On the Planner's Loss Due to Lack of Information in Bayesian Mechanism Design*

José R. Correa and Nicolás Figueroa

CEA, Departamento de Ingeniería Industrial, Universidad de Chile
{jcorrea,nicolasf}@dii.uchile.cl

Abstract. In this paper we study a large class of resource allocation problems with an important complication, the utilization cost of a given resource is private information of a profit maximizing agent. After reviewing the characterization of the optimal bayesian mechanism, we study the informational cost introduced by the presence of private information. Our main result is to provide an upper bound for the ratio between the cost under asymmetric information and the cost of a fully informed designer, which is independent of the combinatorial nature of the problem and only depend on the statistical distribution of the resource costs. In particular our bounds evaluates to 2 when the utilization cost's distributions are symmetric and unimodal and this is tight. We also show that this bound holds for a variation of the Vickrey-Clark-Groves mechanism, which always achieves an ex-post efficient allocation. Finally we point out implementation issues of the considered mechanisms.

Keywords: Mechanism Design, Information Cost.

JEL Classification: C60, C72, D44.

1 Introduction

A wide class of problems of the form $\min\{c^t x | x \in \Gamma\}$ have been analyzed in the literature and their applications to real world problems are vast. In this paper, we consider such a class of problems with an important and realistic complication, the utilization cost of a given resource x_i is private information of a profit maximizing agent.

For example, let us consider a natural situation in supply chain management. A large company needs to procure quantities D_i of a given good for its various locations $t_1, \dots, t_k$. The good is produced at various locations $s_1 \dots, s_l$, each of them with a maximum production capacity Q_j. The delivery of the goods is done through a transportation network in which each link has a cost that is publicly known. This problem, when the production facilities are owned by the company, reduces to a standard minimum cost flow through a network. If, however, the production facilities are owned by private contractors, whose production cost is private information, there is an added layer of difficulty to the problem. Now the company must design a mechanism to minimize

* This work was partially supported by FONDECYT through grants 11070159 and 1090050 and by the Instituto Milenio en Sistemas Complejos de Ingeniería.

M. Mavronicolas and V.G. Papadopoulou (Eds.): SAGT 2009, LNCS 5814, pp. 72–84, 2009.

expected procurement cost, subject to the feasibility constraints on the network, and inducing contractors to reveal their costs in exchange of a profit.

The main contribution of this paper is to study the informational cost introduced by the presence of private information. Such a consideration is important, because it lies at the heart of an old economic question: To make or to buy? If this cost is small, the organizational cost of acquiring small producers may be high and not worth it. If big, such an acquisition may turn out to be profitable for the company trying to procure goods or services. Our main result is to provide an upper bound for the ratio between the cost under asymmetric information and the cost of a fully informed designer. Specifically, we show that for a large class of distributions, containing those that are symmetric and unimodal, the expected cost of an optimal mechanism is at most twice the cost of an optimal solution obtained by a fully informed planner. Neither this bound nor its tightness depend on the combinatorial nature of the problem, but only on the statistical distribution of private information. The latter bound holds for a variation of the Vickrey-Clark-Groves (VCG) mechanism as well, and becomes significantly better in some special situations.

Related questions were studied by Bulow and Klemperer [2], who analyze the suboptimality, in terms of revenue, of VCG for a single unit auction. They show that one extra bidder in a VCG format gives more revenue than the Myerson mechanism. Recently, Aggarwal et al. [1] study the suboptimality, in terms of efficiency, of the Myerson auction, showing that $\Theta(\log k)$ extra bidders suffice to match the efficiency of a VCG mechanism with k bidders, and generalize the result to multiunit auctions. Also, Elkind et al. [4] establish bounds on the payments of the VCG and optimal mechanisms in path auctions, and point out that these may differ significantly. Finally, Hartline and Roughgarden [5] consider similar issues in the context of money burning mechanisms.

We also study the computational cost of calculating an optimal mechanism. We show that such a problem is equivalent to performing parametric linear programming over the set Γ, which is in general of exponential complexity, even if optimization over Γ is simple. For the important class of problems where the set Γ is a 0-1 polytope, however, we give a simple algorithm with the same complexity of the original optimization problem with complete information. For the other problems, we point out that a simple sampling technique, which takes advantage of the owners' risk neutrality, gives a random mechanism yielding the same expected cost as the deterministic one.

The paper is organized as follows. In section 2 we quickly review the characterization of the optimal bayesian (i.e., utilization costs are random variables) mechanism for the whole class of problems with a linear cost function and a fixed constraint set. Our main results concerning the informational cost are found in section 3, while the computational considerations are discussed in section 4.

2 The Model

2.1 The Environment

We consider a setting in which scarce resources must be allocated to carry out a given project. The cost of each resource may be public or private information. Depending on the situation, the planner's goal is to minimize her own expected cost, or the social cost of the project. To this end, she can design a mechanism where the owners with private information have incentives to reveal their private information.

In our framework each resource $a \in A$ is represented by a variable x_a, and is associated with a marginal cost of utilization c_a. The set A is partitioned into two sets A_1 and A_2. Costs of resources $a \in A_1$ are private, and thus c_a is private information and is distributed according to F_a, whose bounded support is the interval $[\underline{c}_a, \overline{c}_a] \subset \mathbb{R}_+$. The distribution F_a is assumed to have a density f_a which is continuous and strictly positive in $[\underline{c}_a, \overline{c}_a]$. For simplicity we also assume that $F_a(c_a)/f_a(c_a)$ is nondecreasing (satisfied among others by the family of logconcave distributions). Costs of resources $a \in A_2$ are public information and equal c_a. Resources are scarce and subject to an exogenous feasibility constraint $x \in \Gamma \subseteq \mathbb{R}^{|A|}$, which we assume compact. Therefore, if all costs c_a, $a \in A$ were known, the planner would solve $\min\{c^T x : x \in \Gamma\}$. However, costs of resources in A_1 are unknown and thus the planner must design a mechanism to elicit this information in order to achieve her goal.

We now give a key property that holds in this environment. It states that if the cost of a resource increases, the value of the corresponding variable, in a cost-minimizing solution, does not increase. This intuitive and simple result turns out to be critical for characterizing the optimal mechanism.

Lemma 1. *Let $x(c) = \{x_a(c)\}_{a \in A}$ be the minimum cost assignment in Γ for a cost vector c. Then $x_a(\cdot, c_{-a})$ is non-increasing for all $a \in A$.*

Proof. Consider a cost vector c and let c' be defined as $c'_e = c_e$ for all $e \in A - \{a\}$ and $c'_a = c_a + \varepsilon$ for some $\varepsilon > 0$. From the definition of $x(c)$ we have that: $c^T x(c) \leq c^T x(c')$ and ${c'}^T x(c') \leq {c'}^T x(c)$. Summing both terms we obtain $(c^T - {c'}^T)[x(c) - x(c')] \leq 0$ which is equivalent to $x_a(c) \leq x_a(c')$. □

As usual, if $x \in \mathbb{R}^n$, x_{-i} denotes the vector in which the $i-$th component is removed. We also define: $f(c) = \prod_{a \in A_1} f_a(c)$, $f_{-a}(c) = \prod_{e \in A_1 - \{a\}} f_a(c)$, $C = \prod\limits_{a \in A_1} [\underline{c}_a, \overline{c}_a]$, $C_{-a} = \prod_{e \in A_1 - \{a\}} [\underline{c}_a, \overline{c}_a]$.

2.2 Mechanisms

In order to achieve her objectives, the planner designs a mechanism. In other words, the planner chooses a message space M_a for each $a \in A_1$, together with assignment and payment rules. Given messages from resource owners, these rules determine the amount of each resource used by the planner and the payment received by each owner. Due to the revelation principle it is enough to set $M_a = [\underline{c}_a, \overline{c}_a]$ and consider truthful mechanisms.

Therefore, a mechanism is given by assignment rules $\{x_a\}_{a \in A}$, indicating how much of resource a will be used, and a family of payment rules $\{t_a\}_{a \in A_1}$, indicating the total payment to the owner of resource $a \in A_1$. Naturally, these values depend on the cost revelations of each owner, therefore $x_a : C \longrightarrow \mathbb{R}$ and $t_a : C \longrightarrow \mathbb{R}$. Our framework allows the payment received by the owner of resource a, given revelations c, to be random. If this is the case, t_a denotes the total expected payment to the owner of resource $a \in A_1$. The payoff of the owner of resource a, with cost c_a, when reporting a cost c'_a is given by:

$$U_a(c_a, c'_a) = \int_{C_{-a}} [t_a(c'_a, c_{-a}) - c_a x_a(c'_a, c_{-a})] f_{-a}(c_{-a}) dc_{-a}. \tag{1}$$

The payoff of a resource owner with cost c_a is then:

$$V_a(c_a) = \max_{c'_a \in C_a} U_a(c_a, c'_a). \tag{2}$$

We must also consider mechanisms that give a positive utility to owners and satisfy the feasibility constraints. We can thus give the following definition.

Definition 1. *A mechanism $(x, t) \equiv (\{x_a\}_{a \in A}, \{t_a\}_{a \in A_1})$ is feasible if and only if for all cost realizations c the following hold:*

$$\begin{aligned} &\text{(IC)}\ \ V_a(c_a) = U_a(c_a, c_a)\ \textit{for all}\ a \in A_1, \\ &\text{(PC)}\ V_a(c_a) \geq 0 \qquad\qquad \textit{for all}\ a \in A_1, \\ &\text{(F)}\quad x(c) \in \Gamma. \end{aligned}$$

The Optimal Bayesian Mechanism. With the previous definition, we can write the problem of a cost-minimizing designer as

$$\min \left\{ \int_{c \in C} \left(\sum_{a \in A_1} t_a(c) + \sum_{a \in A_2} c_a x_a(c) \right) f(c) dc \ :\ (x, t) \text{ is feasible} \right\}. \tag{3}$$

Using by now standard arguments introduced by Myerson [9] and extended among others by Elkind et. al. [4] (see [10, Chapter 13] for a detailed treatment) we can characterize the optimal Bayesian mechanism relying on Lemma 1. A proof can be found in the full version of this paper. Indeed, the optimal mechanism can be written as the solution to the following control problem

$$\begin{aligned} &\min_{\{x_a(c)\}_{a \in A}} \int_{c \in C} \left(\sum_{a \in A_1} x_a(c) \left[c_a + \frac{F_a(c_a)}{f_a(c_a)} \right] + \sum_{a \in A_2} x_a(c) c_a \right) f(c) dc \\ &\text{s.t.} \qquad x(c) \in \Gamma \text{ and } \nu_a(c_a) \text{ non-increasing for all } a \in A_1. \end{aligned}$$

Here, $\nu_a(c_a) := \int_{C_{-a}} x_a(c_a, c_{-a}) f_{-a}(c_{-a}) dc_{-a}$ is the expected utilization of resource a for $a \in A_1$.

Because of Lemma 1, and the assumption that $F_a(c_a)/f_a(c_a)$ is increasing, we can relax the constraint asking for $\nu_a(c_a)$ non-increasing and solve the above problem pointwise to obtain a feasible solution. Therefore, we can characterize the optimal mechanism.

Proposition 1. *The optimal assignment rules $\bar{x}(c) = \{\bar{x}_a(c)\}_{a \in A}$ are those solving, for each cost revelations $\{c_a\}_{a \in A_1}$, the following optimization problem:*

$$\min_{y \in \Gamma} \sum_{a \in A_1} \left(c_a + \frac{F_a(c_a)}{f_a(c_a)} \right) y_a + \sum_{a \in A_2} c_a y_a,$$

and an optimal payment rule is given by $\bar{t}_a(c) = c_a \bar{x}_a(c) + \int_{c_a}^{\bar{c}_a} \bar{x}_a(t, c_{-a}) dt$. In other words, $x(c)$ is the minimum cost assignment in Γ with virtual costs $c'_a = c_a + F_a(c_a)/f_a(c_a)$ for all $a \in A_1$, and $c'_a = c_a$ for all $a \in A_2$.

The Truncated Vickrey-Clark-Groves Mechanism. On the other hand, a planner interested in achieving ex-post efficiency may consider the standard VCG mechanism, which solves, for every cost realization c, the problem $\min\{c^t x | x \in \Gamma\}$ and assigns according to the solution rule $x_a^V(c)$: It pays agent $a \in A_1$, $t_a(c) = c_a x_a(c) + (\sum_{b \in A} c_b x_b^{-a}(c) - \sum_{b \in A} c_b x_b(c))$, where $x^{-a}(c)$ is a solution of $\min\{c^t x | x \in \Gamma, x_a = 0\}$. It is well known that such a mechanism is incentive compatible, but can involve infinite costs. However, if the support of the cost distribution is known, payments can be bounded without losing incentive compatibility (and thus efficiency). We denote such a mechanism, with payments given by $t_a(c) = \min\{c_a x_a(c) + (\sum_{b \in A} c_b x_b^{-a}(c) - \sum_{b \in A} c_b x_b(c)), \bar{c}_a x_a(c)\}$, the Truncated Vickrey-Clark-Groves (TVCG) mechanism.

3 Loss Due to Lack of Information

The presence of private information among resource owners increases the cost of performing a given task. A natural problem, with relevant practical implications, is to quantify the relationship between the cost under complete and incomplete information. The former corresponds to a situation where the planner owns the different resources and the technology needed for their production, therefore knowing exactly the production costs. The latter corresponds to a decentralized situation, where the planner has outsourced the production of necessary inputs, and therefore does not know precisely their production costs. Since outsourcing can imply important savings in terms of managerial effort, it is critical to know how much is a firm losing by spinning off some of its components, or how much is a central planner losing by privatizing some key components of a planned economy. Moreover, with incomplete information, a cost-minimizing planner does not necessarily assign resources efficiently (since he considers modified costs), so we consider the question of the expected cost of an efficient mechanism, the TVCG, and its comparison to the cost-minimizing one and the fully informed solution.

Interestingly, both comparisons can be done independently of the combinatorial structure of the problem (given by the set Γ), and depend only on the nature of the incomplete information (given by the distribution functions F_a). The critical lemma is the following:

Lemma 2. *If the distribution F, with $F(a) = 0$ and density f, satisfies that $\mathbb{E}(X | X \leq y) \geq y/\alpha$, where X is drawn according to F, then for $[a, b] \subset \mathbb{R}_+$ and $g(\cdot)$ a nonnegative, non-increasing real-valued function defined on $[a, b]$ we have:*

$$\int_a^b g(c)F(c)dc \leq (\alpha - 1)\int_a^b g(c)cf(c)dc.$$

Proof. Let $g(\cdot)$ be any nonnegative non-increasing real-valued function and F be a distribution, with density f, satisfying the conditions in the proposition. Note that as $g(\cdot)$ is monotone, it is differentiable almost everywhere [7], thus $g'(c) \leq 0$ a.e., implying that

$$\int_a^b g(c)(F(c) - (\alpha-1)cf(c))dc = g(b)\int_a^b (F(s) - (\alpha-1)sf(s))ds \\ - \int_a^b g'(c)\int_a^c (F(s) - (\alpha-1)sf(s))dsdc,$$

is nonpositive if $\int_a^y F(c)dc \leq (\alpha-1)\int_a^y cf(c)dc$ holds for all $y \in [a, b]$. This latter condition is equivalent to $\mathbb{E}(X|\, X \leq y) \geq y/\alpha$, since integrating by parts

$$\int_a^y F(c)dc - (\alpha-1)\int_a^y cf(c)dc = yF(y) - \alpha\int_a^y cf(c)dc,$$

which is nonpositive so long as $\mathbb{E}(X|\, X \leq y) \geq y/\alpha$. □

3.1 Cost Loss Due to Lack of Information

We now turn compare the planner's expected cost when using the cost-minimizing and the TVCG mechanisms to that in case she had complete information. From the description in Section 2.2 (see full version for details), and noting that the worst type $\bar{c}_a$ gets 0 rents in both the cost-minimizing and the TVCG mechanism, we can write the expected cost of both mechanisms as:

$$\mathcal{C}_I = \min_{x(c)\in\Gamma} \int_{c\in C}\left(\sum_{a\in A_1} x_a(c)\left[c_a + \frac{F_a(c_a)}{f_a(c_a)}\right] + \sum_{a\in A_2} x_a(c)c_a\right) f(c)dc. \quad (4)$$

$$\mathcal{C}_{VCG} = \int \left[\sum_{a\in A_1} x_a^V(c)\left[c_a + \frac{F_a(c_a)}{f_a(c_a)}\right] + \sum_{a\in A_2} x_a^V(c)c_a\right] f(c)dc, \quad (5)$$

On the other hand, when complete information is available to the planner, her cost is given by:

$$\mathcal{C}_C = \min_{x(c)\in\Gamma} \sum_{a\in A_1}\int_{c\in C} c_a x_a(c)f(c)dc + \sum_{a\in A_2}\int_{c\in C} c_a x_a(c)f(c)dc. \quad (6)$$

Observe that if $A_1 = A$, that is all costs are private information, and F_a is uniform in $[0, s]$ for all $a \in A$, the planner's problem given by (4) is exactly the same as that in (6) with the costs doubled. Therefore the planner's expected cost in the optimal mechanism is twice as much as that in the complete information setting. Moreover, since the assignment rules of TVCG coincide with the fully informed solution, in this setting the cost of the TVCG mechanism is also twice $\mathcal{C}_I$. In what follows, we extend this result to a very general class of distribution functions, and prove that such a bound is also true for the comparison between the TVCG (which in general has a higher cost than $\mathcal{C}_C$) and the complete information mechanism.

With Lemma 2 at hand, the proof of the next result becomes remarkably simple. Its full significance though, will be evident in the next section, once we establish that large and natural classes of distributions satisfy the hypothesis.

Proposition 2. *If for all $a \in A_1$ the distribution F_a satisfies that $\mathbb{E}(X|\, X \leq y) \geq y/\alpha$, where X is drawn according to F_a, then $\mathcal{C}_I \leq \mathcal{C}_{VCG} \leq \alpha \cdot \mathcal{C}_C \leq \alpha \cdot \mathcal{C}_I$.*

Proof. The first and last inequalities are direct since we first compare the optimal mechanism to TVCG, and the fully informed optimal solution to an optimal mechanism. For the second one, we apply Lemma 2 to expression (5). Note that Lemma 2 holds even if the function $g(\cdot)$ is not continuous, as it may be the case for $x^V(\cdot)$, for instance, when the underlying set Γ is polyhedral or discrete. Thus we can write:

$$\begin{aligned} \mathcal{C}_{VCG} &= \int \left[\sum_{a \in A_1} x_a^V(c) \left[c_a + \frac{F_a(c_a)}{f_a(c_a)} \right] + \sum_{a \in A_2} x_a^V(c) c_a \right] f(c) dc \\ &\leq \alpha \int \sum_{a \in A} x_a^V(c) c_a f(c) dc = \alpha \mathcal{C}_C. \end{aligned}$$

The last equality holds since TVCG assigns efficiently. □

Note that the previous bound is related *only* to the distribution of private information about costs, and not to the particular problem Γ being considered. As we already pointed out, in *any* instance of a combinatorial problem defined by Γ, when all resources are private and the information is distributed uniformly on $[0, a]$, this bound is tight, since for a $\mathbb{E}(X|\, X \leq y) = y/2$.

Observation. A natural question is whether there is a better bound for the comparison between $\mathcal{C}_I$ and $\mathcal{C}_{VCG}$ than just $\mathcal{C}_{VCG} \leq \alpha \cdot \mathcal{C}_I$. If such a bound holds true when we consider the full information mechanism, is it possible to do better when considering the optimal mechanism under incomplete information? The answer is no, as sometimes the incomplete information planner has the same cost as the fully informed planner, while the TVCG mechanism performs badly at a cost $\alpha \cdot \mathcal{C}_C$. Consider for example the case where the planner must send one unit of flow between two nodes, in a two link network. One of the links is public while the other is private, i.e., $A_1 = \{a_1\}$, $A_2 = \{a_2\}$, and $\Gamma = \{(x_{a_1}, x_{a_2}) \geq 0 : x_{a_1} + x_{a_2} = 1\}$. Consider $\{F_{a_1}^{(n)}\}$ a family of symmetric and unimodal distributions for resource a_1, and assume that their support is the full interval $[0, 1]$ and that $F_{a_1}^{(n)} \longrightarrow \delta_{1/2}$, where $\delta_{1/2}$ is the mass distribution putting probability 1 to $c_{a_1} = 1/2$. Assume also that $c_{a_2} = 1$. Then, we have that $\mathcal{C}_{VCG}^{(n)} = 1$, but $\mathcal{C}_I^{(n)} \longrightarrow \mathcal{C}_C \equiv \frac{1}{2}$, while the value of α for these distributions is, as we will see next, 2. Therefore our bound is tight.

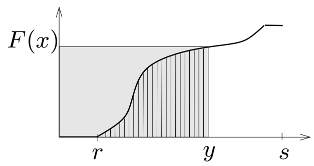

Fig. 1. For $\alpha = 2$ the condition of Proposition 2 states that the area under the curve is at most half of the gray area

3.2 Distributions

Having established that Proposition 2 holds independently of the combinatorial structure of the problem, the main question is thus to determine the distributions satisfying the hypothesis, and how small their corresponding value of α is. Note first that the proposition can be applied to densities which are non-decreasing with $\alpha = 2$. Therefore, for situations where agents are concentrated among "bad" providers, we can do as well as in the case with a uniform distribution.

Let us give a geometric interpretation of the inequality $\mathbb{E}(X| X \leq y) \geq y/\alpha$, where X is a random variable drawn according to a distribution F defined in an interval $[r, s]$. Writing down the expression and integrating by parts we note that the condition is equivalent to

$$(\alpha - 1)yF(y) \geq \alpha \int_r^y F(x)dx. \tag{7}$$

Thus the condition states that for any y in $[r, s]$ the area defined by the rectangle of width y and height $F(y)$ is at least a fraction $\alpha/(\alpha - 1)$ of the area comprised under the curve $F(x)$ between r and y. Figure 1 depicts the situation.

With the intuition provided by the above interpretation we are able to find a number of distributions for which Proposition 2 can be applied. A particularly relevant example occurs when the distribution which is a minimum between m draws of a uniform distribution. Here, we capture a situation where providers have a try at m different technologies and select the best of them. Such an environment is biased towards "good" providers through "natural selection", but even in this case we can provide a tight upper bound.

Proposition 3. *Consider agents whose cost is given by the minimum of m draws from a uniform distribution in $[0, 1]$. Their cost distribution is then given by $F(x) = 1 - (1 - x)^m$, and it satisfies $E(X|X \leq Y) \geq Y/(m + 1)$.*

Proof. Note first that, using condition 7, the inequality $E(X|X \leq y) \geq \frac{y}{m+1}$ is equivalent to

$$\max_{y\in[0,1]} \frac{\int_0^y (1 - (1 - s)^m)ds}{y(1 - (1 - y)^m)} \leq \frac{m}{m + 1},$$

which in turn can be rewritten as $y + (1 - y)^{m+1} + my(1 - y)^m \leq 1$ for all $y \in [0, 1]$. Further cancelations lead to $(1-y)^m + my(1-y)^{m-1} \leq 1$ for all $y \in [0, 1]$, and using the change of variables $s = 1-y$, we obtain that $G(s) = s^m + mys^{m-1} \leq 1$ for all $s \in$

$[0,1]$. Noting that $G'(s) = ms^{m-1} + m(m-1)s^{m-2} - m^2 s^{m-1} = 0$ implies $s = 1$, and that $G(0) = 0$, $G(1) = 1$, the result follows. □

Our result, as we show next, can also be applied to an important class of symmetric distributions, which in particular includes those that are symmetric and unimodal (SUD). A distribution function is unimodal if it has a unique local maximum.

Proposition 4. *Suppose that for all $a \in A_1$ the distribution F_a with density f_a has support $[0,1]$, is symmetric, and satisfies $F_a(y) \leq y f_a(y)$ for $0 \leq y \leq 1/2$. Then we have that $\mathcal{C}_I \leq \mathcal{C}_{VCG} \leq 2 \cdot \mathcal{C}_C$.*

Proof. Because of Proposition 2 we just need to show that if X is a random variable drawn from a symmetric distribution F, whose density f has support $[0,1]$, and satisfies $F(y) \leq y f(y)$ for $0 \leq y \leq 1/2$, then $E(X|X \leq y) \geq y/2$. Using condition 7 this is equivalent to showing

$$yF(y) \geq 2 \int_0^y F(x)dx \quad \text{ for all } y \in [0,1].$$

Note first that the inequality holds for $y = 0$. Furthermore, we know that for all $0 \leq y \leq 1/2$ we have $2F(y) \leq (F(y) + yf(y))$. The latter is equivalent to saying that the derivative of the left hand side of the condition above is larger than the derivative of the right hand side. Thus the condition holds for all $y \in [0, 1/2]$.

Furthermore, note that $F(y) \leq y$ for all $y \in [0, 1/2]$. Indeed, by contradiction assume that $F(z) > z$ for some $z \in [0, 1/2]$. In this case let $1/2 \geq z' > z$ be a real for which $F(z') \geq z'$ and such that $F'(z') = f(z') < 1$ (which has to exist since $F(1/2) = 1/2$). Now $F(z') \geq z' > z' f(z')$ which is a contradiction. Using the symmetry of F, this implies that $F(y) \geq y$ for all $y \in [1/2, 1]$.

Now, let $y \in [1/2, 1]$ and observe that

$$\int_0^y F(x)dx = \int_0^{1-y} F(x)dx + \int_{1/2-(y-1/2)}^{1/2+(y-1/2)} F(x)dx.$$

Using that $F(x) \leq x$ for $0 \leq x \leq 1/2$ to bound the first term and the symmetry of F to evaluate the second, we can write:

$$\int_0^y F(x)dx \leq \int_0^{1-y} xdx + \frac{2y-1}{2} \leq \frac{(1-y)^2}{2} + \frac{2y-1}{2} \leq y^2/2.$$

Using again that $F(y) \geq y$ for all $y \in [1/2, 1]$ we conclude that $y^2/2 \leq yF(y)/2$, which completes the proof. □

It is straightforward to extend the previous proposition to the case when the support of the distributions F_a is an interval $[r, s]$ with $r \geq 0$ and still satisfy the conditions of the proposition. If this is the case the bound becomes $\mathcal{C}_I \leq \frac{2s}{r+s} \cdot \mathcal{C}_C$. The intuition behind this result is natural. For instance, if r is very close to s, the cost under incomplete information approaches that of a fully informed planner. Also, if $s = r + K$, for constant K, the bound also goes to one as r goes to infinity. This is because the amount of information the planner ignores is irrelevant when compared to the total cost of the project.

Furthermore, if the distributions F_a for all $a \in A_1$ are symmetrical and unimodal, then f_a is nondecreasing in the interval $[0, 1/2]$. This implies that $F_a(y) \leq y f_a(y)$ for all $y \in [0, 1/2]$. Thus we have the following corollary.

Corollary 1. *If F_a is SUD for all $a \in A_1$, then $\mathcal{C}_I \leq 2 \cdot \mathcal{C}_C$. Moreover, if F_a is SUD on $[r, s] \subset \mathbb{R}_+$ for all $a \in A_1$, then $\mathcal{C}_I \leq \frac{2s}{r+s} \cdot \mathcal{C}_C$.*

Finally, observe that unfortunately, one cannot expect to obtain a general bound for any class of distributions, and this is particularly bad in situations where most providers are "good". For some decreasing distributions, the bound becomes arbitrarily bad. Indeed, consider the case where the planner must send one unit of flow from an origin to a destination, in a two link network. One of the links is private information with cost distribution proportional to $f(c) = 1/(c + \varepsilon)$ in $[0, 1]$, while the other is public and its cost equals 1 (so $\Gamma = \{(x, y) \geq 0 : \; x + y = 1\}$). A simple calculation shows that

$$\mathcal{C}_I > 1/2 \quad \text{and} \quad \mathcal{C}_C = (\ln(1 + 1/\varepsilon))^{-1} - \varepsilon.$$

Thus, the ratio can be made arbitrarily large for small enough ε.

Furthermore, even for symmetric distributions the ratio can be arbitrarily large. To see this, consider a single good procurement auction with n sellers, i.e., $\Gamma = \{(x_1, \ldots, x_n) \geq 0 : \sum_{i=1}^n x_i = 1\}$, where each seller has a symmetric distribution putting half of the mass at or close to 0, and half at or close to 1. In this situation $\mathcal{C}_C$ is approximately $(1/2)^n$, while $\mathcal{C}_I = \mathcal{C}_{VCG}$ is roughly $(n + 1)/2^n$. The ratio grows to infinity with n.

4 Computation and Implementation

In general, implementing TVCG is no harder than solving $|A_1|$ times the original problem $\min\{c^T x : \; x \in \Gamma\}$, with the additional constraint that $x_a = 0$. In some situations this latter problem can be solved even more efficiently [6]. The situation is different for optimal mechanisms. In fact, to implement an optimal mechanism the planner must compute the assignment and the payments only for a specific cost realization. Note that this is simpler than computing the whole assignment and payment *rules*, which require the assignment and payments for every cost realization.

Given a cost realization c, Proposition 1, states that the assignment can be computed as $\min\{c'^T x : x \in \Gamma\}$ for some virtual nonnegative cost vector c'. This problem is the same as solving one instance of the complete information problem. However, to compute the payments $\bar{t}_a(c) = c_a \bar{x}_a(c) + \int_{c_a}^{\bar{c}_a} \bar{x}_a(t, c_{-a}) dt$ for a specific cost realization, in principle one needs to compute $\bar{x}_a(t, c_{-a})$ for all $t \in [c_a, \bar{c}_a]$. That is we need to solve $|A_1|$ parametric optimization problems of the form:

$$g_i(\theta) = \left(\arg\min_{x \in \Gamma} (c + \theta e_i)^T x \right)_i, \tag{8}$$

where $(\cdot)_i$ denotes the $i-$th component. The computational complexity of such a problem heavily depends on the structure of Γ and determines the complexity of computing the optimal mechanism under incomplete information. We now analyze three cases.

4.1 Case I: Parametric Optimization Is Easy

If the parametric optimization problem (8) can be solved in polynomial time, then the whole mechanism can be computed in polynomial time as well. This includes the case in which Γ is the set of all paths from a given source to a given sink, proved to be computationally easy in [4,6].

Observe that a wider class of problem where parametric optimization turns out to be efficient is when $\Gamma = P \subseteq [0,1]^{|A|}$, with P being an integral polytope. Of course a special case of this is $\Gamma = \{x : Ax = \mathbb{1}, x \geq 0\}$, with A totally unimodular. Shortest $s - t$ path is included in this class since it can be formulated imposing that the total flow across every $s - t$ cut equals one. Other problems in this class include minimum spanning tree and minimum perfect matching.

Proposition 5. *If $\Gamma = P \subseteq [0,1]^{|A|}$, with P an integral polytope, then* (8) *can be computed in polynomial time, by solving exactly two linear programming problems over P.*

Proof. From Lemma1 $g_i(\cdot)$ is non-increasing. Also, since Γ is a $\{0,1\}$ polytope we conclude that $g_i(\theta)$ is either 0 or 1, for all $\theta \geq 0$. Therefore, to fully determine $g_i(\theta)$ (and thus obtain the optimal mechanism), it suffices to compute $\theta^* = \max\{\theta : g_i(\theta) = 1\}$.

To this end we first compute

$$Z = \min\{(c + \theta e_i)^T x : x \in \Gamma, x_i = 0\} = \min\{c_{-i}^T x_{-i} : (0, x_{-i}) \in \Gamma\}.$$

Analogously we compute

$$\theta + Z' = \min\{(c + \theta e_i)^T x : x \in \Gamma, x_i = 1\} = \theta + \min\{c_{-i}^T x_{-i} : (1, x_{-i}) \in \Gamma\}.$$

Obtaining that $\theta^* = Z' - Z$. □

Remark that the previous proposition can easily be extended to the case in which Γ is an integral polytope in $[0, K]^{|A|}$ for fixed K, or even for K of polynomial size in the input. in this case we would need to solve K linear programming problems over Γ. Furthermore, even for general Γ but satisfying that the optimal solutions to $\min_{x \in \Gamma} c^T x$ lie in $\{0, \ldots, K\}^{|A|}$, the optimal mechanism can be obtained by solving $(K + 1)|A_1|$ such problems.

4.2 Case II: Optimization Is Easy but Parametric Optimization Is Hard

Even if optimizing over Γ is easy, the parametric optimization counterpart does not need to be so. For instance, for parametric linear programming, i.e., $\Gamma = \{Ax = b, x \geq 0\}$, the function $g_i(\theta)$ can attain exponentially (in $|A|$) many different values [8]. Additionally, even for more structured problems such as minimum cost flow, $g_i(\theta)$ can have a superpolynomial number of values [3].

However, since resource owners are risk-neutral, we can easily obtain a randomized mechanism that is truthful and gives in expectation the same value, therefore it is also optimal.

Indeed, for a given a cost realization c, the assignment $\bar{x}(c)$ is computed exactly as before (i.e., by solving $\min\{c'^T x : x \in \Gamma\}$), but the payments are computed using randomization. The payments to the owner of resource a, is given by $\bar{t}_a(c) = c_a \bar{x}_a(c) + (\bar{c}_a - c_a)\bar{x}_a(Y, c_{-a})$, where Y is a random variable uniformly distributed in $[c_a, \bar{c}_a]$. In expectation, which is all that matters to a risk-neutral resource owner, the latter payment equals

$$c_a \bar{x}_a(c) + (\bar{c}_a - c_a) \int_{c_a}^{\bar{c}_a} \bar{x}_a(t, c_{-a}) \cdot \frac{1}{\bar{c}_a - c_a} dt = c_a \bar{x}_a(c) + \int_{c_a}^{\bar{c}_a} \bar{x}_a(t, c_{-a}) dt,$$

and thus the mechanism is truthful and optimal. We conclude the following result.

Lemma 3. *An optimal and truthful mechanism can be implemented by solving, for each* $a \in A_1$*, two problems of the form* $\min\{c^T x : x \in \Gamma\}$.

Naturally the mechanism just described can be implemented in polynomial time so long as the optimization problem over Γ can be solved in polynomial time. This enables us to implement a desirable mechanism even if the parametric optimization 8 is hard. However, this mechanism introduces high risk for the resource owners. To avoid this issue we could simply take a larger number N of uniform samples Y_i and compute $\bar{t}_a(c) = c_a \bar{x}_a(c) + (\bar{c}_a - c_a) \sum_{i=1}^{N} \bar{x}_a(Y_i, c_{-a})/N$. With this the dispersion of payments will be reduced, though the computational effort will increase with N, leading to a tradeoff between risk and computational efficiency.

4.3 Case III: Optimization is Hard

We now study what happens when optimizing over Γ is NP-hard, which is the case for a large number of combinatorial problems [11]. As one may expect, computing an optimal mechanism in such a case is also hard, so we can turn to search for truthful mechanism that are approximately optimal.

Suppose that we have an algorithm ALG for solving $\min_{x \in \Gamma} c^T x$, with an approximation guarantee of β. That is an algorithm returning a solution whose cost is at most β times the optimal cost. Suppose furthermore that ALG is monotone, that is the returned solution $x_a^{\text{ALG}}(c_a, c_{-a})$ is decreasing in c_a. Then the mechanism that for each cost realization c assigns according to $x^{\text{ALG}}(c)$ is truthful. Moreover the expected cost for the planner of this mechanism is at most $\beta \cdot \mathcal{C}_I$.

References

1. Aggarwal, G., Goel, G., Mehta, A.: Efficiency of (revenue-)optimal mechanisms. In: Proceedings of the 10th ACM Conference on Electronic Commerce (EC), pp. 235–242 (2009)
2. Bulow, J., Klemperer, P.: Auctions Versus Negotiations. American Economic Review 86, 180–194 (1996)
3. Carstensen, P.: Complexity of some parametric integer and network programming problems. Mathematical Programming 26(1), 64–75 (1983)
4. Elkind, E., Sahai, A., Steiglitz, K.: Frugality in path auctions. In: Proceeding of the 15th ACM-SIAM Symposium on Discrete Algorithms (SODA), pp. 701–709 (2004)

5. Hartline, J., Roughgarden, T.: Optimal Mechanism Design and Money Burning. In: Proceedings of the 40th Annual ACM Symposium on Theory of Computing, pp. 75–84 (2008)
6. Hershberger, J., Suri, S.: Vickrey prices and shortest paths: What is an edge worth? In: Proceeding of the 42nd IEEE Symposium on Foundations of Computer Science (FOCS), pp. 252–259 (2001)
7. Kolmogorov, A.N., Fomin, S.V.: Introductory Real Analysis. Dover Publications, New York (1975)
8. Murty, K.: Computational complexity of parametric linear programming. Mathematical Programming 19, 213–219 (1980)
9. Myerson, R.: Optimal auction design. Mathematics of Operations Research 6(1), 58–73 (1981)
10. Nissan, N., Roughgarden, T., Tardos, E., Vazirani, V.: Algorithmic Game Theory. Cambridge University Press, New York (2007)
11. Schrijver, A.: Combinatorial Optimization: Polyhedra and Efficiency. Springer, Berlin (2003)