seller faces a cost c_j to provide the object at the quality level $j = 1, \dots, J$. In the most general version of the model, c_j may depend on the entire type profile $v \in V$. The model can also be adapted, with the required modifications, to monopsony situations, e.g., procurement contexts, in which a single buyer faces N competing sellers.

The seller's problem is to select a pricing schedule that maximizes her expected profit. Invoking the revelation principle (see Myerson 1979), we restrict attention to the set of all "direct" selling mechanisms in which the buyers simply report their types, and reporting truthfully is a Bayesian-Nash equilibrium. Formally, a direct mechanism (q, m) consists of an *assignment* rule

$$q: V^N \rightarrow \Delta(J \times N),$$

where $\Delta(J \times N)$ denotes the simplex on the set $J \times N$, and *payment rules*

$$m = (m^1, \dots, m^N): V^N \rightarrow \mathbb{R}^N,$$

specifying, for any profile of reported types $\omega := (v^1, \dots, v^N) \in V^N$, the probability $q_j^i(\omega)$ that buyer i is awarded the object of grade j , $(i, j) \in N \times J$, and the payment $m^i(\omega)$ that buyer i makes to the seller. The restriction to deterministic payment rules is without loss of generality because all agents are risk neutral.

A mechanism (q, m) satisfies:

(i) *incentive compatibility*, if “truth-telling” is a Bayesian-Nash equilibrium, i.e.,

$$\hat{U}_i(v, v) \geq \hat{U}_i(\hat{v}, v) \quad \text{for all } \hat{v}, v \in V, i \in N; \quad (\text{IC})$$

and

(ii) *individual rationality*, if each buyer has no incentive to decline participation, i.e.,

$$\hat{U}_i(v, v) \geq 0 \quad \text{for all } v \in V, i \in N; \quad (\text{IR})$$

where $\hat{U}_i$ denotes buyer i 's interim expected payoff function

$$\begin{aligned} \hat{U}_i(\hat{v}, v) &\equiv \int_{V^{N-1}} [v \cdot q^i(\hat{v}, v^{-i}) - m^i(\hat{v}, v^{-i})] \\ &\quad \cdot \prod_{k \in N \setminus \{i\}} dF(v^k) \quad v, \hat{v} \in V. \end{aligned}$$

The seller's problem can be formulated as:

$$(P_0) \quad \begin{cases} \max_{q, m} \int_{V^N} \sum_{i \in N} \left[m^i(\omega) - \sum_{j \in J} c_j(\omega) q_j^i(\omega) \right] \prod_{i \in N} dF(v^i) \\ \text{subject to} \\ (\text{IC}), (\text{IR}) \text{ and} \\ (C) \quad \begin{cases} q_j^i(\omega) \geq 0, & \text{for all } (i, j, \omega) \in N \times J \times V^N, \\ \sum_{i \in N} \sum_{j \in J} q_j^i(\omega) \leq 1, & \text{for all } \omega \in V^N. \end{cases} \end{cases}$$

The linear program (P_0) contains functions with domain $V^N \subset \mathbb{R}^{JN}$. Therefore, the size of any discrete approximation of (P_0) grows exponentially with both the number of buyers N and the number of quality levels J . Thus, direct computation of solutions to finite approximations of (P_0) can only be done for relatively small grid sizes. An initial significant simplification comes directly from the observation that because all buyers are ex ante identical (they draw their values from the same distribution), we can restrict attention, without loss of generality, to *symmetric* mechanisms. However, even after imposing symmetry, the growth with N and J remains exponential.

We propose a reformulation of (P_0) in which the dimensionality is independent of the number of buyers—in particular, all variables have V as domain instead of V^N . The

reformulation relies on *interim* variables (Q, U) obtained by integrating out the types of all but one buyer. The interim probability that a buyer is awarded the object of grade $j \in J$ is

$$Q_j(v) \equiv \int_{V^{N-1}} q_j^i(v, v^{-i}) \prod_{k \in N \setminus \{i\}} dF(v^k) \quad \text{for all } v \in V,$$

and the interim expected utility function of each buyer is defined as

$$U(v) \equiv \sum_{j \in J} v_j Q_j(v) - \int_{V^{N-1}} m^i(v, v^{-i}) \prod_{k \in N \setminus \{i\}} dF(v^k),$$

for all $v \in V$.

By appealing to an extension of the results in Border (1991) (see Lemma 13), we can replace the constraints in (C) with the following “Border constraints” written only with the interim probabilities.

$$(B) \quad \begin{cases} Q_j(v) \geq 0, & \text{for all } j \in J, v \in V \\ N \int_A \sum_{j \in J} Q_j(v) dF(v) \leq 1 - \left(\int_{V \setminus A} dF(v) \right)^N, & \text{for all } A \subset V. \end{cases}$$

Moreover, we can rewrite (IC) and (IR) with interim variables as follows

$$U(v) - U(\hat{v}) \geq \langle Q(\hat{v}), v - \hat{v} \rangle \quad \text{for all } (v, \hat{v}) \in V \times V, \quad (\text{IIC})$$

and

$$U(v) \geq 0 \quad \text{for all } v \in V; \quad (\text{IIR})$$

and, by appealing to the following well-known characterization lemma (Manelli and Vincent 2007), we eliminate most of the (IIR) constraints.

LEMMA 1. *If (Q, U) satisfies (IIC), then U is convex, and $Q(v) \in \partial U(v)$ for all $v \in V$. Conversely, if U is convex, then (Q, U) , where $Q(v) \in \partial U(v)$ for all $v \in V$ satisfies (IIC).*

Lemma 1 immediately implies that (IIR) holds if and only if $U(\underline{v}) \geq 0$, where $\underline{v} := (\underline{v}_1, \dots, \underline{v}_j, \dots, \underline{v}_J)$. Clearly, in any mechanism that maximizes the seller's expected revenue, we must have $U(\underline{v}) = 0$.

Thus, the seller's problem of finding a price schedule that maximizes her expected profit subject to incentive compatibility and individual rationality can be stated as

$$(P_*) \quad \begin{cases} OPT_* = \max_{Q, U} \int_V \left[\sum_{j \in J} [v_j - c_j(v)] \right. \\ \quad \left. \cdot Q_j(v) - U(v) \right] dF(v) \\ \text{subject to } (B), (IIC), \text{ and } U(\underline{v}) = 0. \end{cases}$$

As a final remark, note that the formulation (P_*) eliminates the dependence of the number of variables on N , because it is searching for variables that are functions on $V \subset \mathbb{R}^J$. However, this simplification comes at the cost of having to deal with an exponentially large class of Border constraints (B). This structure lends itself to the use of

cutting-plane algorithms, which consider only a subset of the constraints at any iteration. By dynamically adjusting the subset of constraints, it is possible to efficiently solve the original problem (i.e., in polynomial time in the dimension of the problem).

3. Finite-Dimensional Approximation

In this section we study how to approximate the infinite-dimensional problem of interest (P_*) by a sequence of finite-dimensional problems. In particular, we provide an extension of any solution to a discretized problem for which we can guarantee (a rate of) convergence to the solutions of (P_*). Throughout the paper we impose the following regularity conditions.

ASSUMPTION 2. *The data of the problem (P_*) satisfy: (i) the type space V is a compact subset of $\mathbb{R}_{++}^J$; (ii) the probability distribution F has a twice continuously differentiable density function f , which is bounded away from zero; (iii) the mapping $w: V \rightarrow \mathbb{R}_+^J$, defined as $w_j(v) \equiv v_j - c_j(v)$, is twice continuously differentiable.*

In order to approximate an optimal solution of (P_*), we begin by discretizing the type space V . Let T denote a positive integer that will control the precision of the discretization. For each $j \in J$, let $V_T(j)$ denote the discretization of the interval $[v_j, \bar{v}_j]$

$$V_T(j) = \{v_j, v_j + \epsilon, v_j + 2\epsilon, \dots, \bar{v}_j\},$$

where $\epsilon = \min_{j \in J} \{(\bar{v}_j - v_j)/T\}$. Our discretized version of V , parameterized by T , is defined as $V_T := \prod_{j \in J} V_T(j)$. The grid V_T is a set with $O(T^J)$ elements that defines a net over V such that

$$\max_{v \in V} \text{dist}(v, V_T) \leq \epsilon \sqrt{J}.$$

Based on the probability density function f , we can define a probability distribution function on V_T by setting $\hat{f}(v) = f(v)/(\sum_{t \in V_T} f(t))$. Note that $\hat{f}(v)$ approximates the measure on the hypercube $\prod_{j \in J} [v_j, v_j + \epsilon]$.

For each $T > 0$, by replacing V with the grid V_T , we obtain the following (sequence of) finite-dimensional approximation for (P_*).

$$(P_T) \left\{ \begin{array}{l} OPT_T = \max_{Q, U} N \sum_{v \in V_T} \left(\sum_{j \in J} w_j(v) Q_j(v) - U(v) \right) \hat{f}(v) \\ U(v) - U(\hat{v}) \geq \langle Q(\hat{v}), v - \hat{v} \rangle \\ \quad \text{for all } (v, \hat{v}) \in V_T \times V_T, \\ U(v) = 0, \\ N \sum_{v \in A} \left[\sum_{j \in J} Q_j(v) \right] \hat{f}(v) \\ \leq 1 - \left(\sum_{v \in V_T \setminus A} \hat{f}(v) \right)^N, \\ \quad \text{for all } A \subset V_T, \\ Q_j(v) \geq 0, \sum_{j \in J} Q_j(v) \leq 1, \\ 0 \leq U(v) \leq \bar{U}, \quad \text{for all } v \in V_T, \end{array} \right.$$

where $\bar{U}$ is an upper bound on U that can be derived from the fact that the set V is compact and Q is bounded from above. The “additional” upper bounds on Q are already implied by the Border constraints; they are included explicitly to highlight the fact that any feasible solution is bounded. We denote by (Q^T, U^T) an optimal solution of (P_T).

Because our “target” program (P_*) is infinite dimensional and allows discontinuous functions within its feasible set, the following issues must be confronted. First, in the present setting, the discretization affects not only the computation of the objective function value, but also the constraint formulation. Therefore, intuitive extensions of the solution to a discretized problem might not be near feasible for the continuous problem (P_*). Second, the presence of discontinuous functions Q_j in the feasible set introduces challenges to approximations based on grids. Intuitively, in order for an approximation scheme to be successful, the problem of interest must be stable under small perturbations.

We will work with the following approximation notions. Given a pair of mappings (Q, U) , let $\delta_*(Q, U)$, respectively, $\delta_T(Q, U)$, denote the supremum over all constraint violations in (P_*), respectively in (P_T), by (Q, U) . Analogously, let $OPT_*(Q, U)$, respectively, $OPT_T(Q, U)$, denote the objective function value obtained by (Q, U) in (P_*), respectively in (P_T). Note that (Q, U) is feasible for (P_*) only if $\delta_*(Q, U) \leq 0$.

The following notion of asymptotic optimality is standard: all violations converge uniformly to zero, and optimality is achieved in the limit.

DEFINITION 3. A sequence of mappings (Q^T, U^T) defined over V is said to be asymptotically optimal for (P_*), if $\lim_{T \rightarrow \infty} \delta_*(Q^T, U^T) = 0$ and $\lim_{T \rightarrow \infty} OPT_*(Q^T, U^T) = OPT_*$.

Next, we propose an extension of the finite-dimensional solutions (Q^T, U^T) that has useful properties. Our first result is concerned with the feasibility guarantees of the extension.

THEOREM 4. *Let (Q^T, U^T) be a solution to (P_T) associated with the grid V_T induced by $\epsilon = O(1/T)$. Consider the extensions $(\tilde{Q}^T, \tilde{U}^T)$ of (Q^T, U^T) to the set V defined as*

$$\begin{aligned} \tilde{U}^T(v) &\equiv \max_{\hat{v} \in V_T} U^T(\hat{v}) + \langle Q^T(\hat{v}), v - \hat{v} \rangle, \quad \text{and} \\ \tilde{Q}_j^T(v) &\equiv Q_j^T(\hat{v}), \quad j \in J, \hat{v} \in V_T, \hat{v} \leq v < \hat{v} + \epsilon e. \end{aligned} \quad (1)$$

Under Assumption 2, we have:

- (i) $\tilde{U}^T$ is convex, (ii) $\tilde{Q}^T(v) \in \partial_{2\epsilon} \tilde{U}^T(v)$ for all $v \in V$, and (iii) $\delta_*(\tilde{Q}^T, \tilde{U}^T) \leq O(1/T)$.

Theorem 4 yields nonasymptotic bounds in the maximum violation of the extension. In particular, it establishes

that the extended $\tilde{Q}^T(v)$ is a 2ϵ -subgradient of $\tilde{U}^T(v)$ at $v \in V$; namely, for any $v' \in V$ we have

$$\tilde{U}^T(v') \geq \tilde{U}^T(v) + \langle \tilde{Q}^T(v), v' - v \rangle - 2\epsilon.$$

Recall that the discretization also affects the Border constraints (B). Theorem 4 also establishes bounds on the violations over all measurable subsets of V . However, it will be possible to satisfy all Border constraints of (P_*) by rescaling a feasible solution of (P_T) , or vice versa. A key step in the proof of this result is provided in the following technical lemma, which establishes an isoperimetric inequality relating violations of Border constraints and the measure of the associated subset of the type space.

LEMMA 5. *Consider an arbitrary probability measure F (possibly discrete), and any measurable mapping $Q: V \rightarrow \mathbb{R}^J$ such that $0 \leq \sum_{j \in J} Q_j(v) \leq 1$. Then, for any measurable subset $A \subseteq V$, and any $\eta \geq 0$, the inequality*

$$N \int_A \sum_{j=1}^J Q_j(v) dF(v) \geq 1 - \left(\int_{A^c} dF(v) \right)^N + \eta$$

implies that $\int_A dF(v) \geq \sqrt{2\eta}/N$.

Lemma 5 is instrumental in recovering feasible solutions for (P_*) or (P_T) based on near-feasible solutions that violate only Border constraints. Essentially, it asserts that if a Border constraint is violated, the probability measure of the associated subset of the type space must be relatively large. In turn, this ensures that the right-hand side of the constraint is large (because the probability measure of the complement cannot be close to one). Corollary 6 below shows how all Border constraints can be satisfied without substantially reducing the corresponding objective function value.

COROLLARY 6. *Assume that a pair of mappings (Q, U) satisfies $\delta(Q, U) \leq 1/2$, where $\delta = \delta_*$ or $\delta = \delta_T$. Then, the rescaled pair*

$$(Q^r, U^r) \equiv \left(1 - \frac{\sqrt{\delta(Q, U)}}{\sqrt{2} - \sqrt{\delta(Q, U)}} \right) \cdot (Q, U)$$

satisfies all Border constraints and $\delta(Q^r, U^r) \leq \delta(Q, U)(1 - (\sqrt{\delta(Q, U)} / (\sqrt{2} - \sqrt{\delta(Q, U)})))$.

By applying the previous corollary to the restriction of the optimal solution (Q^*, U^*) to the grid V_T and (P_T) we are able to bound the (unknown) optimal value of (P_*) .

THEOREM 7. *Let (Q^*, U^*) denote the optimal solution for (P_*) and let $(\hat{Q}^*, \hat{U}^*)$ be their restriction on V_T . Under Assumption 2 we have that $\delta_T(\hat{Q}^*, \hat{U}^*) = O(1/T)$ and*

$$OPT_* \leq OPT_T \left(1 - \frac{\sqrt{O(1/T)}}{\sqrt{2} - \sqrt{O(1/T)}} \right)^{-1} + O\left(\frac{1}{T}\right).$$

Theorem 7 provides a numerical bound for the true optimal value of (P_*) that can be used to evaluate the quality of any feasible solution for (P_*) . Finally, we close this section by establishing its main result.

COROLLARY 8. *Under Assumption 2 the sequence of extensions $(\tilde{Q}^T, \tilde{U}^T)$ defined as in (1) is asymptotically optimal for (P_*) .*

The proof of the above corollary simply combines Theorems 4 and 7.

4. Algorithmic Structure

Having established the validity of the finite-dimensional approximations (1), we proceed to address the computational issues that arise when we solve (P_T) numerically.

To have a sense of how quickly the discrete program grows with T , note that the dimensionality of the variables, cardinality of (IIC) constraints, and cardinality of Border constraints are, respectively, $O(JT^J)$, $O(T^{2J})$, and $O(2^{T^J})$. In order to obtain numerical solutions for relatively large values of T , the particular structure of the problem (P_T) must be exploited. In particular, it is computationally intractable to simply enumerate all Border constraints for large values of T . We start by showing an equivalent characterization for the Border constraints.

LEMMA 9. *Consider the Border constraints in (P_*) or (P_T) . For any given $Q: V \rightarrow \mathbb{R}^J$ the sets $A \subset V$ can be restricted to the form of*

$$E_\alpha(Q) = \left\{ v \in V: \sum_{j \in J} Q_j(v) \geq \alpha \right\} \quad \text{for all } \alpha \geq 0.$$

The version of Lemma 9 where J is a singleton was first established by Border (1991). In this paper we fully exploit the computational consequences of this characterization. The following lemma plays a crucial role in our method. It establishes that both separation oracles associated with (IIC) and (B) can be implemented in polynomial time in T^J . That is, given a candidate solution (Q, U) , it is possible to efficiently verify that (Q, U) either satisfies all (IC) and (B) constraints or exhibits one constraint that is violated by (Q, U) .

LEMMA 10. *The separation oracle for the Border constraints (B) can be implemented in $O(JT^J \ln T)$ operations. The separation oracle for the (IC) constraints can be implemented in $O(JT^{2J})$ operations.*

The usefulness of the new representation of the Border constraints derived in (9) is notable. It enables an implementation of the separation oracle for (B) to be more efficient than the separation oracle for (IC) .

In light of Lemma 10, the cutting-plane method can be efficiently implemented to solve (P_T) for any T as follows:

Algorithm 11

Let S^k denote the set of constraints at iteration k .

Step 1: Let $k = 1$, $S^1 = \emptyset$, $\overline{OPT} = \infty$.

Step 2: Solve the linear program associated with S^k .
Let (Q^k, U^k) denote an optimal solution and let OPT^k denote the optimal value.

Step 3: Select a subset $I^k \subset S^k$ of inactive (IC) and (B) constraints for (Q^k, U^k) .

Step 4: Solve the separation oracle for (IC) and (B).
Let A^k denote a subset of the violated inequalities to be added.

Step 5: If $OPT^k < \overline{OPT}$, set $S^{k+1} \leftarrow (S^k \setminus I^k) \cup A^k$ and $\overline{OPT} \leftarrow OPT^k$. Else set $S^{k+1} \leftarrow S^k \cup A^k$.

Step 6: If $A^k = \emptyset$, stop.

Else set $k \leftarrow k + 1$ and goto Step 2.

Algorithm 11 is a cutting-plane algorithm that works with a subset of (IC) and (B) constraints at each iteration. Step 5 is there to prevent the dimensionality of the program from growing unnecessarily by dropping inactive constraints whenever the objective value decreases. The following lemma establishes that the algorithm terminates in a finite number of iterations.

THEOREM 12. *Algorithm 11 terminates in finite steps with the optimal solution.*

Finally, invoking Grötschel et al. (1981, 1988) and Padberg and Rao (1980), because both separations oracles can be solved in polynomial time with respect to the dimension of (P_T) , we have that the linear programming problem (P_T) is solvable in polynomial time (for instance, via the ellipsoid method).

5. Implementation and Computational Results

In this section we briefly discuss some aspects of our implementation and display our computational results in Table 2.

To properly solve any large-scale linear programming problem, considerable effort into the implementation is needed. Throughout the course of the implementation of Algorithm 11, we relied on many existent linear programming solvers, including SeDuMi, SDPT3, LINSOL, and CPLEX. During an initial phase of the implementation, we used MatLAB as the working environment for preliminary testing, and later we built a C++ environment to properly work with CPLEX for better memory management.

Table 2 summarizes the results of the implementation on a variety of instances for different configurations of the grid (T), buyers (N), and quality levels (J). As a by-product of the implementation, by solving instances without considering incentive compatibility constraints, we recover solutions for the first-best problem, whose optimal value

Table 2. Computational results for various model specifications.

J	N	T	OPT VAL	Active		Total	
				(IC)	(B)	(IC)	(B)
Seller's problem							
2	2	5	6.051008	56	9	(25) ⁴	2 ²⁵
2	2	10	5.946440	257	28	(100) ⁴	2 ¹⁰⁰
2	2	20*	5.896710	2,355	118	(400) ⁴	2 ⁴⁰⁰
2	2	30*	5.880743	2,802	920	(900) ⁴	2 ⁹⁰⁰
2	5	5	6.734230	54	9	(25) ⁴	2 ²⁵
2	5	10	6.605535	246	18	(100) ⁴	2 ¹⁰⁰
2	5	20*	6.535155	1,365	398	(400) ⁴	2 ⁴⁰⁰
2	10	5	6.99129	40	8	(25) ⁴	2 ²⁵
2	10	10	6.88799	207	17	(100) ⁴	2 ¹⁰⁰
2	10	20*	6.8185167	1,159	512	(400) ⁴	2 ⁴⁰⁰
3	2	5	4.263083	641	139	(125) ⁶	2 ¹²⁵
3	5	5	7.82672	793	171	(125) ⁶	2 ¹²⁵
3	10	5	8.16429	549	134	(125) ⁶	2 ¹²⁵
First best							
2	2	5	6.540664	—	25	—	2 ²⁵
2	2	10	6.487975	—	100	—	2 ¹⁰⁰
2	2	20	6.452313	—	400	—	2 ⁴⁰⁰
2	5	5	6.901566	—	24	—	2 ²⁵
2	5	10	6.833399	—	86	—	2 ¹⁰⁰
2	5	20*	6.800695	—	2,917	—	2 ¹⁰⁰
2	10	5	7.050792	—	20	—	2 ²⁵
2	10	10	6.991210	—	62	—	2 ¹⁰⁰
2	10	20*	6.956887	—	200	—	2 ¹⁰⁰
3	2	5	8.468104	—	181	—	2 ¹²⁵
3	5	5	9.035501	—	145	—	2 ¹²⁵
3	10	5	9.450985	—	135	—	2 ¹²⁵

Notes. The seller's maximization problem is the model considered in §3. The first-best problem ignores the incentive compatibility constraints. The instances marked with “*” were solved using a simplex method, whereas the others were solved using an interior point method.

is always larger than the value of the seller's maximization problem (P_T) . In all instances, the optimal solution was obtained. This illustrates the finite termination property established in Theorem 12.

The use of general cost functions and probability distributions is allowed by the implementation. Regarding the linear programming solver, as is typical in cutting-plane methods, warm-starting the solver with the previous solution can lead to significant improvements in running time. Although IPM were usually faster than simplex-type methods to solve each LP from the scratch, simplex-type were able to use warm-starts and were more efficient in the larger instances that we considered.

Finally, it is worth pointing out that the number of active Border constraints is substantially smaller than the number of active incentive compatibility constraints in all tested instances.

6. Insights and Discussion of a Particular Example

The computation of discrete approximations also provides an opportunity to detect patterns in the qualitative nature

Figure 1. Buyer surplus function U for $T = 30$.

of the optimal solutions that can help in formulating educated guesses about the functional form of the optimal solutions to the infinite-dimensional problem (P_*) posed by the theory.

In this section we focus on a simple example involving only two quality levels, where the buyers' types are drawn from a uniform distribution. Formally, we have $J = 2$, F uniform on a rectangular support $[v_1, \bar{v}_1] \times [v_2, \bar{v}_2]$. Figure 1 shows the graph of the buyer surplus U in (1), for a grid with $T = 30$.

A reasonable visual "approximation" for the function U is the maximum of two one-dimensional, strictly convex functions: one that depends only on v_1 and the other that depends only on v_2 . Moreover, these two functions meet along a linear path, on which U is nondifferentiable.

The solutions obtained by considering various values for the cost vector c , and various grid sizes, exhibited patterns that have prompted us to formulate the following simple mechanism as a candidate for a good approximation of the optimal solution in the infinite-dimensional problem (P_*) . Consider the following class of mechanisms, parameterized by the "add-on price" p :

$$Q_1^p(v_1, v_2) = F^{N-1}(v_1, v_1 + p)1\{v_1 > v_2 - p\}$$

$$Q_2^p(v_1, v_2) = F^{N-1}(v_2 - p, v_2)1\{v_1 < v_2 - p\}$$

and

$$U^p(v_1, v_2; p) = \max \left\{ \int_{v_1}^{v_1+p} Q_1^p(t, v_2) dt, \int_{v_2-p}^{v_2} Q_2^p(v_1, t) dt \right\}.$$

For any p , the mechanism (Q^p, U^p) is feasible for the infinite-dimensional program (P_*) because it can be implemented by the following auction named "exclusive buyer mechanism" for the case with two quality levels (see Brusco et al. 2008). The buyers compete in a second-price or ascending-bid auction, possibly augmented with a reserve price r , for the right to be the only buyer who gets to choose between buying the object of lower quality (grade 1) at no

additional cost, or the object of higher quality (grade 2) for an additional payment of p . Because the value for buyer $i \in N$ of winning the auction is $\beta_i \equiv \max\{v_1^i, v_2^i - p\}$, it is a dominant strategy for buyer i to bid β_i at the auction. Thus, in equilibrium, buyer i wins the auction only if $\beta_i \geq \max\{r, \max\{\beta_j; j \in N \setminus \{i\}\}\}$, is assigned the low-quality and high-quality object with probabilities $Q_1^p(v_1^i, v_2^i)$ and $Q_2^p(v_1^i, v_2^i)$, respectively, and his expected utility is $U^p(v_1, v_2; p)$.

Let p_* denote the value of p that maximizes the seller's expected profit within the class of exclusive buyers mechanisms, i.e., across all values of $p \in \mathbb{R}$. Table 3 shows how the value of the objective function in (P_*) evaluated at (Q^{p_*}, U^{p_*}) (last row) compares with the values obtained by the objective function in (1) for various grid sizes, and F uniform on $[v_1, \bar{v}_1] \times [v_2, \bar{v}_2] = [6, 8] \times [9, 11]$, $c_1 = 0.9$ and $c_2 = 5$.

As Table 3 shows, the exclusive buyer mechanism performs quite well relative to the numerical optimal solutions, at least for the case in which F is uniform. Moreover, as Figure 2 illustrates, the difference between the U function shown in Figure 1 and the buyer surplus function in the optimal exclusive buyer mechanism, for a grid of size $T = 30$, is relatively small. (The largest deviations occur near the "manifold" of nondifferentiable points of U .)

It is interesting to contrast the nature of the exclusive buyer mechanism with the qualitative features of the optimal mechanisms in the single-buyer models, studied in Armstrong (1996), Rochet and Choné (1998), and Manelli and Vincent (2007). A main result in Armstrong (1996) is that, in any optimal mechanism, it is optimal not to sell with positive probability. In contrast, in the optimal exclusive buyer mechanism with no reserve price, which performs quite well relative to the numerical optimal mechanisms in our numerical simulations, the object is always sold. Rochet and Choné (1998) establish that

bunching is robust in these multidimensional screening problems, even with very regular distributions of types. This comes from a strong conflict between participation constraints and second order incentive compatibility conditions. (Rochet and Choné 1998, p. 783)

Manelli and Vincent (2007) reinforce the point that, in one-buyer settings, optimal mechanisms are generically

Table 3. Optimal values of (P_T) and the objective function value of the exclusive buyer mechanism (*obtained by numerical integration with $T = 1,000$).

J	N	T	OPT VAL
2	2	5	6.051008
2	2	10	5.946440
2	2	20	5.896710
2	2	30	5.880743
2	2	1,000*	5.838323

Figure 2. The difference between the buyer surplus function in (P_T) , for $T = 30$, and the buyer surplus function in the optimal exclusive buyer mechanism.

quite different in nature from the simple “take-it-or-leave-it” mechanisms, which are always optimal in the one-dimensional case:

We find that the set of extreme points contains, in addition to price-posting, many “novel” mechanisms. In particular, extreme points need not be simple functions (Example 2), and even when they are, they may randomly assign objects to consumers (Examples 1 and 3). In contrast to the one-good case, the form of the optimal mechanism is determined by the prior distribution of buyer valuations. (Manelli and Vincent 2007, p. 154)

In contrast, our analysis suggests that the exclusive buyer mechanism with an optimally chosen “add-on” price can perform quite well, and shares many of its defining features with its one-dimensional counterparts, i.e., Myerson’s optimal auctions: there is no “bunching,” and it is defined by simple rules that do not depend on the prior distribution of buyer valuations.

Appendix A. Auxiliary Technical Results

LEMMA 13. *The inequalities in (C) are equivalent to the inequalities in (B).*

PROOF. Using our notation, Proposition 3.1 in Border (1991) states: Let $X: V \rightarrow [0, 1]$ be measurable. Then X is implementable by a symmetric auction if and only if for each measurable set $A \subset T$, the following inequality is satisfied:

$$\int_A X(v) dF(v) \leq \frac{1 - (\int_{V \setminus A} dF(v))^N}{N}.$$

By “ X is implementable,” Border means that there exists a symmetric function $x = (x^1, \dots, x^N): V^N \rightarrow [0, 1]^N$ such that

$$X(v^i) \equiv \int_{V^{N-1}} x^i(v^i, v^{-i}) \prod_{k \in N \setminus \{i\}} dF(v^k).$$

The equivalence between (C) and (B) follows immediately from Border’s Proposition 3.1, by letting $x^i(\omega) = \sum_{j \in J} q_j^i(\omega)$ and $X(v^i) = \sum_{j \in J} Q_j(v^i)$. $\square$

LEMMA 14. *Let h be a Lipschitz function on V , with constant L , and let*

$$\hat{h}(v) = \{h(\hat{v}): \text{for some } \hat{v} \text{ such that } \|v - \hat{v}\| \leq \varepsilon\}$$

be a piecewise-constant function. For any measurable function $g: V \rightarrow \mathbb{R}$, we have that

$$\int_D |h(v) - \hat{h}(v)| g(v) dv \leq L\varepsilon \int_D |g(v)| dv \quad \text{for any } D \subset V.$$

PROOF. Because h is Lipschitz, by definition of $\hat{h}$ we have that $|h(v) - \hat{h}(v)| \leq L\varepsilon$. The result follows by integration. $\square$

LEMMA 15. *For some $T > 0$, consider the grid V_T defined by $\varepsilon = O(1/T)$. Let $\tilde{U}^T$ be defined as in (1). For any $v \in V$, we have that*

$$0 \leq \tilde{U}^T(v) - \tilde{U}^T(v') \leq \varepsilon, \quad v' \in V_T, v' \leq v < v' + \varepsilon.$$

PROOF. Note that $\tilde{U}^T$ is a convex function. By its definition, any subgradient of $\tilde{U}^T$ belongs to the set $\text{conv}\{Q^T(v): v \in V_T\}$. It follows that any subgradient $s \in \text{conv}\{Q^T(v): v \in V_T\}$ is such that $s \geq 0$ and

$$\begin{aligned} \|s\|_1 &\leq \max_{\alpha_v \geq 0, \sum_{v \in V_T} \alpha_v = 1} \left\| \sum_{v \in V_T} \alpha_v Q^T(v) \right\|_1 \\ &\leq \sum_{v \in V_T} \alpha_v \sup_{v' \in V_T} \|Q^T(v)\|_1 \leq 1. \end{aligned}$$

Thus,

$$\begin{aligned} \tilde{U}^T(v) - \tilde{U}^T(v') &\leq \sup_{\alpha \in [0, 1], s \in \partial \tilde{U}^T(\alpha v + (1-\alpha)v')} |\langle s, v - v' \rangle| \\ &\leq \sup_{\alpha \in [0, 1], s \in \partial \tilde{U}^T(\alpha v + (1-\alpha)v')} \|s\|_1 \|v - v'\|_\infty \leq \varepsilon. \end{aligned}$$

Moreover, because $s \geq 0$ and $v - v' \geq 0$, we have

$$\tilde{U}^T(v) \geq \tilde{U}^T(v') + \langle s, v - v' \rangle \geq \tilde{U}^T(v'). \quad \square$$

LEMMA 16. *Under Assumption 2, we have that*

$$\sum_{v \in V_T} \left| \hat{f}(v) - \int_{v \leq v' \leq v + \varepsilon} f(v') dv' \right| \leq 2\varepsilon \sqrt{J} \sup_{v \in V} \|\nabla f(v)\| \text{vol}(V).$$

PROOF. For convenience, let $C_v = \{v' \in V: v \leq v' < v + \varepsilon\}$ where $\varepsilon = O(1/T)$. By the mean value theorem, $\int_{C_v} f(v') dv' = f(\tilde{v})\varepsilon^J$ where $\tilde{v} \in C_v$. The triangular inequality yields

$$\begin{aligned} \sum_{v \in V_T} \left| \hat{f}(v) - \int_{C_v} f(v') dv' \right| &\leq \sum_{v \in V_T} \left(|\hat{f}(v) - f(v)\varepsilon^J| + |\varepsilon^J f(v) - \int_{C_v} f(v') dv'| \right) \\ &\leq \sum_{v \in V_T} (|\hat{f}(v) - f(v)\varepsilon^J| + \varepsilon^J |f(v) - f(\tilde{v})|). \end{aligned}$$

Note that for the first term we have

$$\begin{aligned} \sum_{v \in V_T} |\hat{f}(v) - f(v)\epsilon^J| &= \sum_{v \in V_T} \frac{f(v)}{\sum_{v' \in V_T} f(v')} \left| 1 - \epsilon^J \sum_{v' \in V_T} f(v') \right| \\ &= \left| 1 - \sum_{v' \in V_T} f(v')\epsilon^J \right| \\ &= \left| \sum_{v' \in V_T} f(\tilde{v}')\epsilon^J - \sum_{v' \in V_T} f(v')\epsilon^J \right| \\ &\leq \sum_{v' \in V_T} \epsilon^J |f(\tilde{v}') - f(v')|. \end{aligned}$$

Finally, we bound

$$\begin{aligned} \sum_{v \in V_T} \epsilon^J |f(v) - f(\tilde{v})| &\leq \sum_{v \in V_T} \epsilon^J \sup_{v' \in C_v} |\langle \nabla f(v'), v - \tilde{v} \rangle| \\ &\leq \epsilon \sqrt{J} \sup_{v \in V} \|\nabla f(v)\| \text{vol}(V). \quad \square \end{aligned}$$

Appendix B. Proofs of §3

PROOF OF THEOREM 4. Within this proof we omit the superscript T for notational convenience. First note that $\tilde{U}$ is convex (because it is the maximum of affine functions) and nonnegative in V by construction because $\tilde{U}(v) \geq U(v) = 0$.

To show (ii), consider any two points $v, v' \in V$ and denote by $\hat{v}, \hat{v}' \in V_T$ the points that $\hat{v} \leq v < \epsilon e$, and $\hat{v}' \leq v' < \epsilon e$. Next, recall that all (IIC) are satisfied for all $\hat{v}, \hat{v}' \in V_T$. Because $\tilde{Q}(v) = Q(\hat{v})$, $\tilde{U}(\hat{v}) = U(\hat{v})$, $\tilde{U}(\hat{v}') = U(\hat{v}')$, we have

$$\begin{aligned} \tilde{U}(v') - \tilde{U}(v) - \langle \tilde{Q}(v), v' - v \rangle &= \tilde{U}(v') - U(\hat{v}') + U(\hat{v}) - \tilde{U}(v) \\ &\quad + U(\hat{v}') - U(\hat{v}) - \langle Q(\hat{v}), v' - v \rangle \\ &= -\epsilon + \langle Q(\hat{v}), v' - \hat{v}' + \hat{v} - v \rangle + U(\hat{v}') \\ &\quad - U(\hat{v}) - \langle Q(\hat{v}), \hat{v}' - \hat{v} \rangle \\ &= -\epsilon - \|Q(\hat{v})\|_1 \|\hat{v} - v\|_\infty \geq -2\epsilon, \end{aligned}$$

by Lemma 15, $\tilde{U}(v') \geq \tilde{U}(\hat{v}')$, $\langle Q(\hat{v}), v' - \hat{v}' \rangle \geq 0$, and $\|Q(\hat{v})\|_1 \leq 1$. Thus, $\tilde{Q}(v) \in \partial_{2\epsilon} \tilde{U}(v)$.

Because violations to (IIC) are bounded by (ii), to show (iii) we need to control the violation among all Border constraints. By Lemma 9 it suffices to consider sets of the form $E_\alpha = \{v \in V: \sum_{j=1}^J \tilde{Q}_j(v) \geq \alpha\}$. The definition of $\tilde{Q}$ implies that these sets are of the form unions of cubes of side ϵ (and therefore volume ϵ^J). Let $A = \bigcup_{v \in A_T \subset V_T} \{v' \in V: v \leq v' < v + \epsilon e\}$, then we have

$$\begin{aligned} \left| \int_A \sum_{j \in J} \tilde{Q}_j(v) dF(v) - \sum_{v \in A_T} \sum_{j \in J} Q_j(v) \hat{f}(v) \right| &\leq \sum_{v \in A_T} \sum_{j \in J} Q_j(v) \left| \int_{v \leq v' < v + \epsilon e} f(v') dv' - \hat{f}(v) \right| \end{aligned}$$

$$\begin{aligned} &\leq \sum_{v \in V_T} \left| \int_{v \leq v' < v + \epsilon e} f(v') dv' - \hat{f}(v) \right| \\ &\leq 2\epsilon \sqrt{J} \sup_{v \in V} \|\nabla f(v)\| \text{vol}(V) \end{aligned}$$

by Lemma 16. A similar bound applies to the right-hand side of the Border constraint. Therefore, the maximum Border violation is bounded by $O(\epsilon) = O(1/T)$. $\square$

PROOF OF LEMMA 5. Let $\alpha = \int_A dF(v)$. First note that $N \int_A \sum_{j \in J} Q_j(v) dF(v) \leq N\alpha$. This implies that $N\alpha \geq 1 - (1 - \alpha)^N + \eta$.

Using

$$(1 - \alpha)^N \leq e^{-N\alpha} \leq 1 - N\alpha + \frac{N^2 \alpha^2}{2} \quad \text{for } \alpha \in (0, 1),$$

we have that

$$N\alpha \geq 1 - 1 + N\alpha - \frac{N^2 \alpha^2}{2} + \eta = N\alpha - \frac{N^2 \alpha^2}{2} + \eta.$$

This implies $\alpha \geq \sqrt{2\eta}/N$. $\square$

PROOF OF COROLLARY 6. Consider an arbitrary violated Border constraint such that

$$N \int_A \sum_{j \in J} Q_j(v) dF(v) = 1 - \left(1 - \int_A dF(v)\right)^N + \eta,$$

where we have $\eta \leq \delta(Q, U)$ by definition of δ . For notational convenience, let $\alpha = \int_A dF(v)$ denote the probability measure of A .

Next, consider scaling the mappings (Q, U) by

$$\begin{aligned} \frac{1 - (1 - \alpha)^N}{1 - (1 - \alpha)^N + \eta} &= \frac{1}{1 + \eta/(1 - (1 - \alpha)^N)} \\ &\geq 1 - \frac{\eta}{1 - (1 - \alpha)^N}, \end{aligned}$$

which would make it satisfy the Border constraints for this particular set A .

By Lemma 5 we have that $\alpha \geq \sqrt{2\eta}/N$, which implies that

$$1 - (1 - \alpha)^N \geq N\alpha - \frac{N^2 \alpha^2}{2} \geq \sqrt{2\eta} - \eta.$$

This yields a bound on the scaling

$$\frac{1 - (1 - \alpha)^N}{1 - (1 - \alpha)^N + \eta} \geq 1 - \frac{\eta}{\sqrt{2\eta} - \eta} = 1 - \frac{\sqrt{\eta}}{\sqrt{2} - \sqrt{\eta}}.$$

Because this last quantity is monotone in $\eta \leq \delta(Q, U)$, it suffices to scale (Q, U) by

$$t := 1 - \frac{\sqrt{\delta(Q, U)}}{\sqrt{2} - \sqrt{\delta(Q, U)}}$$

so that (tQ, tU) satisfies all Border constraints.

Regarding (IIC) if for every v, v' in the domain we have

$$U(v) - U(v') \geq \langle Q(v'), v - v' \rangle - \delta(Q, U),$$

because $t \geq 0$ it holds that

$$tU(v) - tU(v') \geq \langle tQ(v'), v - v' \rangle - t\delta(Q, U).$$

Finally, for (IIR) note that if $U(\underline{v}) \geq 0$ we have $tU(\underline{v}) \geq 0$. On the other hand, if $U(\underline{v}) < 0$, multiplying by t reduces the violation as claimed.

Therefore, $\delta(tQ, tU) \leq t\delta(Q, U)$ and all Border constraints are satisfied by (tQ, tU) . $\square$

PROOF OF THEOREM 7. Let $(\hat{Q}^*, \hat{U}^*)$ denote the restriction of (Q^*, U^*) to V_T and $(\tilde{Q}^*, \tilde{U}^*)$ the extension of $(\hat{Q}^*, \hat{U}^*)$ to V (as in (1)). First note that Lemma 5 and Corollary 6 can be applied to $(\hat{Q}^*, \hat{U}^*)$ in order to obtain a feasible solution for (P_T) . Let $t = 1 - (\sqrt{\delta_T(\hat{Q}^*, \hat{U}^*)}/(\sqrt{2} - \sqrt{\delta_T(\hat{Q}^*, \hat{U}^*)}))$ so that $(t\hat{Q}^*, t\hat{U}^*)$ is feasible for (P_T) because all (IIC) are already satisfied.

In order to bound $\delta_T(\hat{Q}^*, \hat{U}^*)$, note that by convexity for any $h \geq 0$ we have

$$\frac{U^*(v) - U^*(v - he)}{h} \leq \sum_{j \in J} Q_j^*(v) \leq \frac{U^*(v + he) - U^*(v)}{h}$$

for all $v \in V$,

where e denotes the vector of all ones. Next we apply Lemma 9 to restrict attention to the class of Border constraints with sets E_α . Without loss of generality, we can restrict E_α to be a compact set.

Once again, by convexity we have that $\langle Q^*(\hat{v}) - Q^*(v), \hat{v} - v \rangle \geq 0$. Applying this to $\hat{v} = v + he$ for $h \geq 0$ we have that $\sum_{j \in J} Q_j^*(v) \leq \sum_{j \in J} Q_j^*(v + he)$. Therefore, if $v \in E_\alpha$ we have that $v + he \in E_\alpha$ for any $h \geq 0$ such that $v + he \in V$.

Define $\mathcal{L}(E_\alpha) = \{v \in E_\alpha : \nexists h > 0 \text{ such that } v - he \in E_\alpha\}$. For each $v \in \mathcal{L}(E_\alpha)$, let $h(v) = \max_{h \geq 0} \{h : v + he \in V\}$. Therefore, we can rewrite Border constraint as

$$\begin{aligned} & \int_{E_\alpha} \sum_{j \in J} Q_j^*(v) f(v) dv \\ &= \int_{\mathcal{L}(E_\alpha)} \int_0^{h(v)} \sum_{j \in J} Q_j^*(v + he) f(v + hv) dh dv \\ &= \int_{\mathcal{L}(E_\alpha)} \left[U^*(v + he) f(v + he) \Big|_0^{h(v)} \right. \\ & \quad \left. - \int_0^{h(v)} U^*(v + he) e' \nabla f(v + he) dh \right] dv \\ &= \int_{\mathcal{L}(E_\alpha)} \left[\tilde{U}^*(v + he) f(v + he) \Big|_0^{h(v)} \right. \\ & \quad \left. - \int_0^{h(v)} \tilde{U}^*(v + he) e' \nabla f(v + he) dh \right] dv + O(1/T) \end{aligned}$$

by Lemma 14 because U^* is Lipschitz, V is compact, and f has bounded first and second derivatives. Convergence is uniform over all sets E_α because U is Lipschitz (constant at most 1), V is compact, by Lemma 14. Combining this result with Lemma 16 yields the rate of convergence $\delta_T(Q^*, U^*) = O(1/T)$ and

$$\begin{aligned} OPT_T &\geq \sum_{v \in V_T} (\langle t\hat{Q}^*(v), v \rangle - t\hat{U}^*(v)) \hat{f}(v) \\ &= \int_V (\langle t\tilde{Q}^*(v), v \rangle - t\tilde{U}^*(v)) dF(v) + O(t/T) \\ &= t \int_V \langle Q^*(v), v \rangle - U^*(v) dF(v) + O(t/T) \\ &= t OPT_* + O(t/T). \quad \square \end{aligned}$$

Appendix C. Proofs of §4

PROOF OF LEMMA 9. We will show that for any measurable set $A \subseteq V$ there is a set of the form E_α whose violation of Border is at least as large as the violation associated with A . Let α be the largest value such that

$$\int_{E_\alpha} dF(v) \geq \int_A dF(v),$$

which exists because the mapping $\alpha \mapsto \int_{E_\alpha} dF(v)$ is continuous from the left. We have that

$$\begin{aligned} & \int_A \left[\sum_{j \in J} Q_j(v) \right] dF(v) \\ &= \int_{A \setminus E_\alpha} \left[\sum_{j \in J} Q_j(v) \right] dF(v) + \int_{A \cap E_\alpha} \left[\sum_{j \in J} Q_j(v) \right] dF(v) \\ &\leq \int_{A \setminus E_\alpha} \alpha dF(v) + \int_{A \cap E_\alpha} \left[\sum_{j \in J} Q_j(v) \right] dF(v) \\ &= \alpha \int_{A \setminus E_\alpha} dF(v) + \int_{A \cap E_\alpha} \left[\sum_{j \in J} Q_j(v) \right] dF(v). \end{aligned}$$

Note that if we can take $\int_{E_\alpha} dF(v) = \int_A dF(v)$, we also have $\int_{E_\alpha \setminus A} dF(v) = \int_{A \setminus E_\alpha} dF(v)$ and E_α leads to a larger violation than A (because the respective right-hand sides are equal). Therefore, we can assume $\int_{E_\alpha} dF(v) > \int_A dF(v)$. In turn, this implies that

$$H_\alpha = \left\{ v \in V : \sum_{j \in J} Q_j(v) = \alpha \right\}$$

has positive measure with respect to F , that is,

$$\int_{H_\alpha} dF(v) > \int_{E_\alpha} dF(v) - \int_A dF(v) > 0.$$

(Otherwise, we would be able to choose a larger α .)

Consider the violation of the Border constraint associated with A

$$\delta_A := \int_A \left[\sum_{j \in J} Q_j(v) \right] dF(v) - 1 + \left(1 - \int_A dF(v) \right)^N \quad \text{and}$$

$$\delta(\mu) := \alpha\mu + \int_{E_\alpha \setminus H_\alpha} \left[\sum_{j \in J} Q_j(v) \right] dF(v) - 1 + \left(1 - \int_{E_\alpha \setminus H_\alpha} dF(v) - \mu \right)^N,$$

where $\mu = \int_H dF(v)$ with $H \subset H_\alpha$. The latter is a convex function in the scalar $\mu \in [0, \int_{H_\alpha} dF(v)]$ so that it is maximized at one of the extremes (i.e., the extremes have a violation at least as big as δ_A). Note that there might be no set H associated with intermediate values of μ (say $\mu = \int_A dF(v) - \int_{E_\alpha \setminus H_\alpha} dF(v)$), but the maximum is achieved in the extreme. This allows us to obtain at least the same violation as δ_A , either with $H = \emptyset$ or $H = H_\alpha$. These cases correspond to sets of the form E_α . $\square$

PROOF OF LEMMA 10. Regarding the separation oracle for (HC), note that we can simply enumerate them. Each of the $O(T^{2J})$ constraints requires $O(J)$ operations.

Next we turn to Border constraints. Let (Q, U) be the point to be considered by the oracle. First construct the $O(T^J)$ -dimensional vector $\bar{Q}$ defined as

$$\bar{Q}(v) = \sum_{j \in J} Q_j(v),$$

which accounts for $O(JT^J)$ operations. Then we sort the vector $\bar{Q}$ in descending order using $O(T^J \ln(T^J))$ operations to create $\hat{Q}$. Next note that we can generate all the $O(T^J)$ sets E_α associated with Q by considering the following sets of components

$$\{\{1\}, \{1, 2\}, \{1, 2, 3\}, \dots, \{1, 2, \dots, O(T^J)\}\}$$

of $\hat{Q}$. Because this can be done incrementally, the computational cost to search over all these sets is $O(T^J)$. $\square$

PROOF OF THEOREM 12. Let S^k denote the set of indices of included Border and IC constraints at the k th iteration. Because there are a finite number of constraints, there is a finite number of possible sets S^k , each of them containing a finite number of extreme points or faces. Therefore, there is only a finite number of different objective function values that can be achieved at any iteration. Because inequalities are discarded only if an improvement on the objective function is observed, either we keep accumulating inequalities or a improvement in the objective function is made. $\square$

Endnotes

1. In mechanism design terminology, “ex post” variables depend on all buyers’ privately known types; whereas “interim” variables are obtained after integrating out all but one buyers’ types.

2. A mechanism $\{q^i, m^i; i \in N\}$ is *symmetric* if

$$q^i(v) = q^{i'}(\tilde{v}) \quad \text{and} \quad m^i(v) = m^{i'}(\tilde{v})$$

for all $i, i' \in N$, $v, \tilde{v} \in V$, such that $v^i = \tilde{v}^{i'}$, $\tilde{v}^i = v^{i'}$, and $\tilde{v}^k = v^k$ $k \neq i, i'$. If (q, m) is optimal and asymmetric, then its “mirror image” (q', m') , obtained by permuting buyers, is also optimal. Because the feasible set is convex, and the objective is linear, the mechanism $1/2(q, m) + 1/2(q', m')$ is also optimal, and symmetric. The above argument mimics the steps in Maskin and Riley (1986, footnote 11).

3. Following standard notation in convex analysis (Rockafellar 1970), we denote the subdifferential of a convex function $U: V \rightarrow \mathbb{R}$ at $v \in V$ by $\partial U(v) = \{s: U(\hat{v}) \geq U(v) + \langle s, \hat{v} - v \rangle \text{ for all } \hat{v} \in \mathbb{R}^n\}$.

4. For the case with $J > 2$ quality levels, the winner of the auction selects a price-quality pair from a menu of up to J choices.

Acknowledgments

The authors thank the National Science Foundation for support under grant SES-0849349. They gratefully acknowledge the excellent research assistance by Peter Franklin and useful comments by Leslie Marx and Peng Sun.

References

- Armstrong, M. 1996. Multiproduct nonlinear pricing. *Econometrica* **64**(1) 51–75.
- Border, K. 1991. Implementation of reduced form auctions: A geometric approach. *Econometrica* **59** 1175–1187.
- Brusco, S., G. Lopomo, L. Marx. 2008. The economics of contingent re-auctions. Working paper, Duke University, Durham, NC.
- Grötschel, M., L. Lovász, A. Schrijver. 1981. *Geometric Algorithms and Combinatorial Optimization*. Springer-Verlag, New York.
- Grötschel, M., L. Lovász, A. Schrijver. 1988. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica* **1** 169–197.
- Hiriart-Urruty, J.-B., C. Lemaréchal. 1993. *Convex Analysis and Minimization Algorithms*. Number 305–306, Grundlehren der math. Wiss. Springer-Verlag, Berlin (two volumes).
- Manelli, A., D. Vincent. 2007. Multidimensional mechanism design: Revenue maximization and the multiple-good monopoly. *J. Econom. Theory* **137** 153–185.
- Maskin, E., J. Riley. 1984. Monopoly with incomplete information. *RAND J. Econom.* **15**(2) 171–196.
- Maskin, E., J. Riley. 1986. Optimal auctions with risk averse buyers. *Econometrica* **52**(6) 1473–1518.
- Mussa, M., S. Rosen. 1978. Monopoly and product quality. *J. Econom. Theory* **18** 301–317.
- Myerson, R. 1979. Incentive compatibility and the bargaining problem. *Econometrica* **47** 61–73.
- Myerson, R. 1981. Optimal auction design. *Math. Oper. Res.* **6** 58–73.
- Padberg, M. W., M. R. Rao. 1980. The Russian method and integer programming. Working paper, New York University, New York.
- Rochet, J., P. Choné. 1998. Ironing, sweeping, and multidimensional screening. *Econometrica* **66**(4) 783–826.
- Rockafellar, R. T. 1970. *Convex Analysis*. Princeton University Press, Princeton, NJ.
- Wilson, R. 1993. *Nonlinear Pricing*. Oxford University Press, Oxford, UK.